Multi-criteria decision support system and Data Warehouse for designing and monitoring sustainable industrial strategies an Italian case study

C. Di Mauro, J.P. Nordvik, and A.C. Lucia

European Commission - Directorate General JRC - Joint Research Centre Ispra
Institute for the Protection and Security of the Citizen (IPSC)
Technological and Economic Risk Management Unit
I-21020 Ispra (VA), Italy

Abstract: Decision-making processes for managing and defining policies related to social and environmental impacts from industry are becoming more and more complex. Raising public concerns on environmental issues requires public authorities to collect more accurate information related to impacts of industry. Overlaps between environmental, social and political interests, and rapid speed of change contribute even further to the complexity. Unfortunately, this multifaceted information is spread throughout multiple repositories. Thus, there is a need to integrate the available information within a structured network and to design a decision support system suitable for helping decision makers to define sustainable policies. This paper reports the experience gained in this field during the implementation of the SIMAGE project. The project is financed by the Italian Ministry for Environment and aims to improve the risk management of industrial areas in Italy. In particular, it describes how an architecture based on a data warehouse system can support the efficient implementation of the multi-criteria decision support system. The basic requirements of the prototype have further been evaluated by means of a forum consisting of major representatives of public bodies, acting as a proxy for the policies-makers.

Keywords: Environmental management; sustainability; decision support tools; industrial risk; data warehouse.

1 INTRODUCTION

The development of industrialised countries during the last century has resulted in, among other aspects, a number of risks and negative effects on the natural and social environment. The issues concerning rational decision-making have become complex and, consequently, such decisions are becoming more and more difficult to make. Overlaps between environmental, industrial, social and political issues and rapid speed of change contribute even further to the complexity.

In recent years, many experts have highlighted that incorporating sound information about risks (e.g., the probability and likely consequence of various hazards) into the environmental decision making process is the best way to ensure the protection of public health and the environment, at the lowest cost and with the greatest benefits to society [JIEE, 1998]. The definition and comprehension of the concept of "risk" is rather complex. Risk is not strictly measurable and no consensus has emerged on what value of "plausible upper bound on risk" should be taken generally to represent an acceptable level [Putzrath, 1999]. The notion of acceptable risk cannot be considered of as "acceptable" or "not" in isolation, but only in combination with the costs and benefits that are attendant to that risk. Considered in isolation, no risk is acceptable [Kaplan, 1997].

Risk management would be easy only if it was a measurable and quantifiable element. But aspects like "societal benefits" or "policy efficiency" are not strictly measurable. The development of mathematics of probability and statistics has tended to obscure the fact that precise quantifiable risk is possible only in a minority of cases [I. Shepherd, 2000]. Thus, it may be argued that uncertainty can't be considered as a deterrent for sound decision-making process [Funtowicz, 2000]. Policy and decision-makers need to design strategies on the basis of what is known. They can't wait until the future brings a better and clearer understanding of phenomena. Interpretation of the degree of uncertainty and risk management become two relevant aspects of the decision-making process. Shepherd [2000] considers these elements as the "fundamental components of relationship between science and governance".

Therefore, it may be argued that the interpretation and expertise devoted to accounting for economic and social needs in risk management is weak. We are strongly convinced that a better interpretation of such issues requires a more efficient approach for collecting and managing data. Since data related to the impact of industrial activities (e.g. environmental, economical, social), present a very high level of heterogeneity, and are stored in a multitude of repositories, it is straightforward to deduce that the information required by an efficient decision-making process needs to be based on a distributed architecture [Kelly, 1996]. Moreover, the information required by the decision-making process could change according to the requirements expressed by the involved stakeholders. Thus, the decision-making process needs to be supplied with up-to-date information and managed by a flexible system able to satisfy the requirements of stakeholders [Arndt, 2001].

This paper reports the experience gained in this field during the implementation of the SIMAGE project. The project is financed by the Italian Ministry for the Environment and aims to improve the risk management of industrial areas in Italy. In particular, it describes how an architecture based on a data warehouse system can support the efficient implementation of the multi-criteria decision support system (DSS).

2 BACKGROUND

Based on the expertise of the Institute for the Protection and Security of the Citizen (IPSC) on the identification and assessment of risks arising from natural and technological hazards over large territorial areas, the Italian Ministry for the Environment has established a Framework Agreement with the Joint Research Centre that aims to develop an integrated information exchange network that covers the main existing Italian risk areas. The agreement includes a group of projects in the provinces of Brindisi, Taranto and Venice that aim to realise and install a local integrated system for monitoring and preventing pollution, risks related to industrial activities, and accidents deriving from transport of hazardous and toxic substances. The local provincial systems will be integrated in a network at national level. The overall system will be completed by June 2003 is called SIMAGE. The major objectives of the project are:

- The creation of harmonised air-quality networks for the industrial areas of Brindisi and Taranto, including the integration and optimisation of existing air-quality networks, the installation of new monitoring stations and air pollutants instruments, and the definition of quality-control procedures and associated control laboratories;
- The design and establishment of a national coordination centre for environmental information exchange concerning air, water and soil quality, networked with the major existing Italian risk areas, in particular, the areas of Brindisi, Taranto, Porto Marghera, Priolo-Augusta, Gela, Milazzo;
- The development of a pilot system for the monitoring and control of the transportation of dangerous substances mainly via road, including a benchmark exercise of tracking and mobile communication technologies, and the realisation of local systems in Brindisi, Taranto and Porto Marghera for traffic control and emergency management.

According to the objectives of SIMAGE, a consistent amount of information will be collected in the next few years. In order to exploit the available information and to strengthen the action of the Ministry of the Environment, IPSC will also provide the Ministry of the Environment with an IDST with the capability to help decision and policy makers on the definition of policies for sustainable industrial development.

3 DESIGNING AN INTEGRATED DECISION SUPPORT TOOL – IDST

Since the project aims at evaluating the needs for designing supporting tools about industrial sustainable policies, the approach that has been adopted respects a strategic level, which considers the risk posed by industry as part of the problem related to a complex regional system. This strategic level requires a comprehensive vision of the effect over time and space of the policies under definition in terms of environmental impacts and socio-economical consequences. Within the context characterised by multiple and conflicting objectives, it is particularly interesting and effective to apply Multi-Criteria Decision Aiding Analysis (MCDA). This approach enables evaluations of alternatives involving multiple objectives, taking into account the different priorities of the involved parties (e.g. authorities, population, industry, employees, and other groups of interest) [Roy, 1994].

As the concept of sustainability related to industry implies a high number of aspects and evaluations (e.g. environmental impacts, social benefit, risk control and management), the typical approach for designing a MCDA based on a static database and a computer system, can shows some limitations. Moreover organisation, government and companies accumulate and store data with objective to exploit this data and drove an advantage at later stage. Unfortunately, neither the accumulation, nor the storage process, seems to be completely reliable for supporting the policy design and the decision-making process.

Because of the high heterogeneity and multitude of distributed repositories of information required by such a complex decision-making process, a MCDA should be based on a distributed structure of which information can be assessed dynamically according to requirements of the decision-making problem. The definition of such a complex approach for the definition of policies for sustainable development of industry derives from the fact that decision-making is an iterative cognitive process that is not possible to describe rigorously. The evaluation of these scenarios requires an interaction between criteria and alternatives as the decision-makers get better insight into their preferences and further expand the set of alternatives. For this reason, decision support applications must be built in a manner that permits changes to occur easily and quickly. Decision-makers and stakeholders need to collaborate using an iterative process that involves continuous changes [Arndt, 2001]. For such reasons we propose to improve the architecture of DSS, designing a system based on a data warehouse.

3.1 Data warehouse

Data warehouse can be defined as "subject-oriented, integrated, time-varying, non-volatile collections of data that is used primarily in organisational decision making [P.Vassiliadis, 2000]. A data warehouse is a collection of technologies aimed at enabling the decision maker to make better and faster decisions. Data warehouse architecture can be considered as layers of data on top each other in which data from one layer are derived from data of the lower layer.

A data warehouse is a complex system; the volume of recorded data is vast and the processes employed for extraction, transformation, storage and aggregation are numerous, sensitive to changes and time-varying. The metadata repository serves as a roadmap that provides a trace of all the interactions and changes.

Considering such structure (Figure 1.), we can consider that the lowest layer of the data warehouse is directly related to the environment of source data. The extraction or the aggregation process can be performed defining ad hoc query procedures, or data mining techniques. Another form of aggregation can be the implementation of models that take their inputs via the different operational database and store their outputs into the reconciled data repository. Consequently, only the top layer is related to the decision process. As a DSS, a data warehouse must provide high-level quality of data and quality of service. Coherency, accuracy, accessibility, availability and performance are among the features required by the end users of the data warehouse. Also in this case, the information transition front the Data warehouse to end-users can be performed using data query techniques, data mining techniques or modelling approaches.

3.2 Data Quality

Normally, stakeholders involved in a decision or negotiation process are concerned about the quality of data but it could be argued that quality of data is strictly related to use of the data [Jeusfeld, 1998]. Quality of data is of highly subjective nature according to the expectations of each stakeholder. In these terms, data quality is strictly related to the satisfactory degree of data accuracy, i.e. the usefulness of information. Contrarily, data deficiencies, non-availability or accessibility problems are definitely objective, and depend mostly on the information system definition and implementation, i.e. the *responsiveness* of the system.

Therefore, the design, administration and the evolution of the IDST, are critical aspects that can't be solved univocally, which have to be consistent with the end-user requirements.

Hence, the quality of data can be considered in the perspective of the Life Cycle of the information. Starting from a demand of information, the raw data can be collected and throughout the Data warehouse transformed into information specifically tailored for end-users. Only considering the user's expectations and the significance of the emerging information, it can be analysed if the system respond to the end-user demand. Thus, data quality is strictly correlated to Life Cycle Analysis (LCA) of generation process of information. LCA become important in order to optimise the IDST based on a Data Warehouse.

In this sense, IDST goes beyond the typical DSS because it is not simply the integration of a class of model, which helps in information synthesis. For a typical DSS based on static databases and a set of linked models the needs of DSS user are normally well known. Instead, the requirements for the definition of industrial policy can differ according to stakeholder's goals and can change over time [Di Mauro, 2000].

At the moment the IDST prototype system is under the development phase. The tool will be based on the MCA methodology, specifically tailored for aiding decision-makers to cope with sustainability of industrial areas [Di Mauro, 2001]. The main purpose of this phase will be the systematic organisation and analysis of overall information required for the decision process and the comparison of both real and simulated scenarios. The tool will be consistent with most common regulatory standard and will help decision-makers to evaluate scenarios according the in-force regulatory issues. The conceptual general scheme of the prototype is reported on the Figure 2.

3.3 Feasibility

In order to validate the approach and the characteristics of the IDST prototype, a participatory approach has been assumed, which involves proxies for policy and decision makers, and stakeholders. According to the main objective of SIMAGE, a forum [FARI, 2001] has been organised in order to have international, interdisciplinary and transdisciplinary exchanges between scientists, stakeholders, policy-makers, and NGOs about alternatives for industrial sustainable development in Italy and consequently in Europe.
The first meeting confirmed that sustainable development is presently one of the most important subjects on political, technical and socio-economic level and that some regional authorities in Italy have already tried to face this problem.
Experts argued that normally policy-makers and authorities perceive environment not globally but only in terms of its principle components (e.g. soil, air, water) and consequently environmental management reflects this "un-integrated vision". Consequently, the definition of an IDST based on the data warehouse system may be a successful approach even if still many problems need to be solved. It seems that stakeholders and decision-makers, even assisted by experts, have quite limited experience with software tools to support the policy and the decision-making process. In any case, everybody agreed upon the necessity to have more advanced tools but it may be difficult to define a more consistent approach for managing data and information. Participants considered juridical norms as the most important aspect to look at when approaching environmental problem and dissemination of information.

4 CONCLUSIONS

Nowadays, environmental management and related policies are being revisited with a higher level of awareness and from a different perspective where the understanding of the human-environment interactions and the application of science and common sense have been introduced. According to this perspective, there is an increasing interest in integrating the various dimensions of the industrial impacts within a common framework, considering at the same level the environmental effects and the socio-economic effects.

In this regard, the data warehouse is not a revolution in technical terms, but its implications in terms of new knowledge creation could steer the decision-making process on the field of environmental protection and sustainable industrial development, because it is the vehicle to discover threats and opportunities hidden within the large amounts of data. Data exists in abundance but it is often unusable to support decision-making because it is unstructured, unintegrated, aged or inaccurate. The data warehouse allows data to be transformed into information that could become knowledge once it is processed by human intervention. However, the final burden of transforming the information into knowledge depends on the acumen of the decision-makers and stakeholders. The effective use of this latest technology can only be achieved with the creative and innovative capacity of the human components.

Thus, data warehouse may be the ideal tool for successful decision-making. A data warehouse provides the requisite architecture to transform data from different sources into a single image of reality that is easy to analyse and then pro-act accordingly. In a data warehouse environment, all data are uniquely defined for the entire organisation and structured to be combined or aggregated. Data warehousing represents a map for a new way of looking at data and information.

The main purpose of the ongoing project is the systematic organisation and analysis of overall information required for the decision process and the comparison of both real and simulated scenarios.

5 REFERENCES

Arndt, H. K., M. Christ, O. Günther, Environmental Reporting in the Public Interest, in: *Environmental Information Systems in Industry and Public Administration*. Edit by C. Rautenstrauch. Idea Group Publishing, Hershey, PA., 2001.

Barrow, C. J., Environmental Management: Principles and practice", Routledge Editor, London, 1999.

Di Mauro, C., T. Teunenbrooek, A.J.C. Sinke, Time and Land-use as interpretation keys for the evaluation of different remedial scenarios. Proceedings of Consoil Conference on Soil Pollution, September 18-24, Leipzig, Germany, 2000

Di Mauro C., F. Mazzeo Rinaldi, J. P. Nordvik, Integrated approach for effective policy planning and monitoring in industrialised areas - an Italian case study, Proceeding of *54th Meeting of European Working Group - Multicriteria aid for decision*, Durbuy, October 4 and 5, 2001.

FARI, A discussion forum for participated design of an Integrating Decision Support Tool for effective policy planning and monitoring in industrialised areas, Technical note, edited by C. Di Mauro, F. Rinaldi Mazzeo, J. P. Nordvik, JRC Ispra, Italy, 2001.

Funtowicz, S., I. Shepherd, D. Wilkinson & J. Ravetz, European Policy-making - Science and Governance in the European Union: a contribution to the debate, Science and Public Policy, 25(5), 327-334, Beech Tree Publishing, Guildford, UK, 2000

Jeusfeld, M.A., C. Quix, M. Jarke, Design and analysis of quality information for data warehouses, Proc. 17th International Conference on Conceptual Modeling (ER'98), Singapore, Nov 16-19, 1998.

JIEE, How policy makers and the public use risk information in environmental decisions: lessons from three cases, JIEE Report No. 98-07, The Joint Institute for Energy and Environment, Knoxville, TN, June, 1998

Kaplan, S., The Words of Risk Analysis. Risk Analysis, 17(4), 407-417, 1997

Kelly, S., *Data Warehousing : The Route to Mass Customization*, John Wiley & Son, New York, US, 1996.

Putzrath, R. M., Fundamentals of health risk assessment. Use, derivation, validity and limitation of safety indices, Risk Analysis, 19(2),231-247. 1999

Roy, B., M. Paruccini , Multiple Criteria Aid for Decision in Environmental Management, in: "Applying Multiple Criteria Aid for Decision to Environmental Management". Edit by M. Paruccini, EUROCOURSES: Environmental Management, Vol. 3, Kluwer Academic Publishers, Dordrecht, NL, 1994.

Shepherd, I., Science and Governance in the European Union, EUR 19554 EN, JRC-Ispra, March 09, 2000.

Vassiliadis, P., *Data Warehouse Modeling and Quality Issues*, Ph.D. Thesis, National Technical University of Athens, Athens, Greece, June 2000.

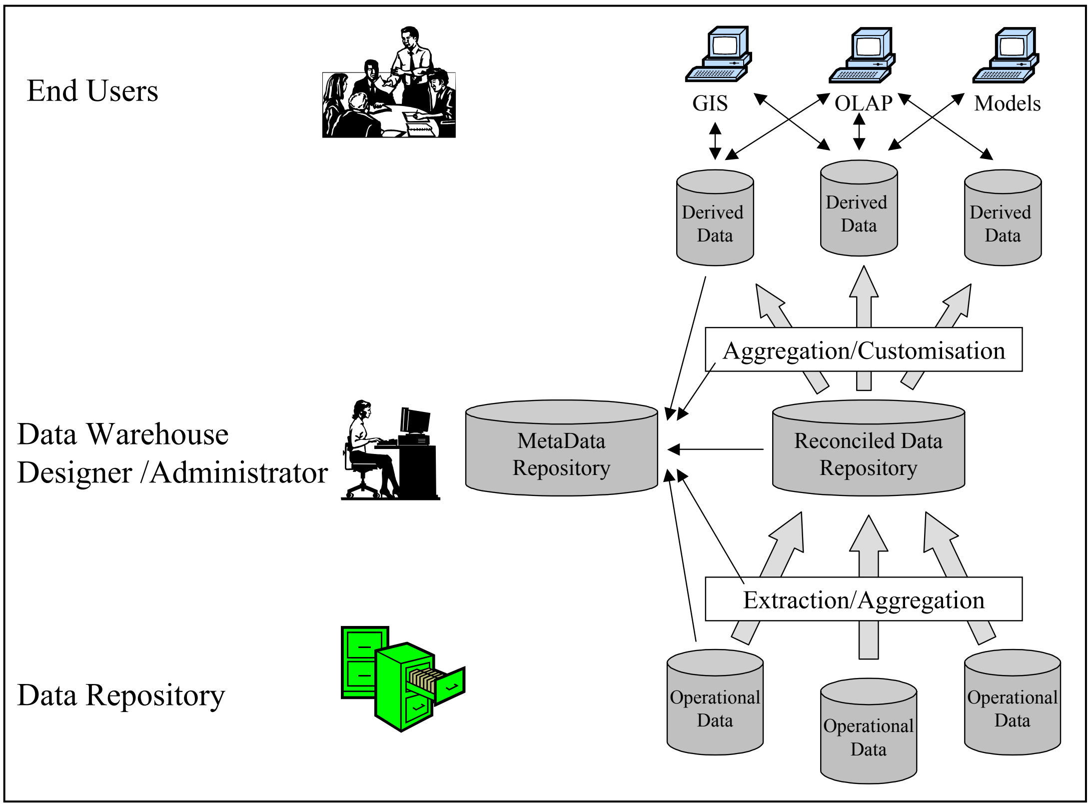

Figure 1. A generic architecture for a data warehouse

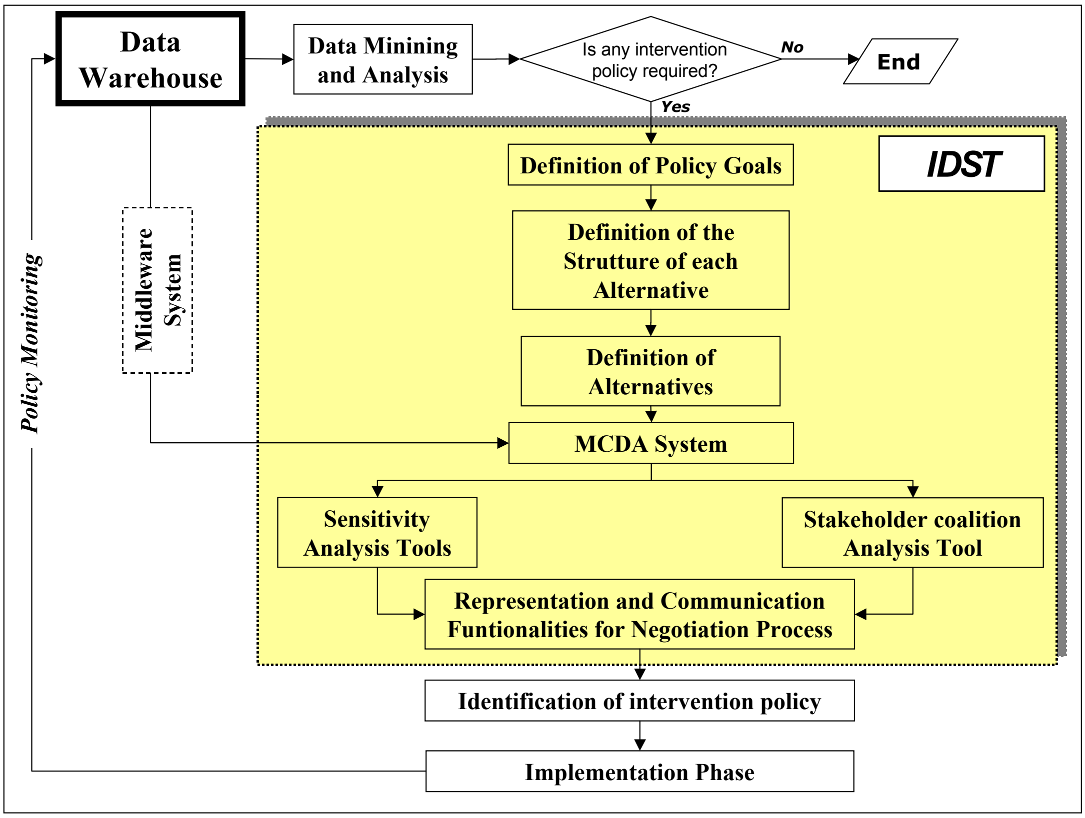

Figure 2. Conceptual scheme of the IDST prototype