

傳統醫學 분야 知識情報化 사업과 古文獻 DB 구축

安相佑¹⁾

The web service system of traditional knowledge about
oriental medicine and database of old traditional documents

Ahn, Sangwoo²⁾

Traditional Knowledge[TK] is a concept which summings up all the knowledge of traditionally inherited, includes traditional medicine, therapy, food and environments. Such traditional knowledge has been perished and neglected through industrial development. We, Korean Institute of Oriental Medicine, promotes digitalization of this traditional knowledge to support and maximize its application, principally in Traditional korean medicine.

We can get opportunity of reorganize the traditional knowledge to more practical and available form by this project. And also we introduce briefly its meaning and worth, and seek for application to the study of medical history.

Key Words : Traditional Knowledge[TK], Database, Oriental medicine, Medical history, Traditional korean medicine

傳統智識(Traditional Knowledge, TK)은 전통적으로 계승되어 온 모든 지식을 총망라하는 개념으로서, 傳統醫藥, 傳統治療技術, 食品, 農業, 環境 등에 관한 지식뿐 아니라, 문학, 음악, 미술 등을 포함한 것이다. 전통지식의 특성상 상당수가 개인 또는 특

1) 韓國韓醫學研究院

2) KIOM : Korean institute of oriental medicine

정집단이 소유하는 것이 아니라 토속사회의 지식인 관계로 지식의 소유자의 설정이 곤란하고 또한 그 가치에 대한 인식이 부족하여 많은 전통지식들이 급속하게 멸실되고 사장되어왔다. 하지만 일련의 국제적 사건들을 계기로 이러한 전통지식에 대해 새로운 시각으로 바라보게 되었고 이러한 전통지식이 한 개인이 소유한다는 개념을 떠나 국가적 차원으로 보호하고 그 권리를 주장해야 된다는 추세 속에서 각종 국제적 논의가 진행되어 오고 있다. 2001년 12월 2차 WIPO 정부간위원회에서는 전통지식을 검색 가능한 선행기술에 포함시킬 수 있는 새로운 기준 개방에 대해 논의하였고, 그 결과 국제적인 TK Datedase 및 전자도서관 구축이 전통지식의 선행기술 활용방안으로 제시되었다. 이러한 추세에 따라 기존에 구축하여온 中國의 TCM DB, 印度의 전통지식디지털라이브러리(TKDL), 베네주엘라의 Biozulua DB등을 모델로 하여 각국이 자국의 지식정보 자원들에 대한 디지털도서관의 구축에 박차를 가하고 있다.

유구한 역사를 가진 우리나라의 경우 수많은 전통지식들을 보유하고 있고 그 가치를 전 세계적으로 인정받고 있다. 하지만 뒤늦게 국제적 TK논의에 참여하여 아직까지는 법적으로나 제도적으로 부족한 현실에서 정부의 주도하에 각종 전통지식정보 자원들에 대해 그것들을 보호하고 영구한 자료로서 보관하고자 하는 작업들이 진행 중에 있다. 특히 디지털 기술의 발달로 멸실위기에 있던 많은 지식정보들이 보존과 동시에 대중화가 가능하게 되었다. 이러한 대표적인 사례가 『朝鮮王朝實錄』이나 『八萬大藏經』 등의 고문헌 자료들이나 사라진 역사유물 유적들의 CG작업들로 전자의 경우는 원문에 대한 DB구축과 상용화로 많은 이들의 지적 욕구를 충족시켜 주고 있다.

특히 이러한 전통지식자원 중 우리나라는 타 국가들보다는 傳統醫藥부분에서 비교우위에 있다고 평가되고 있다. 하지만 지금까지 그것들에 대한 권리화에 소홀하였고, 또 실제로 권리를 주장하기에는 많은 난관을 가지고 있다. 한의학분야의 지식들은 다른 분야에 비하여 문헌을 통한 전승이 많이 이루어져 왔고 고문헌에 대한 비중이 다른 어느 분야보다 높다고 할 수 있을 것이다. 따라서 韓醫學 분야의 古文獻의 보존문제는 곧 韓醫學 전통지식의 보존과 직결된다고 할 수 있을 것이다. 이에 한의학연구원에서는 개원 초부터 전통의약분야의 지식정보자원의 보호에 대하여 여러 가지 법적 제도적 방안의 연구와 동시에 그것들을 영구히 보존할 수 있는 디지털화 작업을 서두르고 있다. 그 결과 『醫方類聚』와 『鄉藥集成方』의 진산화, 『醫部集成』이나 『濟衆新編』 등의 상용화 및 사라져버린 『御醫撮要』, 『備豫百要方』의 복원 등 많은 가시적인 성과가 있었다. 이러한 작업을 통해 실용성을 가진 전통의약부분의 지식은, 향후 객관화와 과학화를 거칠 수 있는 기회를 제공받을 수 있을 것이다.

최근에는 이러한 노력들의 가치를 인정받아 정부지원으로 한의학 지식정보자원 디

지털화 사업을 수행하게 되었고 이를 통해서 전통의학 분야 지식정보화 사업의 의의와 韓醫古文獻 DB의 가치에 대해 평가를 할 수 있는 기회가 되었기에 간략하게 소개하고자 한다.

I. 傳統醫學 知識의 情報化 必要性

- 한의학은 수천 년을 통해 이론과 기술이 전수되어 온 체계화된 우리나라 전통 의학으로써 세계 여타 나라에서도 그 유래를 찾기 어려운 고유기술이면서 현대 의학이 당면하고 있는 제반문제를 해결할 수 있는 분야로 그 유효성이 설득력 있게 부상되고 있음.
- 세계적으로 전통의학 관련 시장규모는 수십조 원에 이르고 있으나 아직까지 한국의 전통의학은 그 효용가치가 큰 의미를 지니고 있음에도 불구하고 정보의 활용과 시장 내에서의 비중은 극히 낮은 상태임.
- 한의학 분야에서 학술정보를 공유, 제공하는 서비스는 미비한 수준이고, 그나마 운영면에서도 체계적으로 분류되지 않아 한의학 정보서비스 이용이 어려운 실정임. 따라서 한의학 분야의 정보환경 조성이 시급한 실정임.
- 특히 국민들의 한방정보에 대한 관심이 고조되고 있는 가운데, 전통적인 한의학의 고문헌 및 치료기술에 대한 정보의 체계적인 수집과 정리, 그리고 디지털 이징과 정보서비스 시스템 구축이 절실히 요구되는 상태임.
- 우리나라의 경우 한의학이 정식적인 의학체계에 포함되어 있고, 인터넷과 통신망도 발달되어 있어, 국제적으로 충분히 주도적인 경쟁력을 가지고 있음.
- 1차적인 자료의 보존, 복원 및 활용도 제고를 위한 체계화된, 포괄적인 정보관리가 절대적으로 필요하며, 그를 기반으로 한 학술적인 연구의 활성화를 도모해야 하는 시점임.

1. 韓醫學 知識情報資源의 種類와 價値

1) 韓醫學 知識情報資源의 分類

- 고문헌/ 고의서, 잡지류 등
- 치료기술/ 침구, 설진 등
- 임상정보/ 치료경험 등
- 약재정보/ 화상정보 등
- 동양 의학 정보/ 중국, 일본의 의서 및 치료기술 등
- 학술정보/ 논문, 연구보고서 등

2) 保存價値

- 고문헌 보존의 문제
- 치료기술 보존의 문제/ 경험방의 측면에서 한의사 사후 전달 체계 단절
- 임상정보, 약재정보 등 시의적으로 변화하는 내용을 보존해야할 필요

3) 活用價値

- 난해한 고문헌의 정확한 검색 및 열람 제공 필요
- 특히 처방, 병증, 약재 등의 통합 검색 등 활용도 제고
- 흩어져 있는 치료기술, 임상정보, 약재정보 등의 포털사이트 구축 필요
- 한중일 3국의 학술정보 및 동양의학 정보의 교환 필요

II. 韓醫學研究院 情報化 事業概要

1. 主要事業內容

○ 韓醫學 古典文獻 디지털라이징 및 DB 구축 (XML)

- 한의학 고전문헌 원문 텍스트(XML) 구축
 - 한의 고전문헌 원문의 XML 구축(Full-Text)
 - 지식정보자원관리 고전문헌 표준 DTD 및 XML 작성지침 준용
- 한의학 고전문헌 원문 이미지(TIFF, JPEG) 구축

- 한의 고전문헌 원문의 TIFF, JPEG 제작(이미지)
- 사용자 편의성을 고려하여 원본과 최대한 유사한 상태의 고품질 유지
- 한의학 고전문헌 메타 DB 구축
 - 한의 고전문헌 서지정보 메타데이터 구축(서명, 저자, 발행년도, 목차 등)
 - 고전문헌 메타데이터 표준 및 지침 적용
- 한의학 고전문헌 해제 정보 구축
 - 한의학 고전문헌 관련 전문가에 의한 해제 작성, 입력, 검증
 - 고전문헌 163종, 의방유취 총론, 병증문 91문, 인용서목 153종 해제
- 韓醫學 治療技術의 디지털화 및 DB 구축(XML)
 - 한의학 우수경험방 메타데이터 구축(질환별, 처방별, 사용약재별, 처방자, 특색, 사용시기 등)
 - 한의학 표준 처방을 기본형식으로 구축
 - 한의학 치료기술 관련 각종 정지화상 정보 DB 구축
 - 한의학의 치료기술 가운데 선별하여 디지털화
- 韓醫學 知識情報資源 서비스 시스템 구축
 - 한의학 지식정보 검색 서비스를 위한 시스템 구축
 - 구축 DB의 서비스를 위한 색인 DB구축 및 기타 지원
 - 한의학 정보에 최적화되고, 기 구축 자료와 당해사업자료의 통합 검색 기능 구현 및 검색엔진 개발
 - 사용자 이용편의를 고려한 최적의 검색기능 및 홈페이지 디자인 제공
 - 한의학 정보자원의 활용성 향상을 위한 홍보기능 개발
 - 한의학 지식정보 데이터베이스 관리 시스템 구축
 - 한의학 지식정보 데이터베이스 관리를 위한 도구 개발
 - 신규자료 및 기구축 자료에 대한 입력, 수정작업이 용이한 운영시스템의 개발
 - DB 구축을 위한 시스템 구축
 - DB 구축의 효율성 제고 및 작업 진도 관리를 위한 한의학 지식정보자원 관리 시스템 구축
 - 기타 DB 구축을 위한 응용 소프트웨어 개발
 - 서비스 관리 및 서비스 이용현황 등 통계 기능

2. 디지털화 對象資料 內域

대상 자료	세부 내용	자료 형태	자료량
韓醫學 古典文獻	한의학 고전문헌 원문 텍스트 (醫方類聚, 簡易辟瘟方 등 160여종)	Text(한문)	87,000면
	한의학 고전문헌 해제 (醫方類聚 등 300여종)	Text(국한문)	1,800매
	한의학 고전문헌 원문 이미지 (家庭救急方, 本草幽涵 외 160여종)	Image(흑백)	50,000면
韓醫 傳統治療技術	한의 전통치료기술 텍스트 (韓方治療技術 외 20종)	Text(국한문)	5,000면
	한의 전통치료기술 이미지 (鍼灸經絡經穴圖譜 외)	Image(칼라포함)	10,000면

3. DB 구축 방향

- 분산 DB기반 통합검색시스템 구축에 적합하고 자동백업에 편리한 DBMS를 구성한다.
- 유니코드 3.0을 지원할 수 있도록 DB를 구성한다.
- DB구축 대상자료는 관련기관(국립중앙도서관 외)에 구축된 자료(DB)와 중복되지 않아야 한다.
- DB구축은 저작권 문제가 해결된 자료를 가지고 디지털화를 추진하며 구축된 DB는 이 용자에게 무료로 서비스를 한다.
- 국가정보통합검색시스템 및 관련기관 종합정보센터의 연계 검색을 위해 지식정보자원 관리사업의 메타데이터 표준을 참조·적용하며, 한의학분야 메타데이터 표준을 제정, 적용함으로써 정보공동이용을 위한 데이터 호환성 확보
- 메타데이터 및 해제, 목차, 원문 XML, 원문 TIFF, JPEG 등에 대한 명확한 입력 기준 및 상세지침을 이용하여 데이터 구축의 일관성을 확보함.
- 유사사업공정을 참고하여, 시험 공정운영을 통해 효과적인 작업단위를 구성하고, 작업 단위별 업무 범위와 내용을 명확히 하여 작업 능률 극대화
- 중복 및 오류데이터의 정확한 확인과 수정을 위해 다단계 수동 검증, 별도의 품질보증 팀 운영 등 품질보증 활동 강화
- 자료별 특성을 고려하여 원활한 XML 데이터 서비스가 가능한 방안을 마련한다.
- 자료의 오류를 검사하기 위해 자료별, 작업공정별 효율적인 검증방법(Tool, 인력 등)을 이용하여 검증을 실시한다.

4. DB 構成要件

대상 자료	DB 구축 요건	자료 형태	비 고
韓醫學 古典文獻	<ul style="list-style-type: none"> - 전문을 입력하고 XML 태깅을 한다. - 각 서명의 목록은 서명별/저자별로 분류될 수 있어야 한다. - 기본 검색 외에 확장 검색 기능을 추가하여 제공한다. - 원문의 주석은 [주]라고 표기하고, 파란 색으로 본문보다 한 급수 작게 디스플레이하며, 주석만을 대상으로 검색 가능하게 한다. - 본문 내 이미지는 본문과 함께 디스플레이 하되 이미지 포함 문자는 텍스트 검색과 동일하게 검색할 수 있도록 한다. - 고전문헌 텍스트 중 이미지로 처리한 부분에 글자가 있을 경우 태그를 부착하여 글자를 입력한다. 	Text	한문 자료
	<ul style="list-style-type: none"> - 해제는 해당 자료에 연결, 별도 창으로 디스플레이 한다. 	Text	국한문 혼용 자료
	<ul style="list-style-type: none"> - 원문 이미지는 텍스트의 제목 단위로 링크, 별도 창으로 디스플레이한다. - 기타 정의하지 않은 사항은 주관 기관과 협의한다. 	Image	사진 자료
韓醫 傳統 治療 技術	<ul style="list-style-type: none"> - 전문을 입력하고 XML 태깅을 한다. - 각 자료의 목록은 서명별/저자별로 분류될 수 있어야 한다. - 우수 경험방 메타 데이터 구축(질환/처방/약재 등) - 한의학 표준 처방을 기본 형식으로 구축 - 기타 형식 요건은 주관 기관과 협의 후 진행 	Text	국한문 혼용 자료
	<ul style="list-style-type: none"> - 치료 기술의 내용 분류에 따라 대상 자료를 디지털화한다. - 자료 검색 시 유관 텍스트에 링크되어 검색, 열람되도록 한다. - 기타 형식 요건은 주관 기관과 협의 후 진행 	Image	사진 자료 (컬러 포함)

5. 대상자료 통합 · 연계활용

- 한국한의학연구원의 학술논문 메타 데이터 110,000건, 학위논문 메타 데이터 3,000 건 등의 기 구축 DB를 XML로 변환하여 통합 검색되도록 연계
- 관련기관에 기 구축되어 있는 서비스 중인 정보와의 연계 방안 제시
 - 국가지식정보통합검색시스템 및 KISTI와의 정보 연계
 - KERIS, 한의학 관련 12개 대학과의 링크 정보

6. 新出字 管理方案

과 정	내 용
漢字入力	○유니코드 3.0을 기본 코드로 하여 입력 ○이형자는 기본자형으로 입력하고 이체자는 자형 그대로 입력 ○유니코드 내에 포함되어 있는 이체자는 그대로 반영
新出漢字 여부파악	○신출한자로 추정되는 한자에 대하여 신출 여부를 확인 - 입력 중 유니코드 3.0에 정의되지 않은 문자 - EXT_A 영역 검색 - EXT_B 영역 검색(해당 코드값을 반영한 인코딩 스키마 적용) ○한국역사정보통합시스템 구축사업(역통 사업)에서 기 등록된 신출한자와 비교하여 중복되지 않게 처리한다. ○중복된 한자의 경우 역통 사업과 연계하여 통합, 관리한다. ○유니코드에서 발견되지 않는 한자는 임시 식별 코드 부여, 폰트 제작
新出漢字 여부확인	○신출한자에 대한 기준과 최종판단은 주관기관에서 수행 ○신출한자는 기본적으로 자형 이미지, 부수, 획수 정보를 관리
주관기관 검수	○발생한 신출한자에 대하여 주관기관 신출한자관리팀에서 신출자 여부를 검수
임시 식별 코드 부여	○발생한 신출한자에 대하여 임시코드를 부여하고 신출한자 발생을 통제
확정 코드 부여	○임시코드로 부여된 신출한자를 통합하여 하나의 목록으로 만들고 확정된 코드를 부여
폰트제작 및 데이터 갱신	○ 주관기관 검수가 끝난 한자는 코드를 부여하고 트루타입 폰트 제작 ○ 제작된 폰트로 데이터 수정 반영 ○ 역통 사업의 신출한자와 호환되도록 관리

7. 최종목표

1) 조직목표

- 한의학 관련 지식정보 자원을 효율적으로 통합 관리, 유통하는 한의학 지식정보 공동활용체제 구축
- 한의학 관련 기초자료부터 정보수요자의 자료 획득까지의 전과정을 유기적으로 연결하는 한의학 지식정보자원의 Single-Gate 서비스 구현
- 유관사업을 추진하고 있는 공공기관 및 민간업체와의 정보공동이용을 통한 중복구축 방지 및 정보화투자 효과 극대화

- 아직까지 미흡한 국내 의학정보망의 첫 구축을 통하여 올바른 의학정보의 신속, 정확한 제공 및 연구원내 정보 서비스 체제의 구축

2) 정보화 목표

○ 중장기 목표

- 한의학 관련 정보의 유형, 생산기관, 분야, 보유기관 등에 관계없이 모든 국내외 한의학 정보에 대한 통합 검색 및 원문 제공
- 국내외의 대학 및 학회, 연구기관 등 유관기관, 민간업체와의 연계 확대를 정보공동활용으로 서비스 범위 확대, 정보에 대한 접근도 및 활용도 제고
- 대학 및 학회, 연구소 등에 대한 지속적인 정보화지원을 통하여 한의학 정보 생산기관에 대한 정보화 분위기 조성을 유도하고, 이를 연계 이용하는 메타데이터 중심의 공동이용체계 조성

○ 단기 목표

- 1차적인 자료, 즉 한의학 고문헌과 치료기술의 정보화에 역점을 두며, 지식정보의 디지털화 과정의 시스템화로 시간과 비용 절감, 신규 발생 정보의 신속한 제공 기반 마련
- 대학 및 유관기관, 민간업체를 통해 구축된 한의학 관련 정보의 메타데이터 구축을 통한 연계활용 및 서비스 범위 확대
- 한의학 연구자는 물론, 일반 국민들도 활용할 수 있는 차별화된 콘텐츠 서비스를 통한 한의학 보건의료 정보의 활용도 제고

III. 韓醫學 古文獻의 DB構築

1. 韓醫學 古文獻 디지털화의 意義

1) 學術的 意義

- 引用書目들을 통한 실건의서의 복원
- 板本간 비교를 통한 정확한 교감
- 효율적인 검색을 통한 학문경쟁력 제고
- 자료 탐색에 소요되는 시간 및 비용 절감

2) 産業的 意義

- 약재정보의 디지털화를 통한 신약제의 개발
- 치료기술정보의 디지털화를 통한 한의학 산업의 활성화

3) 社會的 意義

- 정보화 소외 분야의 해소
- 과학기술/ 의학 분야의 정보화 촉진
- 전통의학과 첨단 정보기술의 접목

2. 古文獻 DB의 活用 방안

1) 古文獻의 價値

- 민족문화의 우수성의 선양: 국민적 자부심
- 뛰어난 치료술과 의학정보
- 서양의학과 구별되는 선조들의 지혜

2) 古文獻의 現在

- 체계화된 관리체계의 필요성 대두
- 고문헌이 담고 있는 우수한 정보들이 공유되지 못하여 폄하되는 실정
- 고의서에 대한 물이해의 보편화

3) 古文獻의 DB화 目標

- 제대로 된 해제의 작성/ 기존 국립중앙도서관, 규장각 해제 등의 문제점
- 집필의 방향
- 해제와 고문헌 DB의 학문적인 기대효과

4) 古文獻 DB의 活用 方案

- 고문헌 가치(보존, 활용)의 극대화
- 전승 중단된 의서의 복원
- 처방, 치료기술 등의 연구 및 발전에 도움
- 약재정보, 체질의학 등의 연구 및 발전에 도움

3. 古文獻의 解題

1) 2003 韓의학 知識정보자원 디지털화 사업 古文獻 解題

순 번	구 분	해 제 면 수
1	古文獻 對象資料 163종	1,141 면
2	醫方類聚 總論 및 病症門 91문	353 면
3	醫方類聚 인용서 153종	306 면
계	총408건	1,800 면(200자 원고지 기준)

2) 解題作成 기본 틀

아래 그림은 현재 韓의학 知識정보 디지털화 사업에서 적용한 고문헌 해제의 작성 원칙이 담겨진 기본 틀을 보여주고 있다.

```

<해제>
  <해제서지>
    <목록번호> . . . </목록번호>
    <제목> . . . </제목>
    <제목_한글> . . . </제목_한글>
    <작성자> . . . </작성자>
    <작성일> . . . </작성일>
  </해제서지>
  <본문>
    【정의】
    【서지사항】
    【체제 및 내용】
    【자료적 특성 및 가치】
    【소장현황 및 영인관계】
    【참고문헌】
    【형태사항】
  </본문>
</해제>
    
```

3) 解題 本文의 構成原則

순서	포함 내용	설명 방식	비고
정의	책의 정의를 간략하게 명사구로 제시	例) 正祖의 命에 의하여 1790년(正祖 14)에 具允明 등이 편찬한 『增修無冤錄大全』을 徐有隣 등이 언해한 法醫學書.	필수
서지사항	편찬시기, 편찬자, 편찬과정, 편찬목적, 이본 등에 관한 정보를 제시	책의 간행에 관한 꼭 필요한 정보는 卷首·卷末의 註記, 刊記나 筆寫記 內賜記, 序文, 跋文, 印記 등을 직접 인용부호를 써서 원문을 그대로 제시한다. 단, 서문이나 발문이 너무 긴 경우에는 서지와 관계 없는 부분은 중략표시를 하고 생략할 수 있다.	필수
체재 및 내용	책의 체재 및 내용을 소개	서, 발, 목차, 범례 등을 포함한 서적의 일반적 특징과 주요 내용을 간략히 서술한다.	필수
자료적 특성 및 가치	자료적 특성 및 자료적 가치를 설명	일반 서지와 달리 역사적 관점에서 의학적 가치를 논술하고 기존 의방서와의 차이점과 특징을 기술한다.	필수
소장현황 및 영인관계	원본의 소장현황과 영인관계를 제시	주요 판본의 소장기관 혹은 전본의 출판사를 기록하여 연구에 필요한 기초정보자료로 제공	선택
참고문헌	해제를 작성하는데 이용된 기존 해제 및 논문 서명 등을 기재	집필자의 작성시 참고문헌 혹은 참조 연구논문을 기록	선택
형태사항	크기, 재질, 서체, 어미 등 형태적 요소와 관련된 사항 기술	고문헌 서지학 전공자의 내용감수와 교정을 거친다.	선택

4. 韓醫學 古文獻 主題分類 (試案)

다음은 한의학 고문헌 DB의 구성과 이용자 편의를 도모하기 위한 대상자료의 주제분류를 정하기 위하여 수 차례 전문가 자문회의의 선정 조정과정을 거쳐 설정한 시안의 내용을 보여주고 있다. 현재 14개의 소주제별 분류의 항목간 등위는 없으며 소주제 아래 임의 배열한 내용주제는 대상자료의 분류 배치를 위한 설정으로 실제 검색 서비스에서는 제공되지 않는 항목이다. 항목 아래 예시된 대상서는 정보화 사업에 주요 대상자료로 선정되었거나 향후 대상자료에서 우선하여 열거하였으며 몇몇 항목에서는 대상서가 아니라 할지라도 해당 항목의 이해도를 돕기 위해 임의로 배치한 곳도 있다.

1) 基礎理論

- 內難醫經: 新刊補註釋文黃帝內經素問, 合注素問靈樞, 八十一難經
- 醫經註釋: 素問大要, 素問節要, 素問句讀俗解
- 基礎理論: 病源, 醫易, 運氣, 藏象 등 의학이론서류
 張仲景五臟論, 五臟六腑圖, 小兒巢氏病源候論, 素問立式運氣論奧
- 後世醫論: 格論, 統論, 東垣十書, 扶陽論, 醫宗損益
- 傷寒金匱: 傷寒類要, 入門傷寒, 活人心

2) 診斷/脈法

- 脈法: 纂圖方論脈訣集成, 仁齋直指方論醫脈眞經, 脈訣理玄秘要, 幕經
- 四診類: 察病諸症, 察病用藥訣

3) 本草藥物

- 本草: 湯液本草, 食物本草, 本草類涵要領, 本草精華
- 藥性歌: 損益本草, 入門醫鑑本草總括
- 現代藥理: 鮮漢藥物學

4) 經驗方書

- 肘後方, 太平聖惠方, 千金方, 聖濟總錄, 醫家必用
- 四醫經驗方, 良方金丹, 經效方, 丹谷經驗方抄
- 單方類: 單方備要, 單方藥, 博施良方, 本草附方便覽, 通俗藥處方

5) 綜合醫書

- 醫方集略, 編註醫學入門, 名方類證醫書大全
- 東醫寶鑑: 醫鑑重磨, 春鑑錄, 醫鑑集要, 廣濟秘笈, 濟衆新編, 仁濟志
- 鄉藥醫書: 鄉藥集成方, 醫林撮要, 青囊訣, 一金方, 舟村新方
- 方藥合編: 醫方合編, 醫方活套, 漢方醫學指南

6) 防疫/救急

- 瘟熱病論: 分門瘟疫易解方, 溫疫論, 簡易辟瘟方, 辟疫新方, 鳳城新方
- 瘡疹: 治疹指南, 瘡疹集, 痘疹集要, 麻疹奇方, 經驗痘方, 痘科彙編

7) 婦人/小兒

- 婦人大全良方, 達生編, 產方隨錄, 生產秘方, 胎產要錄
- 及幼方, 小兒醫方. 丁茶山小兒秘方, 經驗秘方小兒保鑑, 漢方醫學小兒專科

8) 四象體質

- 體質醫學: 東醫壽世保元, 東醫壽世保元甲午舊本
- 臨床/醫論: 普濟演說, 格致藁. 東武遺稿, 明善錄, 四象金櫃秘方

9) 皮膚/外傷

- 外傷科類: 治腫指南, 外科精要, 醫腫金鑑, 若山好古腫方撮要
- 眼耳鼻咽喉科類: 醫眼方, 銀海精微補, 目科一覽

10) 鍼灸經絡

- 鍼灸銅人: 銅人經, 重刊神應經, 銅人鍼灸圖, 經絡學總論
- 推拿按摩: 小兒推拿秘訣

11) 養生食餌

- 養生導引: 養生類要, 道藏輯要抄, 保養志, 退溪活人心法
- 食餌食療: 食療纂要, 食醫心鑑, 四時纂要抄, 餌方
- 家庭救護: 攷事撮要, 山林經濟, 修養叢書類輯, 胎教新記
- 飲食調治: 閨閣叢書, 飲食知味方, 釀酒方

12) 醫史筆記

- 歷史說話: 醫說, 歷代醫學姓氏, 壺山外史
- 醫學關聯 文獻雜記: 宜彙, 濟生篇. 解惑辯疑
- 醫案醫話: 錦囊, 丁氏經驗方, 谷青冗語
- 醫學問答: 朝鮮醫答, 醫學問答
- 醫哲學類: 鑿圃問答, 演小天地問答
- 醫學論評: 醫零, 神氣踐驗

13) 韓醫學雜誌

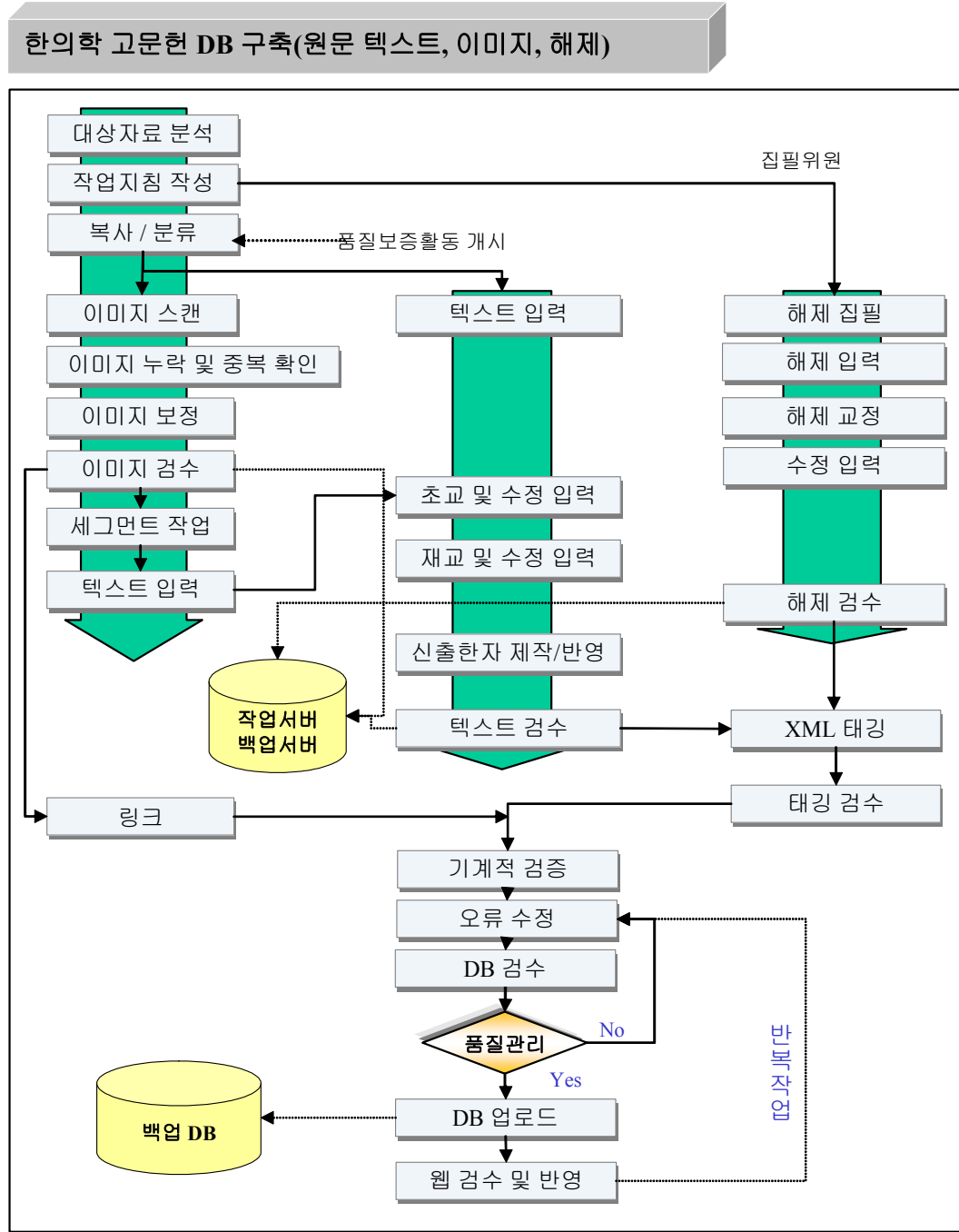
- 韓方雜誌: 醫藥月報, 朝鮮醫學界, 東方醫藥, 東醫報鑑, 東洋醫藥

14) 其他: 전문의학서는 아니지만 의학 관련내용이 포함된 것

- 法醫書: 新註無冤錄, 檢屍狀式, 審理錄, 檢要
- 獸醫書: 馬醫方, 水牛經, 古本鷹鵠方
- 東西醫學類: 東西醫學要義, 家庭救急方
- 亡失類: 濟衆立效方, 新羅法師方, 鄉藥惠民經驗方, 新集御醫撮要方
- 雜著類: 鏤板考, 醫生必攜, 藥房謄錄, 醫科八世譜, 醫登第譜, 太醫院先生案

5. 한의학 고문헌 DB구축 실례 (한의학지식정보자원 디지털화 사업)

1)DB구축의 개요도



2)한자입력방법

입력교정 솔루션

SSC 한적솔루션

- 고문헌 한적 자료 전산화업체에서 개발한 모델로 대용량 한적 자료를 빠르고 정확하게 데이터베이스로 구축하기 위해 필요한 기능들을 작업 단계별로 모아 프로그램화한 한적 전용 입력교정 솔루션임

주요 기능

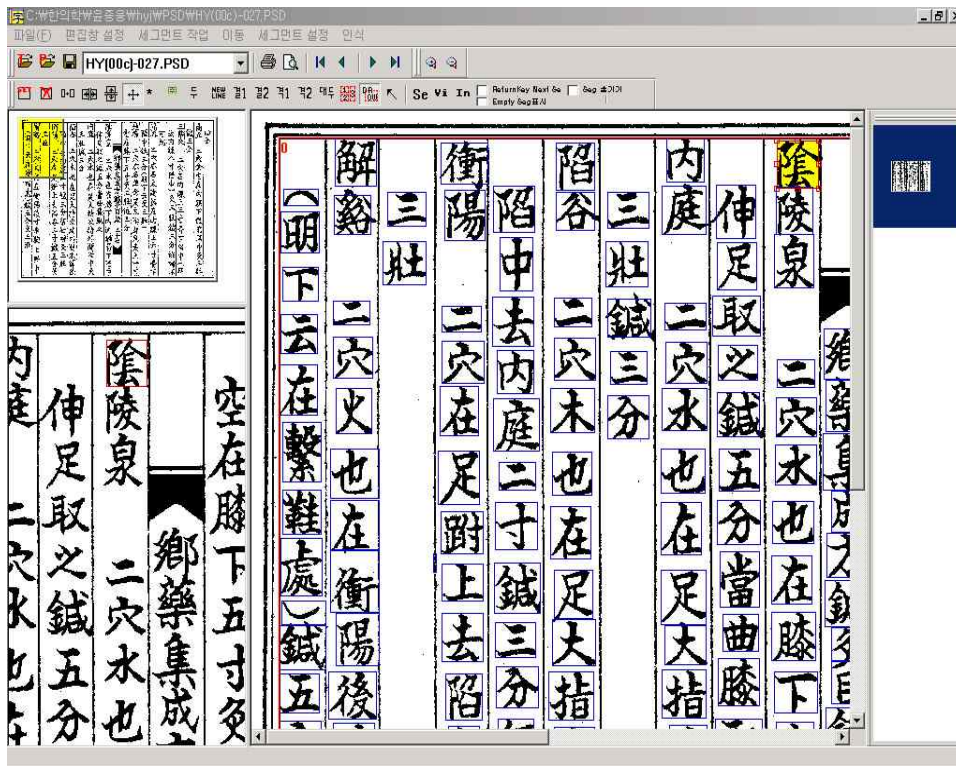
- 한적 자료를 디지털 이미지로 전환시켜 글자 단위로 이미지를 자동 분할
- 분할된 이미지를 자동 인식
- 인식된 글자를 유사한 자형별로 묶어 일괄 입력(Image Clustering Method)
- 직접 입력을 위한 다양한 한자 입력 방법(Unicode Based) 지원
 - 초고속한자입력법(뿌리법), 음가변환, 부수-획수 등 다양한 입력 방법 지원
- 선택한 낱자의 위치정보 제공으로 입력/교정 시 오류를 최소화
 - 원본 이미지 파일, 파일에서의 위치정보(페이지/행/열) 확인
- 실제 사업에서 입증된 편리하고 효율적인 사용자 인터페이스

시스템 구성

	<p>이미지자동 분할기</p>		<p>유사도 입력기</p>
	<p>1차 교정기</p>		<p>2차 교정기</p>

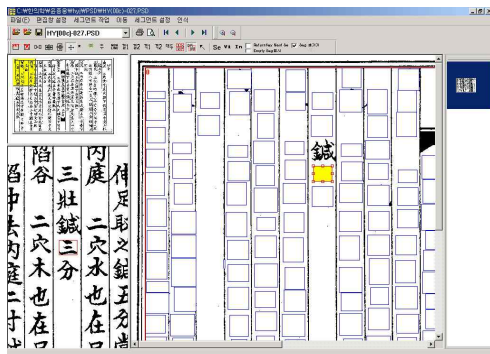
자동 분할기(세그먼트 및 입력)

- 자동분할기는 스캐닝된 고문헌의 이미지를 글자 단위로 하나씩 분할해서 코드화된 문자로 입력할 수 있도록 작업을 하는 프로그램
- 일괄처리에 의한 대용량 자동 분할 기능
- 미세한 분할을 위한 수동 분할 기능
- 소자쌍행(小字雙行)과 같이 한적 자료의 형태상의 특징까지도 마크업을 부가 가능
 - 격자, 결자, 별제(원주), 두주, 주석 처리 기능
 - 다양한 검수 기능으로 세그먼트 누락을 최소화
- 입력 공정에서 기본적인 형식태그를 처리하고 교정 단계에서 중점적으로 검토

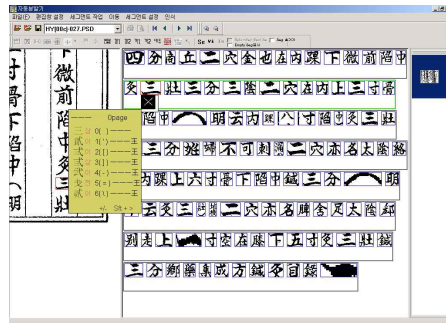


자동 분할기(계속)

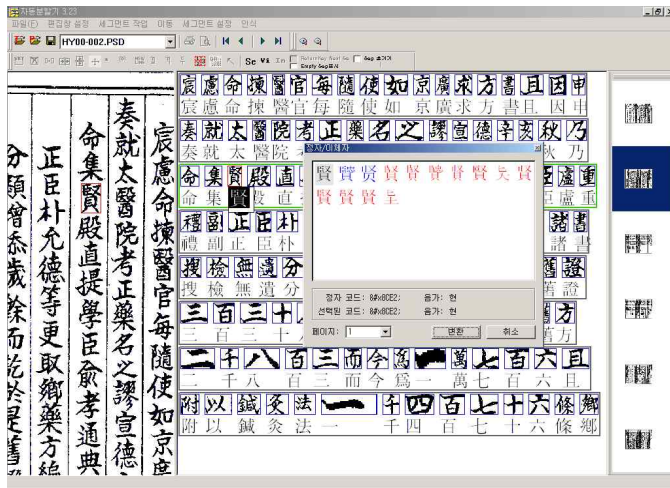
- 세그먼트 이미지 감추기 기능으로 순도율에 결정적 영향을 미치는 세그먼트의 누락을 손쉽게 검수할 수 있는 강력한 검수 기능
- 분할된 낱자 이미지를 직접 입력 한자 초고속입력법을 이용한 입력 진행
- 입력시 가장 심각한 문제가 되는 이체자 부분은 본 솔루션에서 제공하는 이체자 사전을 참고하여 근원적으로 해결할 수 있음



이미지 감추기 기능



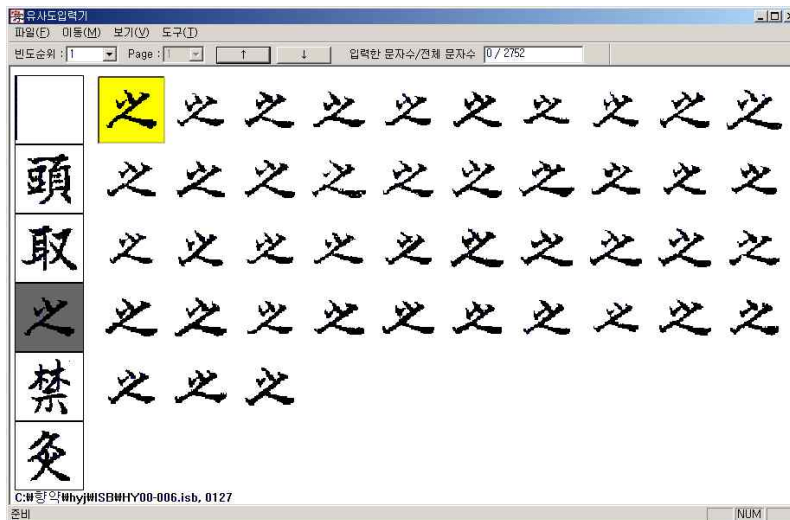
낱자 입력



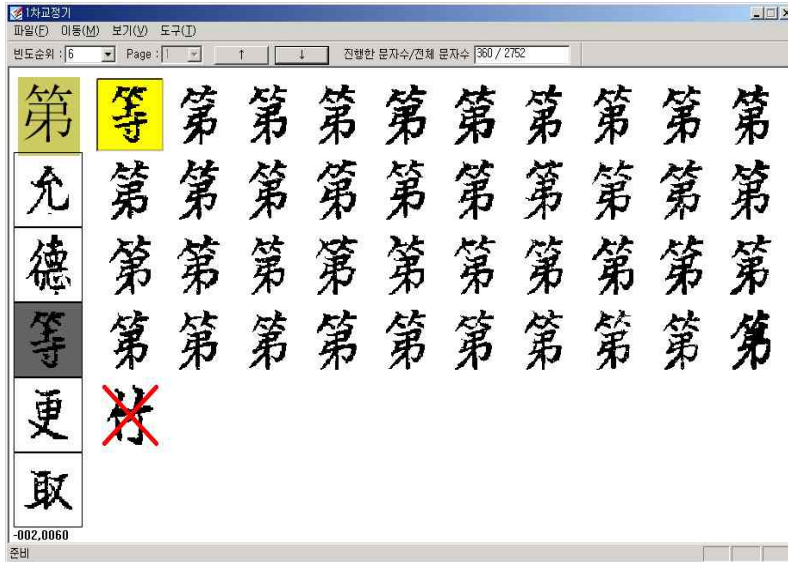
이체자 사전 기능

유사도 입력기

- 유사도 입력기는 글자 단위로 세그먼트된 이미지를 유사한 자형별로 묶어서 한 화면에 모아놓고 이질적인 요소를 제거한 후 나머지 글자를 단 한번의 입력으로 그룹핑 된 전체 글자를 일괄 입력하는 새로운 개념의 입력 시스템이다.
- 표준화된 자형과 유사도를 계산하여 유사한 이미지 형태별로 그룹핑
- 유사도별로 그룹핑된 원자료에 대한 이질적인 요소를 체크하여 다음 공정에서 처리
- 작업의 임계치를 적절하게 설정하여 가장 효율적인 수치로 설정
- 한자 초고속입력법(뿌리법), 부수획수 입력, 음가 입력을 지원
- 옛한글을 포함한 유니코드에서 표준화된 27,484자의 한자 입력 방법 지원
- 일본어 및 한어 병음에 의한 입력 등 다국어 지원
- 전체글자 선택, 해당열 선택 등의 편리한 교정 기능 제공



1차 교정기(코드별 교정)



- 1차 교정기는 전 공정에 날자 입력과 이미지 유사도 입력으로 입력된 데이터를 대상으로 입력된 코드별로 이미지를 나열하여 일괄 교정을 진행하는 시스템이다.
- 나열된 동일한 형태 이미지 중 상이한 이미지만 선택해서 교정
- 비전문가를 고려한 프로그램이므로 별도의 교육이 필요치 않아 소프트웨어 즉시 사용이 가능
- 동일한 코드값의 문자들만을 나열, 대조하므로 오류코드 식별이 용이
- 글자 찾아가기 기능은 ㄱ, ㄴ, ㄷ 등 문자의 차이가 없이 문맥으로 판단해야 하는 글자들을 집중적으로 확인할 수 있음

2차 교정기(일자 교정)

0 竹頭下行一寸是次灸兩第

1 大指節上下各七壯治脾

2 風一俗呼爲八風一灸脾

灸兩大指節

HY00-010, 197

준비

- 2차 교정기는 전 공정에서 처리하지 못한 특수한 문자나 한자 등을 전문 인력에 의해 처리하는 시스템이다.
- 1차 교정 솔루션에서 해결할 수 없는 한자만 모아서 집중적으로 교정
- 1차 교정에서 중급의 인력을 활용하여 교정을 진행하고 1차 교정에서 해결할 수 없는 문제를 전문가에 의해 2차 교정 단계에서 해결
- 오류 글자에 대해 한자 초고속입력법(뿌리법)을 적용해서 입력
- 입력이 불가능한 한자는 신출자로 등록한 후에 전문가의 검수를 거쳐 최종적으로 신출자로 확정

2차 교정(일자 교정)

名 名 名 名 名 名 名 名 名 名

府 一 名 舌 本

分得氣則瀉
神聰 四穴在百會四
後頂 一名交衝在百
壯鍼 二分(明云四分
風府 一名舌本在項
宛宛 中疾言其肉立
失音鍼 三分(明云四
瘕門 (一作瘕) 一名
央入髮際五分宛宛
瘕鍼 三分(素註云上

□ 원문 보기 기능을 통해서 의문이 나는 글자에 대한 재확인 작업 진행

- 단순화된 교정 알고리즘 적용으로 교정 소요 시간 단축
- 수동 교정 대비 작업능률 50% 이상 향상

Neo-UNI

확장 한자 입력 가능

- 유니코드 3.0에서 지원하는 27,484자를 포함하여 “한국한자어사전” 편찬을 위해 개발한 도합 72,000자 한자를 뿌리법으로 입력 가능

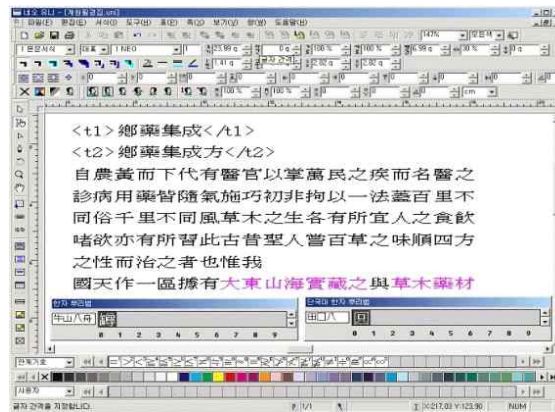


도표 편집/이미지 삽입

- 작성한 도표를 HTML 문서로 저장하거나 HTML 문서로 작성한 표를 호출하는 가능
- 이미지를 삽입하거나 이미지의 정보만을 디스플레이 하는 기능



이미지 보정 솔루션

- 본 사업에서 세그먼트 기반 SSC한적솔루션을 적용하기 위해서는 원문 이미지를 스캐닝할 때 일부 자료의 경우 자료의 훼손상태가 심하여 일반 스캐너로 작업을 하였을 때는 이미지 세그먼트 작업이 어려운 경우가 있음
- VRS(Virtual rescan)는 그림자나 어두운 백그라운드 이미지를 보정작업을 처리하여 세그먼트 분할 및 인식을 용이하게 함.
- VRS는 잘라진 글자의 edge를 연결해 줌으로써 OCR 엔진의 인식률을 향상 시켜 줌



시범서비스 화면(해제포함)



시범서비스 화면(원문이미지포함)

The screenshot displays a web browser window with the following elements:

- Browser Title:** 한의학 지식정보자원 웹사이트 - Microsoft Internet Explorer
- Address Bar:** http://218.48.79.209:8181/
- Navigation Bar:** Home, Search, Favorites, Media, etc.
- Search Bar:** Search for '고문헌' (Old Documents) with filters for '전체' (All) and '명칭' (Name).
- Left Sidebar:** A tree view showing the hierarchy of '향약집성방' (Hangeupjeongbeop), including '風門' (Wind Gate) and various '風' (Wind) related terms like '風痰', '風冷失聲', etc.
- Main Content Area:**
 - Header:** home > 고문헌 > 서명찾아가기(하) > 향약집성방 1
 - Buttons:** 확대 (Enlarge), 축소 (Shrink), 저장 (Save), 인쇄 (Print), 원문 이미지보기 (View Original Image), 해제보기 (Release View), 매일보내기 (Daily Send).
 - Title:** 風痰
 - Text:** 聖惠方論曰夫風痰者是血脈壅塞飲水積聚而不消散故成痰也其候或冷或熱或結實食不消化胸膈痞滿短氣好眠頭眩目暗常欲嘔逆者是也
 - Image Viewer:** A window titled '원문 이미지 보기 - Microsoft Internet Explorer' showing the original text in a calligraphic font. The text includes '風痰' and various characters in a vertical column.

IV. 期待效果 및 向後 研究方案

1. 기대효과

1) 지식정보화 효과측면

- 수천 년간 전수되어온 방대한 양의 한의학 고문헌 자료를 체계화
- 한의학 자료를 첨단 저장매체에서 저장, 활용성을 극대화
- 세계 각국의 연구기관과 정보자료 등을 공유
- 각종 한의학 연구에 관한 객관적인 검증가능
- 한의학의 저변확대와 체계적 한의학 정보 보급을 위한 기반조성
- 한의학 관련 기술의 지적재산권 문제를 해소하고 특허 등의 방법을 통해 지적재산권을 확보할 수 있는 기반 마련

2) 경제활성화 측면

○ 정량적 측면

- 한의학 연구 및 개발 정보 획득에 소요되는 시간 및 비용 절감효과
- 고문헌의 처방 연구를 통한 복합 한약제 신약 개발 효과
- 고문헌의 처방 및 치료기술연구에 기초한 대체의학 개발 효과
- 한약, 침술 등 세계 최고 수준의 한의학 기술 개발은 전통 고문헌의 경험 이 집적되어 있는 처방 및 치료기술 연구가 선행되어야 함.

○ 정성적 측면

- 초고속 정보통신망을 기본으로 한 멀티미디어 형식의 한의학 정보제공으로 한의학 관련 산업을 현대화하고 세계화
- 사회 각 분야에서 요구되고 있는 전문디지털도서관 관련요소 기술의 확보
- 한의학 전체의 연구방향, 발전전략에 관한 기준을 확립하여 연구중복을 방지
- 한의학과 서양의학의 협동학술연구 및 협진체계를 통한 의료비용을 감소

- 원격진료시대에 대비한 한의학 정보서비스 제공

3) 국가위상도 측면

- 고유전통 한의학을 세계화함으로써 국가 경쟁력 제고
- 전자매체 활용을 통한 정보화 사회의 구현
- 良質의 대국민 한의학정보 서비스 제공
- 한의학의 이미지 향상을 통한 국가위상 확립

2. 사업수행 중 문제점 및 대책

- 현재 2003년도에 한의학 고문헌 디지털링, 주요고문헌의 DB구축 중복판정으로 인한 대상제외로 『東醫寶鑑』, 『方藥合編』 등 필수 의학서들이 타기관의 기구축(이미지)을 사유로 본 사업에 제외됨으로써 서비스에 결함이 발생
 - 기관간의 양해와 협조를 통해 한의학 고문헌의 경우 한의학 전문기관에서 서비스할 수 있는 장치 마련 시급
 - 기존 자료의 미흡함(이미지서비스 만의 제공, 전문가의 한계적 활용)
- 개인 또는 기관 소장 고문헌 자료의 수집 미흡으로 인해 디지털화 대상 자료의 한계
 - 체계적이고 집중적인 수집활동 필요. 이를 위해 사업의 중요성 및 최소한의 예산 확보 필요
 - 자료확보 방안을 위한 기관차원의 계획 수립(滅失가능성이나 保存價値, 學術價値, 순으로 우선순위 부여)
- 이미지 기반의 서비스는 원본의 보존차원, 텍스트 기반의 서비스는 내용의 활용도 차원에서 불가결한 작업
 - 다양한 서비스를 위한 정부의 예산 배정 현실화 필요

○ 번역 서비스의 필요

- 국역, 영역, 多國語 서비스 지향, 국제화 코드 적용.