

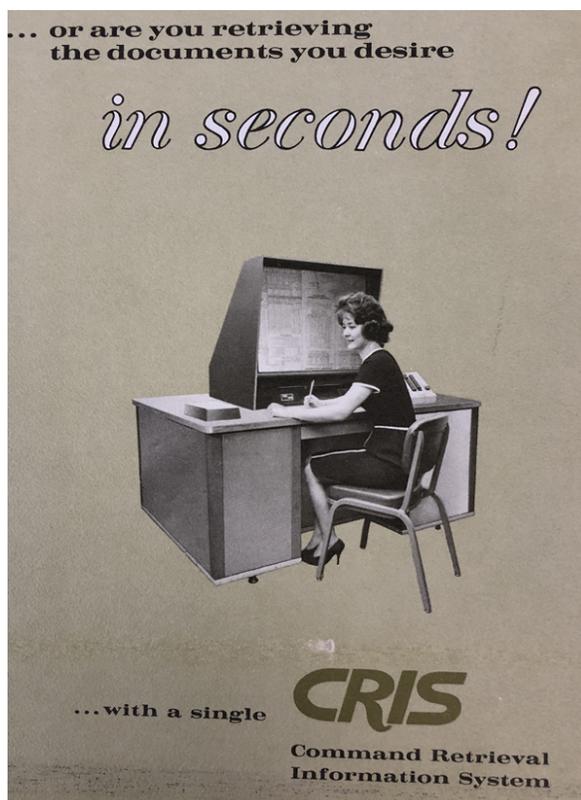
News from Nan

COMPUTER HISTORY

THE LONG HISTORY OF NATURAL LANGUAGE QUERIES AND SEARCHES

JULY 7, 2018 | DICK

From 2014 to 2020 we researched and cataloged over 5000 items (hardware, software, and text) at the Computer History Museum in Mountain View, California. We sometimes blogged about our experiences and the fascinating items we encountered; this post is part of that set.



CRIS Information Retrieval System

I recently read a book by Charles P. Bourne and Trudi Bellardo Hahn titled *A History of Online Information Services 1963-1976*. The notes and source materials for the book are now part of the Computer History Museum archive, which we are helping to preserve and catalog for use by researchers.

I was struck by the fact that during the time period covered by the book some remarkably “modern” technologies had already been developed and put into practice, including the following:

- search queries in natural language form
- relevance scoring and ranking of the output of search results
- the use of stop lists for scanning text
- stem searching and “wild card” searching
- search term weighting
- incorporating synonyms automatically in searches
- the use of alternate search terms
- proximity searches
- phonetic searches

However, in spite of the algorithmic wizardry, most of the early projects described in the book were modest by today’s standards, because the computers of the day were expensive, had slow CPUs, main memory measured in kilobytes, and external disk storage measured in megabytes.

Much of the data to be searched resided on magnetic tape or even punched cards, and remote access was limited and expensive. Getting an answer to a query was not a matter of typing it into Google and instantly getting back thousands of “hits,” as is true today.

Because of the lack of computer resources, the systems required for information retrieval tasks might be available only at certain times during the day, and an individual query might take minutes or hours to complete. The answer might be typed on paper, line by line, on a teletype terminal.

If you were remote to the computer complex, you would be operating over a phone line running at about 30 characters a second.

These limitations didn’t stop the early information retrieval pioneers from making many clever inventions, many of which form the basis of today’s advanced artificial intelligence (AI) systems!

◀ 2019 ◀ A HISTORY OF ONLINE INFORMATION SERVICES 1963-1976 ◀ CHARLES P BOURNE ◀ COMPUTER HISTORY MUSEUM
◀ INFORMATION RETRIEVAL