

연구중에 생산된 과학기술 참조데이터 관리에 관한 연구

A Study on Management of Scientific Technology Reference Data that is Produced during Research

채 균 식(Kyun-Shik Chae)*
이 응 봉(Eung-Bong Lee)**

목 차

- | | |
|------------------------|-----------------------------|
| 1. 서론 | 3. 2 측정표준 분야 참조 데이터 확보 |
| 1. 1 연구필요성 및 목적 | 4. 실험데이터 관리시스템 설계 - 측정표준 분야 |
| 1. 2 연구내용 및 방법 | 4. 1 데이터 처리 과정 |
| 2. 참조표준 정의 및 연구동향 | 4. 2 실험데이터 구성요소 |
| 2. 1 참조표준 동향 | 4. 3 실험데이터 분류 |
| 2. 2 국내동향 | 5. 실험데이터 관리시스템 구현 |
| 2. 3 해외동향 | 6. 국가참조표준 구축 체계 방안 |
| 3. 실험 데이터 관리 시스템 구축 | 7. 결론 |
| 3. 1 지적자산으로서의 실험데이터 관리 | |

초 록

신뢰도가 공인된 과학기술데이터는 같은 분야 연구자들이 지속적이면서 반복적으로 사용할 수 있도록 물리화학적 상수, 공인된 물성값, 또는 공인된 과학기술적 통계 값을 가지고 있어야 한다. 이러한 측정데이터는 정보의 정확도와 신뢰도를 과학적으로 분석 평가하여 공인함으로써 국가적으로 모든 분야에 널리 사용할 수 있는 참조표준데이터(Standard Reference Data)로 관리되어야 한다. 본 논문에서는 연구중에 생산된 다양한 실험데이터를 국가 참조표준 데이터로서 관리되어야 하는 필요성과 국가참조표준 센터 설립을 위한 방안을 제시해 본다. 그 시도로 측정표준 연구정보를 하나의 참조정보로 이용할 수 있도록 측정표준 연구정보 시스템을 개발하였고 이 시스템에 대한 운영체계를 자세히 소개함으로써 국내 참조정보 관리에 대한 방안을 제시하고자 한다.

ABSTRACTS

Scientific/technical data whose reliability is authorized have to include physical/chemical constants, authorized property value or authorized scientific statistical values so that the scientists in the same field use them continuously and repeatedly for their research. The data also have to be managed as Standard Reference Data that will be used in other area nationwide by analyzing, evaluating and authorizing the accuracy and reliability of the information with scientific methods. This study suggests a plan for establishing the National Center for Standard Reference Data as well as a necessity for managing the data produced in research process as a national Standard Reference Data. Also, a prototype of research information system for measurement standards is developed and described in detail for research information for measurement standard to be used as a reference data, then a plan to managing national reference data is suggested, in addition.

키워드: 과학기술데이터, 참조표준데이터, 참조데이터, 측정표준, 연구정보시스템
Scientific/Technical Data, Standard Reference Data, Reference Data, Measurement Standard, Research Information System

- * 한국표준과학연구원 정보전산그룹(cks@kriss.re.kr)
** 충남대학교 사회과학대학 문헌정보학과 부교수(eblee@cnu.ac.kr)
논문접수일자 2003년 11월 17일
게재확정일자 2003년 11월 26일

1. 서론

1.1 연구 필요성 및 목적

인터넷의 급속한 확산으로 이용자들은 과거에 비해 훨씬 다양한 정보를 시간과 공간의 제약 없이 언제든지 사이버 공간에서 이용할 수 있게 되었다. 인터넷에서 제공하는 다양한 정보들은 우리의 일상생활 환경을 바꾸어 놓을 뿐만 아니라 전문적인 지식을 습득하는데도 필수적인 수단으로 자리잡아 가고 있다. 그러나 이용자들이 인터넷에서 보다 전문적인 정보를 얻고자 할 때 마땅히 이용할 전문정보가 절대 부족하다고 느끼고 있으며, 설령 제공되는 정보가 있을지라도 이용료를 요구하거나 제한된 정보만 제공하는 경우가 많다. 특히, 국내의 경우 과학기술자들의 의해 생산되는 다양한 과학기술정보가 있음에도 불구하고 관리와 제공을 위한 체계적인 환경이 마련되어 있지 않아 연구 기록의 손실을 가져오고 있으며, 이러한 정보들이 국가 과학기술 지식정보 관리 차원으로 시급히 관리되어 과학기술자들이 공유함으로써 연구생산성 향상과 국가 경쟁력 향상에 기여하도록 해야 한다.

과학기술정보는 두가지로 크게 나누어 볼 수 있다. 학술지를 비롯한 학술회의, 기술보고서 등에 수록된 텍스트 위주의 연구논문 정보와 연구과정 중에 발생한 수치나 도표 등으로 보다 전문화된 실험데이터인 사실정보(Factual Data)로 나누어 볼 수 있다. 연구논문이나 회의록과 같은 텍스트 위주의 학술정보원은 출판사와 제공기관의 노력으로 디지털

화되어 전자저널이나 특정 연구사이트를 통해 제공되고 있으므로 라이선스만 있으면 언제든지 이용할 수 있다. 그러나 과학기술 전문 연구기관에서 생산된 다양한 실험데이터는 그 중요성에 비해 관리가 거의 되지 않고 있다. 만약 연구 중에 생산되는 실험데이터가 체계적으로 관리되어 제공된다면 같은 분야 연구자들은 서로 정보 공유가 가능하게 되고 활발한 정보교류와 중복연구 방식을 통해 연구의 질적 수준을 향상시킬 수 있을 것이다. 특히 과학기술자들이 생산한 연구정보의 공유체계가 단절됨으로 인해 같은 분야 연구자들이 같은 결과 값을 얻기 위해 또 다시 많은 시간과 노력을 기울이게 됨으로 국가 전체적으로 볼 때 비경제적인 연구 활동이 될 수밖에 없다. 따라서 어떤 연구 분야의 실험데이터가 같은 분야 연구자나 기술개발자들이 지속적이면서 반복적으로 사용할 수 있도록 물리화학적 상수, 공인된 물성값, 또는 공인된 과학기술적 통계 값에 근접할 수 있는 실험 데이터를 구축하는 것이 매우 중요하다.

이러한 측정데이터들은 정보의 정확도와 신뢰도를 과학적으로 분석·평가 하여 공인함으로써 국가적으로 모든 분야에 널리 사용할 수 있는 참조표준 데이터(Standard Reference Data)로 관리되어야 한다. 실험데이터가 참조표준 데이터로 인증 받기 위해서는 국제적으로 요구되는 실험조건에서 생산되어야 하고, 그 분야에 세계적으로 인정을 받는 전문가의 평가가 뒷받침되어야 한다. 그러나 국내에서 생산되는 과학기술 정보는 아직까지 선진국 수준의 실험 환경과 조건에 적합한 환경에서 생산되는 정보라고 보기 어렵기 때문에 초기 시도에서는

참조표준데이터(Standard Reference Data) 보다는 참조데이터(Reference Data)에 가까운 데이터로 관리하여 점차적으로 그 범위와 체계를 확대하여 참조표준 데이터 구축을 해 나가야 한다. 그렇게 함으로서 국내 과학기술자 및 산학연 잠재적 이용자들로부터 필요성에 대한 인식이 생겨나게 하여 다양한 분야의 국내 순수 과학기술 참조표준 데이터로 발전해 나갈 수 있도록 해야 한다.

1.2 연구내용 및 방법

한국표준과학연구원에서는 국가표준을 확립하여 산업체에 보급함으로써 제품의 국가 경쟁력 향상을 목표로 하는 대표적 측정표준 연구기관이다. 길이, 질량, 온도를 비롯한 다양한 기초과학 분야의 측정표준 연구능력은 국제적 수준에 도달해 있지만 각 분야별로 측정표준 확립 위해 생산되는 다양한 실험데이터는 아직까지 체계적으로 관리되고 있지 않고 있다. 따라서 국제적으로 신뢰도 있는 측정표준 능력을 인정받기 위해서는 해당 분야의 측정표준 기술 능력과 더불어 최종 기술이 얻어지기까지 과정도 매우 중요하게 관리되어야 한다. 그 중에 하나가 신뢰성 있는 데이터 값이다. 특히 정밀 정확도 측정을 위해 연구되어지는 측정표준 연구는 하나의 측정값을 얻기 위해서 반복적으로 실험이 이루어지는 경우가 많다. 그 결과 값의 가치를 신뢰받기 위해서는 실험 중에 생산된 데이터를 체계적으로 관리하여 관련 전문가에게 인정을 받음으로써 보다 신뢰도 높은 연구결과를 얻을 수 있다.

이러한 맥락에서 한국표준과학연구원에서는 국제수준의 측정표준 기술을 산업체에 제공함과 동시에 국가 참조표준 확립을 위해서도 많은 노력을 하고 있다. 그 일환으로 연구원이 국가 참조표준 센터로서 역할을 수행 위해 연구원 내부에서 생산된 다양한 연구 실험 데이터를 수집, 관리, 제공할 수 있는 체계 구축을 위해 제도를 정비하고 관련시스템을 구축하여 운영하고 있다.

본 논문에서는 참조표준에 대한 정의와 국내외 연구동향을 살펴보고, 참조표준 데이터 관리의 일환으로서 측정표준 실험데이터의 관리 방법을 위해 한국표준과학연구원에서 개발한 연구정보관리시스템을 설명한다. 또한, 연구원에서 추진하고 있는 국가 참조표준 정보 관리의 체계적 운영방안을 소개하고자 한다.

2. 참조표준 정의 및 연구동향

2.1 참조표준 정의

참조표준이라 함은 측정 데이터 및 정보의 정확도와 신뢰도를 과학적으로 분석·평가하여 공인함으로써 국가사회의 모든 분야에서 널리 지속적으로 사용되거나 반복 사용이 가능하도록 마련된 자료로서 물리 화학적 상수, 공인된 물성값, 공인된 과학기술적 통계 등을 말한다(국가표준기본법 제3조 7항).

참조표준 데이터는 물리화학뿐만 아니라 지리학, 천문학, 생명과학, 재료공학, 환경공학, 등에 관한 Quantitative Numerical Data가 대

〈표 1〉 과학기술 데이터 분류

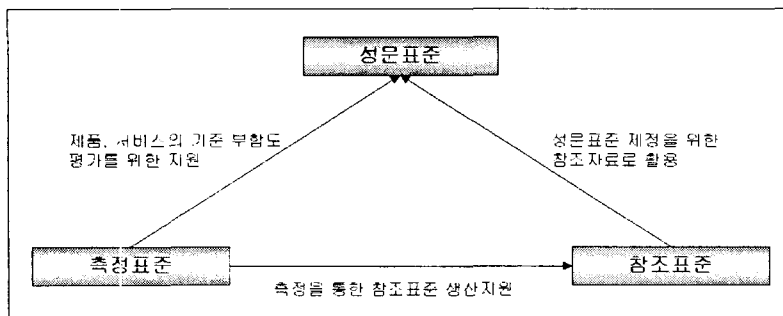
종류	특성	예
Class A	잘 정의된 시스템에 의해 얻어지는 데이터로서 재현성이 있다(Repeatable measurement on well defined system).	물리, 화학, 재료 등
Class B	관측에 의해 얻어지는 데이터(Observational data)	생물, 천문, 지구과학 등
Class C	조사에 의해 얻어지는 통계적 데이터(Statistical data)	인구 통계, 인체골격 등

부분이며 데이터가 생산되기까지 측정수행과정, 대상재질에 대한 세분된 성질, 취급된 현상 또는 특수한 조건(온도, 압력 등)을 동시에 수록해 놓음으로써 차후 수행되는 다른 연구들에 중요한 기초 정보로 활용할 수 있도록 하기 위함이다.

참조표준 데이터가 국제적으로 중요하게 다루어지는 이유는 여러 가지가 있겠으나 우선 생각해 볼 수 있는 것이 첫째 과학기술 선진국가로서 위상에 관한 것이다. 즉 국제적으로 공인될 정도로 신뢰도 높은 측정 데이터를 생산할 수 있는 나라는 이미 선진국의 반열에 오른 국가들이다. 따라서 참조표준체계가 확립되어 있다고 한다면 그것은 그 나라에서 생산되는 데이터의 신뢰도에 대한 보증이 되며 국가적인 이미지의 개선에 크게 기여하게 된다. 둘째 산업제품에 대한 품질 보증 역할을 한다. 참조표준은 국가적 권위의 제고라는 무

형의 산업 경쟁력 확보에 기여할 뿐만 아니라 국내에서 생산되는 데이터들의 신뢰도 향상이라는 효과를 거두게 되어 각종 산업제품의 설계, 재료 선택 등에 유용하게 활용될 수 있으며 국가적으로 국산제품의 대외경쟁력 향상에 도 기여하게 된다. 셋째 국가적 차원에서 산업 기술 보호 기능이 있다. 최근 들어 선진국들의 제조 기술은 후진국에 이양하고 자신들은 설계 기술과 품질관리 기술을 운용하여 고 부가가치 산업형태로 바뀌고 있다. 수출품의 특성 데이터에 대한 측정 자료가 국가적으로 공인을 받은 경우와 그렇지 않은 경우는 대외경쟁력에서 많은 차이가 있다.

측정표준, 성문표준, 참조표준은 국가표준을 이루는 하나의 주요 요소로서 다루어지고 있으며, 참조표준 데이터로서 가치를 인정받기 위해서는 신뢰성 평가가 이루어진 데이터만이 인정받을 수 있는 것이다. 따라서 한국표



〈그림 1〉 국가표준상관관계도

준과학연구원에는 측정표준 연구를 기반으로 다양한 분야의 참조표준 정보를 생산할 수 있으며 동시에 전문가에 의한 신뢰성 평가 체계도 갖추고 있다.

2.2 국내동향

국내에서 필요로 하는 참조표준 정보는 미국을 비롯한 선진국가에서 참조표준 데이터를 구입하거나 인터넷을 통해 이용하는 실정이다. 과학기술 일반에 대한 참조표준 정보는 비용을 지불하고서 해외에서 구입할 수 있지만 국내에서 사용되어야 할 참조표준 정보는 구할 수가 없다. 예를 들면 한국인의 인체치수나 한국인에 필요한 의학 신물질 등은 국내에서만 필요한 정보이므로 반드시 표준화된 정보가 아닐지라도 참조정보로만 사용할 수 있다면 산업파급 효과는 매우 클 것이다.

국내에서는 한국표준과학연구원과 한국과학기술정보원에서 참조표준에 대한 관심을 가지고 국가 참조표준 센터 설립을 위해 많은 노력을 해오고 있다. 참조표준 정보가 신뢰도 공인 여부에 비중을 두고 데이터로서 관리가 되어야 하기 때문에 데이터의 신뢰도를 평가할 수 있는 연구력을 가진 기관에서 주도적으로 이끌어 가야한다. 한국표준과학연구원의 경우 CODATA의 기본물리상수 Task Group에서 발간하는 자료목록에 14편이 인용되는 등 미국 NIST와 비슷한 정밀측정능력에 대한 인식이 국제적으로 알려져 있다.

따라서 국내에서 확보한 참조표준 데이터만으로 새로운 산업시대를 준비할 수 없으며, 이를 보완하기 위해 선진국에서 보유한 고급 과

학기술정보를 이용해야 할 경우 정작 필요한 데이터는 철저히 보호되어 이용하기가 어렵다. 이제 과학기술정보는 더 이상 국제사회에서 공유의 대상이 아니라 지적소유권의 대상이 되고 있다. 산업의 고도화로 기술이 축적되면서 자체 설계 능력이 고양되었고, 이에 따라 각종 플랜트설계, 산업제품디자인 등에 있어서 기본적으로 필요한 수치 데이터에 대한 수요가 점차 증가하고 있다. 우리나라는 제조업 수준은 향상되었으나 고부가가치의 제품 설계 기술은 상대적으로 취약하다. 따라서 설계 기술에 기본이 되는 수치 및 물리 상수 참조표준 데이터가 필요한 것이다.

2.3 해외동향

선진국들은 측정표준과 성문표준에 대한 기초를 확립하여 이를 기반으로 기술이 발달하였고 관련 데이터의 생산량은 늘어났지만 신뢰성 있는 데이터 확보의 필요성은 더욱 절감하고 있다. 이에 따라 수치 데이터를 생산하고 공인하는 기간들을 분산형으로 지정하여 참조표준 데이터를 운영하고 있다.

초창기 참조표준 데이터 관리를 시도한 미국의 경우 과학기술분야의 데이터가 축적되면서 데이터 품질에 대한 문제가 제기되었다 (David R. Lide, Jr, 1984). 그 중 가장 큰 문제는 어떤 물질에 대한 데이터의 내용이 서로 달라 어느 것을 사용할 지에 대한 기준이 마련되지 못했다. 이를 위해 국립표준원(NIST-National Institute of Standards and Technology)에서는 참조표준(Standard Reference Data) 센터를 설립하였고, (Brady,

Edward L 외 1964; Franz L, 1966) 수치 데이터의 신뢰도를 평가하여 산업계 및 과학기술계에서 필요로 하는 공인된 데이터를 공급하고 있다(<http://www.nist.gov>). NIST 참조표준 자료에 대한 운영의 기본은 생산과 관리가 분산형으로 이루어진다는 것이다. 데이터생산 및 배포는 NIST에서 주관은 하지만 미국 전역에서 다양한 기관들이 참여한 데이터센터 개념으로 운영하고 있다. 분산형 데이터 센터를 운영하고 있지만 가장 중요한 것은 참여기관들이 데이터 생산에 적합한 연구 환경과 관련분야 저명한 전문가들에 의한 데이터 평가가 얼마나 신뢰도 있게 평가되는지에 초점을 두고 있다. NIST SRD 센터에서 데이터 평가 기준은 1) 데이터의 생산과정이 규격과 얼마나 잘 부합하는가? 2) 생산과정이 잘 명기되어 있는가? 3) 물리적 법칙과 잘 일치하는가? 4) 다른 측정방법이나 계산결과와 비교하여 결과값이 어떤가? 에 평가 기준을 삼고 있다.

이밖에도 유럽과 일본에서도 국가차원에서 지원하고 있으며 일본의 경우 국립금속재료 연구소(NRIM), 국립무기재료연구소(NIMC)의 재료 및 화학분야 과학기술자들이 재료데이터 관리자로 활동하고 있으며 초전도, 피로, 세라믹, 강도 등 66개 분야에서 활동하고 있다. 유럽국가들도 EU를 중심으로 CRIDS(Community Research and Development Information Service)를 설립하여 CODATA DB의 지적재산권 부여에 관한 것을 CODATA Working Group에 상정하는 등 활발한 활동을 하고 있다.

3. 과학기술 실험데이터 관리시스템 구축

3. 1 지적자산으로서의 실험데이터 관리

정밀 측정분야를 연구하는 연구원들은 하나의 값을 얻기 위해 반복적인 실험이나 측정 활동을 한다. 따라서 연구원들은 반복적으로 측정된 데이터가 장차 유용한 데이터로서 가치가 있을 것이라는 생각을 하고 있지만, 이를 축적, 관리 할 수 있는 기반 여건이 마련되어 있지 않아 유사한 연구 분야에 활용되지 않고 있다. 또한 연구원들이 특정분야의 연구를 수행하다가 근무처를 옮겨갈 경우 기관에서는 새롭게 관련 연구환경을 셋업해야 하고 오랫동안 축적해온 경험과 노하우가 손실되는 경향이 있다. 따라서 이러한 문제점을 극복하여 실질적인 연구정보 공유와 지적 자산을 보호하기 위해서는 실험데이터 관리가 중요하다.

측정표준 연구 분야는 최종적으로 정확한 측정값을 얻기 위해 반복적인 실험을 거쳐야 하는 대표적인 연구 분야이다. 따라서 한국표준과학 연구원의 주 연구 분야인 측정표준에 대한 실험데이터의 중요성을 인식하여 데이터를 축적할 수 있도록 시스템을 개발하였고, 연구자 스스로 참여할 수 있는 제도적 장치를 마련하였다. 보다 구체적으로 살펴보면 1) 데이터 신뢰도 확보를 위해 내부 전문가에 의한 데이터 검증 - 데이터 관리자가 되어 제공된 입력 필드에 정보가 정확히 입력되었는지, 입력한 정보가 측정표준 참조 정보로서 가치가 있는 것인지를 검증한다. 2) 측정표준 참조 정보로서 가치가 인증되면 연구자들의 개별 업적 고과 점수로 인정해 주는

시스템을 도입하여 운영하고 있다.

3. 2 측정표준 분야 참조 데이터(Reference Data) 확보

정밀정확도가 요구되는 산업분야에서 제품의 불량률을 최소화시켜 소비자들이 원하는 제품을 생산하기 위해서는 신뢰도 있는 측정표준 참조정보를 활용해야 한다. 연구자들은 최종 데이터가 생산되기까지 연구환경을 비롯한 데이터 생산조건 등을 상세히 기술해 줌으로써 데이터 신뢰도를 향상시킨다. 실험실내의 환경상태를 최적의 상태로 유지 관리할 수 있도록 온도, 습도, 먼지, 전자기차폐, 진동, 소음 등의 연구실 환경기준을 관리하고 수시로 점검하도록 한다(ISO/IEC17025, 1999).

이용자들은 연구원에서 구축한 신뢰도 있는 측정표준 연구정보를 활용함으로써 추가적인 실험이나 연구 없이도 만족한 연구결과를 얻을 수 있도록 지원한다. 구축된 데이터가 참조표준 데이터로서 인정받기는 미흡한 부분이 있지만 연구결과로 얻어진 데이터이기 때문에 참조데이터(Reference Data)로서 가치는 있다고 판단된다. 연구원들은 자신이 수행하는 연구분야의 생산 데이터가 이용자들에게는 매우 중요한 데이터로 사용된다는 가정 아래 신뢰도 있는 데이터 생산을 할 수 있도록 유도한다. 이러한 신뢰도 있는 데이터가 축적된다면 연구원 내부에서는 측정표준 분야의 보다 전문적인 데이터를 확보하게 됨으로 연구원 이직에 따른 연구 중단이나 지적자산의 유실을 최소화할 수 있을 것이다. 연구원에서는 제도적으로 신뢰도 있는 데이터 검증과 데이터 제공자에 대한 적

절한 보상으로 측정표준 참조데이터의 질적 양적 확대를 기대할 수 있을 것이다.

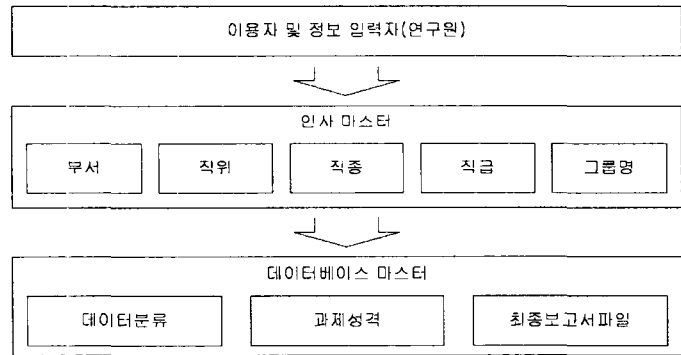
4. 과학기술 실험데이터 관리시스템의 설계 - 측정표준분야

측정표준 연구정보관리시스템의 주요 특징은 1) 관리자에 의한 데이터 신뢰도 확보 2) 전문적 지식에 의한 데이터 분류체계 구축 3) 데이터 입력자에 대한 보상 체계 도입 4) 웹 기반 정보처리 및 제공 시스템 5) 연구자별로 입력정보 파악기능 등이다. 연구자들이 웹 환경에서 데이터를 입력하고 관리자에 의해 입력된 데이터를 확인하여 최종 데이터로 저장된다. 데이터 처리의 모든 과정이 시스템에 의해 자동화가 되어 있지 않고, 관리자에 의해 부분적으로 수동적인 처리가 된다. 이는 데이터의 신뢰도와 품질관리를 위한 방안으로써 연구자들이 입력한 데이터가 데이터로서 가치가 있는 정보인지, 제시된 포맷에 맞게 입력되었는지를 확인하여 최종 저장하게 된다. 본 시스템이 지속적으로 발전하기 위해서는 기능적인 보완도 필요하지만 연구자들이 참여로 데이터 양적, 질적 확보가 무엇보다 중요하다. 따라서 연구원에서는 데이터의 질적, 양적 수준 확보를 위해 정책적으로 데이터를 입력한 연구자에게 입력권을 환산하여 인사고과점수로 반영하는 제도를 마련하였다.

측정표준 연구정보시스템은 입력된 데이터를 신속히 접근할 수 있는 것에 초점을 맞추었으며 중복데이터 입력 방지, 데이터 품질유지를 위한 정보등록 및 수정, 게시관, 공지사항,

〈표 2〉 측정표준 연구정보 시스템 테이블 스키마

테이블명	설 명	비 고
insa	인사마스터	연구원 정보 관리 테이블, 관리자에 의한 관리
b_buseo	부서	소속 부서, 관리자에 의한 관리
b_jikwi	직위	연구원의 직위, 관리자에 의한 관리
b_jikeub	직급	연구원의 직급, 관리자에 의한 관리
b_jikjong	직종	연구원의 직종, 관리자에 의한 관리
b_sil	그룹명	연구자가 소속된 그룹명 관리
p_data	데이터마스터	연구정보가 입력될 데이터 마스터 테이블
b_category	데이터분류	데이터마스터
p_charact	과제성격	연구원 과제 관리 DB와 연계
p_report	최종보고서파일	해당 데이터의 보고서 정보
게시판	게시판	관리자에 의한 초기화 가능, 수정
공지사항	공지사항	관리자에 의한 작성
FAQ	찾은 질문	관리자에 의한 작성



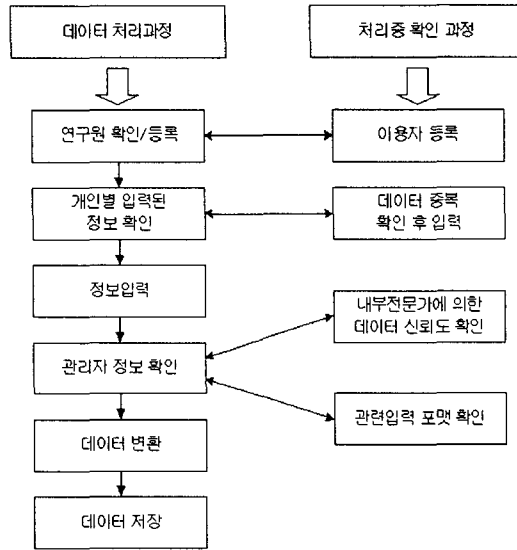
〈그림 2〉 시스템 주요 구성 정보

FAQ등으로 구성하여 설계하였다. 본 시스템에 사용된 컴퓨터 환경은 서버로서 Linux, 360Mhz x 2, RAM512MB이고, RDBMS는 MySQL이며 브라우저는 Explore5.5 이상을 기반으로 구축하였다.

4. 1 데이터 처리과정

측정표준 연구정보가 저장되어 이용되기까

지는 입력한 데이터의 신뢰도와 검증을 거쳐 관리자에 의해 PDF 포맷으로 변환하여 이용자들에게 제공된다. 데이터 처리 과정에서 다소 많은 부분을 전문가와 관리자에 의해 수동적으로 다루어진다. 이는 데이터 품질과 신뢰도와 관련된 것으로서 측정표준 연구정보 데이터로서 가치와 통일된 데이터 포맷을 유지 시킴으로서 데이터 신뢰도 향상에 많은 노력을 기울였다.



〈그림 3〉 데이터 처리과정

데이터 처리과정은 〈그림 3〉에서 보는 것과 같이 입력된 데이터는 관리자와 내부 관련 전문가에 의해 인정을 거쳐 최종 데이터로 저장된다. 데이터의 양적인 면도 중요하겠지만 신뢰성 있는 데이터를 제공함으로써 이용자들은 측정표준 연구정보에 대한 신뢰성을 가지게 될 것이다.

입력된 데이터가 무분별하게 유출되거나 배포되는 것을 최소화하기 위해 근본적인 데이터 보호 역할은 할 수 없지만 PDF 파일로 변환하여 관리함으로써 연구원 내부 지적자산을 보호할 수 있는 최소한의 기능은 한다고 본다.

4. 2 연구정보 데이터 구성요소

실험중에 생산된 데이터가 데이터로서 가치를 인정받기 위해서는 적절한 데이터 생산 환경에서 일정한 포맷을 가지고 전문가 평가에 의해서 만이 신뢰도를 유지 할 수 있다. 데이

터에 대한 신뢰도 평가는 첫째 보다 정확한 데이터를 관련 이용자들에게 제공할 수 있으며, 둘째 보다 정확한 데이터를 데이터베이스에 저장 보관하고, 셋째 실험자들은 자신의 연구결과에 대한 신뢰성을 스스로 보장 할수 있고, 넷째 이를 바탕으로 실험, 분석방법 등에 대한 개선 사항을 제시 할 수 있다.

한가지 예로 물성 자료에 포함되어야 할 정보는 실험방법, 데이터의 정확성, 오차한계, 추정오차, 참고문헌, 문헌의 신뢰도, 주석, 데이터의 신뢰도 평가 정보(평가여부)를 일반적으로 고려 해 볼 수 있다. 본 측정표준 연구관리시스템에서는 이러한 데이터 정보를 기초로 데이터 분류를 비롯한 보다 상세한 정보가 기술 되도록 〈표 3〉 과 같이 구성하였다.

4. 3 데이터 분류

본 시스템을 설계할 때 가장 많은 노력을

〈표 3〉 데이터 구성 요소

입력대상 정보	내 용	입력조건	필드 구성
제목	입력할 데이터의 내용을 가장 잘 나타낼 수 있는 함축적인 내용	필수	
데이터 분류	제시된 분류항목을 참고하여 입력할 데이터가 가장 적절히 속할 수 있는 것 선택	필수	
데이터형태	수치데이터, 이미지데이터, 기타	필수	한 개 이상 선택 가능
대상 물질	실험데이터에 사용된 물질		
데이터생산조건	데이터 생산에 사용된 여러 조건 (특정 연구장비, 온도, 압력, 습도등)	필수	
상세정보보기(URL)	상세한 연구결과를 가진 전용사이트 연결을 위한 URL 기록		
불확도 범위	실험한 최종데이터의 불확도 범위		
적용분야	데이터의 응용 및 적용 분야		
과제명	입력한 데이터가 생성되기까지 수행한 과제명		
연구기간	yyyy.mm.dd-yyyy.mm.dd		
과제성격	기간고유, 수탁, 특정, 공업기반...		
참고문헌	참고문헌		
최종데이터	그림(그래프), 테이블, 수치	필수	
공개여부	공개, 비공개 선택		
입력일자	시스템에서 자동 부여		

기울인 부분이 데이터 분류이다. 데이터 분류는 기본적으로 한국표준과학연구원에서 수행하는 연구 분야를 어느 정도 망라할 수 있어야 하고, 물질의 특성에 따라 체계적으로 저장되어 관리될 수 있도록 하였다. 데이터 분류의 구성은 〈표 4〉에서 보는 바와 같이 4개 분야로 크게 나누고 각 분야별로 하위 주제 분류로 이루어져 있다.

5. 실험데이터 관리시스템 구현

본 연구에서 개발한 측정표준 연구정보시스템은 데이터 제공, 입력, 관리에 대한 모든 과

정을 웹 인터페이스로 구성하였다. 이용자 측면에서 데이터를 편리하게 이용할 수 있도록 웹 원문처리 어플리케이션으로 많이 사용되는 PDF 포맷으로 제공하고, 관리측면에서는 데이터 등록은 곧 개인의 인사고과 점수와 연결되기 때문에 중복 방지와 데이터 평가를 통한 품질관리에 많은 노력을 기울였다. 데이터 입력자인 연구원들은 수시로 자신이 입력한 데이터를 전체적으로 확인해 볼 수 있으므로 스스로 중복입력 방지와 기존데이터 수정, 삭제 기능을 동시에 확인해 볼 수 있도록 하였다. 연구원들이 직접 입력한 데이터는 곧바로 이용자들에게 제공되는 것이 아니라 내부 전문가의 적절한 평가 및 부족한 사항을 보완한

〈표 4〉 데이터 분류표

1차 Property	2차 Field
Physical and Chemical Properties	Acoustics and Vibration Clinical analysis Electrochemical analysis Environment and Food analysis Fluid and Flow Force and Gravity Frequency Gas Composition Humidity and Moisture Inorganic Analysis Laser Applications Length and Dimensional Metrology Mass and Density Nano technology Nuclear Metrology Optical Properties Organic Analysis Pressure and Vacuum Thermometry Thermophysical Properties Time Torque
Electric and Magnetic Properties	Biomagnetism Capacitance and Inductance Colossal Magnetoresistance Electromagnetic Field Electromagnetics Electrostatic Discharge Josephson standard and application MEMS devices Magnetic measurements Millimeterwave Nano technology RF and Microwave SET, DNA, NSOM, TMR devices SQUID and its application Superconductive Electronics Voltage/Current Thin Film Process
Materials and Mechanical Properties	Ceramic membrane characterization Microstructure characterization Nano technology Non-destructive Evaluation Strength and Hardness Structure analysis by X-ray Thin film and surface properties
Human and Information Technology	Anthropometry and Biomechanics Computing Technology Environmental Technology Ergonomic Topics Information Technology Instrument Technology Nano technology Public welfare Statistics & Survey

후 최종 데이터로서 활용할 수 있도록 하였다.

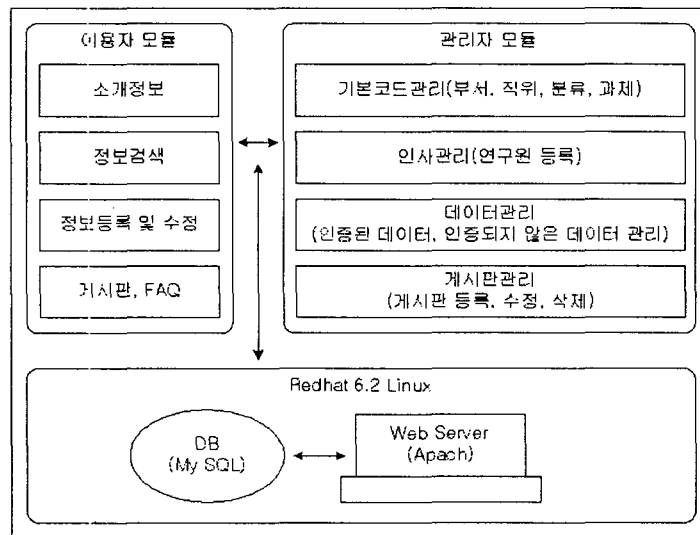
데이터는 입력자에 의해 공개와 비공개로 제한을 둘 수 있도록 하였으며 비공개인 경우 데이터의 일반적인 서지정보만 제공되고 원문 정보는 제공되지 않기 때문에 볼 수가 없다. 이럴 경우 데이터 입력자와 직접 연결하여 원하는 정보를 얻을 수 있도록 하였다.

시스템의 기본 구조는 사용자모듈과 관리자 모듈로 크게 나누어진다. 사용자 모듈에서는 일반이용자가 구축된 데이터를 검색해보는 기능과 입력자가 데이터를 직접 구축하는 부분으로 나누어진다. 관리자는 시스템 운영에 필요한 모든 정보관리뿐만 아니라 입력자가 올려놓은 데이터를 평가하고(데이터 인증 부분), 인증하는 작업을 관리자 모듈에서 실행한다. 각 모듈에서 처리된 데이터는 구조상 무관하게 연결되어 있지 않고 상호 작용으로 데이터가 저장되고 관리된다. 입력된 데이터는 서버 내의 데이터베이스에 저장되어 제공 관리된다.

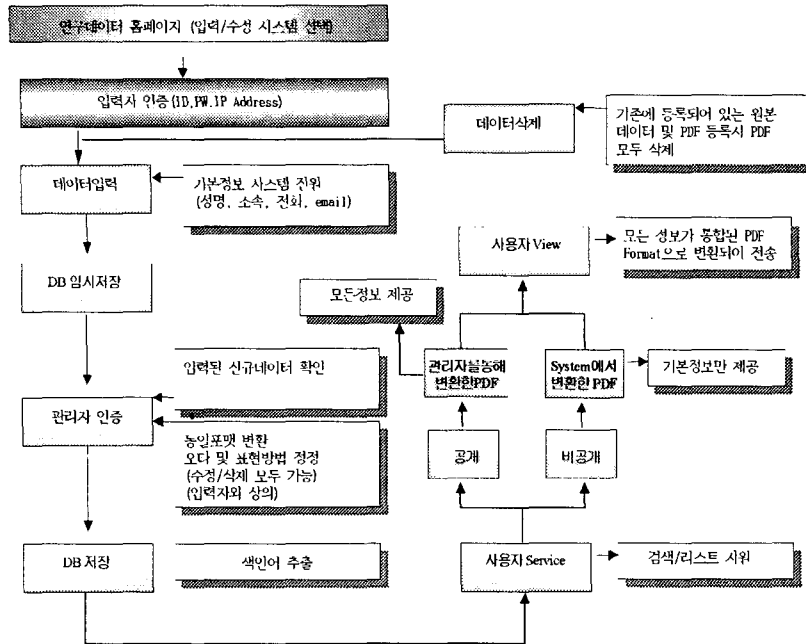
측정표준 연구관리시스템은 입력한 정보를 관리자가 확인하고 관리자는 데이터에 문제없으면 인터넷상에서 제공될 수 있도록 수동적인 데이터 변환과정을 거쳐 최종 데이터로 이용된다. 데이터는 한국표준과학연구원에서 수행한 연구결과물을 관련 수요자들에게 보급 확산하는 차원에서 무료로 공개한다.

측정표준 연구정보 사이트의 초기화면에서는 이용자들이 구축되어 있는 데이터를 신속히 검색해 볼 수 있도록 검색창을 제공하고 있다. 그밖에 기능은 일반적인 사이트와 비슷한 형태로 측정표준연구정보 소개, 게시판, 공지사항, FAQ로 구성되어 있다.

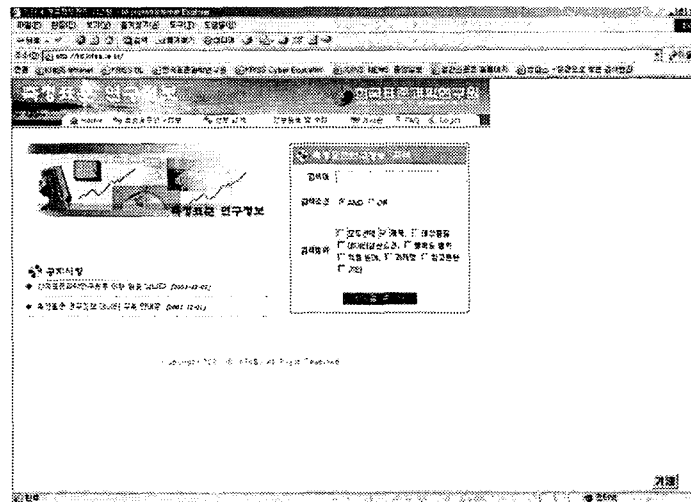
정보입력을 위해서는 연구자가 자신의 정보를 등록하여 로그인 할 수 있는 정보를 부여 받아야 한다. 시스템에 로그인이 되면 자신이 현재까지 입력한 정보 리스트를 보여주게 되는데 이 기능은 데이터의 중복입력을 방지하기 위해서이다. 정보 입력자는 입력에 앞서 도



<그림 4> 시스템 구성도



〈그림 5〉 데이터 입력된 데이터의 시스템적 처리 과정



〈그림 6〉 초기화면

음말 기능을 이용하여 정확한 입력방법을 숙지하게 된다.

입력된 정보는 직접적으로 시스템에 반영되는 것이 아니라 관련 전문가와 관리자의 인증

을 거친 후 시스템에 등록된다. 데이터 등록은 제공된 항목에서 반드시 입력해야 할 항목을 구분해 줌으로서 데이터 질적 수준을 유지할 수 있도록 하였다. 정보 입력자는 대상정보의

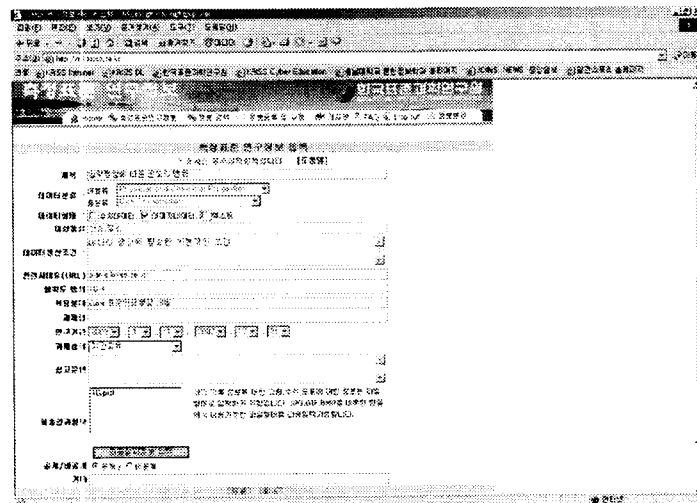
기본적인 요소를 입력한 후 본 시스템에서 가장 핵심적인 데이터인 수치 및 그래프, 테이블 형태의 최종 데이터를 입력하게 된다. 최종 데이터는 대부분 파일 형태로 소유하고 있기 때문에 파일을 등록할 수 있도록 하였다.

입력시스템에서는 데이터의 품질과 신뢰성을 위해 기본적으로 입력해야 할 필수입력사항이 제공된다. 따라서 연구원들은 입력대상 정보중 가장 기본적인 정보인 제목, 데이터분류, 데이터형태, 데이터생산조건을 입력해야 한다. 데이터 생산조건에 대한 기록은 가능한 면 간략하면서 명확한 정보 입력을 위해 데이터 입력 길이를 제한하고 있다. 입력데이터가 공개 가능한지를 선택하게 함으로써 연구원 지적 자산 보호에 할 수 있도록 하였다. 비공개를 선택하면 입력한 정보중 가장 중요한 최종 정보인 첨부 화일(수치, 이미지, 그래프)에 대한 정보를 볼 수가 없고 필요하다면 직접 저자와 개별 접촉으로 데이터를 얻을 수 있도록 유도하고 있다. 입력한 정보는 관리자 및

전문가에 의해 최종 확인하기 전까지 수정할 수 있도록 하였다.

<그림 10>에서 보는 것과 같이 입력한 데이터가 공개일 경우 관리자에 의해 PDF로 저장하여 관리되지만, 비공개일 경우 PHP의 PDF Library를 이용하여 자동적으로 보여 주며 사용되는 library는 PDF, JPEG, TIFF library를 이용하여 정보를 변환한다.

인터넷에서 효율적인 정보검색을 위한 많은 연구가 있어 왔다. 정도율과 재현율이 보장된 성능 좋은 검색엔진을 사용하여 이용자가 요구하는 질의에 신속히 정보제공을 할 수 있다면 완벽한 정보검색시스템으로 각광받게 될 것이다. 그러나 아직까지 이러한 환상적인 검색엔진이 개발된 것이 없기 때문에 정보검색 시스템과 보완적인 방법으로 수작업에 의한 분류/요약 작업을 함으로써 검색 효율성 향상을 도모해야 할 것이다. 정보에 대한 분류/요약이 시스템적으로 자동화되기 위해서는 완벽한 인공지능을 갖춘 시스템이 요구되어야 하



<그림 7> 정보입력 화면

기 때문에 문헌정보학에서 다루는 분류기술은 검색성능을 향상시키는데 중요한 역할을 한다 (Eugene Garfield, 2001).

기존의 웹 정보검색시스템은 수없이 많은 가공되지 않은 웹 정보들을 웹 정보검색시스템 데이터베이스에 수집하여 분석한 후 검색에 편리하도록 저장하였다가 이용자의 정보요구가 있을 때 적합한 정보를 검색하여 제공하는 방식을 사용한다(백준호 외 1999).

검색시스템은 단순한 키워드 검색을 기반으로 각 검색필드별로 이용자가 질의한 내용을 용어별 좌우 절단에 의해 질의어와 매칭하여 정보를 찾아낸다. 아직까지 많은 정보가 수록되어 있지 않아 정보검색에 대한 시스템적인 부담은 크게 느끼지 못하고 있지만 데이터 량이 많이 늘어나더라도 충분히 현재 검색기능으로 감당할 수 있을 것이라 생각된다. 정보검색을 위해 제공되는 기능은 간략정보와 상세

정보 검색 두가지를 함께 이용할 수 있도록 하였다.

본 사이트의 정보등록은 한국표준과학연구원 연구자들이 주로 입력한다. 정보입력은 최종 결과정보를 기준으로 공개와 비공개를 입력할 때 지정하게 되어 있다. 공개일 경우 관리자가 등록자가 입력한 최종파일을 기존의 서지정보와 함께 변환하여 기록하게 된다. 비공개일 경우는 시스템내에 정보로서만 관리되고 일반이용자는 최종 데이터를 볼 수 없다. 비공개일 경우 시스템 내에서 PDF Library를 이용한 자동 PDF 변환 기능을 제공한다.

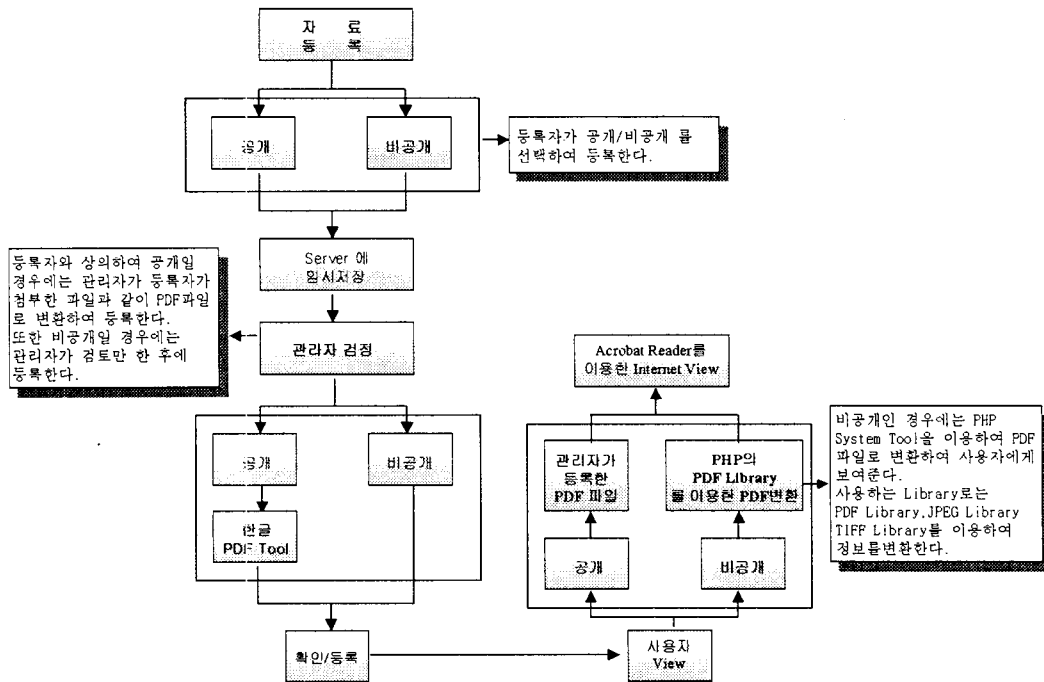
현재까지 본 시스템에서는 정보등록을 위해 관리자에 의한 수동적인 방법이 많이 동원되기 때문에 데이터 관리에 인력과 시간이 많이 소요되므로 원문정보 자동처리 기능을 보완함으로써 발전적인 시스템으로 개발해 나가야 할 것이다.

The screenshot shows a search interface with the following sections:

- 특정표준 연구정보 검색**: Search criteria section with fields for search terms, conditions (AND/OR), and various filters like '모두선택', '제목', '대상물질', etc.
- 데이터 검색리포트**: A table displaying search results with columns for '순번' (No.), '제목' (Title), '참조데이터' (Reference Data), '(비)공개' (Public/Private), '등록자' (Registered User), and '등록일' (Registration Date).

순번	제목	참조데이터	(비)공개	등록자	등록일
21	21 K2O의 열화상도	수치데이터	공개	김종철	2003-01-02
20	20 SnS의 열화상도	수치데이터	공개	김종철	2002-12-31
19	19 SnO의 열화상도	수치데이터	공개	김종철	2002-12-31
18	18 In의 열화상도	수치데이터	공개	김종철	2002-12-31
17	17 Kovar의 열화상도	수치데이터	공개	김종철	2002-12-31
16	16 W-Cu의 열화상도	수치데이터	공개	김종철	2002-12-31
15	15 알루미늄의 열화상도	수치데이터	공개	김종철	2002-12-31
14	14 자외선광학의 열화상도	수치데이터	공개	김종철	2002-12-31
13	13 알루미늄의 열화상도	수치데이터	공개	김종철	2002-12-31
12	12 Cu의 열화상도	수치데이터	공개	김종철	2002-12-31
11	11 주석도판의 열화상도	수치데이터	비공개	최병일	2002-12-31
10	10 U-2wt%Mo의 열화상도	수치데이터	비공개	최병일	2002-12-31
9	9 Ni-10Cr의 열화상도	수치데이터	공개	최병일	2002-12-31
8	8 Ni-10Cr의 열화상도	수치데이터	공개	최병일	2002-12-31
7	7 Ni-10Cr의 열화상도	수치데이터	공개	최병일	2002-12-31

<그림 8> 검색 참 및 검색 결과



〈그림 9〉 원문정보 처리과정

- 【 제목 】 승리부서 권열 실장 일제승리화대 개수
- 【 데이터분류 】 Materials and Mechanical Proper
- 【 데이터형태 】 수치데이터
- 【 대상물질 】 type 403 스테인레스강
- 【 데이터생산조건 】 열처리: 1050C 90분 austmitc: 액세서 rising load test
- 【 불확도 범위 】 None
- 【 적용범위 】 내식부품의 설계, 재료의 개발, 열처
- 【 과제명 】 승리 부식 시험 기술 개선
- 【 연구기간 】 2001-01-01 ~ 2002-11-31
- 【 과제성격 】 특경
- 【 참고문헌 】 한국표준과학연구원 보고서 KRISs
- 【 공개/비공개 】 공개

〈그림 10〉 PDF로 변환된 원문정보

6. 국가참조표준 구축 체계 방안

참조표준 데이터는 새로운 연구나 기술개발에 공업규격이 준용되어 제품관리, 공정관리

의 효율화를 이루는 것처럼 실험에 참고되도록 범 국가적 차원에서 준비해야 한다. 이렇게 함으로서 연구기간 단축, 경비절감이 이루어지고, 결과치와 비교에 의해 측정결과의 보증

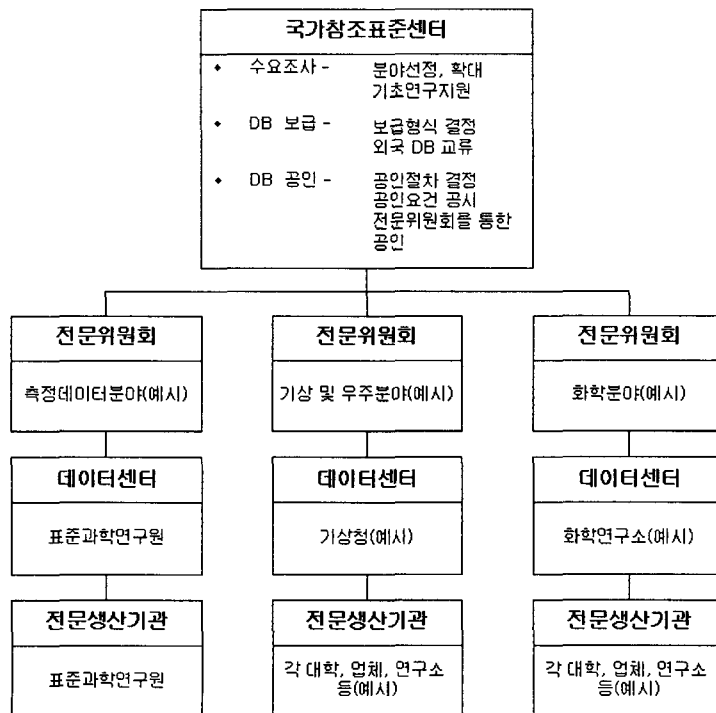
을 얻을 수 있는 traceability(소급성)을 통해 막대한 효과를 얻을 수 있다.

앞서 밝힌바와 같이 선진국에서 생산되는 일반 과학기술 분야 참조표준 정보는 예산만 있으면 얼마든지 국내에서 이용할 수 있다. 그러나 국내에서만 필요로 하는 참조표준 정보는 반드시 국내 과학 기술인에 의해서 만들어 지고 축적되어야 한다. 이러한 정보들은 반드시 표준화 정보가 아닐지라도 최소한 참조할 수 있는 정보로 이용할 수 있도록 체계를 구축해야 한다.

순수 국내 참조 정보를 확보하기 위해서는 국가산업 전 분야에서 필요로 하는 과학기술 정보를 1-2년내에 양질의 DB로 구축하는 것은 어렵다고 생각된다. 오히려 짧은 기간에 확

립하려는 것은 많은 시행착오로 실패할 가능성이 높다. 따라서 참조표준체계를 공인된 데이터 생산, 수집, 평가, 보급 등이 자동으로 진행되는 것을 전제로 소규모 단기목표를 두고, 이의 운영을 통해 확보된 Know-How를 기초로 범국가적 “국가참조표준체계”를 구축하는 것이 바람직할 것이다.

이를 위해서는 다음과 같은 추진계획을 가지고 체계를 구축해야 할 것이다. 1) 데이터의 생산현황 및 수용조사 2) 데이터 생산, 수집, 평가 절차 확립 - 자체 생산 데이터 항목에 대해, 공인을 받기 위해 필요한 데이터 생산 요건을 작성한다. 3) DB 구축 및 보급 4) 분야별 전문 데이터 관리 센터 지정 운영으로 국가 참조표준체계를 확립해야 할 것이다.



<그림 11> 국가참조표준센터 설립방안

국가참조표준센터를 통한 참조표준체계 구축 사업은 실질적으로 DB를 생산하는 것이 아니라

DB를 체계적으로 생산할 수 있는 시스템을 구축하는 것이 중요하다. DB 구축 방법에 있어 참조표준 DB의 구축 유형은 전략적인 DB 구축 사업을 통하는 방법과 불특정 다수의 연구 결과를 수집, 평가하여 축적하는 형태의 두 가지로 구분된다. DB구축 초기 단계에는 전자의 유형이 중심이 될 수 있으나, 이후 운영 단계에서는 후자의 경우가 중요하고 항구적인 방법이 될 수밖에 없다. 국가참조표준센터의 기능은 전자보다는 후자의 경우에 대해 비중을 둔다.

범국가적인 참조표준 DB 생산 시스템을 과학기술 전 분야에 걸쳐 일시에 구축하는 것은 많은 시행착오의 소지를 안고 있다. 특히 개개 분야의 데이터에 대해 신뢰성을 공인한다는 것은(공인 절차, 요건 등을 정하는데 있어서) 치밀한 사전 준비를 필요로 한다. 따라서 본 사업에서는 참조표준 데이터의 생산 능력과 평가능력을 갖춘 1개 기관을 지정하고 분산형으로 기관 내부의 "참조표준센터"를 설립하여 자체 생산 데이터의 수집, 평가, 가공, 보급 체계를 우선 확립하는 방식을 택하고자 한다. 1) 제 1단계 추진 전략 - 참조표준 데이터 생산 능력과 평가 능력을 동시에 확보하고 있는 유일한 국가 연구기관인 한국표준과학연구원을 대상으로 "참조표준센터"를 설립 운영하고, 추후 "국가 참조표준 센터" 설립에 필요한 Know-How(업무 체계, 시간, 경비, 인력 소요)를 확보한다. 2) 제 2단계 추진 전략 - 분야별로 데이터 생산 과정에 대해 품질 관리체계 공인(KOLAS 또는 ISO)을 획득한

국가연구기관을 대상으로 데이터센터를 지정하고, 표준연 "참조표준센터"를 모델로 분야별 데이터 생산 및 평가 체계를 확립한다. 공신력 부여를 위한 데이터 공인 절차 및 이에 필요한 생산 요건을 전국의 연구기관에 공시한다. 3) 제 3단계 추진 전략 - "국가참조표준센터"를 설립하고 전국의 연구기관에서 생산되는 과학기술 기초 데이터를 수집, 평가하여 DB로 구축하고 보급할 수 있도록 한다.

7. 결 론

실험중에 생성되는 과학기술 데이터는 데이터별로 물리화학적 상수, 공인된 물성값, 공인된 과학기술 통계값을 가지고 있게 되면 연구자들은 이들 데이터를 지속적이며 반복적으로 사용할 수 있을 것이다. 이러한 데이터는 과학기술분야에 종사하는 많은 연구기관들이 데이터 중요성을 인식하여 적극 데이터 구축에 협력해야 한다. 이를 위해서는 정보주도로 새로운 법적 제도적 장치를 마련하여 국가 참조표준 보급 체계를 구축해야 할 것이다.

과학기술 연구에 필요로 하는 많은 데이터는 아직까지 선진국 주도로 관리, 제공되고 있다. 이러한 데이터는 예산만 있으면 대부분 구입가능하기 때문에 이용에는 큰 어려움이 없을 것이다. 그러나 국내 과학기술 연구에 필요한 참조표준 데이터는 국내 과학기술인들만이 생성할 수 있는 것이다. 예를 들면 한국인의 체위 조사, 한국인체질에 맞는 의학 신물질, 다양한 국내식품, 국내 산업체에서만 이용되는 각종 자본재들은 반드시 표준화된 정보

가 아닐지라도 최소한 참조할 수 있는 정보로만 존재해도 가치는 있다고 본다.

본 연구에서는 이러한 국내 참조정보 구축의 일환으로 한국표준과학연구원에서 생산되는 측정표준 참조정보(Reference Data)를 구축할 수 체계적으로 구축할 수 있도록 시스템을 개발하여 운영하고 있다. 구축된 이러한 데이터는 아직까지 참조표준 정보로서 인증받기는 어렵지만, 연구자와 이용자들의 호응도 유도, 미흡한 시스템 기능 보완, 데이터 구축 및 제공에 대한 체계적인 보상체계 확립 등에 많은 노력이 필요하다. 또한 연구원 내부 전문가들에 의한 데이터 신뢰도 검증으로 질

적 수준 유지와 지속적인 구축을 통한 양적 성장으로 조만간 순수 국내 측정표준 과학기술 정보로서 가치를 발휘할 것으로 기대한다.

무엇보다도 측정표준 참조정보 확보를 위한 측정표준 연구정보 시스템 구축은 순수 국내 과학기술정보 확보를 위한 새로운 시도라고 할 수 있으며, 국내 과학기술 연구기관들을 중심으로 특성있는 과학기술 정보를 구축해 나간다면 몇 년 내로 가시적인 효과를 거둘 수 있을 것이라고 생각된다. 이를 위해서는 정부 주도로 국내 참조표준 센터를 설립하여 데이터의 생산과 보급 활용에 이르는 종합적인 체계를 마련하는 것이 매우 중요하다.

참 고 문 헌

- 백준호, 최준혁, 이정현. 1999. 한국어 웹 정보 검색시스템의 정확도 향상을 위한 연구와 피드백 에이전트 『한국정보처리학회』, 6(7): 1832-1840.
- 국가표준기본법 제3조 7항, 참조표준의 용어 정의.
- David R. Lide, Jr. 1984. "The National Standard Reference Data System of the United States" *Computer Physics Communications*, 33: 207-210.
- Franz L. 1966. "Information Handling in the National Standard Reference Data System," *NBS Technical Note* 290, 24.
- Edward L Brady & Wallenstein Merrill B. 1964. "National Standard Reference Data System Plan of Operation," *NSRDS-NBS* 12.
- Eugene Garfield. 2001. "A Retrospective and prospective View of Information Retrieval and Artificial Intelligence in the 21st Century," *Journal of the American Society for Information Science and Technology*, 52(3): 18-21.
- ISO 17025. 1999. "General Requirements for the Competence of Testing and Calibration Laboratories First Edition; Cancels and Replaces", 5.3
- <http://www.nist.gov/srd/onlinelist.htm>. "NIST Database List"