

데이터베이스 출판시스템 관련 백과사전 색인

— 동아.백과사전 사례를 중심으로

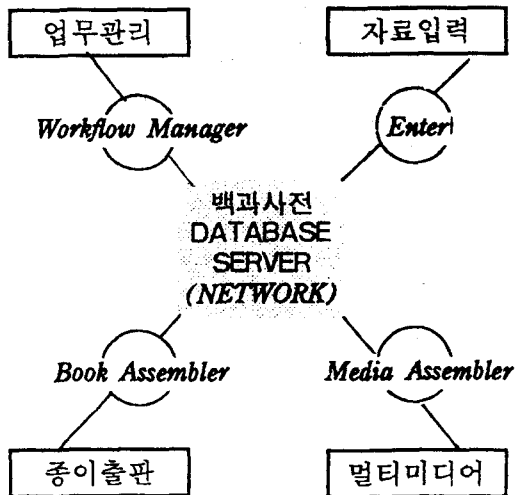
동아출판사 백과사전연구소 이신영

백과사전 자료를 항목별로 구조화된 데이터베이스로 구축하여 백과사전 본책 및 색인 출판, 멀티미디어 출판, 정보망 구축 등에 전환되도록 실무 시스템을 구성하였다.

I. 백과사전시스템 구성 배경

- 자료 속성
  - 자료량 방대
  - 정형화된 자료 형태
  - 지속적인 수정·갱신 필요
- 업무 형태
  - 다수의 담당자
  - 주제별 업무 분장
  - 장기 작업기간
  - 토털 전자출판시스템 구축
  - 편집경험 보존
- 활용 형태
  - 색인 구성
  - 멀티미디어로의 전환
  - 전문사전 출판자료로 활용

II. 백과사전시스템 개념도



III. 색인 기획

cf. 백과사전 이용목적

- 용어 사전적 지식을 얻고자 하는 경우
- 특정어를 매개로 한 관련되는 폭넓은 지식을 얻고자 하는 경우
- 특정어를 매개로 도판, 사진, 수치 등 개별 정보 소재를 알고자 하는 경우

1. 구성

(1) 색인표제어의 문자에 따라 한글색인, 로마자색인, 한자색인 편으로 구성한다.

- ① 한글색인 : 본책 항목, 한글 약·별칭, 본문 색인어
- ② 로마자색인 : 로마자가 병기되어 있는 본책 항목의 로마자, 로마자 약·별칭, 학명
- ③ 한자색인 : 중국·일본·기타 한자문화권 항목 중 현지음으로 읽힌 본책 항목의 한자

(2) 한글색인 표제어는 본책 항목수 대비 150%, 표제어당 평균 색인관계 2.5건, 표제어당 평균행수는 2행으로 한다.

cf. 신판 본책항목수 130,000 (28권)

▶ 한글색인

- 색인표제어수 : 195,000 항
- 색인관계 : 455,000 건
- 색인행수 : 390,000 행

▶ 로마자·한자색인

- 로마자색인수 85,000 항
- 학명색인수 : 9,000 항
- 한자의래어색인수 : 2,000 항

(3) 별권 2책으로 묶는다. 한글색인편은 분권하지 않는다.

2권 총 1,600면 (4단 조판)

▶ 권1

한글색인 / 일러두기

▶ 권2

로마자색인 / 한자색인 / 일러두기  
 집필자·편집위원·자료협조처 명단  
 부록 : 문화사 연표, 각종 일람

(4) 독자의 검색 효율성 및 본책의 소항목주의 보완, 독립된 자료로서의 색인 등에 중점을 두어 구성한다.

2. 체 제

**색인표제어 한자·로마자 표기 (부연부)**

수록권 : 면, 참조항목명 수록권 : 면, ...

(자료형태표지) 참조항목명 수록권 : 면 제목

(1) 표제부

① 표제어

본책 항목/비항목 표지

② 한자·로마자 표기

표제어에 대한 정보성이 큰 표기 택일

③ 부연부

인명*	.....국명, 분류
지명	.....국명
문화재	.....문화재
기관	.....국명
단체	.....분류
문헌	.....분류
작품	.....국명, 분류
생물	.....분류
일반	.....분류

(2) 기술부

① 수록 항목명

표제어와 동일한 경우에는 생략  
 대항목의 특정 부문을 참조할 경우, 중  
 제목명까지 기술

② 항목 소재

항목 시작 권, 면 순으로 기술  
 대항목의 특정 부문을 참조할 경우, 중  
 제목이 시작하는 위치 지시

③ 가보라 항목 표지

가보라 지시된 항목명 및 수록 권·면  
 기술

④ 자료형태 표지

주요 사진, 도판, 표만을 대상  
 색인표제어 항목에 자체수록된 경우 제외  
 수록 권·면은 해당 자료가 실린 항목을  
 기준으로 지시

⑤ 제 목

사진·도판·표의 이름, 단, 수록 항목  
 명과 동일한 경우는 생략

⑥ 부제목(sub-title)

필요에 따라 색인어를 부문별로 묶어 제  
 시하고, 해당 부문을 기술

예) 오그덴 Ogden, C. K. (영국, 심리학) 21 : 244, 리처즈 5 : 253  
 오그덴 Ogden (미국) 21 : 245  
 오그번 Ogburn, W. F. (미국, 사회학) 21 : 245, 기딩스  
 4 : 180, 문화지체 7 : 268, 기술사관 4 : 324  
 오그보모쇼 Ogbomoshu (나이지리아) 21 : 245  
 오극관 五極管 (전기) 21 : 245, 그리드 3 : 630  
 오금선 吳綾善 (의학·교육) 21 : 245  
 오기 五伎 (민속) 21 : 245, 금환 1 : 561, 대연5 : 213, 산예  
 11 : 224, 속독5 : 468, 월전10 : 410, 향약잡영  
 25 : 250  
 오기노규사쿠 荻野久作 (일본, 의학) 오기노다실 21 : 246  
 오기노히수태조절 (생활) ← 수태조절 15 : 324  
 반데르발스 van der Waals, J. D. (네덜란드, 물리) 13 : 250  
 반데르발스 (네덜란드, 물리) → 반데르발스  
 라이트 Wright, F. L. (미국, 건축) 11 : 100  
 鋼 주벽건축 25 : 57 낙수장  
 네덜란드 Netherlands, The 5 : 257, .....  
 (산업) 유리공예 22 : 190

예) Aude (프랑스) 오드현 21:200  
 Aude R. (프랑스) 오드강 21:200  
 Audemars Piguet (공학) 오데마피제 21:254  
 Auden, W.H. (미국, 문학) 오든21:200  
 Audibert, J.S. (프랑스, 문학) 오디베르티 21:200  
 Audimeter (기계) 오디미터 21:200  
 audio (공학) 오디오21:200  
 Auditorium (공학) 오디토리엄 21:200  
 Audubon, J.J. (미국, 동물) 오듀본 21:200  
 auramine (화학) 오라민 21:200  
 Aurangabad (인도) 오랑가바드 21:200  
 aureole (미술) 오레올 21:200  
 circle of fifth (음악) 오도권 21:200  
 Coelotes bicaudatus (동물) 오대산가계저미 21:245

예) 岡崎 오카자키 21 : 373  
 岡山 오카야마 21 : 372  
 岡山縣 오카야마현 21 : 373  
 大江健三郎 오에겐자부로 21 : 349  
 帶廣 오비히로 21 : 309

3. 배 열

(1) 한글색인

- ① 본책의 항목 배열순에 준한다.
- ② 표제어 내의 띄어쓰기 및 각종 문장부호는 무시하고 배열한다.
- ③ sub-title로 묶은 경우에는, 각 부문 안에서 가나다순으로 배열한다.

(2) 로마자색인

- ① 관사는 뒤로 보낸다.
- ② 전치사는 원래 위치에 둔다.
- ③ 시양인명은 성, 이름 순으로 둔다.
- ④ 지형부, system of ~, university of ~ 등 형태가 굳어진 고유명사, 전문용어 등은 의미를 살펴 의미부를 앞세운다.

예) Ogburn, W. F.  
 Old Man and the Sea, The  
 Ivan Mt. Michigan, Lake  
 Luke, Saint Cluny, Musee de  
 Michican, University of

- ⑤ 표제어 내의 띄어쓰기, 문자부호(약상기호 등)는 무시한다.
- ⑥ 대소문자는 무시하되, 동일표기인 경우에는 대문자를 우선한다.
- ⑦ 숫자가 포함되어 있는 경우에는 A에 앞서 배열한다.
- ⑧ 철자가 동일하면 본책 항목 배열순에 따른다.

(3) 한자색인

- ① 한자 한글음순으로 배열한다.
- ② 동일한 발음인 경우에는 본책 항목 배열순에 따른다.

IV. 색인 관련 데이터베이스 구조 및 출판

1. 데이터베이스

- (1) 항목자료, 항목구성자료, 분류자료, 별약칭, 가보라항목, 색인 등을 항목번호로 연계한 관계형 DB를 구성한다.
- (2) 색인자료 관리 및 본책과의 연동성(선정상황, 가보라.참조, 수록소재 등)에 중점을 둔다.

[항목DB]

항목NO.  
 표 제 /한글  
           /한자  
           /로마자  
 분 류 /주제  
           /속성  
           /지역  
           /언어권  
 선정상황  
 가보라. 참조상황  
 기 타 /별이칭  
 수록상황/권  
           /면

[구성DB]

항목관련NO.  
 본 문 / 사 진  
 도 판 / 표

[분류DB]

[가보라·참조DB]

[별·약칭DB]

[색인DB]

색인어 NO  
 참조항목 NO  
 참조항목관련 NC  
 부제목명

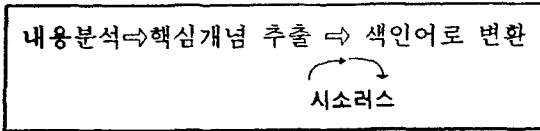
- 색인어는 모두 항목자료화 한다.
- 항목자료는 고유번호를 가지며, 색인자료는 고유번호 간의 참조 관계로 구성되므로, 항목자료의 수정 및 선정상황 변경시 연동된다.
- 각종 분류는 별도의 DB를 두어, 각각 대/중/소의 3단계로 하고, 약어 등을 규정하여, 변경시 항목자료의 분류와 연동된다.
- 선정상황, 즉 본책의 항목여부에 따라 항목·비항목 색인어를 표지한다.
- 별·이칭 DB를 두어, 항목자료와 동일한 분류를 적용시키고, 기본 색인어로 추출한다.
- 가보라·참조 DB를 두어, 기본 색인어로 추출한다.
- 출판상황에 따라 갱신되는 수록권, 면자료로 참조항목 소재를 표지한다.
- 표제어의 한자, 로마자 상황에 따라 한자·로마자 색인어 구성 여부를 자동인지한다.
- 기본색인어(본책 항목의 표제부, 별·이칭, 가보라·참조 등)는 자동 생성하고, 추가 색인은 수동 추출한다.

2. 출 판

- (1) 출판에 관련된 데이터베이스 필드별로 각각 표지(Mark-up)를 지정한다.
- (2) 지정된 표지별로 출판에서 표현될 서체, 단락규격, 간격 등을 지정한다.
- (3) 레이아웃 프로그램에서 표지를 해석하여, 지정된 양식을 표현한다.

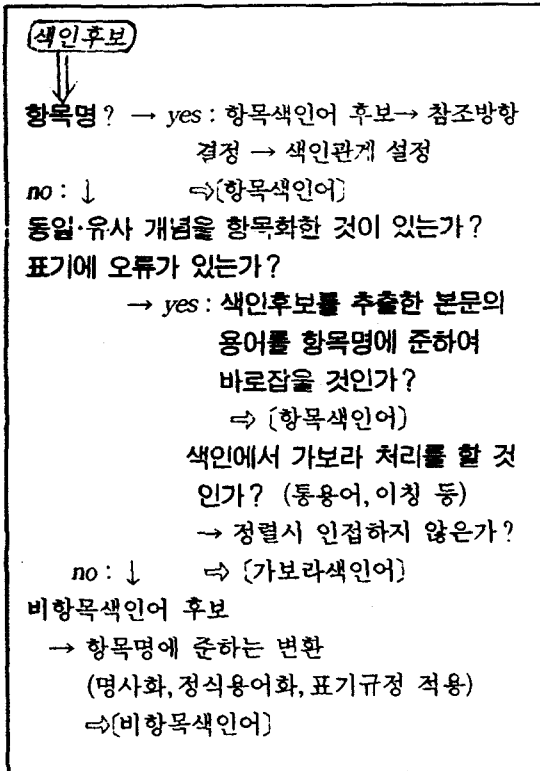
V. 색인 처리

cf. 색인처리 과정



백과사전의 경우, 본책 항목목록이 일종의 시소러스로서 작용하여, 추출된 핵심개념은 다음 과정을 거쳐 색인어로 변환된다.

또한 본책과의 내용 및 표기 연계성, 항목 선정의 균형 등 본책 출판과도 밀접한 연동성을 갖는다.



VI. 향후 방향

1. 색 인

- 색인어 관계에 따른 자료 구조화
- 시대 등 기타 특성 분류 보강
- 자연어 검색 보강
- 자동색인 적용
- 독자층에 따른 선별 색인

2. 백과사전

- 현재시점에서의 정보 제공
- 독자층(연령, 필요정보)에 따른 개별 대응 출판
- 다양한 멀티미디어 매체로 전환 : 패키지형(CD-ROM), On-line 통신서비스 등
- 도서관전산망 등 외부정보망과의 접속

# 참고문헌 #

김석영(1992) 초록 및 색인론, 산업기술정보원

최석두(1993) 무전거시스템에 관한 연구. 한국문헌정보학회지 제 25집

김두식(1994) 전자출판과 멀티미디어의 이해와 활용. 타래

日本索引家協會(1983) 索引作成マニュアル

稻村徹元(1977) 索引の話. 日本圖書館協會

ANSI American National Standard for Library and Information Sciences and Related Practices - basic criteria for indexes. Cleveland, D. B., Cleveland, A. D. (1983) Introduction to indexing and abstracting, Libraries Unlimited.