August 1991

Bringing It All Back Home: U.S. Government Efforts to Improve Access to Foreign Research Information

Robert Billingsley

Follow this and additional works at: https://jdc.jefferson.edu/scitechnews

Let us know how access to this document benefits you

Recommended Citation
Available at: https://jdc.jefferson.edu/scitechnews/vol45/iss3/4

This Article is brought to you for free and open access by the Jefferson Digital Commons. The Jefferson Digital Commons is a service of Thomas Jefferson University's Center for Teaching and Learning (CTL). The Commons is a showcase for Jefferson books and journals, peer-reviewed scholarly publications, unique historical collections from the University archives, and teaching tools. The Jefferson Digital Commons allows researchers and interested readers anywhere in the world to learn about and keep up to date with Jefferson scholarship. This article has been accepted for inclusion in Sci-Tech News by an authorized administrator of the Jefferson Digital Commons. For more information, please contact: JeffersonDigitalCommons@jefferson.edu.
BRINGING IT ALL BACK HOME:
U.S. GOVERNMENT EFFORTS TO IMPROVE ACCESS TO FOREIGN RESEARCH INFORMATION
BY
ROBERT BILLINGSLEY

U.S. GOVERNMENT'S NEED FOR FOREIGN-LANGUAGE INFORMATION
For anyone involved in providing scientific and technical information (STI) services, the following scenario is all too familiar. A scientist, engineer or other type of researcher has located an article in a periodical, needs the information yesterday, but cannot read it. The source of the needed information is not written in English. Has the article been translated? Can anyone on the library or information center staff provide a gist of the information? Often the answer to both of these questions is no.

In the global body of STI, nearly 50 per cent is presently published in foreign languages. The following description of foreign information activities in the U.S. Government may be only the tip of an iceberg in terms of the nation’s real demand for translations. The flow of foreign STI now includes important languages that U.S. translators are largely unable to translate: Arabic, Farsi, Chinese, and Korean. If these languages could be more easily and economically translated, there would be a significant increase in translations. The true demand for foreign S&T literature is hard to measure.

Of course, industry and academia also have to deal with foreign STI, but the purpose of this article is to describe what some elements of the U.S. Government are doing to improve access to foreign S&T literature.

PATENT & TRADEMARK OFFICE (PTO)
According to policy makers at the PTO, an average of 70 per cent of the information in a given technology can be found only in patents. The rest is found in more traditional literature, such as technical publications and scientific papers. Unfortunately, over 8 million foreign patents filed abroad are not available in English, and must be translated to determine the patentability of foreign applications filed at the PTO.

Because of potentially profitable markets, intellectual property law has recently burgeoned and become decidedly international. This trend is evident in 1989 PTO figures: of the 102,692 U.S. patents granted, 46.7 per cent were issued to inventors living outside the United States. Significantly, organizations receiving the most USPTO patents were Japanese corporations (Hitachi, Toshiba, Canon, and Fuji Photo Film); 21 per cent of the U.S. patents were issued to residents of Japan.

Kathleen D’ell Orto of the USPTO’s Translation Section says to approve or decline a foreign application for a U.S. patent, USPTO examiners must check abroad to assess the status of the patent
there or the application’s patentability, if new. In the early 1980s a translation company was contracted by this office to supplement the in-house staff of five translators.

Since then, the Translation Branch’s workload has grown dramatically, wherein the unit faced a written volume of 9,948,079 words in FY 1990 (more than six times the 1982 demand). Full time translating staff consists of six translators (three solely for Japanese) and four contract translation firms.

NATIONAL TECHNICAL INFORMATION SERVICE (NTIS)

Of the documents NTIS acquires each year, 30 per cent (23,000) reports are from foreign sources, about one-third of those (10 per cent of the total acquisitions) are not in English. The agency performs only limited translation, because the cost must be amortized over a very small number of copies sold.

NTIS is cooperating with the Air Force Foreign Technology Division using SYSTRAN machine translation (MT) to translate Russian biotechnology information. Tests on a sample of journal articles had good results. Plans call for several Russian databases and many journal articles to be translated in the next two years.

NTIS, along with the Japan Technology Program, is active on the U.S.-Japan Task Force on Scientific and Technical Information, under the 1988 U.S.-Japan Agreement on Science and Technology. This group was formed to negotiate with counterparts in Japan to enhance the flow of Japanese STI.

DEFENSE TECHNICAL INFORMATION CENTER (DTIC)

DTIC is the Department of Defense’s central collector and distributor of STI resulting from DOD-funded research and development. Also, DTIF has Memoranda of Understanding with some NATO countries and Sweden to exchange science and technology (S&T) literature. From 7 to 10 per cent of DTIC’s collection (2,000 to 3,000 documents per year) comes from foreign sources. While some are in English, many documents arrive with no English titles, text, or abstracts to indicate the subject content. Rough translations of such documents, generated by a machine translation system such as SYSTRAN, would help to process these documents more effectively. Users would have quicker, more accurate access to foreign S&T information. DTIC plans to install an interface to connect users with SYSTRAN services at the Foreign Technology Division and with Latsec (the U.S. SYSTRAN developer) services in San Diego.

DEFENSE INFORMATION ANALYSIS CENTERS (IACS)

DTIC manages 12 IACS, which are contracted to review S&T literature worldwide. While information may appear in foreign languages, the IACS often cannot afford to obtain translations. S&T materials in Russian, German, Japanese, and French need to be translated to track S&T development. This situation reduces the scope of information that the IACS can collect and evaluate. Lower translation costs would increase the IACS’ use of foreign information.
AIR FORCE FOREIGN TECHNOLOGY DIVISION (FTD)

Located at Wright-Patterson Air Force Base, Ohio, FTD monitors S&T development abroad and maintains databases to organize and distribute records of foreign S&T. Between 50,000 to 60,000 pages of Russian text at FTD is machine-translated. Military and civilian intelligence agencies use CIRC as a primary source of foreign S&T information.

Over 20 years ago, DOD chartered FTD to develop and operate a machine translation system, SYSTRAN (through a contract with Latsec), to translate S&T documents from Russian into English. SYSTRAN’s FTD version can translate Russian, German, and French quite successfully. Recently, they began developing Japanese and Spanish MT. SYSTRAN also offers these three language pairs commercially, as well as others. Other languages under development include Japanese and Korean, as well as English into Greek and Scandinavian languages.

During Operation Desert Storm, SYSTRAN was used to translate Russian and French materials, when the information had to be translated quickly. Since no organization had an Arabic-English MT system, any Arabic materials had to be sent to outside contractors. By the time the translations were received, the war was over.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA)

NASA’s Center for AeroSpace Information (CASI) compiles and publishes bibliographies of aerospace information from Japan, the Soviet Union, and Europe. Each bibliography focuses on aerospace S&T items that appear in NASA’s public databases.

The Center contracts about 3 million words per year of translated S&T information. The international program acquires foreign S&T information by means of bilateral, trilateral, and direct agreements with foreign organizations and countries. The STI Program has exchange agreements with over 43 countries at the institutional level, and has third-party agreements with the United Kingdom, Sweden, Belgium, the Netherlands, and Italy through the European Space Agency (ESA). It also has direct, negotiated agreements with Australia, Canada, and Israel, and is negotiating similar agreements with Japan, India, and Hungary. It exchanges tapes, hard copy, and microfiche documents under these agreements and receives about 5,000 reports per year from ESA. This rate of foreign-report acquisitions is 25 per cent higher than three years ago. The Center collects and enters these journal articles, conference proceedings, and technical government reports into the NASA/RECON aerospace database.

Engineers and scientists at NASA R&D Centers access this foreign information in the NASA/RECON aerospace database on a regular, usually time-dependent, basis. If a translation cannot be completed within the allotted time, requestors tend to cancel the actions, rather than pay for information that would be delivered too late to contribute meaningfully to their projects. Clearly, there are many current and potential users in the NASA community who require translation services, but only if the information can be delivered in a timely fashion.
NASA is installing SYSTRAN software on mainframe equipment to begin testing an MT system on a global network.

**DEPARTMENT OF ENERGY (DOE)**

Through exchange agreements with foreign governments and international organizations, foreign research information is collected and made available to DOE through its Office of Scientific and Technical Information, and to the public through the Commerce Department’s National Technical Information Service. Bibliographic citations and abstracts in English of this literature are available in publications and in on-line databases, such as the Energy Science and Technology Database (EDB). Of about 200,000 document citations DOE processes annually, about 50,000 are available only in foreign-language text—predominantly Russian, German, and Japanese.

Last year there were over 200 requests for translations at each of the DOE National Laboratories, but fewer than half were actually performed. As in many other US Government organizations, DOE scientists often cancel their requests for translation due to delays or lack of funding. Over 80 per cent of the translation requests are for Russian, German, Japanese, and French. If less costly translations were available in a timely manner at less cost, DOE scientists would request and use translations more often.

**NATIONAL LIBRARY OF MEDICINE (NLM)**

Administered by the National Institutes of Health under the Department of Health and Human Services, NLM has two full-time staff who translate from Russian, German, French, Spanish, and Italian into English, and occasionally from English into French or Spanish. They also contract additional translation services, mostly for Japanese, Chinese, Finnish, and Swedish. This unit has received increasing requests for Japanese and Chinese translations in recent years. If requestors lack funds for contract translations, they either wait for the in-house translators or recall the requests.

In-house translation has decreased while dependence on contractors has increased. Yearly translation totals for contractors and staff are 800,000 to 900,000 words. When a lengthy published work is translated, a copy of the translation is sent to the National Translation Center, but no record is kept of translated articles.

**NATIONAL TRANSLATIONS CENTER (NTC)**

Under the Library of Congress, NTC provides no translation services. It serves as a clearinghouse for translations done throughout the United States. It receives donations of full-text translations and titles of works translated, which it shares with its subscribers when they request searches. Each year, NTC receives over 12,000 translations from public- and private-sector organizations. Over 75 per cent of these deposits are from American businesses, and most of the collections is S&T-oriented. NTC has a relatively small budget and a manual indexing system, but a tape distribution service for their indexes is forthcoming.
U.S. GOVERNMENT ROLE

This is not a comprehensive listing because of the limitations of space here. Almost every major government agency must deal with foreign S&T literature. Translation of STI is not managed under a government-wide strategy. While the U.S. Congress gave the Department of Commerce (through NTIS) the responsibility for handling foreign STI, the task cannot be undertaken effectively without more cooperative efforts by all involved agencies.

MT may not be equal in readability to human quality translation, but the state-of-the-art in this technology is still able to convey the subject content of articles, abstracts, and extracts. The U.S. Government has had a significant role in supporting MT development over the last 20 years, for more needs to be done.

Because language changes as rapidly as the technologies it describes, MT offers a “real-time” solution that can help American S&T development and ultimately national and industrial policies. Human translation may be more readable but may also be slow, expensive, and inaccurate, with respect to new technical terminology. A coordinated investment of effort and funding in MT by government, academia, and industry could unlock what is essentially half the world’s S&T knowledge for the American researcher.

In November 1990, with the National Science Foundation and Department of Defense agencies, the Japan Technology Program at the Department of Commerce sponsored a visit to Japan by an expert team to investigate Japanese MT accomplishments. The visit was coordinated under contract by the Japan Technology Evaluation Center (JTEC) of Loyola College, Baltimore. That team discovered that Japan is ahead of the U.S. in developing Japanese to English systems. But they also concluded that the U.S. was in a better technological posture than Japan for translating European languages into the native tongue. Simply put, the U.S. may have better automated access to more foreign languages than the Japanese.

CONCLUSION

In a global economy, American businesses must know where they can invest in product development with some certainty. The growing proportion of foreign patents, especially Japanese, suggests that foreign sources of STI could play a major role in maintaining America’s international scientific and technological position. The strength of the U.S. national security depends a great deal on knowledge created here as well as abroad. The challenge for the S&T library or information center is to develop services that will assure that their patrons have access to complete, timely, and accurate foreign STI.