

NTIS의 바람직한 분류체계에 관한 연구

- 국가과학기술표준분류체계의 유용성을 중심으로 -

송충한*, 설성수**

I. 서론

과학기술혁신본부는 국가 인프라 구축 차원에서 2009년까지 향후 4년간 총 1,216억원의 예산을 투입하여 국가과학기술종합정보시스템(NTIS)을 구축할 계획으로 있으며, NTIS가 완성되는 경우 기획, 수행관리, 성과 확산 등 국가 R&D의 전주기관리가 가능할 것으로 기대하고 있다. 국가 연구개발 과제는 그동안 과기부를 비롯하여 산자부, 정통부, 교육인적자원부 등 19개 부처가 개별적으로 관리해 오면서 체계적인 관리가 이뤄지지 못했고 정보의 공동활용 미흡, 정보 품질 저하 등이 문제점으로 제기되어 왔다. 이에 따라 과기혁신본부는 지난 2004년 국가과학기술위원회에서 국가과기종합정보시스템 구축을 위한 기본 방침을 확정하고 지난해 사업기관 선정작업을 거쳐 올해부터 사업에 착수하고 있다. NTIS는 범부처적으로 추진되는데다 국가과학기술정보의 큰 틀을 잡기 위한 대규모 투자라는 점에서 기존 정보 포털 시스템과는 차별화된다.(전자신문 '06.4.25)

NTIS의 중점 추진과제는 첫째, 국가 장비·기자재 통합 라이프사이클 관리 둘째, 과학기술 인물 관리 셋째, 성과물의 체계적인 통합 관리 넷째, 과기정보 수집 및 유통 관리 플랫폼 구축 다섯째, 과제 관리 표준화 등 다섯 가지로 구분되고 있다. 과학기술혁신본부는 이러한 NTIS의 중점 추진과제가 완성되는 경우 다양한 정책 정보 제공 등 연구과제의 사전 스크린과 협동연구 추진 및 투명한 연구과제 관리가 가능해져 중복 사업 해소 등에 따른 예산 배분의 효율이 현행 예산기준 1%에서 0.1% 이하, 장비의 중복 구매율은 현행 5% 수준에서 0.5% 이하로 크게 개선될 것으로 보고 있다. 또 폐쇄적으로 유통되던 고급 과학기술정보가 기업에 손쉽게 제공됨에 따라 중소기업의 사업화 성공률은 현재 14%에서 30%까지 향상될 것으로 기대하고 있다.(전자신문 '06.4.25)

이러한 전망은 NTIS가 차질없이 구축됨은 물론 NTIS가 그 역할을 충분히 수행하는 것을 전제로 한다. NTIS가 차질없이 구축되려면 표준화 과제가 선결되어야 하며, 이러한 표준화는 표준분류체계, 업무절차표준화, 코드표준화의 3가지 요소로 구분될 수 있다.(국가과학기술종합정보시스템(NTIS) 구축을 위한 정보화전략기획(ISP) 최종보고회, 2005.6.29) 본 연구에서는 이러한 표준화 중에서 과학기술표준분류체계와 관련된 부분에 대해 살펴보고자 한다. 국가차원에서 다양한 목적을 갖는 각 정부부처 및 산하기관이 생성한 정보를 서로 공유하기 위해서는 정보의 유통과 분류가 가능토록 하는 과학기술표준분류가 가장 핵심적인 기반이 되기 때문이다.

II. 기존 과학기술분류에 대한 고찰

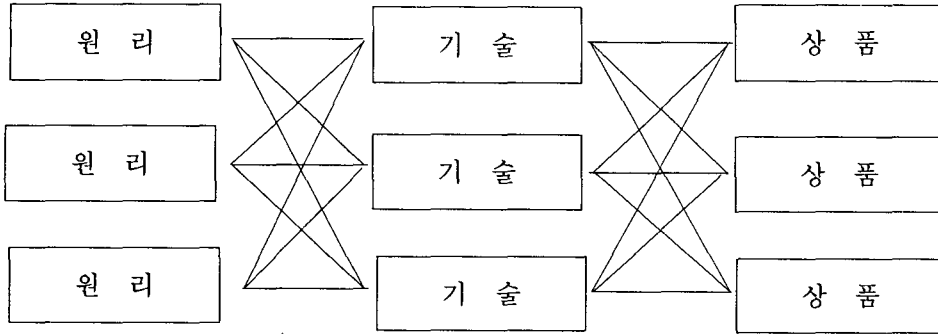
우리나라의 과학기술분류는 크게 연구분야에 대한 분류(연구분류)와 기술분야에 대한 분류(기술분류)로 구분된다. 연구분류는 학문분류라고도 할 수 있는 것으로 원리들이 결합된

* 송충한, 한국과학재단 정책연구실장, 042-869-6400, chsong@kosef.re.kr

** 설성수, 한남대학교 중국통상·경제학부, 042-629-7608, s.s.seol@hannam.ac.kr

서 과학적 원리를 기준으로 분류하는 것이고, 기술분류는 하나 혹은 다수의 과학적 특정한 기술영역을 기준으로 분류하는 것이다. 연구분류와 기술분류와의 관계는 다음과 같이 나타낼 수 있다.

[그림 1] 과학적 원리와 기술, 상품의 관계



자료: 설성수, 송충한(2000) p43

우리나라에서 연구분류를 사용하는 대표적인 기관은 기초연구를 지원하는 한국과학재단이다. 기초연구는 그 특성상 연구분류의 적용이 보다 적합하고 동시에 위상수학 등과 같이 기술분류로 표시할 수 없는 분야가 존재하기 때문에 원리를 중심으로 한 연구분야를 사용하고 있다.¹⁾ 한국과학재단의 분류는 다른 연구지원기관의 분류와 매우 상이하다. 첫째, 연구분야와 적용분야를 병행하여 기재하도록 하고 있다.(설성수, 송충한 1999) 일반적으로 연구분야는 원리측면에서 접근하기 때문에 연구 자체만을 위해서는 좋은 분류이지만, 이러한 연구가 어느 분야에 활용될 수 있는지를 파악하는 것이 매우 어려워지게 된다. 이러한 문제점을 보완하기 위해 연구분야와 함께 이 연구가 활용될 수 있는 적용분야를 병행하여 기재하도록 하고 있는 것이다. 현재 적용분야의 대분야는 지식의 진보, 문화, 환경, 공공복지, 국방, 산업 등으로 구분되고 있어 순수한 기초연구에서부터 산업적 응용이 가능한 경우까지 모든 연구의 적용분야가 고려되고 있다. 이러한 방식의 장점은 연구활동을 원리측면에서 분류하므로 모든 연구활동에 대한 분류가 가능할 뿐 아니라, 이 연구활동이 적용되는 분야가 동시에 파악됨에 따라 연구와 이의 활용 간에 연계가 가능하게 된다는 것이다. 단순히 표현하여 연구분야와 산업분야간 연결고리가 파악되므로 연구개발정책과 산업정책의 연계가 가능하게 된다. 이는 한걸음 더 나아가, 연구분류에 의한 인적자원의 파악은 곧 산업정책을 위한 인적자원의 수급과도 연계가 가능하게 된다는 것을 의미한다. 즉, 연구개발정책, 인력양성정책, 산업정책 간의 연계가 가능하게 된다는 것을 의미한다. 둘째, 연구분야와 적용분야 기재시 하나의 분야만을 기재하는 것이 아니라 사업에 따라 3개에서 5개까지 기재가 가능하고 각 분야가 해당 연구에서 차지하는 비중을 기재하도록 하고 있다. 이는 과학기술의 발전에 따라 새롭게 발생하는 융합분야 등 신규 분야에 대한 표현이 가능하다는 것을 의미한다. 연구분야를 하나만 기재하는 과거의 방식에서는 새로운 분야가 발생하였을 때 이 분야에 대한 분류가 추가되어야 새로운 분야의 기재가 가능하지만, 연구분야를 복수로 기재하는 경우 분류 표상에 추가 없이 새로운 분야의 기재가 가능하게 된다.

기초연구를 지원하는 또 하나의 대표적 기관인 한국학술진흥재단의 경우에는 원리와

1) 2005년 KISTEP에서 관리하던 특정연구개발사업의 사업관리가 한국과학재단으로 이관되면서 한국과학재단에서는 기초과학연구사업에는 기존의 과학재단 분류를 사용하고 있고, 특정연구개발사업에는 KISTEP에서 작성한 기술분류를 사용하고 있어 두 개의 분류를 사용하고 있으나, 분석의 편의상 과학재단이 사용하여 오던 연구분류를 과학재단분류라고 지칭하고자 한다.

학과 개념이 혼합된 분류를 사용하고 있는데, 형태상으로는 학과분류에 보다 가깝다. 이에 따라 동일한 내용이 여러 곳에 중복되어 분류되는 현상이 발생하게 된다. 예를 들어, 반도체의 경우 전기공학, 금속공학, 재료공학, 전자공학의 분류에 중복적으로 포함되어 있다. 이러한 분류는 중복을 가능한 회피해야 한다는 분류의 기본원칙을 충족시키지 못하고 있어 분류로서의 기본의미를 상실하고 있다.(설성수, 송충한 2001; 박병무, 2002)

우리나라에서 대표적인 기술분류는 KISTEP에서 작성한 국가과학기술표준분류이다. 국가과학기술표준분류는 신태영 외(1994)의 기술분류에서부터 출발하고 있다. 이 기술분류는 국가의 기술개발사업의 분류를 위해 작성된 것이므로 분류자체가 기술의 사용/용도별로 분류되어 있어서 기술개발에 적용하기가 매우 적합하게 되어 있다. 그러나 특정기술의 성격 중심이 아니라 사용되는 내용을 중심으로 한 분류이므로 구체적이고 세부적인 대상을 가질 수 있다는 장점을 가진 반면 세부적인 내용이 끊임없이 자주 수정되어야 한다는 단점을 갖고 있다. 다시 말해서 세부적인 내용은 그 기술을 구성하는 원리가 약간만 변해도 전혀 다른 형태로 나타나기 때문에 기술변화가 빠른 분야에서는 세부적인 내용을 파악하기 위해 끊임없는 노력이 필요할 뿐 아니라 그러한 노력들 들여도 그 결과가 곧 진부해진다는 문제점이 있다. (설성수, 송충한 2001)

신태영 외(1994)의 기술분류는 과학기술기본법의 공표와 더불어 새로운 국면으로 전환하게 된다. 즉, 2001년에 제정된 과학기술기본법에 의거 국가과학기술표준분류를 작성해야함에 따라 STEPI에서 국가과학기술표준분류를 작성하기 위한 작업이 수행되었고 그 결과 발표된 것이 조황희 외(2001)의 연구결과이다. 그런데 조황희 외(2001)에서는 기술분야만 분류되어 있고 자연과학분야는 제외되어 있어 국가과학기술표준분류로 채택되지 못하였다. 조황희 외(2001)에서 자연과학분야가 포함되지 못한 이유는 기술분류를 중심으로 한 분류에 대해 자연과학 전공 연구자들의 반발이 매우 거세었고, 분류의 속성상 연구분류와 기술분류를 혼용하는 경우 중복의 문제가 심각하게 발생하기 때문인 것으로 판단된다. 이후 KISTEP에서 조황희 외(2001)의 연구결과를 이어받아 여기에 자연과학분야를 추가하여 국가과학기술표준분류를 작성하였다.(과학기술부, 2002) 그러나 과학기술부(2002)의 표준분류가 연구분류와 기술분류의 동시사용에 따른 중복문제를 해결한 것은 아니었다. 다만, 과학기술기본법의 시행에 따라 국가과학기술표준분류가 마련되어야 했고 이를 위해 단순히 자연과학분야의 연구분류를 기존의 기술분류에 추가하고 분야간 상호 참조를 표시하여 가능한 혼란을 줄이고자 하였을 뿐이다. 현재의 과학기술표준분류체계는 2005년도에 개정하였으나, 2002년도에 작성된 분류체계와 기본원칙이 크게 다르지 않다.(과학기술부, 한국과학기술평가원, 2005) 결국 국가과학기술표준분류체계는 분류의 기본방향으로 제시한 배타성의 문제를 여전히 해결하지 못하고 있다.²⁾

이 이외에도 각 연구관리기관들은 자신의 고유사업 수행을 위해 각기 나름대로의 과학기술분류표를 보유하고 있다. 이들 기관의 분류는 위에서 언급한 3개의 대표적 분류와 또 다른 양상을 보이고 있는데, 가장 큰 특징은 소관사업이 과학기술의 일부 분야에 국한되어 있어 해당 분야에 대한 상세한 분류체계를 도입하고 있고 소관분야 이외에는 분류자체가 존재하지 않는다는 점이다.

III. NTIS 측면에서 본 국가과학기술표준분류 및 호환표의 유용성

2) 박병무(2002)는 국가과학기술표준분류체계의 개요에서 분류의 기본방향으로서 포괄성, 배타성, 유사성, 규모성, 보편성을 제시하고 있는데, 포괄성은 과학기술활동의 모든 분야를 포함한다는 것이고 배타성은 분류항목들이 서로 최대한 독립적이고 구별이 가능하도록 하는 것을 말한다.

앞서 언급한 바와 같이 정부는 과학기술 연구개발활동과 관련된 정보 즉, 인력, 연구비, 연구장비, 과제수행현황, 연구성과 등 모든 정보의 수집 및 유통을 위해 NTIS를 추진하고 있다. 여기서는 NTIS에의 적용을 염두에 두고 국가과학기술표준분류 및 호환표의 유용성에 대해 살펴보려고 한다.

정부가 막대한 예산을 들여서 NTIS를 추진하는 이유는 우리나라의 한정된 자원(resources)과 연구개발활동(activity) 그리고 그 성과(output)를 체계적으로 파악하여 국가정책들 간의 연계성을 확보함으로써 정부예산의 효과성을 높이기 위해서이다. 즉, 과학기술 관련 활동과 자원 현황을 국가차원에서 파악하여 과학기술정책 나아가 교육 및 인력양성정책, 산업정책 등의 수립을 위한 기본적인 자료로 활용하기 위한 것이며 이를 위해 과학기술 활동과 관련된 정부부처, 지원기관, 연구기관 등에서 발생하는 각종 정보를 서로 공유할 수 있도록 국가차원의 정보유통체계를 구축하고자 하는 것이다. 그런데, 과학기술 연구개발활동과 관련된 각각의 기관이 생성한 정보(DB)를 국가차원에서 공유하기 위해서는 생성된 정보를 내용별로 구분하여 정보유통이 가능하도록 하는 것이 바로 분류체계이다. 따라서 국가과학기술표준분류체계의 유용성은 결국 NTIS의 유용성을 결정하는 가장 기본적인 요소가 되는 것이다.

1. 국가과학기술표준분류의 유용성

여기서는 연구지원기관들이 국가과학기술표준분류를 사용하지 않고 각 기관의 독자적인 분류표를 사용하는 이유에 대해 살펴본 후 국가과학기술표준분류 자체의 유용성에 대해 살펴보려고 한다.

현재 과학기술기본법에 근거한 “국가과학기술표준분류”가 있음에도 불구하고 각 연구지원기관들은 고유의 사업을 수행하기 위해 독자적인 분야분류를 사용하고 있다. 각 연구지원기관들이 “국가과학기술표준분류”를 사용하지 않고 각기 고유의 분야분류를 사용하는 이유는 “국가과학기술표준분류”를 사용하는 경우 각 기관의 연구지원 및 관리 업무의 원활한 추진이 불가능하기 때문이다.

한국보건산업진흥원의 경우 소분야 숫자는 625개 이지만 이들 소분야와 관련된 “국가과학기술표준분류”의 소분야 숫자는 56개에 불과하다. 즉, 한국보건산업진흥원의 경우 “국가과학기술표준분류”중 보건의료와 관련된 분류를 10배 이상 상세하게 분류해야 사업관리가 가능하게 된다는 것을 의미한다. 다음의 <표 1>에서 보는 바와 같이 표준분류상에서는 하나의 소분야에 불과하지만 한국보건산업진흥원에서의 사업관리를 위해서는 ‘미생물학·기생충학·면역학 기술’이 10개의 소분야(기타 제외)로 구분되어야 하는 것이다. 따라서, 국가과학기술표준분류를 모든 연구관리기관이 사용하기에는 많은 문제가 있다. 그런데, 국가과학기술표준분류가 어떻게 바뀌더라도 이러한 문제는 동일하게 발생할 가능성이 매우 클 것이라 는 점을 간과하지 않아야 한다.

<표 1> 사업성격에 따른 분류의 차이

국가과학기술표준분류	한국보건산업진흥원 분류
M. 보건·의료	11. 기초 및 임상의학
M1. 기초의약학	1107. 면역학
M13. 미생물학·기생충학·면역학 기술	110701. 백신 관련기술
	110702. 자가면역질환 관련기술

	110703. 중앙면역학적 기술 110704. 장기이식 기반 기술 110705. 과민질환 110799. 기타 면역학 관련기술 1108. 감염 및 미생물학 110801. 세균성 질환 관련기술 110802. 바이러스성 질환 관련기술 110803. 진균성 질환 관련기술 110804. 기생충 질환 관련기술 110805. 항생제 내성기전 관련기술 110899. 기타 감염 및 미생물학 관련 기술
--	---

다음으로 국가과학기술표준분류 자체가 갖는 문제점에 대해 살펴보자. 현재의 국가과학기술표준분류가 갖는 가장 큰 문제점은 연구분류와 기술분류라는 차원이 각기 다른 두 가지 기준에 의한 분류가 하나의 틀 안에 동시에 존재한다는 것이다. 따라서 하나의 연구활동을 분류할 때 여러개의 분류가 해당되는 문제점이 발생하게 된다. 박병무(2002)에서 제시한 분류의 기본방향 중 ‘배타성’이 성립되지 않고 있는 것이다. 일반적으로 분류체계에서 배타성이 성립되지 않았을 때의 가장 큰 문제점은 첫째, 연구자가 자신의 연구활동을 분류할 때 어느 곳으로 분류해야하는지에 대한 정확한 판단을 어렵게 하며 둘째, 정책입안자의 입장에서는 연구자의 분포, 연구활동의 현황 등 정확한 연구지도(R&D Map)의 파악이 어렵게 된다는 것이다.

2005년에 개정된 현행 국가과학기술표준분류에는 17개의 대분야, 146개의 중분야, 976개의 소분야(달리 분류되지 않는 경우 제외)가 있다. 이들 각 소분야 중에서 분야가 매우 유사한 경우에는 각 소분야에 관련된 유사분야를 기재하여 분류에 참고하도록 하고 있는데, 이러한 경우가 많을 수록 분류의 원칙인 배타성이 성립되지 않게 된다.

<표 2> 현행 과학기술표준분류의 상호참조 예시

B. 물리학 → B6. 응집물질물리 → B63. 반도체(G51, I41-I46) G. 재료 → G5. 전자재료 → G51. 반도체재료(B63, I44) I. 전기·전자 → I4. 반도체 → I44. 반도체재료(G63, G51)
--

<표>에서 보는 바와 같이 반도체 재료에 관한 연구는 3개의 각기 다른 소분야에 해당될 수 있으며 이와 같은 중복현상은 과학기술표준분류의 정확성을 저해하는 요인으로 작용하게 된다. 이와 같이 상호참조가 되어 있어 분류의 중복성이 발생하는 경우를 각각의 대분야별로 살펴본 것이 다음의 <표 3>이다. 예를 들어, 물리학분야의 경우 기타를 제외한 소분야수는 44개인데 상호참조가 표시된 것은 15개 이다. 이 경우 중복분류된 비율은 $15/44=34.1\%$ 가 된다. 이와 같은 방식으로 대분야별 중복분류된 비율을 살펴보면 수확분야가 0%로서 가장 낮고 전기전자분야가 80.3%로서 가장 높게 나타나고 있으며, 전체적으로는 평균 32.3%의 중복율을 보이고 있다. 즉, 10개의 소분야중 3개 이상이 중복되어 있다는 것을 의미한다.

<표 3> 국가과학기술표준분류의 중복분류된 소분야 비율

대분야	중분야	소분야	상호참조 소분야수	중복분류 소분야비율
A 수학	7	38	0	0.0%
B 물리학	9	44	15	34.1%
C 화학	9	75	23	30.7%
D 생명과학	9	62	26	41.9%
F 기계	9	74	34	45.9%
G 재료	8	53	18	34.0%
H 화학공정	9	64	19	29.7%
I 전기전자	9	61	49	80.3%
J 정보	7	38	8	21.1%
K 통신	8	32	3	9.4%
L 농림수산	9	70	24	34.3%
M 보건의료	9	71	13	18.3%
N 환경	8	56	29	51.8%
O 에너지자원	9	56	25	44.6%
P 원자력	9	53	4	7.5%
Q 건설교통	9	71	17	23.9%
R 우주항공천문해양	9	58	8	13.8%
소계(평균)	146	976	315	(32.3%)

2. 국가과학기술표준분류표상 호환표의 유용성

여기서는 국가과학기술표준분류체계에서 제시하고 있는 ‘각 부처 연구관리전문기관 분류와의 호환표’(이하 호환표로 지칭)의 유용성에 대해 다음의 두 가지 측면에서 살펴보고자 한다. 첫째는 호환표를 사용했을 때의 정보유통이 얼마나 정확성을 유지할 수 있는지에 대한 것이고, 둘째는 호환표 자체의 문제이다.

1) 호환표를 이용한 정보유통의 정확성

국가과학기술표준분류에는 표준분류와 함께 연구관리기관의 분야분류와의 연계를 위한 호환표가 제시되어 있다. 그러나 호환표가 1:1 매핑관계를 형성하지 못하고 있어 호환표를 이용하여 연구관리기관이 보유한 과학기술 관련 정보(DB)를 표준분류로 전환하는 경우 최소 9.1%에서 최대 92.9%까지 1:1 매핑이 불가능하게 된다. 매핑이 불가능하다는 것은 각 기관이 보유한 정보를 표준분류로 전환하였을 때 전환된 정보의 활용이 불가능하다는 것을 의미하므로 이는 NTIS의 성패를 결정적으로 좌우하는 역할을 할 수 있다.

호환표를 이용한 정보유통의 정확성은 연구관리전문기관분류→표준분류, 표준분류→연구관리전문기관분류의 정보유통방향에 따라 각기 달라진다. 우선 연구관리전문기관의 정보를 표준분류로 전환할 때의 문제점을 설명하고 있는 것이 <표 4>이다. 연구관리전문기관의 분류체계를 표준분류체계로 전환할 때 발생하는 문제는 다음과 같이 세 가지로 구분할 수 있다. 첫째는 연구관리전문기관에는 분류되어 있는데 표준분류에는 해당 분류가 아예 없는 경우이다. 둘째는 연구관리전문기관에는 하나의 소분야로 분류되어 있는데 표준분류에는 중

분야로 분류되어 있는 경우이다. 이는 각각의 분류가 갖는 깊이가 달라서 발생하는 문제인데, 연구관리전문기관의 소분야를 표준분류로 전환할 때에는 중분야 아래에 있는 여러 소분야 중에서 어느 것이 적합한 것인지 전산상에서는 판단할 수 없기 때문에 매핑이 불가능하게 된다. 셋째는 연구관리전문기관의 소분야에 매핑되는 표준분류의 소분야가 2개 이상 존재하는 경우이다. 이 경우에도 전산시스템에 의한 매핑이 불가능하게 된다. <표 4>에서와 같이 매핑이 안되는 경우를 살펴보면 환경기술진흥원의 경우 92.9%로 매핑 불가능한 소분야의 비율이 가장 높고, 보건산업진흥원이 9.1%로 가장 낮은 것으로 나타나고 있으며 연구관리전문기관 전체적으로는 평균 51.1%의 소분야가 매핑이 불가능하게 된다. 결국, 연구관리전문기관에서 표준분류로 DB를 전환했을 때의 정확성은 48.9%에 불과하게 된다는 것을 의미한다.

<표 4> 연구관리전문기관분류->국가과학기술표준분류 매핑시 부정확도(소분야 기준)

구 분	지원기관 분류의 소분야수 (A)	호환표에 표준분류 없는것 (B)	표준분류가 중분야인 경우 (C)	표준분류 (소분야)가 2개 이상인 경우 (D)	소분야간 1:1매핑이 안되는 경우 (E=B+C+D)	1:1매핑 불가능비율 (E/A)
한국과학재단	514	23	112	101	236	45.9%
농림기술관리센터	162	0	6	121	127	78.4%
산업기술평가원	367	2	10	97	109	29.7%
정보통신연구진흥원	72	8	3	26	37	51.4%
보건산업진흥원	625	0	1	56	57	9.1%
환경기술진흥원	28	0	2	24	26	92.9%
에너지관리공단	82	0	0	21	21	25.6%
건설교통기술평가원	158	1	10	43	54	34.2%
한전전력연구원	96	0	1	48	49	51.0%
국방과학연구소	69	5	35	24	64	92.8%
평 균						51.1%

- 주1) 학술진흥재단의 경우 중분야까지만 분류호환표가 제시되어 있어 분석대상에서 제외하였음. 학술진흥재단의 경우 학과분류성격이므로 기술분류인 국가과학기술표준분류와 매핑이 불가능하여 중분야까지만 호환표를 제시한 것으로 판단됨
- 주2) 표준분류가 중분야인 것과 표준분류 소분야가 2개 이상인 경우가 중복하여 발생하는 경우에는 표준분류가 중분야인 것으로 카운팅하고 소분야는 제외하여 중복계상을 배제하였음.

다음으로 표준분류→연구관리전문기관분류로 정보를 변환했을 때의 정확성에 대해 살펴보면 다음의 <표 5>와 같다. 이 경우, 표준분류의 전분야를 고려할 필요가 없으므로 각 연구관리전문기관의 세분야와 연계된 표준분류의 세분야에 대해서만 살펴보면 된다. 표준분류에서 연구관리전문기관으로 DB를 전환할 때 표준분류에서의 소분야가 연구관리전문기관에서의 소분야와 1:1 매핑이 이루어지면 정보가 정확하게 전달될 수 있다. 그러나, 표준분류에서의 하나의 소분야에 관련된 연구관리전문기관의 소분야가 복수인 경우에는 1:1 매핑이 이루어질 수 없으므로 정보의 오류가 발생하게 된다. <표 5>에서 보면 표준분류→연구관리전문기관으로 정보를 전환할 때 매핑이 불가능한 비율은 농림기술관리센터가 72.2%로 가장

높고 한국과학재단이 23.4%로 가장 낮게 나타나고 있다.

<표 5> 국가과학기술표준분류->연구관리전문기관 분류 매핑시 부정확도(소분야 기준)

	각 지원기관과의 매핑에 사용된 표준분류의 소분야 수(중복제외) (A)	1개의 표준분류 분야가 2개의 이상의 지원기관 분야와 연계된 경우 (B)	1:1 매핑 불가능 비율 (B/A)
한국과학재단	397	93	23.4%
농림기술관리센터	72	52	72.2%
산업기술평가원	252	95	37.7%
정보통신연구진흥원	77	22	28.6%
보건산업진흥원	109	71	65.1%
환경기술진흥원	41	26	63.4%
에너지관리공단	53	19	35.8%
건설교통기술평가원	144	53	36.8%
한전전력연구원	47	26	55.3%
국방과학연구소	75	18	24.0%
평 균			44.2%

<표 4>와 <표 5>를 종합하여 보면 다음의 <표 6>이 된다. 즉, 호환표를 이용한 정보의 호환을 연구관리전문기관 ↔ 표준분류의 양방향에서 살펴보면 각기관별 호환의 부정확도는 다음과 같다. 여기서 연구관리전문기관→표준분류와 표준분류→연구관리전문기관의 매핑 불가능 비율이 크게 차이가 나는 것은 대체로 분류의 깊이가 다르기 때문이다. 예를 들어, 보건산업진흥원의 경우 보건산업진흥원→표준분류는 부정확도가 9.1%에 불과한데 비하여 표준분류→보건산업진흥원의 부정확도는 65.1%에 달하고 있는데 이는 보건산업진흥원의 분류가 세분화되어 있어 보건산업진흥원에서 표준분류로 전환하는 데에는 별 문제가 없어도 표준분류를 보건산업진흥원 분류로 전환하고자 할 때에는 표준분류 1개에 보건산업진흥원의 분류상의 세분야가 복수로 해당되기 때문이다.

이와 같이, 각 기관의 분류의 깊이가 각기 다른 이유는 각 연구관리전문기관이 고유의 연구지원사업을 수행하기 위한 분류의 깊이와 범위가 각기 다르기 때문에 발생하는 것이다. 이는, 현재의 국가과학기술표준분류를 모든 지원기관이 의무적으로 사용해야 한다면 각 기관의 사업관리에 많은 차질이 발생하게 될 것이라는 것을 의미하는 것이기도 하다.

<표 6> 호환표를 이용하는 경우 1:1 매핑의 부정확도

구 분	연구관리전문기관->표준분류시 1:1 매핑불가능 비율	표준분류->연구관리전문기관시 1:1 매핑불가능 비율	비 고
한국과학재단	45.9%	23.4%	
농림기술관리센터	78.4%	72.2%	
산업기술평가원	29.7%	37.7%	
정보통신연구진흥원	51.4%	28.6%	
보건산업진흥원	9.1%	65.1%	
환경기술진흥원	92.9%	63.4%	

에너지관리공단	25.6%	35.8%	
건설교통기술평가원	34.2%	36.8%	
한전전력연구원	51.0%	55.3%	
국방과학연구소	92.8%	24.0%	
평균	51.1%	44.2%	

앞에서는 각각의 연구관리전문기관별로 호환표를 이용하였을 때 매핑의 부정확성에 대해 살펴보았다. 그러나 대부분의 경우 각 연구관리전문기관들은 다른 연구관리전문기관의 정보가 필요하게 된다. 따라서 각 연구관리기관이 보유한 정보를 호환표를 이용하여 다른 연구관리전문기관이 사용하는 경우 1:1 매핑의 정확성이 어떻게 되는지를 살펴보아야 한다. 이 경우 하나의 연구관리전문기관(A)의 정보가 다른 연구관리전문기관(B)으로 유통되었을 때의 정보의 정확성은 다음과 같이 계산될 수 있다.

호환표를 이용한 정보유통의 정확성 = (A기관→표준분류의 정확성) * (표준분류→B기관의 정확성)

예시: 과학재단의 자료를 표준분류를 이용하여 산업기술평가원에서 활용하는 경우

과학재단→표준분류의 정확도(54.1%=1-45.9%)

표준분류→산업기술평가원의 정확도(62.3%=1-37.7%)

따라서 과학재단→산업기술평가원의 정확도 = 54.1% * 62.3% = 33.7%

이와 같은 방식으로 각 연구관리전문기관들이 보유한 DB를 호환표를 이용하여 다른 기관의 자료로 전환하는 경우 가장 높은 정확도는 보건산업진흥원→한국과학재단의 69.6%이며, 가장 낮은 정확도는 환경기술진흥원/국방과학연구소→농림기술관리센터의 2.0%이다. 이는 호환표를 이용한 각 연구관리전문기관간 정보의 공유에 있어서 정보의 정확도는 가장 높은 경우라고 해도 69.6%에 불과하므로 호환표를 이용한 자료의 공유는 그 유용성이 거의 없다는 것을 의미한다.

<표 7> 연구관리기관간(A기관→B기관) 분류표를 이용한 정보공유시의 정확도

A \ B	한국과학재단	농림기술관리센터	산업기술평가원	정보통신연구진흥원	보건산업진흥원	환경기술진흥원	에너지관리공단	건설교통기술평가원	한전전력연구원	국방과학연구소
한국과학재단		15.0%	33.7%	38.6%	18.9%	19.8%	34.7%	34.2%	24.2%	41.1%
농림기술관리센터	16.5%		13.5%	15.4%	7.5%	7.9%	13.9%	13.7%	9.7%	16.4%
산업기술평가원	53.8%	19.5%		50.2%	24.5%	25.7%	45.1%	44.4%	31.4%	53.4%
정보통신연구진흥원	37.2%	13.5%	30.3%		17.0%	17.8%	31.2%	30.7%	21.7%	36.9%
보건산업진흥원	69.6%	25.3%	56.6%	64.9%		33.3%	58.4%	57.4%	40.6%	69.1%
환경기술진흥원	5.4%	2.0%	4.4%	5.1%	2.5%		4.6%	4.5%	3.2%	5.4%
에너지관리공단	57.0%	20.7%	46.4%	53.1%	26.0%	27.2%		47.0%	33.3%	56.5%
건설교통기술평가원	50.4%	18.3%	41.0%	47.0%	23.0%	24.1%	42.2%		29.4%	50.0%
한전전력연구원	37.5%	13.6%	30.5%	35.0%	17.1%	17.9%	31.5%	31.0%		37.2%
국방과학연구소	5.5%	2.0%	4.5%	5.1%	2.5%	2.6%	4.6%	4.6%	3.2%	

2) 호환표의 정확성에 따른 유용성

앞서는 호환표가 정확하다는 전제 하에 호환표를 이용한 정보유통의 정확성에 대해 살펴보았는데, 여기서는 과연 호환표가 정확한지에 대해 살펴보려고 한다. 과학재단에서 지원하고 있는 특정기초연구사업의 경우 2006년도 신청서를 접수하는 과정에서 과학재단의 연구분류와 현행 국가과학기술표준분류를 병행하여 기재하도록 한 바 있다. 그 결과 1,414개의 연구신청서 중에서 호환표와 동일하게 표준분류를 기재한 연구신청서는 742건이고 호환표와 다르게 표준분류를 기재한 연구신청서는 676건 그리고 연구자의 착오로 호환표가 다르게 기재된 것으로 보이는 연구신청서가 6건으로 나타났다. 따라서 호환표 자체의 정확성은 $52.9\% = (742\text{건} + 6\text{건}) / 1,414\text{건}$ 이고 호환표 자체의 부정확성은 47.1%로 볼 수 있다.

이처럼 호환표 자체의 부정확성이 제기되는 이유에 대해서는 구체적인 사례를 이용하여 살펴보도록 하자. 아래의 <표 6>은 2006년도에 접수된 특정기초연구의 신청서 상에 기재된 내용이다. 호환표에서 과학재단 소분야 광자학(11112)은 표준분류의 소분야 광자학(I32)으로 매핑되어 있다. 실제로 접수된 5개 과제 중에서 표준분류로서 광자학을 기재한 경우는 2건에 불과하고 나머지 3건은 반도체, 양자정보, 광전자 등 호환표와 다른 분야를 기재하고 있다. 결국, 호환표에서는 광자학(11112)을 광자학(I32)로 매핑할 수밖에 없지만 현실적으로는 연구내용에 따라 각기 다른 소분야로 매핑되고 있음을 알 수 있다.

<표 8> 호환표의 정확성에 관한 사례 예시

과제번호	과학재단 소분야(코드)	표준분류 소분야(코드)	정확성
R01-2006-000-10384-0	광자학(11112)	광자학(I32)	정확
R01-2006-000-10900-0	광자학(11112)	광자학(I32)	정확
R01-2006-000-11137-0	광자학(11112)	반도체(G51, I41~I46)	부정확
R01-2006-000-10959-0	광자학(11112)	양자정보(B74)	부정확
R01-2006-000-11197-0	광자학(11112)	광전자(B52~B54, I46, K11)	부정확

IV. 요약 및 결론

정부가 국가과학기술종합정보시스템(NTIS)를 추진하는 이유는 우리나라의 한정된 자원(resources)과 연구개발활동(activity) 그리고 그 성과(output)를 체계적으로 파악하여 국가 정책들 간의 연계성을 확보함으로써 정부예산의 효과성을 높이기 위해서이다. 이를 위해서는 각 정부부처와 연구관리전문기관 등이 보유하고 있는 정보가 원활하게 유통되어 국가 차원에서의 현황파악이 가능해야 한다. 과학기술분류는 이러한 정보의 유통을 매개한다는 점에서 NTIS에서 차지하는 비중이 매우 크다고 할 수 있다. NTIS가 단순히 하드웨어적인 시스템만 마련한다고 해서 정보가 유통되는 것은 아니기 때문이다. 정보의 유통을 위해서는 하드웨어 뿐만 아니라 소프트웨어적인 각종 분류와 표준화도 반드시 필요하기 때문이다.

본 연구에서는 이러한 관점에서 현행 과학기술표준분류체계가 과연 NTIS의 분류로 사용될 수 있는지에 대해 세가지 측면에서 검토하여보았다. 첫째는 과학기술분류체계 자체의 유용성이며 둘째는 호환표를 이용하였을 때 이루어지는 정보유통의 정확성 그리고 마지막으로

막으로 정보유통을 매개하는 호환표 자체의 정확성에 대해 살펴보았다. 살펴본 결과, 현재의 과학기술표준분류를 NTIS의 분류로 사용하는 것에 대해서는 많은 문제가 있는 것으로 나타나고 있다. 우선 표준분류 자체적으로는 중복분류된 소분야가 32.3%에 달하고 있어 분류의 원칙인 '배타성'을 충족시키지 못하고 있다. 둘째로 호환표를 이용한 정보유통의 경우 각 연구관리전문기관의 정보를 표준분류로 전환하였을 때 평균적으로 51.1%가 1:1 매핑이 불가능한 것으로 나타났으며, 연구관리전문기관간의 정보 유통의 정확도는 가장 높은 경우가 보건산업진흥원->한국과학재단의 69.6%이고, 가장 낮은 경우가 환경기술진흥원/국방과학연구소->농림기술관리센터의 2.0%로 나타나 호환표를 이용한 각 연구관리전문기관간 정보의 공유는 그 정확도가 낮아 자료공유로서의 의미가 크게 상실되는 것으로 나타나고 있다. 끝으로 과학재단의 사례를 이용한 호환표 자체의 정확성에 대해 살펴보면 기존 호환표의 정확도는 52.9%에 불과한 것으로 나타나고 있는데, 이는 호환표를 정확하게 작성하였느냐 또는 부정확하게 작성하였느냐의 문제라기보다는 하나의 연구가 다양한 분야(기술)에서 활용될 수 있다는 연구 자체의 특성에 기인한 것으로 판단된다. 이는 호환표의 정확성을 높이는 노력이 정보유통의 정확성을 다소 개선할 수는 있어도 100% 문제를 해결하는 것은 불가능하다는 것을 보여주고 있는 것이다.

NTIS를 구축하여 국가전체의 연구개발현황을 파악하여 과학기술정책 나아가 교육 및 인력양성정책, 산업정책 등의 수립을 위한 기본적인 자료로 활용할 수 있기 위해서는 과학기술분류를 통해 획득한 정보가 체계적으로 각 정책들 간의 연계를 파악할 수 있도록 해야 할 것이다. 즉, 과학기술분류를 통해 파악되는 정보를 이용하여 과학기술정책(연구개발정책) ↔인력양성(교육)정책↔산업정책의 연계가 파악될 수 있도록 분류체계가 새롭게 마련될 필요가 있다. NTIS의 목적에 부합하는 새로운 분류체계가 구축되었을 때 NTIS는 진정한 국가과학기술종합정보시스템으로서의 역할을 수행할 수 있을 것이다.

[참고문헌]

- 조황희 외 (2001), 국가과학기술표준분류표 작성에 관한 연구, 과학기술정책연구소
- 신태영 외 (1994), 연구개발을 위한 한국의 기술분류체계, 과학기술정책관리연구소
- 과학기술부 (2002), 국가과학기술표준분류체계의 수립에 관한 연구.
- 설성수(2006), “학문분류표의 확장”, 과학기술산업분류 관련 연구지원기관 책임자 워크샵 발표자료집, 3월, 한남대 하이테크비즈니스연구소.
- 설성수, 송충한(1999), “연구활동분류의 이론적 검토”, 『기술혁신학회지』 2-3, 11, 19-33쪽.
- 설성수, 송충한(2001), 『지식활동분류의 이론과 실제』, 한남대 출판부.
- 설성수, 송충한 외(1999), 『기초과학연구 분야분류체계 개발연구』, 한국과학재단, 4.
- 송충한(1998), “21세기 대비 기초과학정책의 방향”, 기술혁신학회지, 1-2, 8.
- 송충한(2006), “과학재단 분류체계와 과제”, 과학기술산업분류 관련 연구지원기관 책임자 워크샵 발표자료집, 3월, 한남대 하이테크비즈니스연구소.
- 송충한, 설성수(1999), “새로운 과학기술분류의 철학과 구조”, 『기술혁신학회지』 2-3, 11., 34-47쪽.
- 양창준(2006), “IT산업분류체계”, 과학기술산업분류 관련 연구지원기관 책임자 워크샵 발표자료집, 3월, 한남대 하이테크비즈니스연구소.
- 이병민(2006), “NTIS의 표준화 방향”, 과학기술산업분류 관련 연구지원기관 책임자 워크샵 발표자료집, 3월, 한남대 하이테크비즈니스연구소.