Local entrepreneur experiences rapid advances in the information business

(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

When Bonnie Carroll, founder of a former local information technology company, started as a librarian in the information business in 1971, there were no personal computers, no internet, no World Wide Web. As she puts it, "If you were Rip Van Winkle, and if you fell asleep at the end of the 1960s and just woke up, you would realize that after 50 years just how different the world had become."

In a recent lecture summarized here by Carolyn Krause, Carroll reflected on the history of the scientific and technical information (STI) business in the past half century. STI is defined as "the collected set of facts, analyses and conclusions resulting from scientific, technical and related engineering research and development efforts, both basic and applied."

***

"I don't think President Barack Obama's administration gets enough credit for how they changed the world of science and the world of data," said Bonnie Carroll, founder of the former local company, Information International Associates (IIa), in a recent talk to Friends of Oak Ridge National Laboratory (FORNL). "Obama opened the concept of open government, open access, open science and open data because most information had become digital. It was a democratizing concept because everyone who has a computer can get access to data through the internet. The 2013 memo of the president's science advisor, John Holdren, was a watershed in the federal approach to scientific data. It held that all the results from federally funded R&D should be made publicly accessible."

In combination with the White House Big Data R&D Initiative launched in 2012, the next step was to make open data usable and understandable, she said. Such data was called FAIR data – Findable, Accessible, Interoperable and Reusable. As former secretary general of the Committee on Data for Science and Technology (CODATA), Carroll worked with two critical players in this world of open data. One was Geoffrey Boulton, the CODATA president and lead author of "Science as an Open Enterprise," who wrote, "to fail to communicate the data that supports scientific assertions is tantamount to malpractice." Boulton wrote the report for the Royal Academy in Great Britain on why open data is needed. His successor as the president of CODATA, Barend Mons of the University of Leiden, was one of the biggest movers of the FAIR principles worldwide.

Unfortunately, she said, the United States has lost its leadership role in national policy for open data even though it was initiated in our nation. "Europe has taken the lead in developing principles and implementation mechanisms for FAIR open data. Interestingly enough, China is working hard to gain the world's leadership. Only recently has our nation tried to reclaim a leadership role."

Before she joined ORNL as a librarian in 1971, Carroll had absorbed the lessons of ORNL Director Alvin Weinberg's 1963 report for the White House called "Science, Government, and Information."

"The report stated that transfer of information is an inseparable part of research and development," Carroll said. "The challenge was coping with the growing volume and variable quality of information. Information analysis centers staffed by scientists, documentarians and information specialists were seen by Weinberg as a solution." Carroll's last job at ORNL in the 1970s was to coordinate its 24 information centers that had matured at ORNL.

In the early 1970s when she was at ORNL, she witnessed the first application of computers to information: the "keyword in context" (KWIC) system that generated an index alphabetized on a keyword that appears within an abstract's context, enabling a search for related articles. "The next large step taken in the automation of information retrieval was a result of the RECON system developed for NASA by Systems Development Corporation in San Diego. In the 1970s, RECON became available at ORNL, enabling scientists to be trained as information retrieval specialists. It was the precursor of the Googles of the world."

Local entrepreneur experiences rapid advances in the information business

(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

An STI infrastructure had evolved at the end of the 1960s with the establishment of government agencies, such as the Defense Technical Information Center and the Department of Energy's Office of Scientific and Technical Information in Oak Ridge, where Carroll worked. "We knew how to collect documentation, and we had abstracts and indexes for various scientific disciplines and mission agencies, including nuclear science," she said. "These activities, especially the automation of Nuclear Science Abstracts, were leadership advances in making scientific information digitally accessible."

Here are some highlights of STI history that took place after the initial application of automation to scientific discipline systems such as Nuclear Science Abstracts, Chemical Abstracts, and Index Medicus, according to Carroll. In 1978 the Arthur D. Little report "Passing the Threshold into the Information Age" recommended that insights gained from STI, and societal data are needed for crisis management and decision making. A 1983 article in Science magazine stated that because of electronic access systems, information availability is increasing more rapidly than the ability to absorb it, echoing one of Weinberg's observations. In the late 1990s the Google search engine and the web made it unnecessary to do indexing because of computer advances in free text searching.

Authors experimented with publishing on the web as an alternative or supplement to publishing in scientific journals. In the 1970s typed articles were mailed to journals and in the 1980s electronically published papers were emailed to journals, most of which eventually went online in addition to being printed. Preprints of STI publications on the web have been a growing trend (especially during the COVID pandemic of 2020-2022), because they enable researchers to easily post experimental results prior to peer review.

"A lot of studies have shown that peer review is not the most objective process in the world," Carroll said. "When a scientist puts a preprint on the web, the whole scientific community can make comments. It's peer review by crowdsourcing."

Early this century, Jim Gray of Microsoft envisioned that science had become "data intensive" and that there are four paradigms of science: empirical evidence, scientific theory, computational science, and, now, data science. Because he died suddenly in a tragic accident before his book "The Fourth Paradigm" was finished, it was not published until 2009; one of the three co-authors who completed the book was Tony Hey, an advisor to ORNL.

On Feb. 3, 2003, the National Science Foundation published a book by Daniel Atkins of the University of Michigan on "Revolutionizing Science and Engineering through Cyberinfrastructure." It envisioned an integrated system of hardware and software resources and services that would efficiently connect laboratories over the internet, enabling scientists and engineers to participate remotely in shared experiments, computer modeling and data acquisition, integration, mining and visualization.

Importantly for the development of information was the recognition that scientific data is part of the scientific infrastructure, Carroll said. Between the government and National Academy reports that have been written since this one and the issuance of the Holdren memo, STI and scientific data have been front and center in discussion of the advancement of open data as a critical pillar of science in the fourth paradigm.

Information International Associates (IIa), the company Carroll started in 1988, was originally a firm that managed customers' libraries and produced scientific abstracts for the *Energy Database* (*successor of Nuclear Science Abstracts*). But it evolved into an information technology company that was greatly stimulated by IIa's purchase of a pure-play IT firm. IT uses computers and telecommunications for storing, retrieving, and sending digital information.

IIa delivered information management and information technology services to government agencies and corporations. Information management involves helping customers identify their information needs,

Local entrepreneur experiences rapid advances in the information business

(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

acquire information, organize and store it, develop information products and services, distribute information and use it. Information technology services include developing and managing computer infrastructures for federal agencies. As threats increased, IIa also offered cybersecurity services to its customers to help protect them against unauthorized or criminal use of electronic data.

One of IIa's smaller services was to study how to deal with data privacy, Carroll said. The Internet of Things with their sensors – smart phones, computers, smart speakers, energy meters, cars, baby monitors, ovens and refrigerators – contribute to information overload and data exploitation. Our political views, buying habits and health issues are becoming known to companies and agencies that wish to predict how we will behave, what we will most likely want to buy and whether we fit the profile of being a reliable employee or good student.

Because all data has become digital, "the scientific world is more involved than it used to be in the issues of information management," Carroll said. "The White House Science Office was not interested in STI 30 years ago, but now it has four committees dealing with it."

Thanks to new digital technologies, she concluded, science in the 21st century will be conducted in a fully digital world, and data-intensive science will foster creativity. "Oak Ridge has a long history of involvement in the management of scientific data and is in an excellent position to take a national leadership position," she added.

One reason is that Carroll, who was interim executive director of the World Data System International Program Office, the worldwide consortium of data systems around the world, helped bring it to the Oak Ridge Innovation Institute, a joint venture between ORNL and the University of Tennessee at Knoxville.

That is one of her legacies of local and national importance.

***

Thanks again Carolyn for an in-depth review of Bonnie Carroll's vast array of endeavors worldwide. She is yet another example of Oak Ridge's influence across many aspects of scientific research and development. Her insights into the needs of the information technology services to keep pace with the computerization of all aspects of information gathering and reporting is amazing.

She was able to see the potential for future developments and organize to meet those needs as they arose. Again, an Oak Ridger that many would think of as just Bonnie, is truly a person of broad influence and wide impact in her chosen field.

She is also a most generous person with both her time and financial support to charitable organizations. Bonnie is one of us, but she is also much more. We can be proud she chose the route she has taken and wish her the best in retirement (although, she remains as busy as ever!).

Local entrepreneur experiences rapid advances in the information business

(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

Bonnie Carroll speaks to the Friends of Oak Ridge National Laboratory

Carroll receives an award for IIa given by the Department of Interior

Local entrepreneur experiences rapid advances in the information business

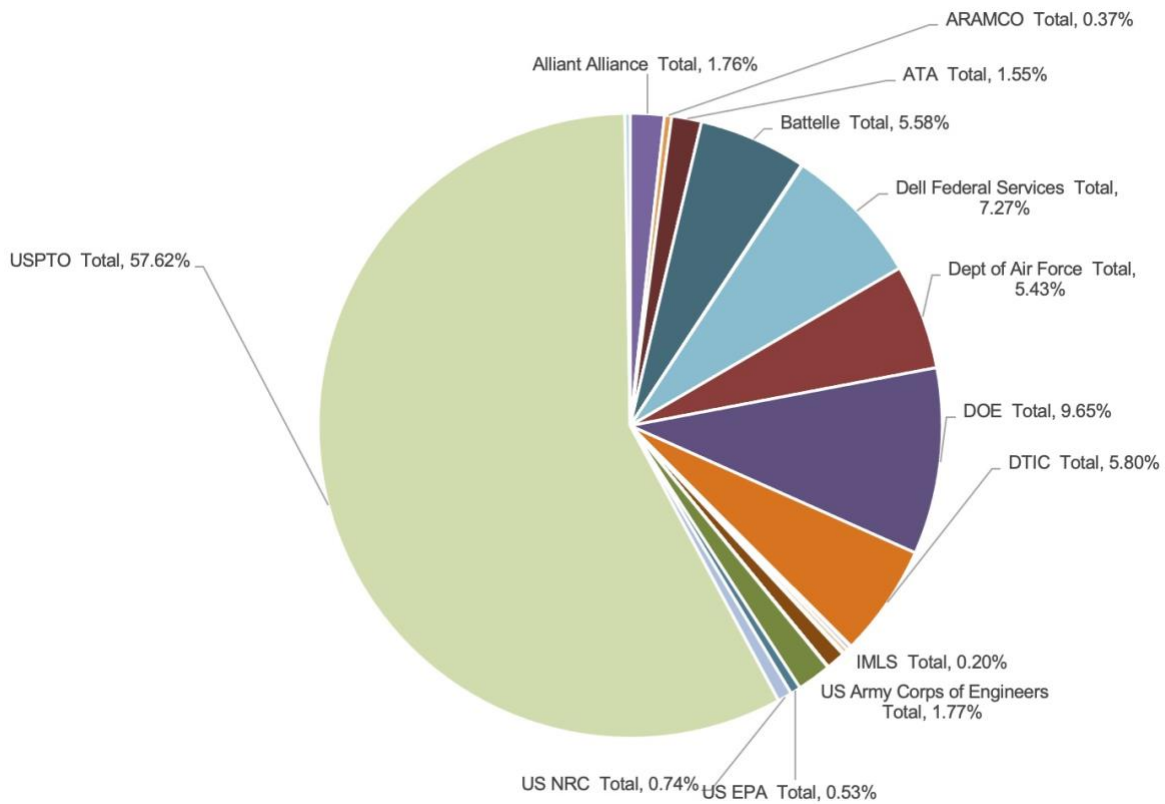(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

Government agencies for which IIa has provided information technology management support

Local entrepreneur experiences rapid advances in the information business

(As published in The Oak Ridger's Historically Speaking column the week of August 22, 2022)

A chart showing revenue sources for IIa.