

데이터베이스 이야기 ①

유경희
(사)한국복지정보통신협의회
부회장

“
CA Condensates란 자기 테이프를 구독 하여 Current Awareness(SDI) 서비스를 실시한바가 있다. 이것이 점차 데이터가 누적되어서 소급조사(Retrospective Search)도 가능하게 되었을 즈음 다이알로그(Dialog)란 기관에서는 과거분 전량으로 소급검색할 수 있는 데이터를 확보하고 있다는 사실을 알았다.
”

1978년이니까 벌써 18년전 일이다. 필자가 주동이 되어 “해외 데이터뱅크” 시찰단을 구성하여 미국과 일본의 데이터베이스 제작기관과 서비스기관을 17개소나 방문해서 시찰 및 공부를 한일이 있다. 방문한 곳중에서 지금도 잊을수 없는 곳은 미국 화학회(American Chemical Society)의 산하단체로서 화학초록 서비스(Chemical Abstracts Service, CAS)란 곳을 방문한 일이 있다. 오하이오주 컬럼버스시. 이 도시를 두번째 방문한 것은 목적이 유명한 컴퓨서브(CompuServe)를 이용해보려고 방문했던 기억이 나고, 세번째 방문은 OCLC를 보기 위해서였다. 그러나 가장 인상적인 것은 첫번째 방문때의 일이었다. 한국인 시찰단 20여명이 한꺼번에 화학정보를 어떻게 수집하고 어떻게 색인하고

어떻게 분류하고 어떻게 입력하고 있는가를 멀리서나마 보려고 일부러 갔었다.

이때 들은 바로는 직원의 총수가 2천명이 넘고 대부분이 화학분야의 학위 소지자이고 한국 사람도 7-10명인가가 근무하고 있었다. 이날밤 한국인 직원의 집에서 저녁초대를 받고 오랫동안 김치를 얻어먹은 기억이 난다.

당시만 해도 1년에 40만건이나 되는 화학문헌을 입수해서 영어로 번역 하여 분류 색인 입력을 하는 방대한 데이터베이스 공장이었다. 기억하건데 이때 이미 일본의 화학자가 몇명이나 파견되어 와서 “인턴”으로 근무하고 있다는 말을 들었다.

당시부터 한국에도 Chemical Abstracts의 자기테이프판을 수입하여 정보 검색서비스를 개시했었다. 초록만 뺀 CA

Condensates란 자기 테이프를 구독하여 Current Awareness (SDI) 서비스를 실시한바가 있다. 이것이 점차 점차 데이터가 누적되어서 소급조사(Retro-spective Search)도 가능하게 되었을 즈음 다이얼로그(Dialog)란 기관에서는 과거분 전량으로 소급검색할 수 있는 데이터를 확보하고 있다는 사실을 알았다. 그래서 시찰단을 구성하였던 것이다.

그후 한동안 이 분야에 관심을 두지 않았었는데, 작년에 이르러서 우연히 그야말로 우연히 (재)일본국제화학정보협회의 전무이사를 만나서 서로 대화를 하게 된것이다. 18년전의 CA에 관한 경험을 털어놓으니...

“유선생의 이야기는 잘 듣고 있어요”

“정말입니까? 그럴리가 없지요. 제가 오랫동안 화학정보와는 손을 떼었어요” “한국에서 처음으로 CA Condensates에 관한 서비스를 실시한것도 잘 알고 있고요. 실은 유선생께서 여러분을 모시고 CAS에 오셨을때 그때는 제가 인턴으로 근무하고 있었습니다. 그리고 오신걸 듣고 알고 있었습니다.” “세상은 좁군요”

“일본국제화학정보협회(JAICI)가 만들어진 것은 실은 일본의 화학문헌을 CA에 입력하는 대표기관으로 발족한 것이

지요. 그런데 CA에서 일본 화학문헌의 경우 5-10%만 CA 데이터로 채택되기 때문에 일본 화학자들의 불만이 이만저만이 아니지요. 그래서 만들기 시작한것이 Chem-J라는 이름의 일본 화학문헌정보이지요. 현재는 Monthly File로서 디스켓으로 배포하고 있습니다.” “한국에서는 이미 자기테이프의 도입은 중단한것 같더군요. 그러나 대부분의 한국 화학자들은 STN International을 통해서 온라인 검색하고 있는 것으로 압니다. 그래서 모두가 CA에 너무 의존하고 있다고나 할까” “다시 말하면 한국에서 혹시 일본의 화학정보가 필요하다면 Chem-J를 한번 이용해보도록 권하고 싶습니다.”

18년전에 한번 스치고간 인연을 17년후에 우연히 만나서 다시금 옛날을 회상하게 된다. 이러한 대화를 하면서 도대체 우리나라의 화학문헌정보는 도대체 누가 모여서 색인해서 입력하고 있는가? 아무도 없다. 자료수집도 분명히 제대로 되지 않고 있을 것이라서 누락된 것이 너무나 많다.

과거분의 모든 데이터를 일목요연하게 만든 데이터베이스가 없는것 같다. 외국의 망라적인 데이터 확보도 중요하지만, 우리의 망라적인 데이터화를 꾸준히 해두어야 할것이다. **QC**

