

INIS에 있어서 데이터 색인

金 成 赫

〈韓國原子力研究所 INIS室〉

오늘날은 1차 및 2차간행물에 포함되어 있는 데이터의 분류방법 및 기술이 중요시되고 있다. 각국의 데이터 분석센터와 국제기구도 문헌속에 포함되어 있는 유효한 데이터를 국제적으로 이용시키기 위해서 많은 노력을 기울이고 있다.

데이터 색인분야(데이터 Flagging과 Tagging*)에서 계획하고 있는 일들은 아직까지 데이터 색인의 기본적인 개념연구 정도이다. 결과적으로 보면 아직까지는 국제적으로 인정된 데이터 분류법이 없으며 이에 대한 준비도 되어 있지 않다.

데이터 Flag는 1차문헌가운데서 데이터의 존재여부를 나타내 주는 것을 말하고 데이터 Tag는 데이터의 특징을 더욱 자세히 표현하기 위한 방법이다.

그러나 최근 데이터 색인의 중요성이 인식되어 몇몇 분야에서 제안 검토되어 왔다. 특히 광범위한 주제분야를 취급하고 있는 INIS(International Nuclear Information System: 국제원자력정보시스템)가 데이터 Flagging을 도입한 것은 이 방법의 선구자로서 주목할만 하다.

本稿에서는 지금까지 몇개의 분야에서 제안된 데이터의 분류방법과 INIS에 도입된 과정, 배경 및 현재 INIS에서 실시중인 데이터 Flagging 순서에 대해 소개한다.

1. 서 론

INIS에 입력되는 문헌의 대부분은 유용한 수

치 데이터를 포함하고 있으며, 이들은 수치 혹은 그래프의 형태로 나타난다. INIS와 같은 서지적 데이터 베이스의 이용자에게 있어서, 이용자가 관심있는 주제, 특히 어떤 현상, 양, 효과, 특성 등이 포함된 수치 데이터가 어떤 문헌에 있는가를 아는 것은 이용자의 관심사이다.

INIS 이용자에게 이러한 정보를 제공하기 위해 1979년부터 국제원자력정보시스템(INIS)에 데이터 Flagging이 도입되었다. 각국의 INIS센터는 1979년 초부터 자국 입력분을 데이터 Flagging하여 입력하고 1980년 1월부터는 전면적으로 실시한다는 계획을 세우고 있다. INIS는 1년에 약 75,000~80,000건의 원자력관계 정보를 수록하는데 이 입력분의 1% 이상을 차지하는 입력국가를 주요 입력국(Major Inputer)이라 한다. 위에서 말한 전면적 실시라는 것은 주요입력국을 의미하며, 우리나라의 경우 입력되는 정보는 1979년에 198건, 1980년에는 200건을 입력하고 있기 때문에 데이터 Flagging에 대하여는 아직껏 준비단계이나 1981년부터는 실시될 것으로 기대된다.

데이터 Flagging과 Tagging에 대해서는 최근 그 중요성이 인식되어 몇몇 분야에서 그 방법이 제안, 검토되어 왔다.

예를 들어 화학분야에 있어서 IUPAC(International Union of Pure & Applied Chemistry)의 제안, 물리분야의 AIP(American Institute of Physics), UNISIST(Universal System for Information in Science and Technology) 등이

* 데이터 Flag는 1차문헌가운데서 데이터의 존재여부를 나타내 주는 것을 말하고 데이터 Tag는 데이터의 특징을 더욱 자세히 표현하기 위한 방법이다.

있지만 아직 시험단계이고 확립된 안(案)이 없다".

2. 데이터의 분류방법

데이터 분류란 어떠한 기준을 만족시키는 Subset를 통해서 이용가능한 모든 데이터 Set를 분산하는 것을 의미한다²⁾.

여러가지 기준에 기반을 둔 Subset사이의 부분적 혹은 전체적인 중복이 인정된다면 같은 기준에 의해 분리된 두개 Subset의 교차는 empty set로 추측할 수 있다.

완전한 데이터 분류법은 하나의 망상에 불과하며, 과거 그리이스 시대부터 지금까지 완전한 과학 분류법이 없다는 것이 그 예이다. 우리가 데이터 분류법에 관심을 갖고 있는 한 그러한 상황은 더욱 힘들게 될 것이다. 그러면 몇개의 기관에서 제안된 데이터 분류법을 살펴보기로 하자.

2.1 UNISIST의 분류법³⁾

- i) 측정시간에 의한 데이터
 - 반복하여 측정할 수 있는 데이터
- ii) 측정장소에 의한 데이터
 - 측정장소에 관계없이 측정할 수 있는 데이터
 - 측정장소에 의해 좌우되는 데이터
- iii) 초기, 추출, 이론 데이터
 - 실험, 관측에 의해 얻어지는 초기 데이터
 - 이론적 모델방법으로 초기 데이터를 결합하여 얻는 데이터
 - 이론적 계산에 의해 얻어지는 데이터
- iv) 확정치, 추정치
 - 주어진 상황하에서 측정될 수 있는 한정치의 데이터(확정치)
 - 주어진 상황하에서 변동하는 데이터(추정치)
- v) 정량적, 정성적 데이터
- vi) 수치데이터 그래프 모형
 - 수치만으로 제시된 데이터
 - 그래프 또는 모형으로 제시된 데이터

- vii) 이용자의 요구에 의해 분류된 데이터
 - 특정한 분야에서 이용된 데이터와 그 분야의 전문가에 의해 독자적으로 만든 데이터
 - 연구원에 의해 사용되는 데이터
 - 광범위하게 이용된 데이터

2.2 IUPAC의 분류법⁴⁾

순수 및 응용화학분야의 데이터를 여러 기준에 의해 90종류로 분류한다. 각각은 2문자 코드에 의해 분류되고 있다. 분류에 의한 중복은 허락되며, 한 문헌에 2개 이상의 코드를 부여해도 된다. 이 분류는 1차간행물에 이용하기 위해 계획되었다.

2.3 AIP의 저자용 데이터 보고용지⁴⁾

이 안은 물질의 집합상태에 관한 데이터를 분류하기 위해 제안되었다. 해당항목을 체크하는 Flagging과 자유로이 기술하는 Tagging과의 2가지 방식이 있으며 다음의 4기준으로 분류한다.

- i) 물질의 집합상태(10구분)
 - 비결정물질
 - 결정물질
 - 액정 및 고분자물질
- ii) 연구된 시스템, 재료의 형태(15구분)
 - 원소
 - 금속 합금 및 금속간 화합물
 - 용융염
 - 전해질 등
- iii) 성질 및 특성(16구분)
 - 역학적 및 유동성
 - 음향학적
 - 열역학적
 - 반도체적
 - 자기적 등
- iv) 환경조건(필요한 경우) (8구분)
 - 고온, 저온, 고압, 저압, 외부전계, 등
 이에 부가해서 4개의 데이터 tags는
- v) 재료, 시스템

- vi) 결정된 성질 또는 가변상수
 - M-측정데이터
 - C-계산 및 추측데이터
 - R-데이터 리뷰 및 편집

vii) 관련 환경

viii) 주석

위의 데이터 색인은 색인자에 의해 실시되는 것이 아니고 저자가 논문을 제출할 때 데이터 보고용지를 작성하여 제출한다.

2.4 IAEA/Nuclear Data Section의 제안⁵⁾

INIS에 대해, 핵 데이터 센터측의 요구로 제안된 것이다. 데이터를 게재하고 있는 문헌뿐만 아니라 측정중인 데이터도 Flagging대상으로 하며, 다음의 4항목을 기준으로 한다.

- i) 새로운 주제에 의한 구분
 - 소립자, 핵반응, 핵구조, 원자반응, 원자구조, 고체, 노물리 등
- ii) 결정방법에 의한 구분
 - 측정치, 평가치, 이론치 등
- iii) 발생국 및 연구소에 의한 구분
- iv) 간행물에 의한 구분
 - 시판 및 비시판자료

2.5 COBIDOC(Netherlands Bibliographical and Documentary Committee)의 제안⁶⁾

COBIDOC은 9개의 데이터 분류기준을 설정했다.

- i) INIS 주제분류표와 디소오러스용어에 나타난 주제에 의한 구분
- ii) 시스템(핵종, 화합물, 바이러스 등)
- iii) 과정(핵반응, 화학반응, 위상전이, 속주세 포배양 등)
- iv) 표준치(각분포, 반응상수, 열전도 등)
- v) 함수(르장드르 다항식, 열전도 등)
- vi) 양의 변화에 따르는 변수의 한계(에너지, 온도, 양의 한계)
- vii) 환경조건(실험에 쓰인 기구, 데이터 환원 방법 등)
- viii) 데이터 결정 및 이용에 의한 구분(실험

치, 평가치, 편집데이터, 요구데이터 등)

- ix) 데이터 제시 형태에 의한 구분(표, 그래프, 컴퓨터매체, 데이터 없음 등)

2.6 INIS 1차안의 분류법

IUPAC, AIP, IAEA/NDS의 안은 주제에 의한 분류법이고 광범위한 주제를 취급하는 INIS에는 위에 열거한 5가지 분류법이 적합치 않은 것이 분명하다. INIS 1차안의 분류법은 데이터의 주제내용에 관한 것과 데이터의 형태, 제시형태에 관한 것으로 나눌 수 있으며, 전자는 최종안까지도 디스크립터에 D를 부여하는 방식이고, 후자는 디스크립터가 아니고 Literary Indicator N밑에 13종의 Sub-Indicator를 부여하는 방식이다.

형태구분은 그 데이터의 결정방법, 이용에 의한 구분 및 CINDA(Computer Index of Neutron Data)구분인 U(인용하여 이용될 때)와 R(새로운 측정과 계산이 필요할 때)이 첨가되어 데이터의 범위를 넓게 잡아주고 있다.

제시구분은 핵 데이터센터측의 요구에 의해 만들어졌다. CINDA에는 이러한 구분이 없기 때문에 주석란의 사용은 어렵다. 실제로 데이터가 없어도 데이터에 관한 연구이면 P(진행중)와 F(완료)를 부여하는 것도 핵 데이터센터측의 요구이다.

CINDA에서는 사신까지도 대상으로 취급하기 때문에 넓은 방법을 취하고 있지만 공개문헌만을 대상으로 하는 INIS에서는 필요성이 희박하다고 생각되어 2차안이후 없어졌다.

3. 데이터 Flagging 도입경위

3.1 도입배경

자연과학분야의 논문은 연구결과에 수치 데이터를 포함하는 경우가 많다. 수치 데이터는 새로운 연구와 개발의 기초가 되기 때문에 이들을 수집, 평가, 집성하여 유효하게 이용시키는 것은 많은 연구자, 학회, 협회의 커다란 관심사이다.

예를 들어 원자로 설계는 팽창 및 고도의 정밀 핵 데이터(원자핵의 성질, 특히 중성자의 핵반응에 관한 데이터)를 필요로 하기 때문에 그 수집, 평가, 유통을 위한 국제협력과 데이터센터의 설치가 비교적 빠르게 실시되었다. 중성자 핵 데이터에 관해서는 현재 4개의 국제센터를 중심으로 하는 조직망이 구성되어 있으며*, 실험데이터와 일부의 평가데이터를 교환함과 동시에 데이터 관련문헌의 색인인 CINDA를 공동으로 간행하고 있다.

협력체제는 1차간행물 및 이 분야의 유력저널 "Nuclear Physics"에 1965년 간행된 실험데이터에 관한 문헌에 한정하여 일정한 형태(Format)에 의한 색인어(Keyword)부여제를 도입했다*. 이것은 일종의 데이터 색인(Data Indexing)이며, 그후 Physical Review C, Journal of Physical Society of Japan 등의 잡지에서도 각각 1973, 1975년에 이 방식을 실시하였다.

핵 데이터 분야에 제한하지 않고 다른 분야에서도 데이터 유효 이용의 움직임이 높아져 이 분야의 국제협력을 추진하기 위해 ICSU(International Council of Scientific Union)는 1966년 CODATA(Committee on Data for Science & Technology)를 설치하고, 각 분야 공동의 과제인 신뢰성 있는 데이터의 생성, 유통을 가능하게 하기 위한 활동을 하고 있다*. 이 활동의 하나로써 1973년 ICSU/AB(Abstracting Board)와 공동협의회(Working Group)를 만들어 데이터의 수집, 평가에 해당하는 데이터센터와 초록지, 색인지를 취급하는 2차정보기관과의 협력관계에 대해서 2회에 걸쳐 조사를 실시하고 1975년 1월 다음과 같은 결론을 얻었다. 즉 ① 2차정보기관과 데이터센터가 공동으로 행하고 있는 서지활동의 상호교환성과 양립성을 높이는 일, ② 1차지, 2차지에 공통으로 적용되는 데이터의 식별, 분류법의 확립을 강조하고 있다. 다시 ②에 관련하여 한 분야에 적용되는 방법은 채택하지 않는다는 것을 첨가하고, 가까운

장래에 데이터 Flagging/Tagging은 피할 수 없기 때문에 적절한 투자를 할 필요가 있다고 결론을 내렸다.

이런 활동이 활발하게 진행되면서 국제적인 문헌정보시스템으로 성공을 거둔 INIS에 대해, 데이터 시스템측으로부터 요청이 있었던 것은 필연적이라 생각된다. 또 INIS 이용자 가운데는 핵 데이터 뿐만 아니라 분자, 원자 데이터를 포함한 문헌을 필요로 하는 이용자들이 많아지고 있다.

3.2 INIS에 도입된 과정

ICSU AB/CODATA 공동협의회의 제언이 있었던 1975년 10월 INIS Liaison Officer 회의에서 핵 데이터 Flagging법을 검토할 전문가 회의가 제안되어 다음해 4월회의가 열렸다. 이 결과 데이터 Flagging의 필요성이 인식되어 핵 데이터 분야에 국한하지 않고 INIS가 대상으로 하는 모든 분야의 데이터에 대해 데이터 색인을 하도록 권고했다¹⁰⁾. 데이터 색인의 방법에 관해서는 CODATA와 특정분야의 제언이 있었기 때문에 이들과 양립시키는 일, INIS 레코드의 구조, 형태를 변경치 않고 초록을 보다 통보적으로 하는 일 그리고 데이터 내용을 자유로이 기술하는 Tag의 추가 등에 대해서 토의했다. 이것을 기초로 하여 INIS 사무국안(1차안)이 작성되어 각국의 INIS센터와 데이터센터 등의 의견을 종합하고 Liaison Officer 회의에서 토의가 시작되었다.

1차안 및 그후의 안에 대한 개략은 표 1에 있다. 1차안의 특징은 그후의 것과 비교해 보면 데이터를 위해 자유로이 기술하는 새로운 Tag의 추가, 초록을 충실히 하는 일, 데이터의 형태구분과 제시구분의 코드화 그리고 이것을 기초로 데이터 참고파일(Data Referal File)¹¹⁾을 작성하는 것이다.

* 데이터의 형태, 주제를 포함하고 있는 문헌이 데이터센터의 어느 파일에 보관되어 있는가 하는 정보를 모은 파일, 데이터 포인트(Data Point)라고도 한다. 중성자 핵 데이터의 문헌색인인 CINDA는 그 예이다.

그후 2차안, 3차안이 제정되고 1978년 제6회 Liaison Officer 회의에서 데이터 Flagging에 대한 구체적인 방법이 승인되어 이번에 실시요령이 제정되었다.

데이터 Flagging이 제기된 이래 3년간에 걸쳐 토의되었다. 도입의 중심적 역할을 한 Gad-jakov*씨는 1차안에 대한 여러 의견을 정리하여 2차안을 작성하는 단계에서 다음의 5원칙을 설정하였다¹²⁾.

이 5원칙은 이번에 확립된 실시요령의 기반이 되고 있다. 즉 ① 종래부터 실시하고 있는 INIS색인의 일반적 요령을 변경하지 않는다. ② INIS레코드의 기본구조를 변경치 않는다. 특히 문헌속에 포함되어 있는 수치 데이터의 증가 또는 소개에 대해서는 기술하지 않는다. ③ 어떤 특정한 주제의 데이터센터와 병행 또는 후에 실시하는 데이터 색인을 대신하지 않는다. ④ INIS 입력단가의 증가가 적정치인 5%를 초과하지 않는다. ⑤ 장래 데이터 Flagging/Tagging 분야의 국제기준을 채택하는데 방해하지 않는다. 이 가운데 ①, ②, ④는 지금까지의 INIS 방식에서 크게 벗어나지 않게 한 것이고 ②, ③은 데이터센터의 본질과 처리할 데이터의 수집, 분석, 평가를 하지 않는다는 것을 의미한다.

데이터센터에서는 좁은 분야의 이용자에게 전문적인 사실, 데이터 그 자체를 정리하여 제공하는 것을 목표로 하고 있고, 입력작업의 중심은 얻어진 데이터를 분석, 평가하는 일 즉 최확치 정도를 구하는 것이다.

한편 INIS에서는 넓은 분야의 문헌정보를 제공하는 것을 목표로 하고 있으며, 데이터센터가 데이터의 「정확함」에 중점을 두고 있는 것에 비해 문헌수집의 「망라성」에 중점을 두고 있다¹³⁾. 데이터의 「정확함」은 데이터 문헌수집의 「망라성」에도 의존하기 때문에 INIS가 이 면에서는 협력한다고 말할 수 있다. 따라서 INIS는 데이터 그 자체의 사실검색(Fact Retrieval)까지는 목표로 하지 않는다는 것을 알 수 있다. ⑤는 ICSU AB/CODATA의 「보편성을 갖지 않는 방

법은 도입치 않는다」는 제언을 따른 것이지만, 지금까지 일반적으로 알려진 데이터의 식별 분류방법은 없다.

광범위한 분야를 대상으로 하고 있는 문헌정보시스템으로서 INIS가 최초이다. 몇개의 분야에서 제안된 식별, 분류방법을 참고로 했지만 INIS에 적합한 것은 없고 결국 독자적인 방법을 만들었다. 이들 5원칙이외에 중요한 점은 이용자에게 어떤 이익을 가져다 주느냐 하는 것이다.

INIS에 데이터 Flagging도입 시간이 걸렸던 것은 데이터 Flagging의 이용효과를 예측, 평가하는 것이 어렵기 때문이었다. 이점은 앞으로의 과제로서 충분히 검토되어야 할 것이다.

4. 입력자료의 데이터 Flagging 실시요령 및 규칙

4.1 데이터의 관점

수치 데이터에는 물질의 물리적 성질, 화학적 성질 등 이른바 물성치와 반응속도 등의 기술적인 특성치, 어떤 사고발생의 통계 등 여러가지 형태가 있다. 이러한 데이터의 특성은 여러 관점에서 파악할 수 있지만, 어떤 데이터든지 공통되는 관점을 정하는 것이 데이터 Flagging에 필요하다. 앞에서 언급한 바와 같이 몇개 학술기관의 방법과 INIS 1차안에 대한 의견을 참고로 하여 다음과 같은 9가지의 데이터 관점이 설정되었다.

- ① 데이터 발생국과 연구기관
- ② 데이터 게재 간행물의 형태(잡지, 레포트 등)
- ③ 데이터 게재논문의 주제분야
- ④ 측정된 양(물성치, 특성치 등)
- ⑤ 연구된 시스템(원소, 화합물, 핵종, 전자회로 등)
- ⑥ 연구에 포함된 과정(화학반응, 핵반응, 식물섭취 등)

* IAEA-INIS 창설이래의 Subject Control Unit의 Head.

1976년 8월에 이임한 이후 IAEA의 의뢰에 의해 이 문제의 고문으로 활약하고 있다.

표 1. INIS데이터 Flagging안의 변천

	1차안 1976. 5	2차안 1977. 9	3차안 1978. 3	실시안 1978. 10
데이터 유, 무	literary indicator N에 의함	좌 동	좌 동	좌 동
데이터 형태구분	literary Sub-indicator에 의함 T(theoretical) V(evaluated) C(compiled) E(experimental) R(requested) W(reviewed) U(used)	디스크립터에 의함 NUMERICAL DATA CALCULATED DATA EVALUATED DATA COMPILED DATA PRIMARY DATA REQUESTED DATA NON-NUMERICAL DATA	디스크립터에 의함 NUMERICAL DATA THEORETICAL DATA EVALUATED DATA COMPILED DATA EXPERIMENTAL DATA	좌 동
데이터 제시구분	literary Sub-indicator에 의함 G(graph) T(table) C(computer media) P(in progress) (데이터 없음) F(finished) (데이터 없음) S(기 타)	디스크립터에 의함 ISOLATED에 VALUES GRAPHS TABLES COMPUTER FILES	디스크립터와 형태코드에 의함 ISOLATED VALUES GRAPHS TABLES (형태코드) T	좌 동
데이터 주제내용	디스크립터에 D 부여 숫자(D ₁ , D ₂ 등) 에 의해 2종류이 상의 데이터를 구별	좌 동	좌 동	디스크립터에 D 부여 숫자가 없으며 2 종류이상의 데이터 를 구별하지 않음
초 록	데이터 기술을 통보적으로 함	—	초록의 경우도 데이터초록 첨부	—
주 석	데이터에 관한 정 보를 자유로이 기 술하는 새로운 Tag을 붙임	—	—	—
기 타	데이터 point 화일 작성	데이터 색인 작성	데이터 색인 작성	—

DATA

- BT1 information
- NT1 NUMERICAL DATA
- NT2 COMPILED DATA
- NT2 EVALUATED DATA
- NT2 EXPERIMENTAL DATA
- NT2 THEORETICAL DATA

그림 1. DATA의 word block

DATA FORMS

- NT1 GRAPHS
- NT1 ISOLATED VALUES
- NT1 TABLES

그림 2. DATA FORMS의 word block

- ⑦ 환경조건(극저온, 초고압 등)
- ⑧ 결정방법에 의한 데이터구분(실험, 이론, 편집 등)
- ⑨ 데이터 제시형태(그림, 도표, 컴퓨터파일 등)

이들 관점가운데 ①, ②, ③은 이미 INIS 서지기술, 주제분류와 디스크립터 세트에 충분히 반영되어 있다. ④~⑦은 데이터의 주제개념을 규정하는 것으로서 대단히 중요하다. 그러나 이 개념을 나타내기 위해서 새로운 코드 등을 도입치 않고 INIS Thesaurus 용어를 사용하는 것으로 충분하다. 그것도 색인시 부여한 디스크립터 속에 이미 기술되어 있다고 생각되어-기술되어 있지 않은 경우는 데이터 Flagging 대상이 되지 않음-이들 ④~⑦의 개념을 나타내는 디스크립터를 3.2.4에 설명한 것 같이 선정하여 D(D는 데이터를 의미함)를 부여하면 된다. ⑧은 디스크립터 DATA의 하위어(Narrower term)가운데 이에 해당하는 특정한 디스크립터를 선정하면 된다. DATA의 어군(Word block)은 그림 1과 같다.

⑨는 디스크립터 DATA FORMS의 하위어(NT) 가운데 그 데이터의 제시형태에 해당하는 디스크립터를 선정하면 된다. DATA FORMS의 어군은 그림 2와 같다.

때때로 데이터의 분석결과는 어떤 양의 계산울 나타내는 방정식의 형태로 주어질 수가 있다. 만약 문헌속에 포함되어 있는 유용한 결과중의 하나가 방정식에 해당하는 어떤 계수의 결정에 관한 것이라면 이 계수는 ISOLATED VALU-

ES로 나타낼 수가 있다.

위의 ⑧, ⑨는 어떤 문헌의 데이터 Flagging 할 때마다 반드시 수행해야 할 필요조건이다. 이제 2차안의 단계에서 위의 9관점외에도 1개의 관점이 고려되었다. 그것도 정량적, 정성적인 관점으로서 정성적 데이터는 비수치 데이터를 의미하며 예를 들어 수치가 부여되지 않은 화학구조와 결정구조도 등이다. 이것에 의해 2차안에서는 그림 1과 같이 데이터 형태구분을 나타내는 디스크립터로서 NON-NUMERICAL DATA가 첨가되었다. 또 데이터가 포함되지 않았어도 이용자로부터 그 측정과 계산의 필요성이 요구되는 가공의 데이터에도 REQUEST DATA를 부여하는 것이 고려되었다.*

그러나 요구(Requested)데이터와 비수치데이터는 3차안이후 그 도입이 필수적인 단계에 이르기까지는 보류하기로 했다.

4.2 데이터 Flagging 순서

4.2.1 지금까지 실시한 색인작업을 완료

데이터 Flagging은 주제전문가인 색인자에 의해 실시되지만, 지금까지 실시해온 색인을 완료한 후에 되어야 한다. INIS에 있어 색인작업은 주제분류, 필요한 경우 제목추가, 제공된 초록의 검토 그리고 주제분석에 의해 디스크립터를 부여하는 일 및 그 디스크립터 속에서 주제색인의 주포목-부포목이 될만한 대상을 뽑아 M(Main heading)-Q(Qualifier)를 부여하는 것이다¹⁴⁾

* 핵 데이터 분야에서는 요구데이터를 수집하여 측정을 촉구하는 체제가 되어 있으며, 이에 관련된 WRENDA(World Request List of Nuclear Data)를 발행하고 있다.

Descriptors	M, Q and/or D Labels
DEUTERIUM TARGET	M2, D
NEUTRON REACTIONS CAPTURE	M1, Q2, D;
ELASTIC SCATTERING	Q1, D
KNOCK-OUT REACTIONS	Q1, D
CROSS SECTION	D
NEUTRON SPECTRA	
ANGULAR DISTRIBUTION	
EV RANGE	D
KEV RANGE	D
MEV RANGE 01-10	D
MEV RANGE 10-100	D
GROUP CONSTANTS	D
EVALUATED DATA	D
GRAPHS	D
TABLES	D

그림 3. 색인과 데이터 Flagging에

그 예를 그림 3에서 볼 수 있다. 그림 3에서 고딕체로 쓰여지지 않은 디스크립터가 지금까지 INIS에서 실시해온 색인작업이다. 이것은 그림 5의 문헌에서 뽑은 것으로 그 내용은 「중수소를 표적(DEUTERIUM TARGET)으로 한 중성자의 포획, 탄성산란, Knock-out 등의 반응(NEUTRON REACTION, CAPTURE, ELASTIC SCATTERING, KNOCK-OUT REA-

CTIONS)에 대한 단면적(CROSS SECTION) 및 반응후의 중성자 스펙트라(NEUTRON SPECTRA)와 각도분석(ANGULAR DISTRIBUTION)에 관한 평가 그리고 중성자 에너지(0.0001eV부터 15MeV)에 대한 군정수(GROUP CONSTANTS)부여」이다¹⁵⁾.

M, Q가 된 DECTERIUM TARGET와 NEUTRON REACTION, NEUTRON REACTION과 CAPTURE 등은 그림 5의 ATOMINDEX 주제색인에서 볼 수 있으며, 그림 4와 같이 ATOMINDEX 본문의 끝에서도 볼 수 있다. 색인완료후 데이터 Flagging의 단계에서 새로운 디스크립터를 첨가하는 것은 3.2.5항을 제외하고는 금지되어 있다. 이것은 데이터 Flagging에 의해서 지금까지 실시해온 색인작업에 영향을 주지 않기 위함이다.

4.2.2 데이터 Flagging할 내용인가의 판단

색인자는 색인을 완료한 후에 그 문헌이 데이터 Flagging할 내용을 가졌는지 그 여부를 판단해야 한다. 판단기준은 「문헌속의 수치 데이터가 이용자에게 유용한가」 「수치 데이터를 포함하는 문헌을 필요로 하는 이용자가 있는가」의 평가가 되어야 한다.

「색인자는 자기 자신을 이용자의 입장에서 자문해 온다」는 것이 판단에 도움이 된다. 이 이

425191 Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV. Bazazyants, N.O.; Zbrodskaya, A.S.; Larina, A.F.; Nikolaev, M.N. International Atomic Energy Agency, Vienna (Austria). International Nuclear Data Committee. INDC(CCP) -107/L. Aug 1978. 63 p. Translated from Russian.

The paper describes the evaluation of deuterium neutron cross-sections, the spectra of neutrons from the reaction $D(n, 2n)P$ and the angular distributions of neutrons from this reaction and of neutrons elastically scattered on deuterium. The evaluation results are presented in the SOCRATOR format. The 26-group system of constants for deuterium is also presented. (author) [Data]
 NEUTRON REACTION: capture, elastic scattering, knock-out reactions;
 DEUTERIUM TARGET: neutron reactions.

그림 4. INIS ATOMINDEX 본문

DEUTERIUM TARGET:

alpha decay

Excited states of ^4He in the ($\text{Li}^6, \text{ } \text{)}$ and ($^3\text{He}, \text{p}$) reactions. (A34). 425274

chemical preparation

Preparation of solid deuterated polystyrene targets (E41). 427273

neutron reactions

→ Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV. (A34, A33). [Data] 425191

NEUTRON REACTIONS:

capture

Neutron capture cross section measurements of ^{99}Tc up to 80 KeV. (A34). 425318

→ Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV. (A34, A33). [Data] 425191

elastic scattering

Neutron cross-sections calculations for ^{239}Pu , ^{243}Pu and ^{235}Pu , ^{237}U and ^{239}U in the 1-150 KeV energy region. (A33). [Data] 425105

→ Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV. (A34, A33). [Data] 425191

knock-out reactions

Integral measurement to test evaluated neutron cross sections for carbon. (A34; 1 to 20 MeV). 425334

→ Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV. (A34, A33). [Data] 425191

그림 5. INIS ATOMINDEX 주제색인

상의 자세한 기준은 부여하고 있지 않기 때문에 유용성의 판단은 색인자에 의존하게 된다. 그러나 지금까지의 과정으로 보아 다음과 같은 것은 알 수가 있다.

① 앞에서 말한 정성적 데이터와 요구 데이터는 대상으로 하지 않는다. 단 앞으로 포함될 가능성이 있다. 물론 그래프도 수치가 있으면 대상이 된다.

② 반복하여 측정할 수 없는 데이터, 측정장소에 의존한 데이터도 대상이 된다.

③ 인용 데이터는 적용되지 않는다. 단 데이터의 평가, 수집 그리고 다른 데이터 센터간의 비교 등은 대상이 된다.

소규모 테스트의 결과에 의하면 입력문헌의 약 반수가 데이터 Flagging의 대상이 되는 것 같다.

4.2.3 Literary Indicator N에 O표

INIS에서 Literary Indicator는 레코드 형태에 대한 실체(Physical entity)를 좀더 구체적으로 지적해 주기 위한 것으로 1개이상 O표 할 수 있다. 해당되는 경우 생략할 수 없는 것과 있는 것 등 10종류가 있다.

N(수치데이터), K(회의자료), L(사전), U(학위논문), W(규격 및 명세서), Z(문헌목록),

Y(진행중인 보고서), E(초록)이 있는데 이들은 문헌검색시 Indicator에 의해 간단히 선별할 수가 있다.

N은 지금까지 수처 데이터 즉 수학 함수표, 물리 데이터표와 같이 표 및 도표로 구성된 문헌에 적용되어 왔지만, 앞으로는 데이터 Flagging에만 사용하기 때문에 INIS구조의 형태(Format)변경은 필요치 않다. 그렇기 때문에 데이터 Flagging되는 입력자료는 Literary Indicator N에 O표를 해주어야 한다.

4.2.4 디스크립터에 D를 부여

지금까지 해온 색인에서 부여된 디스크립터 가운데서 수처 데이터의 개념을 나타내는 디스크립터를 뽑아 그림 3과 같이 D를 부여한다. 여기에서 주의할 점은 linked group이 없을 경우는 아래와 같이 실시하면 되지만, 만약 1개 이상의 linked group이 있다면 색인자는 문헌에서 보고된 데이터가 어느 linked group에 속하는지를 밝혀야 한다.

그림 3에서 D가 붙은 디스크립터를 합치면 [중수소의 중성자 반응에 대한 단면적의 균정수에 관한 데이터]의 의미가 된다. D부여시 다음 사항에 유의하여야 한다.

① 측정된 양(Quantity)에는 반드시(Mandatory) D를 부여한다. 그림 3의 CROSS SECTION, GROUP CONSTANTS가 이에 해당한다.

D는 Worksheet Tag 800에 디스크립터와 같은 줄의 세번째 란에 기입하며 어떤 디스크립터에 M, Q, D가 모두 부여된 경우 쉼표에 의해서 각각을 분리시킨다. 예를 들면 NEUTRON, REACTION: M₁, Q, D; 식으로 기입한다.

② 연구된 시스템, 과정, 환경조건을 나타내는 디스크립터에는 필요에 따라(Optional) D를 부여한다. 그림 3에서 DEUTERIUM TARGET와 NETURON REACTION은 데이터 관점사항의 ⑤에, CAPTURE, ELASTIC SCATTERING, KNOCK-OUT REACTION은 ⑥에 EVRANGE는 ⑦에 해당하는 디스크립터이다.

③ 데이터 Flagging된 문헌의 디스크립터 중에서 M과 D가 동시에 붙어있는 디스크립터는

항상 1개 이상이다. 이것은 주표목 M이 부여된 디스크립터 중에 적어도 하나는 D를 부여해야 한다는 것을 의미한다.

INIS 자기테이프에 기초를 둔 기계검색 및 INIS ATOMINDEX에 의한 수작업(Manual search)으로 수처 데이터를 포함하고 있는 문헌의 소재를 파악하는 것이 데이터 Flagging의 목적이다. ATOMINDEX 주제색인의 주표목은 M이 부여된 디스크립터에 의해 결정되기 때문에 수처 데이터를 포함하고 있는 문헌을 지시하기 위해, M과 D가 동시에 부여된 디스크립터가 적어도 1개 이상 있으면 주제색인부의 문헌제목 다음에 [DATA]가 인쇄된다.

그림 3의 DEUTERIUM TARGET 및 NEUTRON REACTION은 M, D가 동시에 부여되었기 때문에 그림 5의 주제색인에 [DATA]가 나와 있다.

4.2.5 데이터 형태구분, 제시구분을 나타내는 디스크립터 선정

디스크립터 DATA 및 DATA FORMS의 하위어(그림 1, 2의 NT)로부터, 각각 1개 이상의 디스크립터를 선정하여 D를 부여한다. 그림 3의 EVALUATED DATA, GRAPHS, TABLES가 이에 해당하며 [평가치를 갖고 있으며, 표와 도표를 나타내고 있다]를 의미한다. 이들 디스크립터는 데이터 Flagging이외에 부여해서는 안 된다. 이상의 D가 부여된 디스크립터는 지금까지 해온 색인에서 선정된 디스크립터의 Subset에 해당한다. 이 Subset은 INIS 자기테이프에 기록되어 수처데이터를 포함하는 문헌의 기계검색에 제공될 뿐만 아니라 책자형태로는 그림 4, 5에서 보는 바와 같이 ATOMINDEX의 본문, 주제색인에도 [DATA]가 인쇄되어 데이터를 포함하는 문헌인가를 식별할 수 있다.

5. 결 론

지금까지 몇개의 분야에서 제안된 데이터의 분류와 1979년 INIS에 도입된 데이터 Flagging에 대해 소개했다. 광범위한 주제분야를 취급하는 정보시스템인 INIS가 처음으로 실시한 데이

터 Flagging은 앞으로 데이터가 포함되어 있는 문헌의 검색에 어떻게 이용될 것인가가 주목된다. 특히 D 부여에 대한 유효성에 관해서는 여러 관점에서 평가되어야 한다. 그러나 무엇보다도 최대의 과제는 데이터 Flagging여부의 경계를 정할 구체적인 지침을 확립하는 일이다.

프랑스, 서독의 테스트에 의하면 데이터 Flagging되는 문헌의 경우 중성자물리, 고에너지물리가 77%, 원자물리, 고체물리가 58%, 생명과학이 46%라고 하지만 그 판정기준을 명확하게 할 정도까지는 못된다. 따라서 앞으로의 경험을 통해서 이 판정기준을 명확히 하고 각국 INIS 센터간의 통일을 기하여야 한다. 또 이번에 데이터 Flagging의 대상에서 보류된 정성적 데이터와 요구데이터의 확장은 데이터 Flagging 이용효과의 평가가 된 후에 다시 검토되리라 생각된다.

참 고 문 헌

- 1) ICSU AB/CODATA Joint Working Group; "Flagging and Tagging Data to Indicate Its Presence and Facilitate Its Retrieval," Inf. Rept. Bibl., 6, 8 (1977)
- 2) V. Gadjokov "Data Indexing in INIS—Practical Steps Towards Improving the Identification of Data Sources," Institute of Nuclear Research and Nuclear Energy, Sofia, Bulgaria, 12 (1977)
- 3) United Nations Educational, Scientific and Cultural Organization. UNISIST "Study on the Problems of Accessibility and Dissemination of Data for Science and Technology" SC. 74/WS/16, Paris, France, (1974)
- 4) ICSU AB/CODATA Joint Working Group "Flagging and Tagging Data to Indicate Its Presence and Facilitate Its Retrieval," Inf. Rept. Bibl., Vol. 6 (1977), No. 4, P. 7~25.
- 5) J. J. Schmidt (IAEA), Private Communication; 1977 (1977)
- 6) J. F. M. Potters (Compiler). Comments on Data Tagging and Flagging: A Contribution to the Discussion with Respect to INIS Technical Note No. 22; Netherlands Bibliographical and Documentary Committee (COBIDOC), Amsterdam, Netherlands, (1976)
- 7) 更田豊治郎 "核データセンタの活動," 日本原子力學會誌, 20, 323 (1978)
- 8) Important Announcement, Nucl. Phys. 68, 1 (1965)
- 9) "牛島悦子 "科學技術データの入手と流通," ドクメンテーション研究, 25, 9 (1975)
- 10) INIS "Treatment of Nuclear Data Sources in INIS," INIS Technical Note No. 22 (1976)
- 11) V. Gadjokov "Data Indexing in INIS—Problems and Approaches," Proc. 5th Biennial Intern. CODATA Conference, Boulder, Colorado, USA, 203~208 (1977)
- 12) V. Gadjokov "Data Indexing in INIS—Practical Steps Towards Improving the Identification of Data Sources," INIS Technical No. 32 (1978)
- 13) INIS "Procedures for Data Flagging of INIS Input," INIS Circular Letter No. 77 (Oct. 1978)
- 14) 日埜衆司, "シソーラス型 デスク립タの主題標目への流用の可能性," ドクメンテーション研究, 24, 419~423 (1974)
- 15) 横尾 宏, 高橋智子, "INISに ずける データフラッキングの 實施," ドクメンテーション研究, 29, 271~272 (1979)

[부록] 1) 데이터 Flagging실시이전의 INIS Worksheet 절면

INIS WORKSHEET
BIBLIOGRAPHIC AND INDEXING DATA

1 (Use a separate Worksheet for each level circled in the Bibliographic Level box starting with the left most level and enter code box 009 for serial entries use section 2 of this Worksheet Use Abstracts Worksheet for abstracts)

Tag	Data (enter by typewriter only)
100	Bazayants, N.O.; Zabrodskaya, A.S.; Larina, A.F.; Nikolaev, M.N.
110	International Atomic Energy Agency, Vienna (Austria). International Nuclear Data Committee
200	Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV
201	
210	
211	
213	
230	
231	
250	INDC(CCP)-107/L
300	
310	
320	
401	
402	
403	Aug 1978
500	63 p.
600	
610	Translated from Russian
611	
620	
700	
710	3295100 XA

009 S (enter by typewriter only)

230
231
320
403
500
610

Executive Catalogue

Indexer
Date Completed

Puncher

PLEASE TURN OVER

INIS Form 1 (Rev 5)

2) 데이터 Flagging 실시 이전의 INIS Worksheet 후편.

3

009 9 800 Descriptors
Level

Link Indicators	Descriptors	Main Heading and/or Qualifier Labels
1	DEUTERIUM TARGET	M ₂
2	NEUTRON REACTION	M ₁ , Q ₂
3	CAPTURE	Q ₁
4	ELASTIC SCATTERING	Q ₁
5	KNOCK-OUT REACTIONS	Q ₁
6	CROSS SECTIONS	
7	NEUTRON SPECTRA	
8	ANGULAR DISTRIBUTION	
9	EV RANGE	
10	KEV RANGE	
11	MEV RANGE 01-10	
12	MEV RANGE 10-100	
13	GROUP CONSTANTS	
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		

810 Proposed Descriptors

Definition of Inter

g and Suggested Associated Terms

3) 데이터 Flagging실사이절의 INIS Worksheet : Abstracts

001 1A 7803821 002 2/2

**INIS
WORKSHEET**
ABSTRACTS

4 009 / EN (Use a separate Worksheet for each language version of the abstract)

	Tag	Data <small>(enter by typewriter only)</small>
Language	850	English
Abstract	860	The paper describes the evaluation of deuterium neutron cross-sections, the spectra of neutrons from the reaction D(n,2n)P and the angular distributions of neutrons from this reaction and of neutron elastically scattered on deuterium. The evaluation results are presented in the SOCRATOR format. The 26-group system of constants for deuterium is also presented. (Author)

Abstracter
Proofreader
A 16 (May 75)

Puncher
Date Completed

4) 데이터 Flagging 실시후의 INIS Worksheet 전면

001 LA 7803821 002 1/7 003 R W 004 C D 005

006 007

008 A 34 A 33

009 M

INIS WORKSHEET
BIBLIOGRAPHIC AND INDEXING DATA

NON DRAWING
D F O H I J P R T
A M Y C
X L U W Z Y E X O

1 009 M (Use a separate Worksheet for each level circled in the Bibliographic Level box starting with the left most level and enter code in box 009. For serial entries use section 2 of this Worksheet. Use Abstracts Worksheet for abstracts.)

Tag	Data
100	Bazayants, N.O.; Zabrodskaya, A.S.; Larina, A.F.; Nikolaev, M.N.
110	International Atomic Energy Agency, Vienna (Austria). International Nuclear Data Committee
111	
200	Neutron cross-sections of deuterium in the energy range 0.0001eV-15MeV
201	
210	
211	
213	
230	
231	
250	
300	INDC(CCP)-107/L
310	
320	
401	
402	
403	Aug 1978
500	63 p.
600	
610	Translated from Russian
611	
620	
700	
710	3295100 XA

2 009 S

Tag	Data
230	
231	
320	
403	
500	
610	

Descriptive Cataloguer
Proofreader
A 14/Rev. 13 (Oct 76)

Indexer
Date Completed

Puncher

PLEASE TURN OVER

NIS Form 1 (Rev 5)

5) 데이터 Flagging 실시후의 INIS Worksheet 후면 Abstracts은 데이터 Flagging 실시후에도
전과 같다.

3

009 **9** 800 Descriptors
Level

Link Indicators	Descriptors	Main Heading and/or Qualifier Labels
1	DEUTERIUM TARGET	M ₂ , D
2	NEUTRON REACTIONS	M ₁ , Q ₂ , D
3	CAPTURE	Q ₁ , D
4	ELASTIC SCATTERING	Q ₁ , D
5	KNOCK-OUT REACTIONS	Q ₁ , D
6	CROSS SECTIONS	D
7	NEUTRON SPECTRA	
8	ANGULAR DISTRIBUTION	
9	EV RANGE	D
0	KEV RANGE	D
11	MEV RANGE 01-10	D
12	MEV RANGE 10-100	D
13	GROUP CONSTANTS	D
14	EVALUATED DATA ②	D
15	GRAPHS } ③	D
16	TABLES } ③	D ④
17		
18		
19		
20		
21		
22	① Literary Indicator Not 0 ①	
23		
24	② 데이터 평가구분	
25		
26	③ 데이터 제시구분	
27		
28	④ 해당 디스크방식에 D부여	
29		
30		
31		
32		
33		
34		
35		

810 Proposed Descriptors

Definition of Intended Meaning and Suggested Associated Terms

--	--