

인문학자료 정보시스템 구축과 국가지식정보 통합검색시스템 구축

이건식(한국정신문화연구원 전문위원)

1. 한국역사정보통합시스템 구축사업의 교훈

1) 목적

- 한국전산원에서 99년 10월 수립한 지식정보DB 확충 및 연계활용체제 구축 계획의 일환으로 추진
- 데이터 웨어하우스 구축 목적인 정보시스템의 통합을 통한 경비절감, 시너지 효과, 정보공동활용 달성 목적.
- 서울대 규장각, 국사편찬위원회, 민족문화추진회, 한국정신문화연구원 협의체 구성하여 시스템 구축.

2) 장점

- 지식정보 공유체계를 형성하여 중복투자방지, 시너지 효과 창출, 데이터 구축의 표준화, 정보공동활용 달성.
- 타기관의 지식을 공유함으로써 개별 기관의 지식정보 생산 능력 배가.

- 정보생산기관의 주체적인 구축으로 사용자에게 유용한 데이터 웨어하우스 구축.

3) 고려사항

- 이견으로 인한 분야별 종합정보센터 운영협의체의 운영난항이 예견됨.
- 시스템 구축에 앞선 기초연구의 필요

4) 주관사업자 기술력의 미비

- 정보시스템 구축사업 예산의 비합리적 편성으로 인한 주관사업자들의 기술투자 회피.
 - 사업자들의 이윤을 고려한 사업예산의 편성
- 해당분야 전문지식에 대한 이해 부족으로 시스템 구축에 어려움을 겪고 있음.
- 사업자 선정시 이 문제점에 대한 철저한 점검이 필요.

5) 시스템 구축 진행과정

- 정보화사업과제지원 공모(한국전산원) - 사업계획서제출(주관기관)- 과제선정(한국전산원) - 사업시행(전산원,주관기관,주관사업자)
- 위 과정이 문제점은 없음.
- 그러나 정보생산기관의 사업계획서가 대부분 개략적인 수준에 머무르는 것이 현재의 실정임.
- 간혹 준비된 기관들의 완벽한 사업계획서가 있을 수 있으나 이 경우에도 사업 시행에 들어가서는 예측하지 못한 문제점들로 어려움을 겪게 될 것임.
- 따라서 과제 선정후 본격적인 사업시행에 들어가기 전에 아주 적은 예산규모로 예비연구의 진행과정을 두어야 할 것임.

6) 데이터 웨어하우스 DB와 학술연구의 연계

- 데이터 웨어하우스의 DB 선정과 학술연구가 유리되어 있는 실정임.
- 학술연구결과를 DB로 구축하기 위해서는 또다른 학술연구가 수행되어야 함.
- 따라서 DB 구축을 전제로 한 학술연구를 개발할 필요가 있음.
- 이를 위해서 학술연구의 중심기관인 한국학술진흥재단과 협력할 필요가 있음.
- 종합정보센터의 DB 구축을 전제로 한 학술 연구의 경우 기왕의 연구비 이외에 DB 구성 연구비를 추가로 지급하는 방안 강구. 또는 이와 같은 연구를 우선적으로 지원하는 시스템 필요. 단 이것이 학술연구를 방해서는 안될 것임.

2. 인문학자료 정보시스템에 대한 인식의 전환

1) 학술 문화DB의 경제적 효과에 대한 재고

- 한국학 자료의 정보화는 예산 투입 대비 산출이 즉각적으로 이루어지지 않을 뿐더러, 예산이 투입되고 난 후 상당한 시일이 경과한 후 수익이 발생하는 특성이 있음.
- 따라서 이윤을 우선시하는 기업들은 투자를 꺼리고 있는 실정임.
- 이러한 실정을 감안하여 초기 투자는 정부에서 일정부분을 담당할 필요가 있음.
- 본원에서는 그동안 두 차례의 정보화 근로사업과, 서울대 규장각 등과 공동으로 실시하는 '한국역사정보 통합시스템'구축사업을 통해 보유 자료의 디지털화의 기반을 마련하였고, 이용자들로부터 많은 호평을 받고 있음.
- 그럼에도 불구하고 아직 상당량의 자료가 디지털화되기를 기다리고 있으나 예산과 인력이 부족한 실정임.
- 본원이 보유하고 있는 자료들은 학술적, 문화재적으로도 가치가 있기 때문에 온라인 서비스를 할 경우 연구자와 일반인들에게 많은 도움이 될 것임

2) 기구축 DB에 대한 추가 구축 및 유지보수 예산 지원

- 한국정신문화연구원에서 구축한 '한국학전자도서관'과 같이, 국민의 세금으로 구축된 자료가 사장되는 것을 방지하기 위해서는 일정액의 유지보수 예산을 지원해 줄 필요가 있음. 현재 본원에서는 이에 대한 예산을 전혀 지원받지 못하고 있음.
- 한국학 자료 서비스의 질적 향상과 한국문화의 세계적 확산을 위해서도 이것은 필수적인 과제임.
- 많은 공적 자금을 들여 만든 DB를 생명력 없게 만드는 것은 결과적으로 공적 자금의 낭비를 초래함.
- 한국학 자료의 정보화는 우리 나라가 강점을 가질 수 있는 유일한 것이라 해도 과언이 아님. 따라서 정부차원에서 소액이나마 정규예산을 편성하여 지속적으로 지원하든지, 아니면 효과가 좋은 일부 사업에 한하여 프로젝트를 지속적으로 추진할 필요가 있음.

3) 인문학 DB구축의 정책 수립 및 표준화 연구

- 현재 인문학 DB는 기관별로 독자적으로 수행되고 있어서 중복 개발로 인한 투자의 효율성 문제가 일부 있음.
- 또한 독자적인 개발은 표준화된 시스템의 개발을 막고 있음.
- 따라서 초기단계에서 인문학 DB개발과 관련된 전반적인 정책을 마련하고 서비스의 표준화를 연구할 예산을 책정하고, 주요 기관별로 담당자를 선정하여 협의체를 구성하고 국가적으로 지원함으로써, 결정된 표준화안을 확산시키는 조치가 필요함.

3. 국가지식정보 통합검색시스템 구축의 과제

1) DB구축사업의 수행현황 : 비연계적, 비효율적 추진체계

- DB 선정은 고립된 주관기관들이 주로 선정하여 기관별 중복 DB 구축이 발생할 소지가 많으며 기구축 DB들이 상호 연계하여 상승효과를 낼 수 없는 실정임.

- 구축되는 DB들이 통일된 DB 스키마를 가지고 있지 못함. 그리하여 다른 시스템과 자료교환을 상당수 할 수 없는 실정임.
- 우리말 지식정보를 효율적으로 구축할 수 있는 핵심기술과 재료들이 만족할 만하게 개발되지 않고 있음. 더더욱 개발된 기술과 재료마저 전사회적으로 공유되지 않고 있음. 예를 들면 전자통신연구원 연구들은 지나치게 공학적인 기술개발에 치중되어 있음. 내용적인 기술개발도 매우 중요함.
- 컴퓨터 공학자들만이 참여한 협의체에서 우리말 표준코드체계가 정립되어 우리말로 표현된 지식정보를 만족할 만하게 구현할 수 없는 실정임. 따라서 컨버전 등 재입력작업으로 인한 국가 예산의 낭비가 심함. 현재 국립품질기술원과 국립국어연구원이 이 문제에 관여하고 있으나 두 기관간의 뚜렷한 역할 분담이 구획되지 않았음. 그리고 한자코드를 다루는 국립기술품질원 산하 KIRG 위원회는 인문과학자가 대거 참여하고 있으나 한글코드를 다루는 KSG2 위원회는 인문과학자가 거의 참여하지 못하고 컴퓨터 공학자들만이 참여하고 있는 것이 문제임.
- 이상과 같은 개별 분야의 문제이외에 개별분야 상호간의 교류가 있지 않아 각각이 서로 다른 이상을 추구하는 것이 더 큰 문제임.

2) 연계적, 효율적 DB 구축을 위한 대책

(1) DB선정분야

- DB 선정 및 심사기구 구성
 - 정부 중앙부처 중 하나를 지정.
 - 통합선정위원회, 지식분야별 또는 DB 목적분야별 선정위원회, 유사기관 선정협의회 구성
- 목적분야별 선정위원회(1안)
 - 생활 DB : 정부기관, 공공단체, 일반단체
 - 학술·문화DB : 정부기관, 국립대학, 사립대학
- 지식분야별 선정위원회
 - 현재 지식정보연계사업의 분류 수용
- 중복방지를 위한 기 구축 DB 검색시스템 개발

- 기 구축 DB의 활용을 위한 시스템
- 신규 DB 구축시 중복 여부를 검증하는 도구
- DB 구축 승인요청시 이 시스템 인증서 첨부

(2) 통일된 DB 스키마의 채택

- 정보시스템의 분류와 표준화
 - 기 구축 DB를 시스템의 기능으로 분류
 - 시스템별 표준 DB 스키마 적용
 - 상이 시스템과 연계 고려
 - 정보시스템을 우선 기능상으로 분류하고 목적별로 하위 분류
- 정보시스템 기능상 분류
 - 목록시스템, 내용정보시스템 등
- 정보시스템의 목적별 분류
 - 생활 정보시스템, 학술 정보시스템, 문화 정보시스템 등

(3) 핵심기술과 자료의 개발

- 도구개발의 장려
 - 고속 한적입력기 개발
 - 효율적인 입력도구개발(한자의 문자 인식 도구 등)
 - 한글처리 알고리즘
 - 한글 형태소해석
- 자료의 보급
 - 색인어 구성을 위한 용어사전
 - 전자국어사전의 보급
 - 한글 및 한자 옛글자의 Font 데이터 등

(4) 문자 코드의 표준화

- 국립국어연구원에 표준 코드 제정 기능 부여
 - 국립품질기술원과의 협의 필요
- 우리나라에서 사용되는 문자의 수집 발굴 노력 필요

국가지식정보 통합시스템



상호보완적 DB구축 표준적·연계적 DB스키마 생산적 핵심기술 생산적 문자표준코드

정보통신부/ 한국전산원
S/W INFRA

H/W INFRA



DB선정위원회

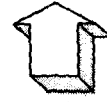
- ▶ 통합/분야별/목적별/유사기관별
- ▶ 한국데이터베이스진흥센터 관장



- ▶ 전자통신연구원
- ▶ 벤처기업
- ▶ 학술기관



- ▶ 국립국어연구원
- ▶ 비표준문자등록센터



3) 종합정보센터의 육성방안

(1) 종합정보센터의 역할과 기능

- 해당분야 DB 선정
 - 중복투자 방지
 - 최상의 시너지를 효과를 얻을 수 있는 DB 선정
 - 중요 지식정보자원 지정 후보 선정

- 학술연구의 선도 : 종합정보시스템에 필요한 DB를 연구과제로 선정

○ 해당분야 데이터 웨어하우스 구축

- 이질적인 DB를 통합하는 메타DB 구축
 - 이질적인 DB를 강제적으로 통합할 경우 진정한 지식 생산을 달성할 수 없음.
 - 이질적인 DB의 장점을 살리되 메타 DB에서 이질적인 DB를 통합하여 DB 연계로 시너지 효과를 창출
 - 기구축 DB를 통합 메타DB로 흡수

○ 데이터 웨어하우스 구축 경험의 축적과 전파

- 과거의 실수가 재현되지 않도록 구축 경험의 공유
- 정보생산기관의 정보화 실무자 정보화 교육을 통한 정보화 역량의 배가와 과거 구축 경험의 공유
- 경비절감과 정보의 공동활용을 위한 다양한 층위에 잇어서의 표준화

○ 해당분야 데이터 웨어하우스 구축을 위한 조정기구의 역할

- 종합정보센터에서 정보생산기관의 데이터를 흡수하여 직접 DB를 구축하는 것은 금기시해야 함.
- 이 경우 종합정보센터 데이터웨어하우스의 DB들이 생명력 없는 시스템으로 전락하게 됨.
- 구축 DB들에 생명력을 줄 수 있는 기관은 정보생산기관임.
- 따라서 종합정보센터에서는 정보생산기관의 정보화 능력을 배가시키는 데에 주안점을 두어야 함.
- 다만 중복투자방지와 시너지 효과를 위해서 정보생산기관들의 DB 구축 방향을 조정하고 각 수준의 표준화를 유도하는 역할을 수행해야 할 것임.

○ 데이터 웨어 하우스 구축 담당 인력의 양성

- 해당 분야 전문가의 정보화 역량 부여를 위한 재교육
 - 전문분야 전문가의 DB 구축과정에 대한 이해

- DB 구축 전문가의 전문분야에 대한 이해
- 통합적 정보화 역량을 가진 인력의 양성
- DB 구축과 전문분야 양 쪽의 소양을 가진 인력의 양성

(2) 종합정보센터의 조직구성

○ 상부조직

- 해당분야를 대표할 수 있는 기관을 선정하여 기본적인 관리와 운영을 맡기되
- 해당분야 주요기관들이 참여한 협의체를 구성하여 종합정보센터의 정책 결정

○ 실무조직

- 종합정보센터 및 종합보시스템의 관리와 운영
- 전문분야 연구원 배치
 - 해당분야 하위 전공별 연구원 배치
- DB 및 기술분야 연구원 배치
 - 네트워크, DB구축, 하드웨어 분야 연구원 배치

- 국가지식정보종합센터에도 DB 구축 전문가만 배치하지 말고 해당 분야 전문가들을 배치해야 할 것임.

<참고문헌>

1. 정보통신부, 知識情報資源管理 基本計劃(案), 2000년 9월
2. 김현, 국가 지식관리의 추진체계 모델, 정보공동활용 심포지엄 지식기반국가의 비전과 실천전략, 2000.11.2.
3. 신상철, 분산된 국가지식정보 통합검색 시스템 구축 방안, 정보공동활용 심포지엄 지식기반국가의 비전과 실천전략, 2000.11.2.
4. 한국정신문화연구원, 1999년도 정보화근로사업 한국학전자도서관구축사업 완료보고서, 2000년 2.1.