

**PANEL SESSION**  
**MACHINE-READABLE DICTIONARIES**

Donald E. Walker  
Natural-Language and Knowledge-Resource Systems  
SRI International  
Menlo Park, California 94025, USA

and  
Artificial Intelligence and Information Science Research  
Bell Communications Research  
445 South Street  
Morristown, New Jersey 07960, USA

**Abstract**

The papers in this panel consider machine-readable dictionaries from several perspectives: research in computational linguistics and computational lexicology, the development of tools for improving accessibility, the design of lexical reference systems for educational purposes, and applications of machine-readable dictionaries in information science contexts. As background and by way of introduction, a description is provided of a workshop on machine-readable dictionaries that was held at SRI International in April 1983.

**Introduction**

Dictionaries constitute a unique resource for a broad range of research involving natural language, information, knowledge, and the analysis of contemporary culture. Although they are often regarded as the special preserve of lexicographers and lexicologists, data contained in dictionaries have significant implications for research in linguistics, computational linguistics, artificial intelligence, information science, psychology, anthropology, sociology, philosophy, education, and probably other fields as well. Dictionaries embody the lexicon of the language. They provide phonological, grammatical, semantic, and historical information relevant for linguists and other language specialists. They are useful adjuncts for the development of natural-language-understanding systems and natural-language-interface technology. They can provide a mechanism for processing full-text data sources and for information retrieval more generally. Dictionary data figure in psychological experiments on language and perception. Semantics and usage are reflected in ways that are factored into ethnosemantic and sociolinguistic research. Philosophical and logical inquiries build on lexical information. For education, dictionaries provide not only reference, but are practical aid for teaching both adults and children reading and writing skills.

Dictionaries have always had these potential attributes, but they are complex structures and difficult to manipulate. Having them available in machine-readable form makes more sophisticated research in lexicology and lexicography possible--and the results of such work feed back into research in the other areas mentioned above. In addition, dictionaries can be utilized in areas like word processing and office automation, where people are currently showing considerable interest in them. A number of dictionaries have now been prepared by computer typesetting, so the tapes used to drive the photocomposer are available. However, there is a significant

difference between having a dictionary in computerized form and having a database embodying its contents which can be accessed in a number of different ways.

**A Workshop**

Recognizing the potential of machine-readable dictionaries and, at the same time, the lack of coordination among people working in the field, Bob Amsler and I organized a *A Workshop on Machine-Readable Dictionaries* at SRI International in April 1983. The National Science Foundation agreed to provide funds (Grant No. IST-8300940; SRI Project 5699), and we succeeded in involving 29 people from Belgium, England, West Germany, Italy, Japan, Sweden, and the United States for a period of three days. The group included research scientists from universities and institutes, publishers, and people involved in marketing dictionary products.

There were a number of objectives that motivated convening the workshop and that served as a guide to its organization and the assessment of its results:

1. Clarification of the research interests and goals of both the participants and the broader community that they represent. Including in the latter are dictionary publishers and the various classes of potential users of machine-readable dictionaries and their by-products.
2. Identification of the resources in the field: for example, dictionaries actually in machine-readable form, the people engaged in research on them, programs developed for processing dictionary data, references to the relevant literature.
3. Examination of the problems entailed in research in this area.
4. Delineation of computational requirements for various research tasks.
5. Specification of guidelines for dictionary design, both form and content.
6. Formulation of a comprehensive plan to coordinate research efforts in the field.
7. Determination of needs and potential sources of funding for research.
8. Arrangements for future workshops or other meetings.

A volume containing a challenge paper prepared by Bob Amsler, contributions from a number of the participants, summaries of the discussions, and an extensive bibliography of work in the field is in preparation.