

**The Application of Expert System and Agent Technology to Water  
Mains Rehabilitation Decision Making**

**Dr Darryl Davis\***

**Dr Bernadette Sharp**

**Artificial Intelligence Group**

**School of Computing,**

**Staffordshire University,**

**Stafford, ST18 0AD**

**e-mail [D.N.Davis@dcs.hull.ac.uk](mailto:D.N.Davis@dcs.hull.ac.uk)**

**\* Now: Department of Computer Science, University of Hull**

## **Abstract**

This paper discusses the challenges encountered in a recent project to develop an expert system to assist water engineers in planning the rehabilitation of water mains. The paper describes the adoption of an emerging agent-based technology in designing the prototype, and highlights the many influential factors and constraints in building the prototype. Many of the challenges faced in developing the system may be typical of the hurdles faced in other commercial domains where intelligent (agent and expert) systems offer solutions to emerging Information Technology problems. These constraints may be particularly acute where the company has no prior experience in applying AI-based technology to its business, or where the domain is new to the AI community.

## **1. Introduction**

This paper presents our work in designing and developing an Agent-based Decision Support Tool (ADST) of use to a U.K. water company in their water mains rehabilitation decision making, and provides a number of perspectives on the project. This work had a number of aims; chiefly to develop a prototype expert system that would model the current decision making processes adopted by the water company and their associates; and to do so within a framework that would allow further decision support capabilities to be added (in future collaborations).

Water mains rehabilitation is a major challenge for (domestic) water supply companies, with a large annual and protracted budget, as to be expected for any company that services a fundamental resource for a number of counties in the U.K. mainland. There are a number of important constraints that impinge on these financial and resource decisions, for example regulatory requirements which are monitored by government established watch-bodies and factors related to customers and users, that affect which areas (i.e. which pipes) are to be maintained or replaced in any one year. Furthermore the decisions made in the

current and previous years affect the decisions to be made about future water supply infrastructure maintenance. This application area is therefore very important and is in need of systems that facilitate more consistent and effective strategic decisions.

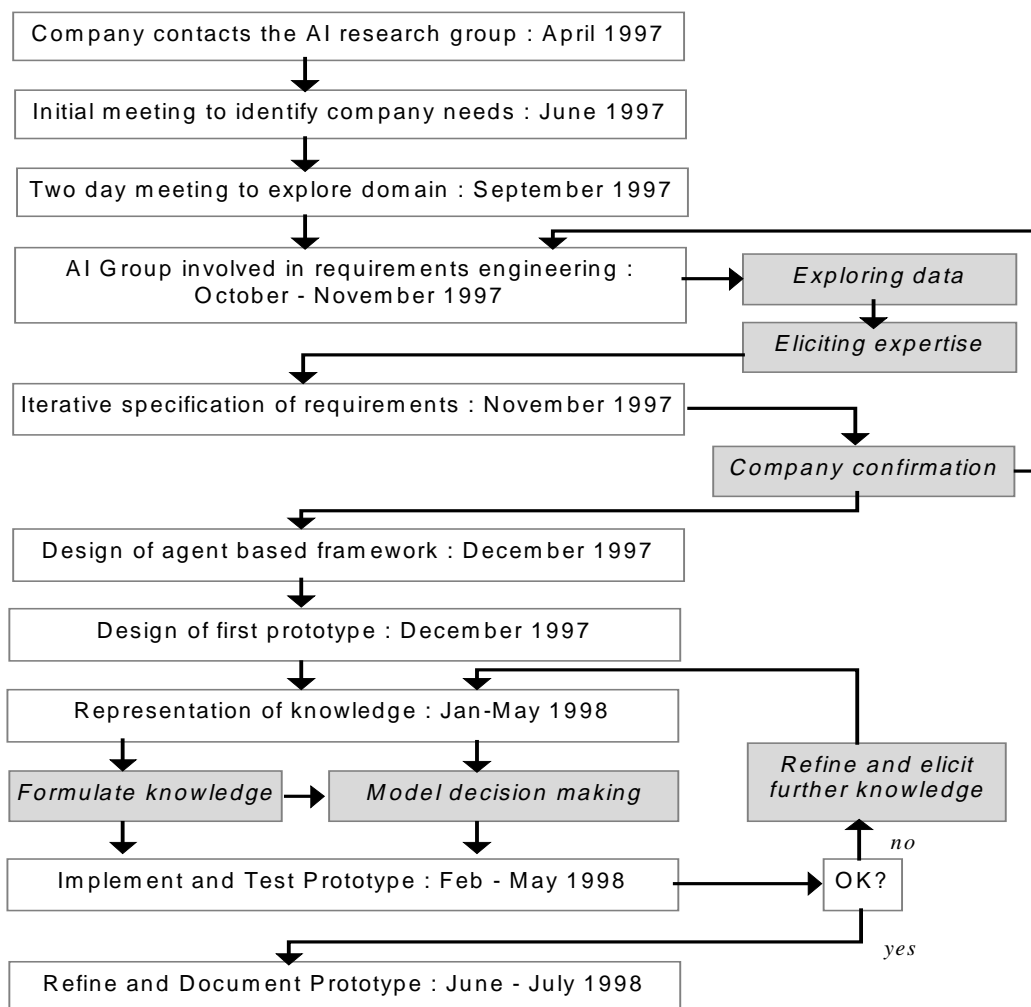
The implemented system models part of the current human strategy used in making rehabilitation decisions; i.e. it embodies some of the expertise of the water company engineers, their consultants and their information handling models. This information handling expertise has been endowed with a degree of extra flexibility that allows the manipulation of the ranking process used in the human model. This enables the production, from any given set of data, of multiple lists, each giving alternative rankings of the potential rehabilitation areas, and so allows the impact of the current rehabilitation strategy and pertinent 'what -if' scenarios (e.g. alternative rehabilitation strategies) to be explored. The proposed full system will extend this further, with additional heuristics, statistical models and other decision support capabilities. In producing this first system, a number of challenges were encountered which may be indicative of the challenges other researchers will have to meet in developing intelligent decision support tools, especially if these systems are to work within an existing IT infrastructure.

## **2. Requirements Engineering and Knowledge Engineering**

The domain encountered seemed, at first, to be a classic scenario for the application of expert systems but due to the complexity of the application domain, an agent based approach to knowledge engineering was identified to be a more fruitful avenue to pursue. The water company required a flexible system to assist water supply experts in analyzing the impact of implementing any changes to the current water mains rehabilitation strategy. The proposed system would need to access a number of existing and new databases held on PCs, plus other potential sources of information, for example a Geographic Information System (GIS) database. In the short term the framework must be able to support a number of functions each of which could be modelled individually using relatively uncomplicated expert systems. The following sections detail some of the methodological caveats being developed by the authors, who are (in

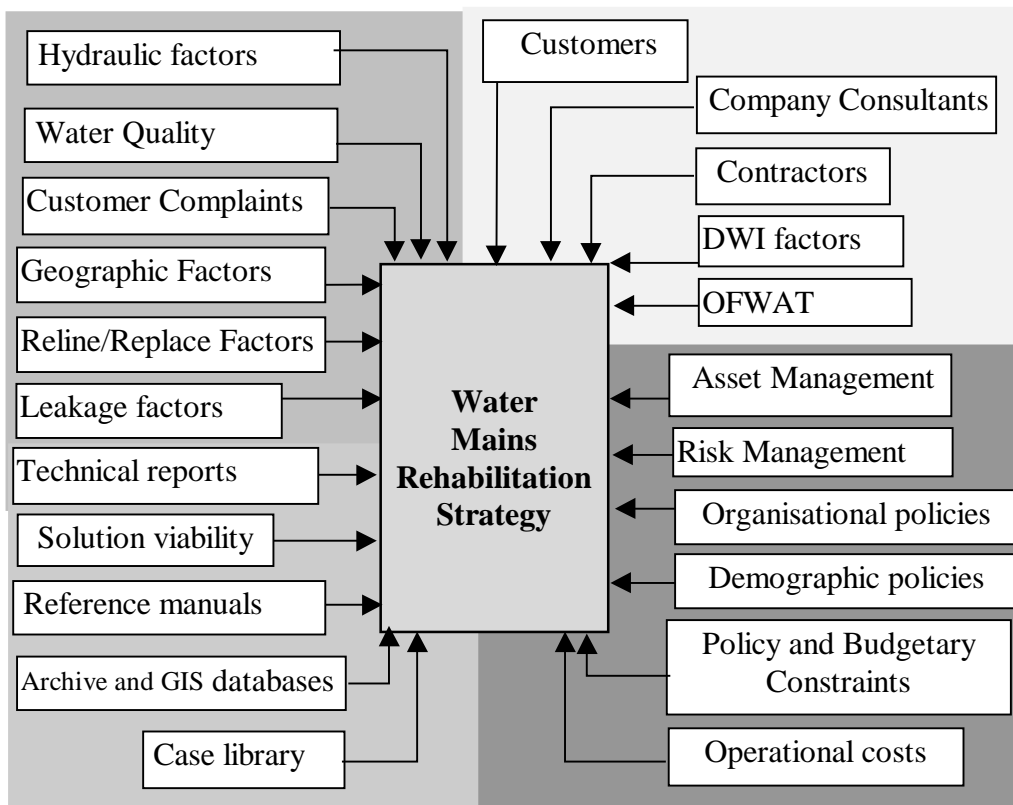
conjunction with other researchers) developing agent technology for a wide range of application domains.

In building the system a number of challenges were encountered. The design of this system required a broad remit, as when our client approached us they had no clear expectations of the benefits of applying expert system technology, other than that it might be of help. They were also very cautious in considering emerging technologies such as agents and data mining. They had heard of our group's collaborations with other water companies (1) and were eager to explore AI technologies; this was at a time when outside (regulatory) pressures required them to be proactive in their planning and future rehabilitation strategies. Early meetings focused on identifying the benefits of developing an expert decision support system and making accessible the human expertise required for the project.



*Figure 1 Overview of phase one of the project*

The next challenge related to the establishment of the system requirements. To achieve this a series of meetings were held, and over a period of six months (within an overall project timetable of fifteen months), a clearer picture of the project objectives was gained. Figure 1 summarizes the process of building an initial decision support system, but fails to convey the amount of communication (in terms of documents, e-mails, meetings and telephone calls) which were necessary to build the prototype. Early attempts at the elicitation of expert knowledge identified an element of fuzziness in the reasoning of the company's experts. It seemed that many cases related to rehabilitation were resolved intuitively because data was either unreliable or incomplete or completely missing. Compliance pressures external to the company were constantly changing, necessitating some existing procedures to be overridden, thus adding another dimension of complexity to the modelling of the decision process.



*Figure 2. Influences to be captured in a complete decision support too.*

A further challenge was introduced through the requirement to use third party software. We found that third-party developed databases were not compatible with standard data communication protocols (e.g. ODBC) requiring us to become unexpectedly involved in data cleaning; one of the more time-consuming processes involved in data mining (2).

Following the many meetings and exchange of documents, a detailed specification was produced identifying the types of data, knowledge, internal and external factors likely to be of relevance in a full decision support tool; figure 2 provides a high level summary of these factors. The four headings used to group the identified factors and constraints are:

- **Engineering factors**, which relate to company's policy on issues such as quality of supplied water, response to customer complaints, water mains leakage, decisions to reline or replace sub-standard water mains, and geographic factors within the area supplied by the water company.
- **Technical factors** such as the viability of a possible water mains rehabilitation plan, the case library related to prior and existing rehabilitation decisions, archive databases which may be of use in determining the effectiveness of past solutions, GIS databases, reference libraries, and databases related to water mains supply and technical reports, for example data from pipe inspections.
- **Organisational policies** involving such factors as risk management, asset management, demographic policies, operational costs, organisational policies related to the rehabilitation criteria, and policy and budgetary constraints.
- **External factors** affecting the company's rehabilitation strategy such as overriding policy statements by the Department of Water Industries, OFWAT (the regulatory body for the water industry), customers, the company's consultants and contractors responsible for carrying out any rehabilitation.

The computational modelling of all these aspects would require several types of functionality and knowledge handling beyond an orthodox (single or multiple knowledge base) expert system, and more

appropriate to the sophistication associated with blackboard systems. Furthermore as an incremental approach to developing a full system was required and the initial implementation did not require an orthodox rule-based system, we thought it wise to consider alternative technologies.

### **3. Expert Systems and Agent Technology**

Agent technology is being used across many different domains, from information retrieval, to image analysis (3) to cognitive science (4,5). Agents offer the flexibility to integrate many different categories of processing within a single system in much the same way as blackboard systems but need not necessarily require a full specification of intended final system at the initial stages. Furthermore the adoption of an appropriate design methodology gives the application builder the opportunity to design a very flexible and incremental system, that should be responsive to the changing and developing requirements and needs of the end user. There are a plethora of agent definitions (6,7,8). These range from descriptions based on a functional analysis of how agents are used in technology to far more ranging expositions based on different interpretations of the role and objectives of artificial intelligence and cognitive science. Artificial Intelligence is a very diverse field and agents can be used as metaphors for work in many areas. This can lead to confusion and devalue the term agent, as other authors have noted (9). This section clarifies the framework of the agent research being highlighted here.

The use of the term agent (in AI and related fields) can be traced back over the last thirty years. A seemingly endless list of agent attributes are possible reflecting these different definitions and the intentions of their creators, and include factors such as intentionality, autonomy reactivity, flexibility, communication, learning, self-actuation etc. Two different but plausible agent definitions (5) are:

**Definition 1:** *An agent is an integrated computational entity with intentionality and some degree of autonomy.*

**Definition 2:** *An agent is a synthetic entity that enables us to study, at a computational, design or theoretical level, what a mind could and can be.*

The first definition equates to the idea of *weak agents*, i.e. agents as (intelligent) information processing systems and is quite open to extension and interpretation. The second definition equates to the notion of *strong agents*, i.e. agents as computational cognitive models that explain and/or simulate, to some degree, reported findings and theories in cognitive psychology, or some other study of minds (or life). While the work, reported here, is very much application driven, and therefore relates to the first definition, the second more encompassing definition has a bearing if we consider an expert system to be an artificial system capable of reasoning about a category of problems in a way that is analogous to the way a human expert resolves similar problems.

A particular class of agent system that is of relevance here are those described as tightly coupled (10). These can be seen as building on the ideas expounded by Medsker (11) but firmly rooted in agent technology. Such systems can be viewed as a development of the ideas embodied in blackboard systems (12), where multiple agencies, each with a set of designated tasks contribute in the formation of solutions to complex tasks requiring many kinds of representation and processing categories.

#### **4. Design of an Agent-Based Decision Support Tool**

Figure 3 shows a schematic of a collaborative society of agent based processes thought suitable for managing the tasks in a decision support system, and enabling an incremental progress from a partial (but functioning) system to a full implementation. Within this overall agent exist a number of task specific agents. For example, the *Heuristic Agent* reasons over current task information using knowledge that represents statistical models or other form of expert advice. Information relevant to the system's current task is held by the *Information Agent* which exchanges information with all the other agents, requesting further information from or depositing new information with the *Database Agent* as required. The *Strategy Agent* performs specific ranking processes on information about water supply zones within the company's regions. While this agent may make use of one of a number of default decision strategies, a custom strategy can be adopted based upon a design arising from the interactions of the user with the

decision support tool via the *Interface Agent*. The *Constraints Agent* reasons about the viability of any tentative rehabilitation decision on the basis of policy type factors as listed above (and shown in figure 2). The influence of cognitive science theories is evident in the grouping of agent-based processes into the three broad categories of: sensing and acting (the *Interface Agent*); task-related cognitive memory (e.g. the *Heuristic Agent*); and Long Term Memory Management, i.e. the agent(s) responsible for managing the database systems used by the decision support tool.

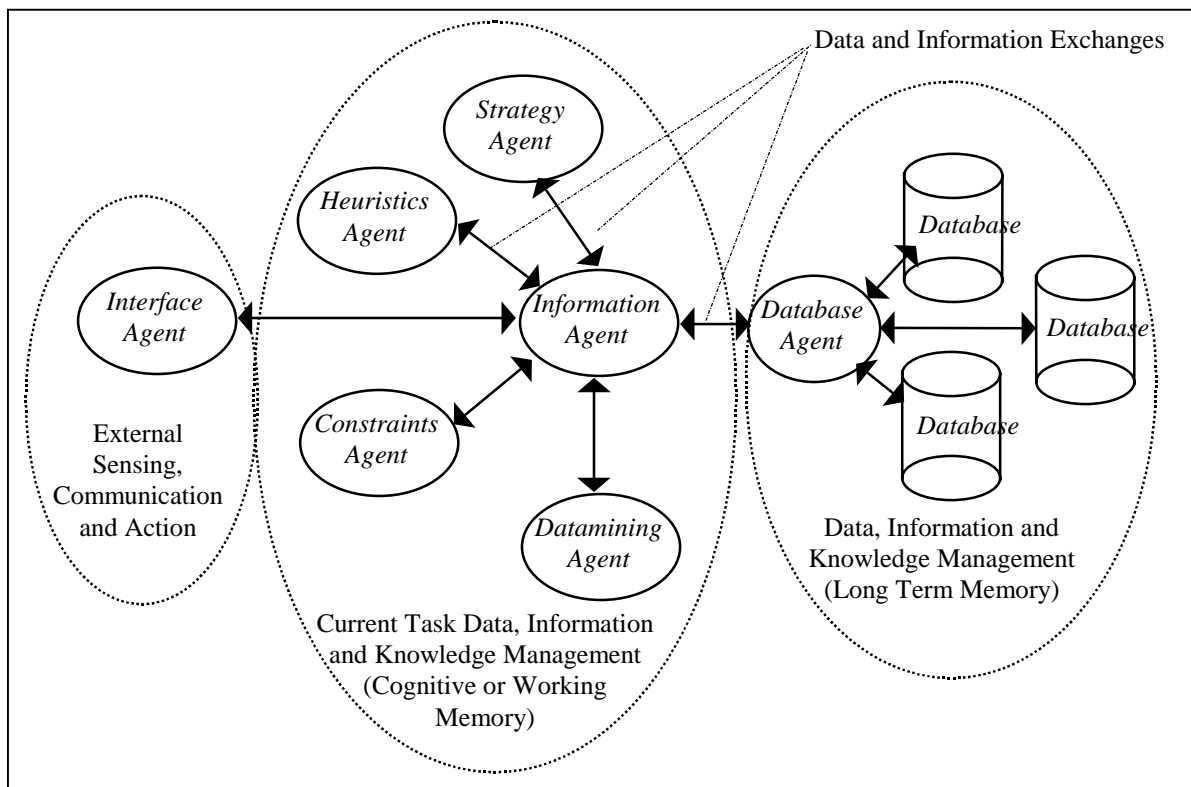


Figure 3. Schematic of proposed tightly coupled agent-based decision support tool.

The nature of the collaboration was such that we could not try and build the full system in the short time initially allocated to the project. However, pressures external to the water company meant that they required some form of decision support tool within a short time span. Furthermore, the development of the full agent-based tool would also require simultaneously solving many research problems, and in particular some very thorny ones related to agent communication and the nature of control processes in such an ambitious business application. For these reasons, phase one of the project tackled a well

established sub-problem (but currently paper-based within the company), albeit in such a way that the developed system would be compatible with the more global aims of the extended project.

Figure 4 shows a schematic, of the implemented (expert) decision support tool, that allows for an extendible functionality and future collaborations. The major components of the first prototype system, which mirrors aspects of the system sketched in figure 3, are:

- **Access Data Manager:** A set of functions and methods that allow data stored in a Microsoft Access database to be transferred into the system and stored as (water supply) zone summary information, using object oriented techniques. This includes constraint handling processes that are used, for example, to check data (for appropriate value and type) coming from the Microsoft Access Database. This is a precursor to the *Database Agent* in figure 3.

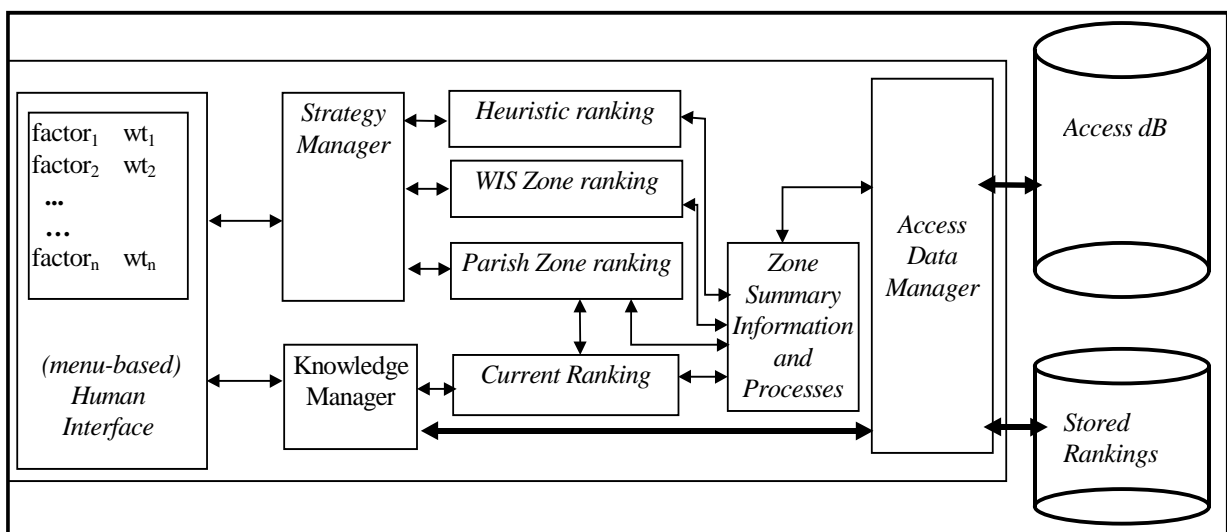


Figure 4. Schematic overview of the Kappa-PC Decision Support Tool

- **Zone Summary Information:** A set of knowledge handling processes, modelled using object classes and methods, that allow water supply (initially Parishes and Water InterZone Supply (WIS) areas) zone information to be collated as attributes of appropriate object class instances.
- **Current Ranking:** A set of object classes and instances that reflect the current ranking.

- **Knowledge Manager:** A set of knowledge handling methods and functions which manipulate the current results; for example, to select the top 90% or first 50 parishes from a water zone ranking. This, together with the Zone Summary Information and the Current Ranking, is a precursor to the *Information Agent*.
- **Strategy Manager:** A menu driven module that allows the selection of existing ranking processes, the creation of (customised) ranking processes, and the management of tasks that together produce rankings based on the water supply zone data and knowledge held in the Decision Support Tool. This module makes use of the *Parish* and *WIS Zone Ranking* processes (again modelled using object technology) that calculate an expertise-based zone ranking of use to the rest of the system. The Strategy Manager, together with more flexible and sophisticated decision-making processes not present in the first system, will form the *Strategy Agent* as described in the text supporting figure 3.
- **Human Interface:** A set of windows containing menus, buttons etc. and information displays that allow the user to navigate within the system to achieve their required objectives. This is a precursor to the *Interface Agent* in Figure 3.

This first system models only a small aspect of the full rehabilitation strategy hinted out in figure 2; specifically it does not include any hot spot analysis, operational factors, financial factors or logistical factors. Table 1 lists the type of factors involved in a water supply zone ranking process, with a nominal set of factor weights. These factors are used in the following modelled process:

- **Step 1.** Select the compliance values prior to database access, if non-default values are required.
- **Step 2.** Build parish zone summary information on the basis of archived material. Dependent upon the database tables accessed this may or may not result in all parishes being represented.
- **Step 3.** Select or create ranking process on the basis of a standard or customised set of factors
- **Step 4.** Perform a ranking over all parishes for which the system has summary information.

- **Step 5.** Select the top  $N$  parishes or top percentages from the parish ranking.
- **Step 6.** Investigate the ranking, and manipulate as required.
- **Step 7.** Perform summary WIS zone ranking if required.

Contributing Factor	Factor Name	Factor Weight
Total Iron	Water Quality	27%
Total Manganese	Water Quality	9%
Total Aluminium	Water Quality	9%
Turbidity	Water Quality	18%
Discolouration	Complaints	9%
Taste and Odour	Complaints	2%
Interruptions	Supply Interruptions	5%
Low Pressure	Pressure	9%
Supply Mains Flushing	Regular Flushing	4%
Supply Mains Filters	Installed Filters	4%
Dosing	Chemical dosing	4%

*Table 1. Typical set of factors involved in a water supply zone ranking process*

The first four steps of this strategy are required to perform the last steps which are optional and can be carried out in any particular order. Once steps 1 and 2 have been performed, any number of different rankings (as created by step 3) can be carried out. Figure 5 shows part of the results from a typical ranking; a full zone report for all zones runs to 250 pages.

The developed tool incorporates a number of other processes which enable the user to investigate and run further analyses on the loaded summary data. For example, to select the top ranked water supply zones or to display information associated with any zone, as well as producing paper-based reports. It also includes a configuration facility that allows the user to create custom decision making processes and store them within the system.

The first phase project produced five tangible deliverables for the collaborating water company:

- An ODBC compatible version of the database supplied to the AI team, with a number of data value and type constraints enforced to remedy oversights of the third party developed database.
- A bare bones (data-less) expert decision support system as described in the text related to figures 3 and 4. This system can be used to investigate water supply zone rankings with any suitable ODBC compatible database.

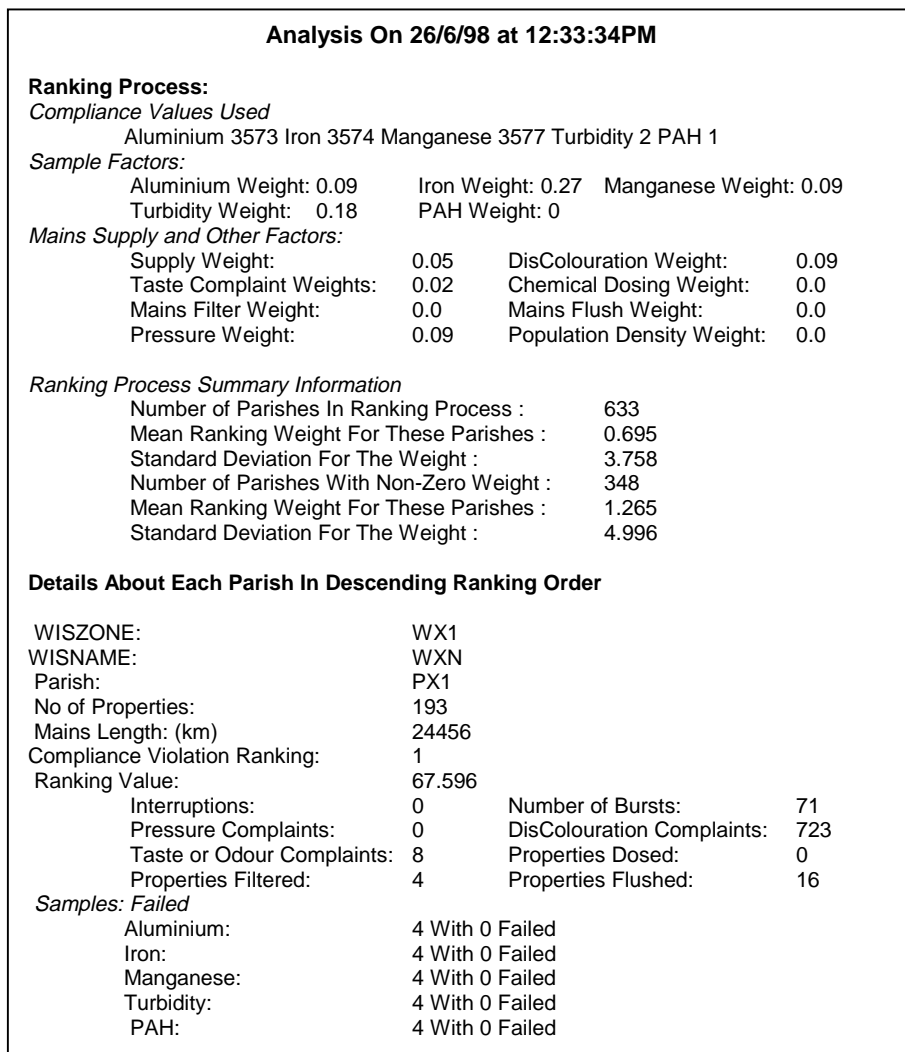


Figure 5. Part of the results from a typical Parish zone ranking process (names made anonymous).

- A data loaded decision support system with summary information from accessing the ODBC compatible database. This system reduces a 16MByte database to a 500Kbyte summary.

- A report which describes all aspects of the project and provides explanations on how to make further ODBC compatible databases and how to use the facilities included in the system.
- Reports detailing the water supply zone rehabilitation rankings from using the system on the database supplied as deliverable one.

## 5. Challenges Encountered in Producing Deliverables

Here some of the challenges, that will need to be met when developing AI based applications for users new to the technology, are highlighted. These challenges belong to three broad categories: orientation challenges; information technology infrastructure problems; and organizational shifts.

- **Orientation Challenges.** It was relatively late on, in the fifteen-month project, that our collaborators began to understand the capabilities of modern AI systems. Similarly they initially held naive expectations about how easy the required capabilities were to incorporate and were often quite surprised at the level of detail required to produce the system specifications. Furthermore, as a recently privatised company, they had been used to dealing with associates who were at one time associated with the former nationalised company. This meant that as an external agency we were at times quite unaware of certain organisational issues, and company policies, which they took to be understood. This unintentional reluctance to inform us of specific developments, particularly in regard to the developing IT infrastructure and proposed changes to the water supply zone demarcation criteria (which occurred during the project), meant that unexpected delays to progress became inevitable. It can be expected that similar challenges lie ahead in other domains where expert and agent-based technologies seem appropriate solutions for emerging IT problems, for example, recently privatised or completely re-organised companies with similar changes in orientation.
- **IT Infrastructure Problems.** These are the unexpected challenges of working with data and information that does not meet expected standards. Before it is usable, the systems holding data and

information must be brought to required and necessary standards. For example, as a result of our initial investigations, we came to understand that the databases that our system now works with, were inconsistent, unreliable and would not meet the required data transfer protocol standard. This required some considerable effort to be spent on addressing these unforeseen problems, many of which were traced to corrupt and missing data entries in the database. It is a mistake to make assumptions about unseen data, and the lack of appropriate (ISO or other standards) certification for software will ensure that similar projects will suffer quite severe delays. For our relatively small (initial) project this extended by one third the project time; it is however difficult to generalise from this.

- **Organisational Shifts.** Due to pressure from external (regulatory) agencies and internal reorganisation, our collaborator has to deal with these requirements as well as managing their own rehabilitation policy. As a result the timetable and contract, following initial decisions, had to be renegotiated on many occasions, deliverables redefined and project targets re-prioritised. This certainly impacted on the knowledge elicitation phase, which is, by definition, is a very complex process. The incorporation of new factors and strategies into a project may not only conflict with previously elicited knowledge but also necessitate a different approach to data design, representation and processing. This shifting goal post problem will be particularly troublesome where the intended collaboration is long-term and the domain of the collaborators dynamic and prone to change.

## 6. Future Work

Proposals for further work in collaboration with the water company are ongoing, but it may be useful to highlight some possibilities here. These possibilities are non-exhaustive and open to negotiation, change and prioritization. The remaining factors associated with figure 2 will have to be incorporated if the system is to be of help in more sophisticated decision making scenarios. We would expect to develop the *Heuristic Agent* to make use of structural models (of water supply pipes) based on data analyses performed by colleagues who have been independently investigating the quality and suitability of further

data. Previous studies (1) have determined that these factors are significant in assessing the life expectancy of the pipe and hence influential in rehabilitation decision making. Such developments would allow a more flexible and informative decision support tool to be constructed.

The development of the *DataMining Agent* will add further data reclamation facilities, as well as data mining capabilities. For example we are developing neural nets that will search databases for factors consistent with past and current decision making. This type of activity may highlight areas of concern that a more orthodox decision support tool would miss; for example, such an agent could look for commonality or differences in the rankings process. In other words given a set of rehabilitation cases the system could analyze the factors influencing the ranking and objectively infer those significant factors which have influenced the decision making process. The cases and their significant factors could then be stored in a comprehensive repository and retrieved and analyzed to recommend a plan of action when new cases are being considered. This would allow the development of a more consistent approach to water mains rehabilitation and investment strategies. A different kind of *DataMining Agent* could search data related to water supply zones or areas for danger areas or hot spots. It could apply current strategy and practices to available databases and select potential zones or areas identified as hot spots given specific requirements related to OFWAT, customer complaints, bursts, and/or leakage data for further consideration. This would allow the water supply company to adopt a more pro-active, strategic and defensible approach to future rehabilitation investments and priorities.

## **Bibliography**

1. Sharp, B., Edwards, B., Dean, A., Davis, D. and Bancroft, G., A knowledge base environment to support water mains rehabilitation decision making, *ERA Technology Conference*, London, October 1998.
2. Fayyad, U.M., Piatetsky-Shapiro, G., Smith, P. and Uthurusamy, R. (Editors), *Advances in Knowledge Discovery and Data Mining*, MIT Press, 1996.
3. Linying, S., Davis, D.N. and Sharp, B. Neural Network for X-ray Image Segmentation, *ICEIS'99, 1st*

*International Conference on Enterprise Information Systems*, Portugal, March 1999.

4. Franklin, S., IDA: A Cognitive Agent Architecture, *Frontiers of Cognitive Agents, IEEE Symposium on Systems, Man and Cybernetics*, San Diego, October 1998.
5. Davis, D.N. Synthetic Agents: Synthetic Minds? *Frontiers of Cognitive Agents, IEEE Symposium on Systems, Man and Cybernetics*, San Diego, October 1998.
6. Wooldridge, M. and Jennings, N.R. (Editors), *Intelligent Agents*. Springer-Verlag, 1994.
7. Maes, P., Agents that reduce work and information overload. *Communications of the ACM* 37(7):31-40, 1994.
8. Sloman. A., What sort of architecture is required for a human-like agent? *Cognitive Modelling Workshop, American Association of Artificial Intelligence Conference 1996*, Oregon, 1996.
9. Franklin, S.P. and Graesser, A.G., Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents, In: *Intelligent Agents III*, J.P. Muller, M.J. Wooldridge and N.R. Jennings (Editors.), Springer-Verlag, Heidelberg, 1996.
10. Wermter, S. Hybrid Neural and Symbolic Language Processing, University of Sunderland Research Report, (WWW Reference, <http://www.tzi.org/ik98/prog/kursunterlagen/s1/wermter.html>), 1998.
11. Medsker, L.R. *Hybrid Intelligent Systems*. Kluwer Academic Publishers, Boston, 1995.
12. Englemore, R. and Morgan, T., *Blackboard Systems*, Addison-Wesley, 1988.

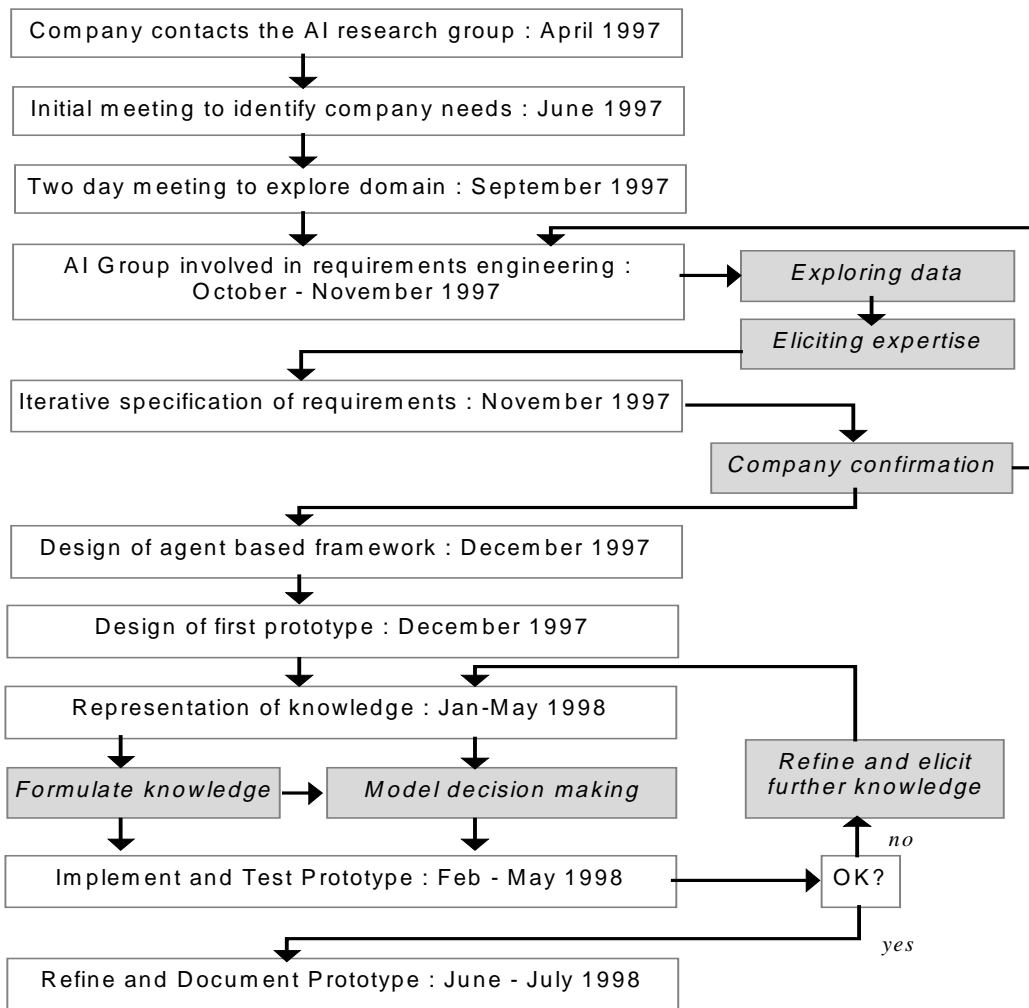


Figure 1 Overview of phase one of the project

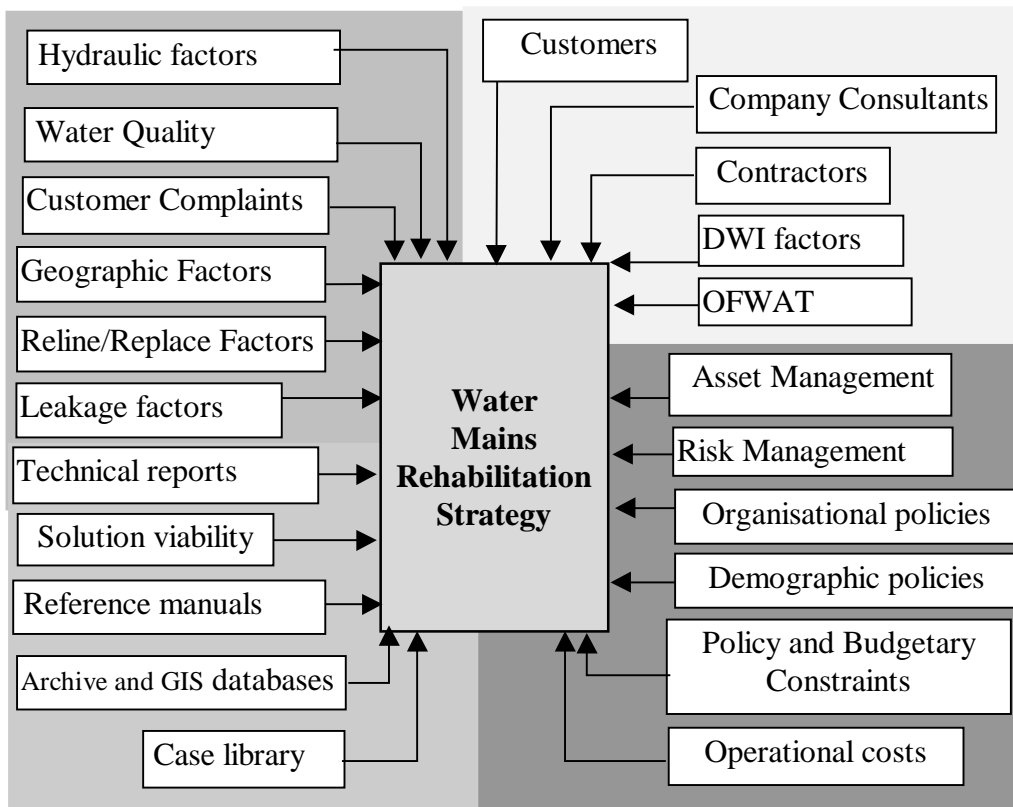


Figure 2. Influences to be captured in a complete decision support too.

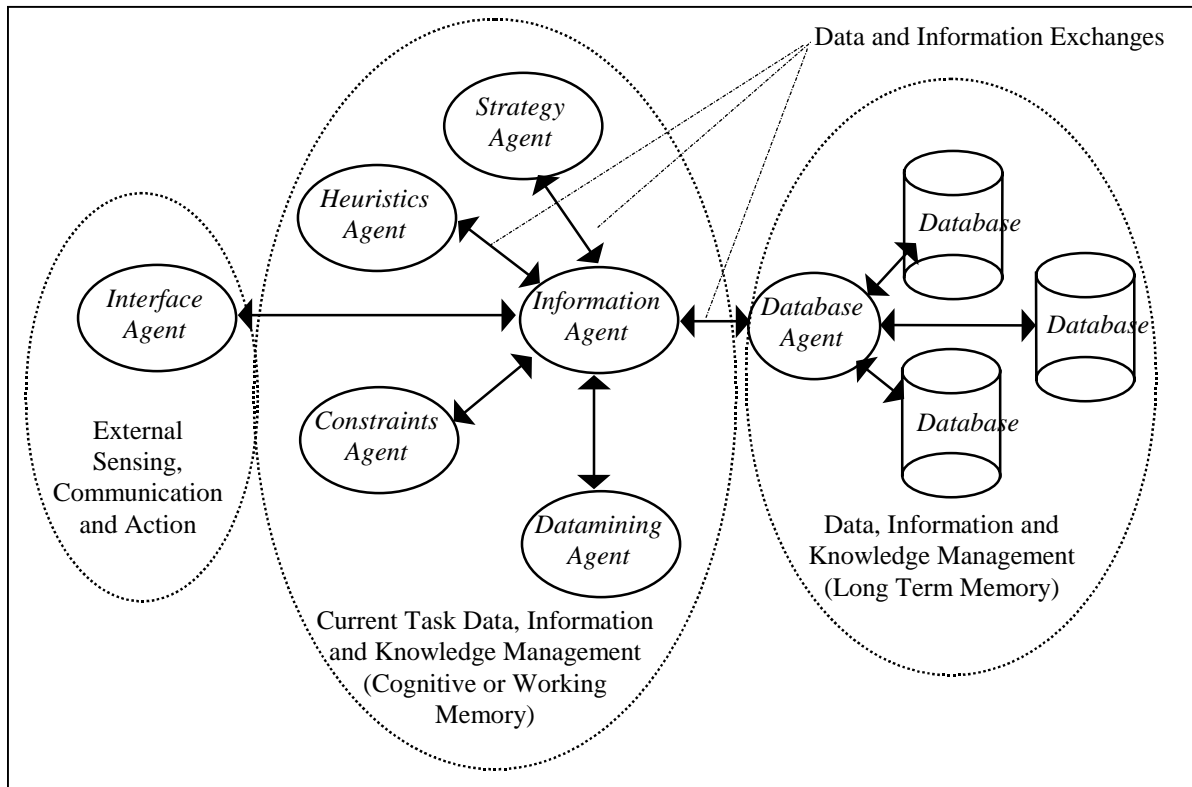


Figure 3. Schematic of proposed tightly coupled agent-based decision support tool.

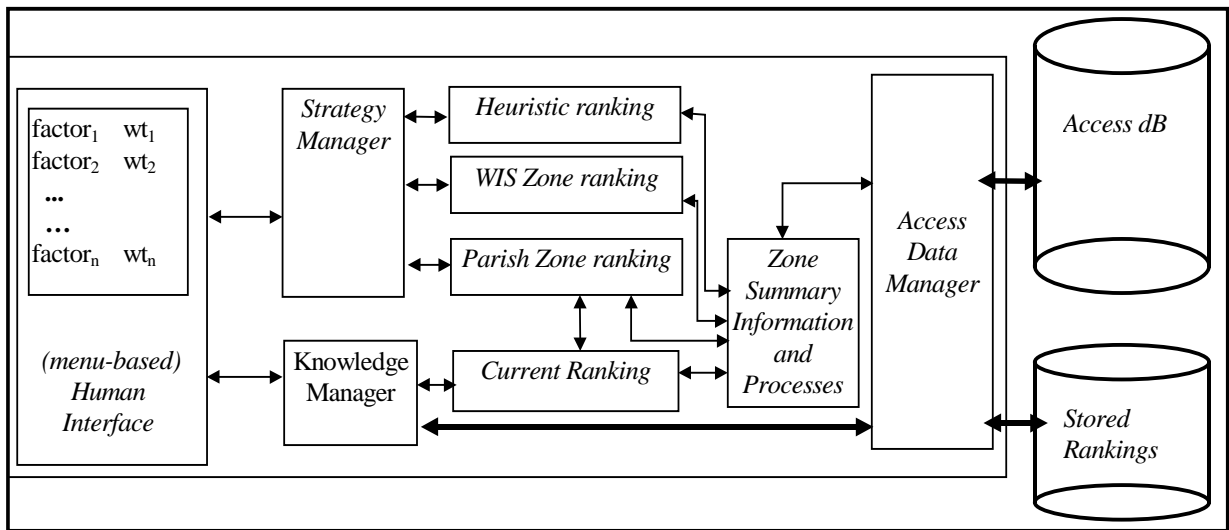


Figure 4. Schematic overview of the Kappa-PC Decision Support Tool

### Analysis On 26/6/98 at 12:33:34PM

#### Ranking Process:

##### Compliance Values Used

Aluminium 3573 Iron 3574 Manganese 3577 Turbidity 2 PAH 1

##### Sample Factors:

Aluminium Weight: 0.09      Iron Weight: 0.27      Manganese Weight: 0.09  
Turbidity Weight: 0.18      PAH Weight: 0

##### Mains Supply and Other Factors:

Supply Weight: 0.05      Discolouration Weight: 0.09  
Taste Complaint Weights: 0.02      Chemical Dosing Weight: 0.0  
Mains Filter Weight: 0.0      Mains Flush Weight: 0.0  
Pressure Weight: 0.09      Population Density Weight: 0.0

##### Ranking Process Summary Information

Number of Parishes In Ranking Process : 633  
Mean Ranking Weight For These Parishes : 0.695  
Standard Deviation For The Weight : 3.758  
Number of Parishes With Non-Zero Weight : 348  
Mean Ranking Weight For These Parishes : 1.265  
Standard Deviation For The Weight : 4.996

#### Details About Each Parish In Descending Ranking Order

WISZONE: WX1  
WISNAME: WXN  
Parish: PX1  
No of Properties: 193  
Mains Length: (km) 24456  
Compliance Violation Ranking: 1  
Ranking Value: 67.596  
Interruptions: 0      Number of Bursts: 71  
Pressure Complaints: 0      Discolouration Complaints: 723  
Taste or Odour Complaints: 8      Properties Dosed: 0  
Properties Filtered: 4      Properties Flushed: 16

##### Samples: Failed

Aluminium: 4 With 0 Failed  
Iron: 4 With 0 Failed  
Manganese: 4 With 0 Failed  
Turbidity: 4 With 0 Failed  
PAH: 4 With 0 Failed

Figure 5. Part of the results from a typical Parish zone ranking process.

Contributing Factor	Factor Name	Factor Weight
Total Iron	Water Quality	27%
Total Manganese	Water Quality	9%
Total Aluminium	Water Quality	9%
Turbidity	Water Quality	18%
Discolouration	Complaints	9%
Taste and Odour	Complaints	2%
Interruptions	Supply Interruptions	5%
Low Pressure	Pressure	9%
Supply Mans Flushing	Regular Flushing	4%
Supply Mains Filters	Installed Filters	4%
Dosing	Chemical dosing	4%

*Table 1. Typical set of factors involved in a water supply zone ranking process*