

‘대량 데이터 처리의 혁신’

-OCR 시스템의 도입과 운용-

권순덕

1. 개요

국내에 컴퓨터가 도입되어 활용되기 시작한 지 불과 5년 정도밖에 되지 않았던 1971년도에 현업에서 발생하는 대량 데이터를 처리하는 방법의 일대 혁신을 일으키려는 시도에서 OCR (Optical Character Recognition) 시스템을 도입하려는 계획이 마련되었다. 인쇄된 양식에 표기된 숫자와 마크를 컴퓨터가 판독하여 처리하게 되면 업무 수행의 신속성, 정확성, 공정성, 경제성을 획기적으로 향상시킬 수 있으리라는 기대의 소산이었다.

이런 계획의 구체적인 실현 과정으로 채택된 것이 문교부의 대입예비고사 채점과 체신부의 시외 및 국제전화 발신증 처리 작업이었다.

대입예비고사의 채점은 예나 지금이나 마찬가지로 그 처리과정의 정확성과 공정성, 그리고 발표 당일까지의 보안 유지가 관건이었다. 또한 출제 문제에 대한 사후 처리 작업으로 문항별 통계분석 처리를 통한 출제 관리에 컴퓨터의 역할이 필수적인 것으로 대두되었다.

체신부의 시외 및 국제전화 발신증은 그 처리과정에 따른 요금징수의 정확성과 신속성이 국고 수입에 막대한 영향을 미칠 수 있는 중대사로 관심을 끌고 있었다.

이런 상황에서 CDC의 OCR 936 시스템을 도입하여 대량 데이터입력처리의 자동화를 추진하여 컴퓨터 활용의 새로운 장을 열게 되었고, 그후 다양한 분야에 있어 입력 자동화나 전표

의 Turn-around 시스템을 개발하는 효시가 되었다.

2. 배경

1971년까지의 컴퓨터 도입 현황은 극히 미진한 상태였을 뿐 아니라 그 활용도에 있어서도 아주 초보적인 수준을 벗어나지 못하고 있는 상황이었다. 대부분이 숫자의 계산, 소트(sort), 프린트 등의 단순 반복 작업을 벗어나지 못하는 실정이었으므로 보다 다양하고 높은 수준의 업무에 활용할 수 있는 방안을 강구하여 컴퓨터 활용 및 보급에 KIST 전산실이 앞장서야 하는 개척자의 입장일 수밖에 없었다.

이런 상황에서 KIST 전산실에 연구비를 제공할 수 있는 조직은, 사실상 정부 및 그 관련 기관이 전부라고 해도 과히 틀린 말이 아니었다. 따라서 당시 전산실 실장 성기수 박사는 정부 기관이 타당성을 부여할 수 있는 컴퓨터 활용 방안을 제시하는 선구자 역할을 떠맡을 수밖에 다른 도리가 없었다.

따라서 문교부나 체신부의 필요성과 KIST 전산실의 선구자적 역할을 수행하기 위한 당위성이 맞아 떨어져 좋게 표현하면 과감하게, 부정적으로 표현하면 무모하게 OCR 시스템을 도입하여 컴퓨터를 이용한 자동입력의 구현을 시도하게 되었다. 필자가 이 프로젝트를 어렵게 평가하는 이유는 그 기술 자체가 보편화되지 않은 초기상태였을 뿐 아니라 시스템에 대한 이해 부족과 더불어 용지 문제, 자료작성 요원의 교육 및 관리 문제 등은 그 해결 방법이 전혀 다른 새로운 차원의 도전일 수밖에 없었기 때문이다.

3. 내용

1971년 하반기 거의 같은 시기에 대입예비고사 채점과 시외 및 국제전화 발신증 처리 업무가 진행되었다. 당시로서는 엄청난 분량인 80만매 정도의 답안지가 처리 대기하고 있었고 매일 수만매의 발신증이 쏟아져 들어왔다. 내용상으로는 별 것이 아닌 입력자료를 읽어 그대로 자기테이프에 담는 일이었으나 갖가지 끔찍한 사건들이 꼬리를 물고 발생하였다.

컴퓨터가 판독하는데 적절한 품질의 용지와 인쇄용 잉크를 준비하기 위해 시험을 거듭하여 규격에 맞추려는 노력을 거듭했으나, 실제 상황에서 벌어진 컴퓨터의 판독 거부반응은 여러 사람의 애를 태울 수밖에 없었다. 용지나 인쇄의 균일한 품질을 유지한다는 것이 당시 국내 현실로는 정말 어려운 일이었다. 또한 수험생들이나 전화국 교환원들이 수성펜을 사용하여 작성한 데이터의 품질도 사안의 중요성에 비추어 볼 때는 턱없이 성의가 부족했다.

더구나 설치된 OCR 시스템의 불안정은 작업 진행에 치명적이었다. 시스템의 작동이 불과한 시간을 넘기지 못하고, 되풀이해서 발생하는 이른바 GI(Ghost Interrupt) 때문에 작업을 중단하고 CE들이 정비하느라 24시간의 대부분을 차지하는 상황에서 관계자들의 초조함은 극에 달했고, 급기야 한밤중에 첨단 컴퓨터에 고사를 지내는 웃지 못할 희극도 벌어졌다. 그 덕분인지는 몰라도 다행히 마감시간에 맞추어 간신히 입력작업을 마쳤을 때 관련자 모두 탈진해 버리고 말았다. 컴퓨터도 달래가며 길들여서 사용해야 한다는 것을 배운 뼈저린 체험이었다.

4. 성과 및 영향

수십만 수험생의 장래가 달린 일이었고 제대로 해내야만 한다는 책임감 때문에 겨우 버티나갔다. 그 중압감으로 참으로 견디기 어려운 기간 동안 성기수 박사가 보여준 의연함에 힘입어 발작을 면하고 무사히 업무를 끝낼 수 있었다. 이제 이보다 더한 어려운 일은 없겠다 싶은 자신감과 성취감은 일생에 커다란 획을 긋는 교훈이 되었다.

작은 시스템 불량으로 CE들과 머리를 맞대고 문제를 추적해가는 과정에서 컴퓨터 작동 원리에 대한 이해가 깊어지는 소득도 올렸지만, 그보다는 사용자에게 제공되어야 할 시스템 안정의 중요성과 그것을 유지하기 위한 보수유지 방법에 눈을 뜨게 된 것이 오히려 더 큰 소득이었다고 할 수 있다.

우여곡절을 많이 겪기는 했으나 만족스러운 결과를 만들어냄으로써 컴퓨터 활용의 새로운 장을 열게 되었고, 대외적인 공신력 증대에 기여하여 결과적으로 그후 다양한 프로젝트 수행의 길을 트는 데 공헌한 바 크다고 생각한다.

5. 맺는 말

이제 20여 년 이상이 지난 현재도 당시의 수준을 벗어나지 못하는 OMR(Optical Mark Recognition) 응용업무에서 더 이상 발전하지 못하고 제자리를 맴돌고 있는 현실이 안타깝기 짝이 없다. 하루 빨리 한글 문자인식이 상용화되어 입력방법의 혁신이 이루어지기를 고대한다.