

# The Application of Artificial Intelligence in the Field of Chemical Regulation

*Giorgio Sorani and Robert A. Lauer*

The Lubrizol Corporation is a specialty chemical company with headquarters near Cleveland, Ohio. Lubrizol was founded over 60 years ago and has enjoyed a reputation for being a leader in the innovative development of specialty chemicals for use in industry and agriculture. A small manufacturing facility is located on the corporate headquarters site. Two large manufacturing facilities complete the domestic operations, one located in Painesville, Ohio, and the other spanning two sites in Deer Park, Texas, and Bayport, Texas. The corporation maintains additional research and manufacturing facilities around the world.

Because it is a worldwide manufacturer and supplier, Lubrizol has an obligation to provide information about the content, hazards, and handling of its specialty chemicals to various organizations, including customers, carriers, communities, and plant personnel. This information is contained in a document known as a *Material Safety Data Sheet* (MSDS). A portion of one MSDS is shown in figure 1.

As a company selling products containing chemicals that can be haz-

Page 1 of 6

The Lubrizol Corporation  
29400 Lakeland Boulevard  
Wickliffe, Ohio 44092  
216/943-4200

MATERIAL SAFETY DATA SHEET

PRODUCT TRADE NAME: FINISHED PRODUCT

CAS NO: Mixture.  
SYNONYMS: None.  
GENERIC/CHEMICAL NAME: Mixture.  
PRODUCT TYPE: Automotive gear oil additive.  
PREPARATION/REVISION DATE: 11/19/1990  
TRANSPORTATION EMERGENCY PH NO (CHEMTREC): 1-800-424-9300.  
NFPA CODE: Health: 2 Fire: 2 Reactivity: 0  
HMIS CODE: Health: 2 Fire: 2 Reactivity: 0  
PRINCIPAL HAZARDS:

WARNING  
- HARMFUL IF INHALED.  
- COMBUSTIBLE LIQUID.  
- MAY CAUSE ALLERGIC SKIN REACTION.

---

SECTION 1 - HAZARDOUS INGREDIENTS

---

- This material is not known to contain greater than 0.1% of any carcinogen required to be listed under the OSHA Hazard Communication Standard (29CFR 1910.1200).  
- From 1 to 5 percent Oleylamine;  
- From 0.5 to 1.5 percent Isopropyl alcohol, CAS no: 67-63-0; OSHA PEL: 400.00 ppm; OSHA STEL: 500.00 ppm ACGIH TLV: 980.00 mg/ cuM., 400.00 ppm; ACGIH TLF STEL: 1225.00 mg/ cu M., 500.00 ppm;  
- From 0.1 to 1 percent 1,2,4-Trimethylbenzene, CAS no: 95-63-6;  
- Please note that the chemical identity of some or all of the above hazardous ingredients is confidential business information and is being withheld as permitted by 29CFR 1910.1200 and various State Right to Know Laws.

---

SECTION 2 - FIRE AND EXPLOSION HAZARDS

---

FLASH POINT: 63 Deg C 145.4 Deg F (PMCC)  
UPPER FLAMMABLE LIMIT: Not determined.  
LOWER FLAMMABLE LIMIT: Unknown  
EXTINGUISHING MEDIA: CO2, dry chemical, alcohol foam. Water can be used to cool and protect exposed material.  
SPECIAL FIREFIGHTING PROCEDURES: Recommend wearing self-contained breathing apparatus.  
UNUSUAL FIRE & EXPLOSION HAZARDS: Toxic fumes, gases or vapors may evolve on burning. Vapors may be heavier than air and may travel along the ground to a distant ignition source and flash back. Container may rupture on heating.

*Figure 1. Example of an MSDS.*

ardous if not handled properly, Lubrizol is required by federal and state regulations to have MSDSs for all products on site. It must also supply MSDS to customers with the shipment of a product they are purchasing for the first time and with the shipment of a product whose MSDS changed since the customer's last order. At a minimum, all customers must receive MSDSs for their products once a year. The generation of this document is a complex process because the chemical make-up of each ingredient of a product must be analyzed to determine potential health and safety effects.

Prior to 1989, the writing of these documents was a manual function

partially automated by the use of a word processing package. Because of the volume of MSDSs required for Lubrizol's 10,000 products, intermediates, raw materials, and research materials, several full-time employees were needed to create and maintain these documents. Because the complex government regulations applicable to the content of MSDSs are subject to interpretation, a sheet written by one individual could contain different wording than a sheet written by another individual for the same product. Although both MSDSs would be in compliance and within the limitations imposed by the government, they might send different messages to the reader. Because of the amount of time and resources necessary to keep Lubrizol's MSDSs current and accurate, it was determined that automating the MSDS process would reduce the associated cost and bring consistency and standardization to the content of the documents.

A project was initiated in January 1987 to develop a system that would be able to store data about all Lubrizol chemicals, utilize these data to determine the content of MSDS for any material, and determine when MSDS had to be sent to a customer. The system would provide for a series of online transactions to be used by the people who formerly wrote MSDSs to capture raw data about each chemical in a relational database. These data would be used to generate the MSDS document. Additional system functions would allow online viewing of the document and printing for shipment to customers. The result would be the complete automation of all MSDS business functions, including the complex process of generating MSDS itself.

### **Decision to Use AI**

As the requirements of the system were defined, several functional aspects presented complex challenges. The major issue was the feasibility of capturing, within the system, the process used by the experts to create MSDS. This highly intricate process involves hundreds of rules and data elements, which are constantly changing to adapt to new types of chemicals and better chemical hazard information. In addition, federal and state government regulations concerning chemicals and MSDS requirements change constantly, causing the process to become even more dynamic. Finally, the sheer volume of chemicals handled by Lubrizol raised issues regarding the performance implications of completely automating the process. The system would need to automatically revise MSDS information for all chemicals containing a particular ingredient if the hazard information on this ingredient changed. These requirements for complexity, ease of maintenance, and the abili-

ty to manipulate large volumes of data indicated that a solution involving AI technology might apply.

At the time, Lubrizol had little experience with AI technology and vendors. An evaluation of five of the more popular expert system packages was undertaken by members of the management information system team. The evaluation included a review of each package's technical architecture, ease of use, cost, and other features. The evaluation took three months and consisted of reviewing literature and visual demonstrations for each product.

The results of the evaluation indicated that none of the packages perfectly addressed Lubrizol's needs. However, it was recognized that the technology was still maturing and that improvements were anticipated in the AI market. Recognizing the risk of pursuing an alternative method to traditional system development approaches, Lubrizol purchased ADS from Aion Corporation. This product appeared strongest in the evaluation and met the majority of Lubrizol's requirements. The intent was to use ADS to develop a workstation prototype of the logic needed to generate MSDS that would be ported to Lubrizol's IBM 3090 mainframe for volume testing and eventual implementation.

### Knowledge Acquisition

The first step in automating the MSDS function was to work with the experts (Lubrizol's toxicologists) to define and understand the process of determining the MSDS content. This process proved to be time consuming primarily because of the great complexity of the process as well as the need to standardize the approach for all types of Lubrizol chemicals. The ideal result was a common set of rules to manipulate the hundreds of data elements and parameters used to determine MSDS content for the more than 10,000 Lubrizol products, intermediates, raw materials, developmental products, research blends, research concentrates, and waste materials.

The complexity of the process as a whole is a result of the nature of MSDS itself. MSDS for each chemical consists of 11 separate sections, each with multiple information items. This arrangement results in nearly 80 separate sets of logic to determine all items. The logic for most of these items is complex in itself, usually involving many data elements, parameters, and rules and having many possible results. For example, the logic to determine the value of the fire codes (figure 1) involves evaluating 13 different pieces of data about each chemical by 25 related rules. The complexity of the entire process is increased by many dependencies between sections. The result of one item (if it ex-

ists) can be used in determining a different item, and multiple items can all be dependent on each other.

Given the complexity of the business rules for writing MSDS, it was necessary to document the experts' knowledge in a format that could easily be converted into expert system rules. A variation of a decision table was used to capture this information. (An example of a simple MSDS decision table is shown in figure 2.) Over 100 decision tables were documented by the experts. These tables were turned into states, objects, and rules in the expert system.

This approach to documenting the logic had a number of advantages. Not only did it help the management information system staff turn business rules into computer logic, it also helped the experts organize their thoughts to better communicate their needs to the management information system staff. It also facilitated the documentation of test conditions and expected results when testing of the knowledge base began. The result was the increased quality of the knowledge-acquisition process and increased speed of prototyping and testing.

### **Building the System**

Concurrent with documenting the experts' knowledge, the surrounding functions of the system were constructed using more traditional development methods. All online conversations and batch functions, except for the MSDS generation knowledge base, were built using the COBOL application generator TELON from Pansophic Systems, Inc. Although they did not directly communicate with it, these system components had to be integrated with the knowledge base to the extent that they shared and processed the same information on Lubrizol's chemicals.

In addition to internally integrating multiple technologies, the system also had to be integrated with other Lubrizol business systems. Formulation information from the manufacturing system was passed to the knowledge base for use in creating new MSDSs or revising existing MSDSs. Information was obtained from the order-handling system to synchronize the printing of MSDS documents with customer orders. The degree of integration with other business systems was highly innovative for an AI application.

On completion of knowledge acquisition, a prototype of the MSDS knowledge base was developed using ADS on an IBM PS/2 workstation. An iterative approach to prototyping was used to confirm the accuracy of the logic documented in the decision tables and the feasibility of deploying the logic as an expert system. The prototyping process was also used to evaluate various approaches to the architecture of the knowl-

Is The Value For Percent Volatile Null?	Is Any Component Prevalence $\geq$ 100%?	Is The Component Value For Percent Volatile Null?	Percent Volatile Field Result
No	–	–	Value of material percent volatile
Yes	No	–	"Unknown"
	Yes	No	Value of component percent volatile
		Yes	"Unknown"

Figure 2. Example of a Decision Table.

edge base, considering issues such as size, efficiency, and ease of maintenance.

The architecture that was chosen for the knowledge base allowed the organization of the logic by MSDS section. The rules for each section were combined in a distinct state, and a controlling state was used to initiate the others, as needed. In addition to the rules related to generating MSDS, the knowledge base also contains components for addressing the complex task of formatting the text into a high-quality document and storing the information in the database for future access. The knowledge base uses integrated forward and backward chaining, as well as heuristic search techniques, to generate MSDS. A portion of the data access needed to support the process is always performed at the start of execution, but various other input-output functions are executed when dictated by inferencing.

At the completion of prototyping, the knowledge base was ported to the mainframe, enhanced, and tuned to operate in its target environment (figure 3). This process of redeployment on the mainframe involved using the ADS *high-performance option*. This facility converts the knowledge base to Pascal and assembler code, allowing normal compiling and linking in the mainframe environment. By using this option, the knowledge base was executed in a fashion similar to traditional mainframe programs.

All data for the system are stored in the DB2 relational database. Although this database management system represented a strategic direction and many benefits for Lubrizol, it posed some important challenges for the system. At the time, few AI products had been integrated with DB2 in production systems. This installation was one of the first of such systems using the ADS product. In addition, the architecture of

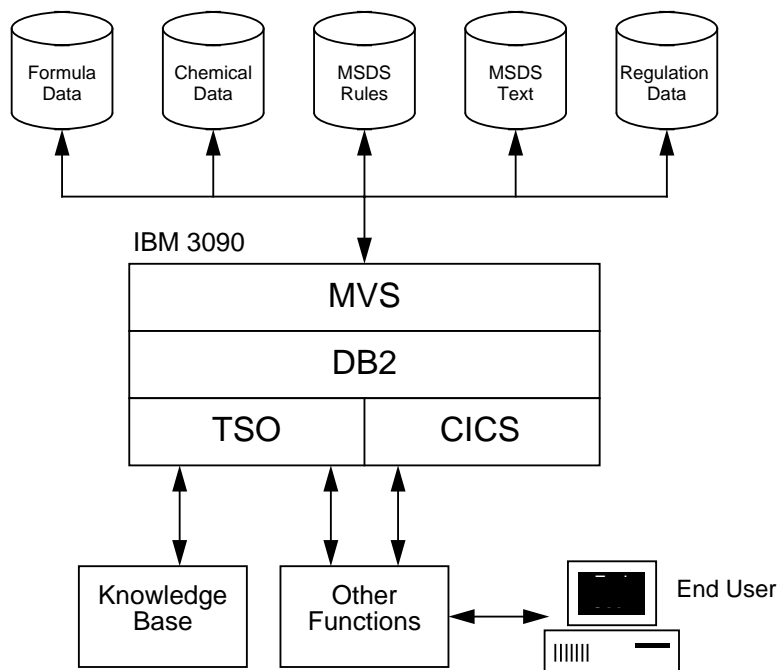


Figure 3. Technical Environment.

ADS-DB2 communications used dynamic SQL for processing DB2 data, which required additional analysis and design of the database structure to avoid possible performance degradation.

Extensive testing of the knowledge base was critical in resolving these issues as well as ensuring the accuracy of MSDS generated by the system. To achieve the greatest degree of accuracy possible, the system was used to generate MSDSs for all Lubrizol chemicals. Seventy percent of these MSDSs were reviewed by the experts. This subset of all MSDSs included chemicals of all types, complexities, and exceptions to assure that the logic was correct. All errors were identified and traced to their source, which was either erroneous logic or incomplete and inaccurate data. The experts corrected the data and worked with management information system personnel to modify the decision tables, as needed, to correct or enhance the logic. This iterative process continued until Lubrizol was satisfied with the accuracy of the information being produced by the system.

The major difficulty encountered during the testing process was the lack of a basis of comparison for the results of the new system. Because the logic used in the knowledge base included many approaches that

were new to Lubrizol, and the current manually generated MSDSs were known to be inaccurate, the experts' review was based on assessing the accuracy of each piece of information in each new MSDS. In addition, MSDSs on materials within a chemical family were compared to ensure the consistency of information across similar chemicals.

The completion of system development and testing occurred close to 2½ years after the project was initiated. During this time, a diverse team of experts, users of MSDS information, and management information system personnel participated in all project phases. The total effort expended on the system was approximately 2,400 workdays, or nearly 11 person-years. Additional cost was incurred to purchase the ADS software for both the workstation and mainframe platforms.

## Production

Because of its volume and complexity, at implementation, the system was impressive for an application using AI technology. The knowledge base used to generate MSDS information and format the documents contains several hundred parameters and rules. The data supporting the generation of all Lubrizol MSDSs are stored in 17 DB2 tables containing nearly 3,000,000 rows of data. The system maintains MSDSs on nearly 10,000 Lubrizol raw materials, intermediates, finished products, and research materials.

The Lubrizol MSDS system has been in full production since October 1989. Lubrizol employees at all domestic locations have the capability to view MSDS documents online and request printed copies. The system generates the proper MSDS documents for shipment to customers with orders of Lubrizol products and on special request. MSDS data are transmitted daily from corporate headquarters to Lubrizol's facility in Bromborough, England, for translation into several languages and to ensure standardization across the corporation. A cooperative processing system that creates Canadian-English MSDSs by running a clone of the knowledge base on a workstation platform was implemented in October 1990. With many other enhancements planned, the system is positioned as a strategic tool for Lubrizol.

Since implementation, the knowledge base has frequently been revised and improved. The major catalysts are changing government regulations, new approaches to the logic arising from better information about the chemicals, and periodic identification by the experts of errors in MSDSs. In addition, various performance issues, such as the overhead of using dynamic SQL, are being addressed through a series of enhancements.



Changes to the logic in the knowledge base originate from the experts. Management information system personnel then work together with the experts to modify the decision tables to reflect the desired changes in the logic. Management information system personnel physically implement the change and generate a variety of example MSDSs, which are reviewed for accuracy by the experts prior to production implementation of the change.

Currently, there are no plans to allow the experts to directly maintain the knowledge base because of the technically complex nature of the process. However, recent developments in the evolution of the ADS software indicate that ease of use might increase in the future to the point where maintenance by the experts is feasible.

### **Benefits**

The system has provided significant benefits to the corporation in a number of areas. Most importantly, Lubrizol is ensured of accurate and timely MSDS information in compliance with government regulations. This ability increases the safety of Lubrizol employees and customers who work with Lubrizol chemicals. Second, the system has given Lubrizol the ability to adapt to constantly changing regulations governing the content and distribution of MSDSs. Third, the cost required to create and maintain MSDSs is drastically reduced. The system is able to handle hundreds of MSDS generations each day, many times the manual capabilities of the experts. Finally, the system standardizes the appearance and content of MSDSs, allowing simplified interpretation by users of the information.

To quantify the value provided by the system, these benefits must first be categorized as tangible and intangible. Tangible benefits include compliance (potential savings are in the hundreds of thousands of dollars in government fines), flexibility (cost of maintaining MSDSs is reduced by hundreds of workdays each year), and speed (in mass revision situations, the system outperforms a manual approach by a tremendous margin, which adds to the labor savings). Intangible benefits include greater accuracy (increasing the level of safety of Lubrizol employees and customers) and standardized MSDS appearance and content (leading to increased customer satisfaction). It should be noted that an exact quantification of the savings provided by the system would be difficult to calculate because of major differences in MSDS format and content between the system and its manual predecessor. The system has allowed Lubrizol to provide significantly greater functions in MSDSs at a much lower cost.

## Summary

AI technology proved to be an excellent solution to the unique requirements and constraints of the MSDS application. The use of an expert system provided the flexibility and ease of maintenance necessary for such a complex application and delivered adequate performance in processing large volumes of data in a dynamic environment. In addition, the use of the AI technology in this manner facilitated the process of knowledge acquisition and prototyping, which improved communication between the experts and management information system personnel. These features contributed to the delivery of a system with higher quality and adherence to user requirements. It is believed that the cost of building and maintaining the system was lower than would have been required for developing the system using purely traditional methods.

The Lubrizol MSDS system is considered innovative for the following major reasons: First, the system is on the leading edge of a new wave of AI applications addressing chemical regulation and is one of the first to automate the MSDS function so completely. Second, at implementation, the system was one of the first large mainframe applications to integrate AI technology with a DB2 relational database. Third, the application of AI technology is highly integrated with existing business systems of more traditional architectures.

For the success of the system and the reasons previously indicated, Lubrizol was granted the 1990 Computerworld Smithsonian Award for technologically advanced applications contributing to the benefit of society. This achievement confirms the applicability of AI technology to the field of chemical regulation and business systems in general.