

가R&D

최기석*[†] · 박만희* · 김영국**

*한국과학기술정보연구원 NTIS사업단
**충남대학교

A Study on Construction of Integrated National R&D Monitoring System

Choi Ki-Seok*[†] · Park Man-Hee* · Kim Young-Kuk**

*NTIS Division, Korea Institute of Science and Technology Information

**Department of Computer Science and Engineering, Chungnam National University

This Study constructs a dashboard system to synthetically and systematically monitor national R&D information based on data warehouse. Managing the national R&D statistics and trend is important since it provides data for policies and decision making for national R&D. Many agencies related to national R&D information collect the basic R&D statistic data which provides the basis of logical decision making and R&D policies. The data has not well been used. The data has not been consistently collected nor managed. The raw data has not been organized nor processed to meet various demands. The needs has been arisen for a consistent national R&D monitoring system to increase the relevance, accessibility and efficiency of data for various users. This study selects 25 key indicators based on the user requirements and designs data warehouse for supporting the indicators using star schema. The dashboard system is developed in this study provides the infrastructure of monitoring national R&D information and analytic environment of supporting statistical analysis and time-series data analysis.

Keywords : National R&D Information, Monitoring System, Datawarehouse, ETL, National R&D Board

1. 서 론¹⁾

세계 각국들은 연구개발 효율성 제고를 위해 국가차원에서 연구개발 동향정보, 연구개발 종합현황과 다양한 분석기법에 기반한 분석정보를 정책적 의사결정에 활용하고 있으며 이를 통하여 국가연구개발의 전략적 방향성을 설정하고 정책의 실효성을 높이려고 노력하고 있다. 정부 연구개발투자의 절대규모와 GDP대비 비중이 지속적으로 증가하는 추세에 있으므로 연구개발의 투자

효율성과 성과지향 연구개발의 중요성은 지속적으로 관심을 받고 있으며 강조되어 오고 있다. 이러한 추세에 발맞추어 교육과학기술부는 2006년부터 국가R&D 투자 효율성 제고라는 목표달성을 위해 국가과학기술종합정보 서비스(NTIS : National science and Technology Information Service, <http://www.ntis.go.kr>, 이하 NTIS) 구축사업을 추진해오고 있다.

국가차원에서 R&D 현황을 종합적으로 모니터링하고 피드백을 통해 정책의사결정을 지원하기 위한 도구내지

는 시스템이 필요한 이유는 다음과 같이 정리할 수 있다.

첫째, 국가차원의 R&D 기획, 집행, 평가, 성과활용에 이르는 개별 단계의 단절현상으로 체계적인 관리가 요구되고 있다.

둘째, 국가R&D 정보간 연계 부족으로 종합적인 현황 파악이 어렵고 정책적 의사결정이 곤란한 경우가 빈번히 발생하고 있다.

셋째, 개별 유관기관에서 작성한 정책자료를 활용하고 있으나 정보의 공유 및 활용도가 낮은 실정이다.

본 연구의 목적은 국가R&D 정보와 관련된 투입요소와 산출요소, 연구활동과 관련된 다양한 정보에 대한 현황과 추세를 종합적으로 파악하고 분석하여 과학기술 정책수립과 투자전략수립에 도움을 줄 수 있는 정보를 제공하고, 정책적 의사결정을 지원할 수 있는 국가R&D 종합모니터링 시스템을 구축하고 구현된 시스템을 소개하는데 있다.

제 2장에서는 국가과학기술종합정보시스템의 개요와 선진국의 유사 시스템의 주요 특징을 살펴보고, 국가 R&D 종합모니터링시스템을 위한 주요 정보항목 및 서비스에 대하여 설명한다.

제 3장에서는 국가R&D 종합모니터링시스템의 구성도, DB 스키마, 메뉴 구성도와 주요 기능, 구축사례에 대하여 설명한다.

제 4장에서는 본 연구의 결과와 추후연구에 대하여 살펴보고자 한다.

2. 국가R&D 종합모니터링시스템 개요 및 정보항목

2.1 국가과학기술정보종합서비스 개요

2007년 과학기술정책연구원(STEPI)에서 총요소생산성의 R&D 투자에 대한 탄력성을 조사하여 R&D 투자 효율성을 측정된 연구결과[1]에 따르면, 우리나라 R&D 투자 효율성(0.182)은 OECD 평균과 비슷한 수준(0.19)이나, 미국(0.22)이나 일본(0.288)보다는 낮아 투자효율성 제고에 대한 노력이 필요한 것으로 분석되었다. 또한 혁신주도형 경제로의 원활한 전환과 잠재성장률의 지속적인 확충을 위해서는 연구개발투자의 확대와 효율성 증대를 위한 정부와 민간의 적극적인 노력이 요망되는 것으로 분석되었다.

특히 국가R&D 사업에 대한 정보가 정부 부처별, 기관별로 관리되면서 발생하는 국가R&D에 대한 중복투자 등의 문제해결과 함께 R&D사업에서 산출된 정보와 자원을 적극적으로 활용하여 연구개발 효율성과 생산성을 제고하자는 요구사항이 증가하였다. 이에 표준화에 따라 R&D 관련 부처(대표연구관리전문기관)와 연계하여 정보를 수집·가공한 후, 공동 활용함으로써 R&D 투자효율성을 제고하기 위해 NTIS를 구축하게 되었다.

2003년 12월 대통령의 지시에 의해 ‘연구개발투자 효율화 방안’이 처음 보고되었으며 2004년 7월 국가과학기술

단계	중점추진과제	세부추진과제
1단계 (’06-’07) 국가R&D사업관련 기반정보시스템 구축	국가연구개발사업관리체계 지원	▪ 국가R&D사업관리시스템
	국가과학기술인력정보 활용체제 구축	▪ 국가R&D참여인력통합관리시스템 ▪ 평가위원통합관리시스템
	장비·기자재 공동활용 관리지원	▪ 장비·기자재 통합관리시스템 ▪ 장비·기자재 전문가관리시스템
	성과종합관리 지원	▪ R&D성과정보관리시스템 ▪ 우수유망기술정보시스템
	과학기술정보공통기반 구축	▪ 국가R&D Board ▪ 과학기술통계시스템 ▪ 과학기술정보유통기반 구축 ▪ 표준과제정보관리시스템 ▪ 정보보호체계 구축
2단계 (’08-’09) 국가R&D사업성과 확산 및 지역정보시스템 구축	기술·산업정보 지원체계 구축	▪ 기술·산업정보시스템
	지역기술혁신지원체계 구축	▪ 지역기술혁신지원시스템

<그림 1> NTIS의 단계별 중점추진과제 및 세부추진과제



<그림 2> 국가과학기술종합정보서비스(NTIS) 개념도

슬위원회에서 ‘국가과학기술종합정보시스템 사업계획’이 확정되었다. 이후 2007년 5월 과학기술관계장관회의를 통해 『NTIS사업추진계획(2007~2009)』을 확정하여 7개 중점 추진과제와 14개 세부추진과제에 대해 2006년~2009년까지 2단계로 구분한 단계별 추진계획을 <그림 1>과 같이 수립하였다. 또한, 각 부처가 선정한 대표연구관리 전문기관이 해당 부처의 국가R&D정보를 NTIS와 연계도록 지원하는 효율적인 정보수집체계를 수립하였다. 그리고 2007년 8월 과학기술관계장관회의에서 341개 항목으로 구성된 R&D정보표준안과 과제고유번호 등을 심의·확정함으로써, 국가차원의 정보공동활용기반을 마련하였다.

2007년 12월에 13개 부처청 산하 12개 대표연구관리 전문기관과 연계하였고, 2009년 3월 현재 <그림 2>와 같이 15개 부처청 산하 16개 대표연구관리전문기관과 연계하여 과제, 인력, 성과, 장비·기자재 등 주요 R&D 정보(341개 항목)를 수집·가공하고 8개의 국가R&D 관련 서비스체계를 구축하여 서비스 중에 있다.

NTIS는 국가 연구개발의 기획에서부터 성과 활용에 이르기까지 전주기에 걸쳐 연구개발의 효율화를 지원하기 위해 구축한 국가R&D 정보 포털시스템이라고 할 수 있다. 지금까지 일본(e-RAD)[8], 유럽(CORDIS)[7] 등 일부 국가가 분야별 과학기술정보시스템을 구축하여 운영하고 있으나 과제, 인력, 성과, 장비·기자재 정보 등이 서로 연계되어 종합 제공되는 포털서비스는 NTIS가 최초라는데 의미가 크다[3].

2008년 3월부터 본격적으로 서비스되고 있는 세부서비스를 간략히 살펴보면 다음과 같다. 국가R&D 사업의 종합적인 현황과 과제정보, R&D사업추진현황에 대한 체계적이고 심층적인 조사·분석을 기반으로 각종 통계정보를 제공하는 국가R&D 사업관리서비스, 국가R&D 사업을 통해 도입된 장비와 기자재의 보다 효율적인 ‘공동활용’을 위해 대학과 연구기관이 보유한 장비와 기자재

를 한눈에 볼 수 있도록 관련 정보를 제공하고, 연구장비를 사용하면서 접하는 애로사항을 전문가에게 직접 질의할 수 있는 장비·기자재 공동활용 서비스, 국가R&D에 참여한 연구자의 전공과 논문, 연구실적 등에 대한 정보를 제공하여 국가R&D 과제 신청 시 인물정보의 중복입력을 피할 수 있으며, 국가R&D 사업평가위원 선정 등 각종 정책결정에 참여하는 전문가 위촉 등에 활용할 수 있는 국가R&D 참여인력정보서비스, 국가R&D 사업으로부터 창출된 다양한 연구 성과를 종합적으로 분석하여 제공하는 R&D 성과정보서비스, 국가R&D 사업의 투자, 참여인력, 장비·기자재, 성과 등의 종합현황을 핵심지표 중심으로 일목요연하게 제공하는 국가R&D Board, 민간 R&D 투자 현황과 기술무역, 특허 등 국가R&D 사업 이외의 각종 통계자료를 알기 쉽게 제공하는 과학기술통계서비스, 연구관리전문기관의 과제 진행현황을 제공하는 표준과제정보관리서비스, 개별 부처별로 상이한 과제관리 프로세스와 정보화 수준에 따른 이용자의 불편을 해소하기 위해 부처로부터 관련 정보를 종합적으로 수집해 가공·정제해서 더욱 의미있는 정보를 제공하기 위한 서비스체제인 과학기술정보유통기반으로 구성되어 있다.

NTIS에서 서비스되고 있는 정보건수는 2008년 기준으로 약 34만 건이고, 2002~2007년 개별 부처별, 기관별로 수행되었던 국가R&D과제 및 현황분석정보 16만 건, 국가R&D 사업에 참여한 약 6만 명의 인물정보 및 현황분석정보, 국가R&D 사업을 통해 취득·보유한 3천만 원 이상 장비·기자재 정보 5만 건(7개 부처 51개 사업의 공동활용필수장비 1만 7천 건 포함), 각 부처 및 기관으로부터 수집한 논문 4만 건, 출원 및 등록 특허 3만 건 등의 성과정보가 포함되며, 국가R&D 사업의 주요 핵심지표는 국가R&D Board로, 민간 R&D 투자 및 OECD국가의 과학기술지표는 과학기술통계서비스를 통해 서비

스 되고 있다.

2.2 선진국 관련 시스템 비교 · 검토

NTIS와 유사한 국가차원의 연구개발정보시스템으로는 미국의 NTIS[11], 일본의 JST[10], 유럽의 CORDIS[7]를 들 수 있다. 미국의 정부차원 연구개발정보시스템은 정부차원에서 지원하는 과학, 기술, 공학 및 기업 관련 정보를 제공하고 있는 국가기술정보서비스(National Technical Information Service)이다. 미국 NTIS는 60년 이상 350개 주제영역에 걸쳐 대략 300만개 정도의 출판물을 기업, 대학 및 공공부문에 제공하고 있다. 제공되는 주요 콘텐츠로는 연구보고서, 컴퓨터 제품, 소프트웨어, 비디오, 오디오 등이며 평균적으로 매년 60,000 레코드 이상이 데이터베이스에 추가되고 있으며 대부분의 레코드는 개요(abstract) 정보를 포함하고 있다. 일본의 연구개발 관련 정보시스템은 과학기술진흥기구(Japan Science and Technology Agency)에서 구축하여 제공하고 있는 JST(Japan Science and Technology) 서비스이다. 정부로부터 승인되었거나 JST에서 자체 추진중인 중장기계획에 부합되는 특정 JST사업이 수행되고 있다. 정부에 의해 수립된 2003년부터 2007년까지 1단계 중장기계획을 달

성하기 위하여 선진기술 창출, 선진기술을 활용한 기업 활성화, 과학기술정보의 보급촉진, 연구자 교류 및 연구 지원, 과학기술에 대한 공공의 이해촉진 등 5개 활동을 주요 목표로 하고 있다. 제공되고 있는 주요 서비스로는 전자저널 및 보고서 서비스, 산학협력 연구성과, 연구활동 및 연구자 경력정보를 제공하는 산학협력 서비스, 생명공학, 화합물, 과학기술 정보자원 등이다. 유럽공동체(EU)의 연구개발 관련 정보서비스는 연구개발 및 기술 혁신 관련 정보를 통합적으로 제공하고 있는 CORDIS(Community Research and Development Information Service)이다. CORDIS의 주요 서비스는 연구과제 공고·지원 및 파트너 검색기능을 제공하는 연구지원 서비스, 연구개발 단계와 성과활용 상태별 연구성과 제공 서비스, 주제인덱스 분류체계를 이용한 연구주제, 국가별 연구개발 정책 및 프레임워크와 연구관련 최신 동향 및 활동 서비스를 제공하는 국가별 서비스, 과제 및 사업정보, 주요 용어와 연구기관 및 전문가 정보를 제공하는 검색 서비스 등이다. 연구과제는 연구내용 및 결과뿐만 아니라 연구과제의 약어 선정, 연구결과의 확산 및 홍보연구도 포함되어 있다. NTIS와 유사한 목적과 성격을 가진 이들 시스템의 주요 서비스 및 특징을 NTIS와 비교하여 정리하면 <표 1>과 같다.

<표 1> 선진국 관련 시스템 비교 · 검토

구 분	한국(NTIS)	미국(NTIS)	일본(JST)	유럽(CORDIS)
담당기관	◦ 교육과학기술부 (www.ntis.go.kr)	◦ 국립과학기술정보국 (www.ntis.gov)	◦ 일본과학기술진흥기구 (www.jst.go.jp)	◦ 유럽연합집행위원회 (cordis.europa.eu)
목적	◦ 연구개발의 기획부터 성과 활용에 이르는 전주기관리를 통한 연구개발 효율성 향상 ◦ 국가 R&D 정보 공동활용을 통한 정책적 지원체제 구축 및 정보활용 양극화 해소	◦ 정부기관 수행 연구정보 제공 ◦ 연구결과 정보제공 ◦ 민간기업의 연구성과 활용 극대화	◦ 과학기술 창조성에 기반을 둔 일본건설 ◦ 산·학·관 제휴를 통한 인적교류 및 공동연구 활성화 ◦ 일본 R&D 활동 모니터링, 연구개발 정책수립 및 조정	◦ 프로젝트 참여 및 파트너 검색 서비스를 통한 연구활동 활성화 ◦ 혁신 아이디어 및 연구성과 전파 ◦ 유럽연합 통합 연구인프라 확충
주요 제공정보	◦ 국가R&D 투자정보 ◦ 국가R&D 과제정보 ◦ 국가R&D 참여인력정보 ◦ 성과정보 ◦ 국가R&D 장비·기자재정보 ◦ 국가R&D 연구성과물 정보 ◦ 과학기술통계	◦ 나노, 농업, 군사 등에 관한 특수정보 ◦ 부처, 국, 프로그램, 프로젝트, 세부과제 등 5단계 수준 분류 ◦ 정책 의사결정지원 정보 ◦ 멀티미디어기반 정보제공	◦ 과학기술정보제공·유통종합시스템(J-STAGE) ◦ 과학기술 보고서서비스(J-EAST) ◦ 특허 및 연구성과종합정보(J-STORE) ◦ 연구개발활동 디렉토리서비스(READ) ◦ 연구자 경력관리 서비스(JREC-IN) ◦ 생명정보서비스(BIRD)	◦ 연구비 지원정보 ◦ 연구주제 및 최신동향정보 ◦ 연구 성과정보 ◦ 과제, 사업, 연구기관, 연구원 정보 ◦ 연구진행 현황정보
정보 구축범위	국가 수준	국가 수준	국가 수준	국가연합 수준
문헌검색기능	보통 (논문·보고서 일부 연계)	강함 (논문·보고서 연계비율 높음)	강함 (논문·보고서 연계비율 높음)	강함 (연구성과 메타정보 제공)
사업·과제 평가정보	있음	없음	없음	없음
정보유통체계	강함	강함	강함	강함
종합모니터링 서비스	있음	없음 (보고서형태 원문 서비스)	없음 (보고서형태 원문 서비스)	없음
상용화 여부	비상용	상용	비상용	비상용

2.3 국가R&D 종합모니터링시스템 구축을 위한 정보요소 및 항목

NTIS의 운영계 시스템에 수집된 국가연구개발 사업, 과제, 참여인력, 성과, 장비·기자재 등의 기반정보를 활용하여 국가R&D 현황과 추세를 모니터링하기 위한 종합상황판시스템(Dashboard system)을 구축하기 위해서는 사업, 과제, 참여인력, 성과, 장비·기자재 등 개별 주제 영역별로 분석 및 서비스가 가능한 정보요소와 정보항목 도출이 필요하다.

2.3.1 국가연구개발사업 조사분석 정보요소

매년 수행되고 있는 국가연구개발사업의 조사·분석은 궁극적으로 국가연구개발사업의 효율성과 생산성 제고 및 국가과학기술경쟁력 향상을 위한 중요한 정책자료로 활용되고 있다[2]. 국가연구개발사업에 대한 전년도 추진 실적의 평가 및 차년도 예산요구의 종합조정을 위한 중요한 참고자료로 활용되고 있으며, 국가 전체의 과학기술 연구개발활동과의 비교분석을 통해 정부의 중장기 국가과학기술정책 방향 수립과 정책추진에 활용되고 있다.

국가연구개발사업 조사·분석기준에 근거하여 국가연구개발사업에 대한 투자가 어떤 분야에, 어떤 형태로, 얼마만큼 투입되었는가를 분석하기 위해 조사·분석의 단위인 세부연구과제들을 경제사회목적, 연구개발단계, 연구수

행주체, 지역, 기술분류(과학기술표준분류, 미래유망기술신기술, 국가기술지도분류), 기술수명주기, 협동연구 여부, 사업목적 등으로 구분하여 분석기준을 선정하였다.

2.3.2 국가연구개발사업 투자현황 정보요소

국가연구개발사업의 투자현황 및 추세정보 제공을 위하여 선정한 분석기준은 부처, 연구개발단계, 연구수행주체, 경제사회목적, 6T, 지역이고 분석대상 데이터는 조사분석이 시행된 2002년 이후 데이터를 범위로 선정하였다. 투자현황 분석을 위해 투자금액, 사업수, 과제수 데이터 항목을 선정하였으며 분석자료의 결과해석 편의를 위해 개별항목의 비중 및 증감율, 연평균증감율을 추가로 선정하였다.

2.3.3 국가연구개발사업 참여인력현황 정보요소

국가연구개발사업에 참여한 참여인력 현황과 추세정보 제공을 위하여 선정한 분석기준은 성별(남, 여), 연령(30세 미만, 30대, 40대, 50대, 60대, 70세 이상), 전공(이학, 공학, 농림수산학, 의약보건학, 인문사회학, 기타), 학위(박사, 석사, 학사이하), 재직기관(대학교, 연구소, 산업체, 관공서, 기술사, 기타), 연구수행주체(국공립연구소, 출연연구소, 대학, 대기업, 중소기업, 정부부처, 기타) 등이고, 분석 데이터 항목으로는 연구책임자수, 여성인력 비중, 증감률, 연평균증감율을 선정하였다.

2.3.4 국가연구개발사업 성과현황 정보요소

국가연구개발사업을 통해 창출된 성과에는 논문(국내, 국외), 특허출원·등록(국내, 국외), 사업화, 인력양성, 장단기연수지원, 산업기술인력양성, 인력교류, 기술무역, 국제학술회의 개최, 국제협력기반, 산업지원, 고용창출 등 다양한 성과항목이 존재하지만 성과통계 검증이 가능한 SCI논문수, 국내특허출원건수, 국내특허등록건수, 기술료징수액, 사업화건수만을 성과항목으로 선정하였다.

국가연구개발사업을 통해 창출된 성과현황 및 추세정보 제공을 위하여 과제수행연도, 부처, 경제사회목적, 연구개발단계, 연구수행주체, 지역, 과학기술표준분류, 6T, NTRM, 기술수명주기를 분석기준으로 선정하였으며 성과통계는 과제시작연도 기준으로 계산하였다.

2.3.5 국가연구개발사업 장비·기자재현황 정보요소

장비·기자재 정보는 범부처차원에서 공동활용을 목적으로 하는 51개 장비구축·운영사업과 대학·연구소 등이 보유하고 있는 장비·기자재의 소재정보 제공을 통해 공동활용 촉진을 목적으로 하고 있다. 국가 R&D 종합모니터링 관점에서 장비·기자재 현황 및 추세제공을 위해 선정한 분석기준은 금액(10억 이상, 5억 이상~10억

<표 2> 국가연구개발사업 조사분석 분류기준

항 목		기 준
경제사회목적		OECD “Frascati Manual”에서 제시하는 기준으로 구분
연구개발단계		OECD “Frascati Manual”에서 제시하는 기준으로 구분
연구수행주체		연구개발예산을 활용하여 실질적으로 연구개발을 수행하는 기관을 의미
지역		16개 광역자치단체별로 구분
기술분류	과학기술표준분류	과학기술기본법 제27조에 의거 국가과학기술위원회에서 확정된 19개 기술분류
	미래유망신기술(6T)	IT, BT, NT, ST, ET, CT 등 6가지로 구분
	국가기술지도(NTRM)	10년 동안 국가경쟁력 확보를 위해 필수적인 핵심기술
기술수명주기		도입기, 성장기, 성숙기, 쇠퇴기로 구분
협동연구여부(참여 연구기관)		기업, 대학, 국공립 출연(연), 외국 연구기관 등 타 연구기관의 참여가 있는 협동연구 여부 구분
사업목적		원천공공복지, 산업기술, 연구기반조성, 연구기관지원으로 구분

미만, 3억 이상~5억미만, 1억 이상~3억미 만, 7천 이상~1억 미만, 5천 이상~7천 미만, 3천 이상~5천 미만, 3천 미만), 지역(수도권, 대전, 지방 및 전체 지역), 기관유형(국립대학, 사립대학, 특정연구기관, 정부출연기관, 국공립기관, 기업(부설연구소 포함))이고 데이터항목으로는 등록건수, 등록금액, DB등록률, 증감률, 연평균증감률을 선정하였다.

3. 국가R&D 종합모니터링시스템의 구성 및 주요 기능

3.1 국가R&D 종합모니터링시스템의 개요

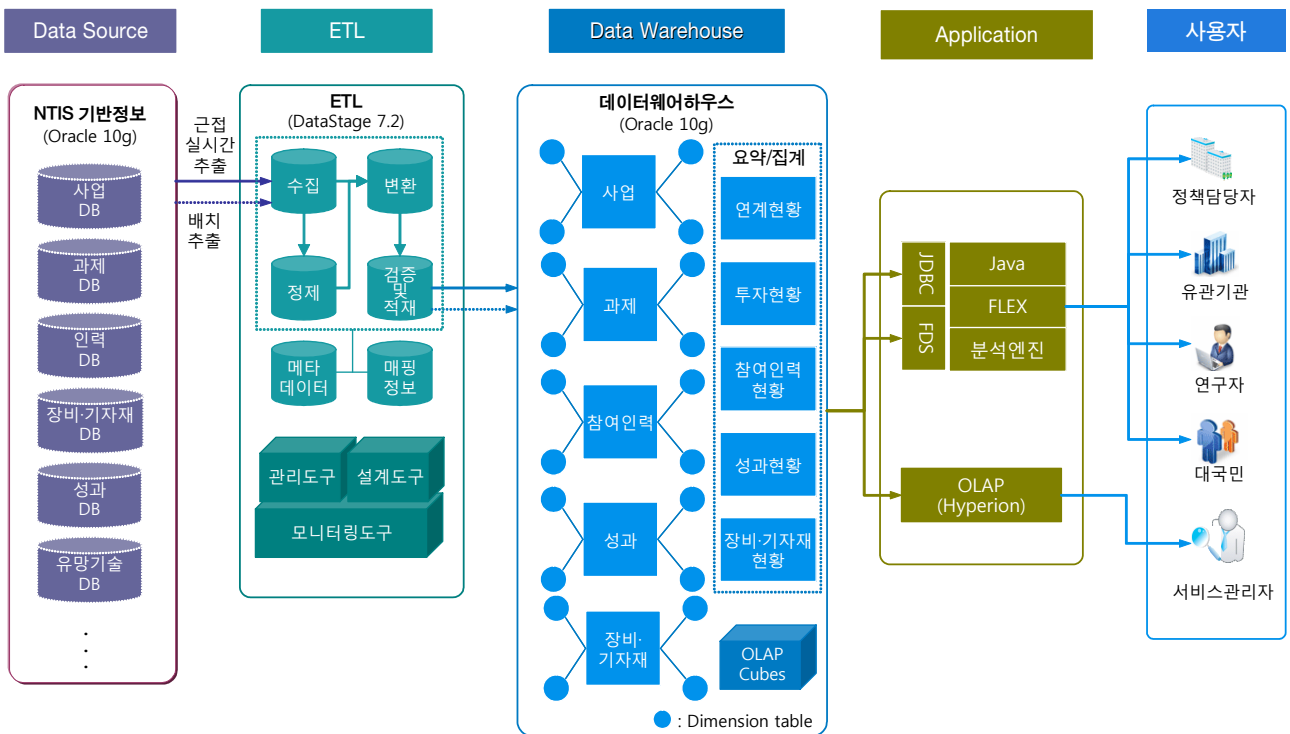
많은 기업과 조직들이 경쟁우위 확보를 위해 새로운 정보기술의 활용에 지속적인 관심을 가지고 있다. 외부 환경 변화에 신속·정확하고 유연하게 대응하기 위해서는 정확한 데이터에 근거한 환경예측과 분석을 통한 신속한 의사결정이 요구된다. 데이터웨어하우스와 데이터 마이닝에 대한 지속적인 관심은 이러한 추세의 일면을 보여주고 있다. Inmon and Hackathorn[5]의 정의에 따르면 데이터웨어하우스는 의사결정을 지원하기 위한 주제 지향적이고 통합되었으며 시계열적인 비휘발성(갱신되지

나 삭제되지 않는) 데이터의 집합이다. 사용자는 데이터 웨어하우스를 통해 양질의 데이터에 접근할 수 있고 이로 인해 데이터 분석능력과 의사결정 능력이 향상된다.

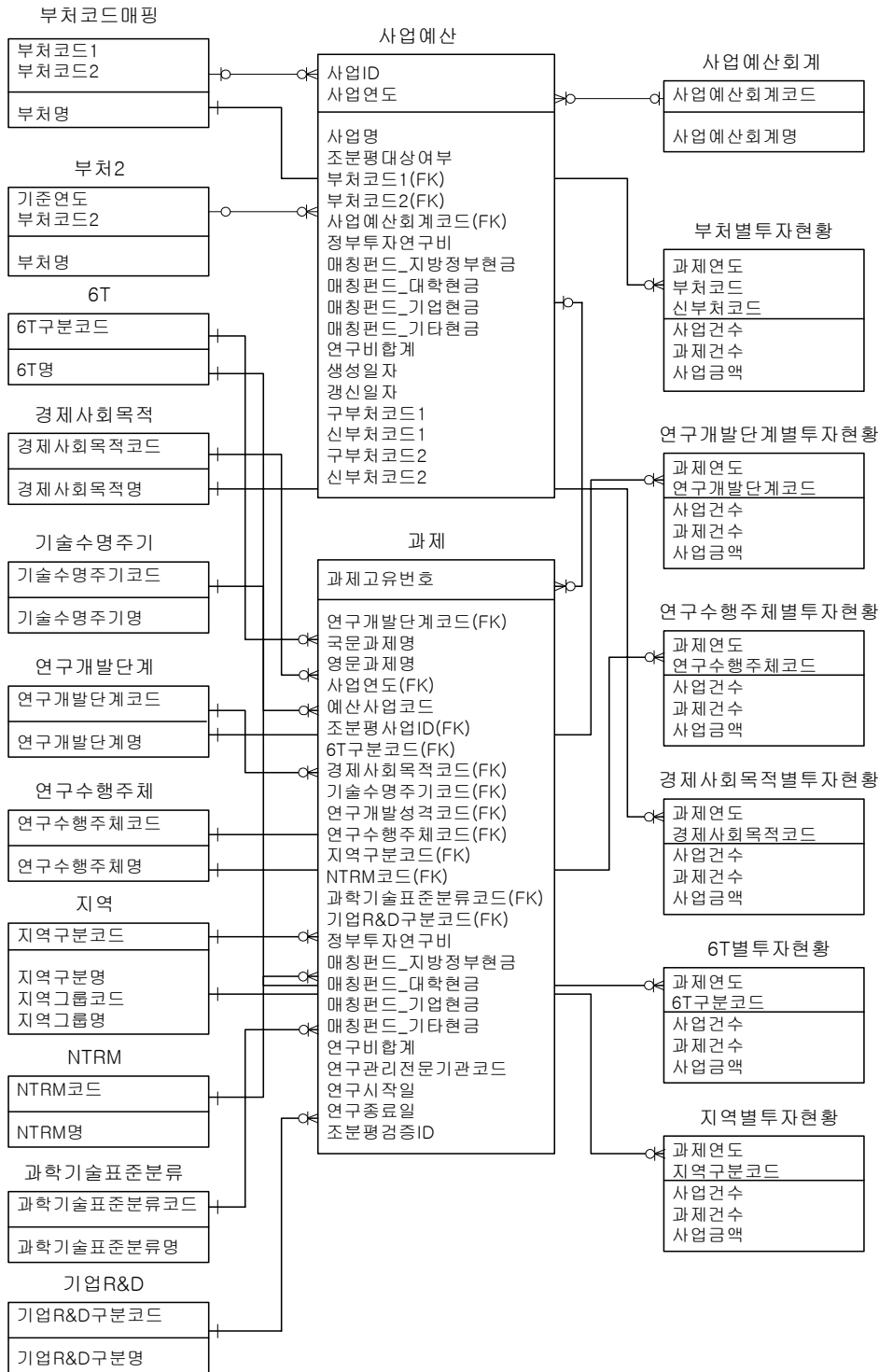
본 연구에서는 15개 부처·청의 대표연구관리전문기관과의 정보연계를 통해 수집된 NTIS 기반정보, 즉, 사업, 과제, 참여인력, 장비·기자재, 성과 정보DB로부터 국가R&D를 종합적이고 체계적으로 모니터링할 수 있는 데이터웨어하우스 기반 종합모니터링시스템(<http://dashboard.ntis.go.kr>, 이하 국가R&D Board)을 구축하였다. 구축된 시스템의 전체 서비스 개념도를 도식화하면 <그림 3>과

<표 3> Job과 Sequence 통계

영역	Job		Sequence
	Dimension	Fact	
과제	9	2	3
사업예산	10	4	3
성과	12	14	5
참여인력	3	2	3
장비기자재	2	-	2
요약통계	-	33	6
Schedule	-	-	3
소 계	36	55	25



<그림 3> 국가R&D Board의 서비스 개념도



<그림 4> 데이터웨어하우스 논리 데이터모델(예시)

같다. 데이터웨어하우스는 Oracle 10g를 이용하여 구축하였고 ETL도구로는 DataStage 7.5[4]를 이용하여 구현하였으며, 사용자의 데이터 가독성과 사용의 편의성 제고를 위해 RIA(Rich Internet Application)기술을 지원하는 FLEX 3.0을 사용하여 구현하였다.

3.2 국가R&D 종합모니터링시스템의 데이터모델링

본 연구에서는 데이터웨어하우스 구축 시 설계작업은 ER 모델링 기법과 다차원 모델링 기법을 통합하여 적용하였고, 데이터간의 관계표현에 중점을 두고 확장성이 뛰

<표 4> Source의 결과생성을 위한 쿼리(예시)

```

SELECT STAN_YY AS STAN_YY -- 사업년도 , PROG_MSTR_O_CD,
COUNT(DISTINCT DECODE(CONFIRM_BZ_YN, 'Y', INV_ANA_EVAL_TRGT_BZ_ID)) AS BZ_CNT -- 사업수,
SUM(CASE WHEN OMMITED_NO_YN = 'N' /* 누락과제 */ THEN NULL
        WHEN CONFIRM_DB_YN = 'Y' /* 유효과제 */ THEN 1 END) AS PJ_CNT -- 과제수,
SUM(DECODE(OMMITED_NO_YN, 'N', NULL, GOV_RSCH_AMT)) AS GOV_RSCH_AMT -- 정부투자 연구비, '0' GB_CD
FROM ST_BZ_PJ_INFO
WHERE CONFIRM_PJ_YN = 'Y' -- 과제확인
AND INV_ANA_TRGT_YN = 'Y' -- 조사분석대상사업
AND CONFIRM_DT_YN = 'Y' -- 초기데이터확정
AND STAN_YY > '2000' GROUP BY STAN_YY, PROG_MSTR_O_CD
UNION ALL
SELECT A.STAN_YY AS STAN_YY -- 사업년도, B.MSTR_N_CD,
COUNT(DISTINCT DECODE(A.CONFIRM_BZ_YN, 'Y', A.div_bz_id)) AS BZ_CNT -- 신사업수,
SUM(CASE WHEN A.OMMITED_NO_YN = 'N' /* 누락과제 */ THEN NULL
        WHEN CONFIRM_DB_YN = 'Y' /* 유효과제 */ THEN 1 END) AS PJ_CNT -- 과제수,
SUM(DECODE(A.OMMITED_NO_YN, 'N', NULL, A.GOV_RSCH_AMT)) AS GOV_RSCH_AMT -- 정부투자 연구비, '1' GB_CD
FROM ST_BZ_PJ_INFO A, ST_BZ_PROG_INFO B
WHERE A.CONFIRM_PJ_YN = 'Y' -- 과제확인
AND A.INV_ANA_TRGT_YN = 'Y' -- 조사분석대상사업
AND A.CONFIRM_DT_YN = 'Y' -- 초기데이터확정
AND A.STAN_YY > '2000'
AND A.DIV_BZ_ID = B.BZ_PROG_CD
AND A.STAN_YY = B.STAN_YR GROUP BY A.STAN_YY, B.MSTR_N_CD
ORDER BY 1, GB_CD

```

어난 ER 모델링 기법을 주로 사용하였으며 성능상의 문제를 고려하여 하나의 테이블에 차원(dimension)의 계층구조 전체가 포함되는 스타 스키마 기법을 추가 적용하였다.

국가연구개발사업은 매년 수행되는 조사·분석을 통해 공식통계로 공표될 때 까지는 일정한 시간지연(2008년 사업에 대한 조사분석은 2009년에 수행)이 발생하므로 본 연구에서는 범부처연계를 통해 실시간으로 수집되는 데이터와 조사분석을 통해 수집되는 데이터를 구분하기 위하여 데이터 구조를 이원화하여 설계하였다.

본 연구를 통해 구축한 데이터웨어하우스는 전체 104개 테이블로 구성되어 있으며, 이중 사실(fact) 테이블이 32개, 차원 테이블이 39개, 요약(summary) 테이블이 31개, 코드관리 및 로그관리를 위한 2개 테이블로 구성되어 있다. 과제현황 및 투자현황과 관련된 논리모델 사례를 그림으로 표현하면 <그림 4>와 같다.

<그림 4>의 논리 데이터모델은 국가연구개발사업의 과제현황과 투자현황 분석을 지원하기 위한 데이터모델을 의미한다. 사실 테이블인 과제테이블이 중앙에 위치하고, 이를 중심으로 6T, 경제사회목적, 기술수명주기, 연구개발단계, 연구수행주체, 지역 등 여러 차원을 배치한 스타 스키마(star schema) 형태를 취하고 있다. 투자현황과 관련하여 부처별, 연구개발단계별, 연구수행주체별, 경제사회목적별, 6T별, 지역별 투자현황 정보와 분류기준별 현황은 서비스속도를 고려하여 개별적인 요약테이블로 구성하였다.

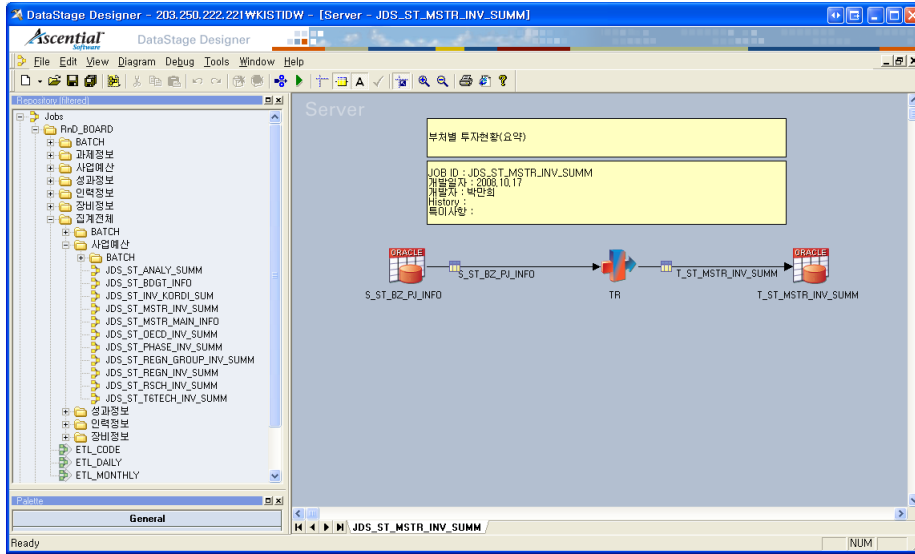
3.3 ETL 도구 및 활용

NTIS의 운영계 시스템으로부터 관련 데이터를 추출하여 변환, 정제, 통합 및 요약작업을 수행한 후 원하는 곳으로 적재해주는 작업수행을 위해 ETL 도구 가운데 하나인 DataStage 7.5를 이용하여 구현하였다.

DataStage의 기본구성은 서버 모듈과 클라이언트 모듈로 나누어지며 서버에 설치되는 모듈은 Universe라는 File DB와 dsrpcd라는 ETL Engine demon이 설치되고 구동된다. Client 모듈에는 DataStage Administrator, DataStage Manager, DataStage Director, DataStage Designer가 설치되고 DataStage의 Job 생성, 관리, 실행, 모니터링 기능을 제공한다.

DataStage Administrator는 ETL 작업에 필요한 프로젝트를 생성, 삭제 및 프로젝트의 전반적인 환경설정을 관리할 수 있는 기능을 제공한다. 본 연구를 통해 구축된 시스템에서는 DB 접속정보(서버, DB, 사용자, 암호 등)와 과제연도 관련정보를 사용자정의 환경변수로 설정하고 프로젝트내에서 생성되는 모든 Job에서 공통으로 이용할 수 있도록 설계하여 환경설정 변경관리의 유연성을 제고하였다. DataStage Manager는 DB의 메타데이터, DataStage의 Job, Routine 관리기능을 제공한다.

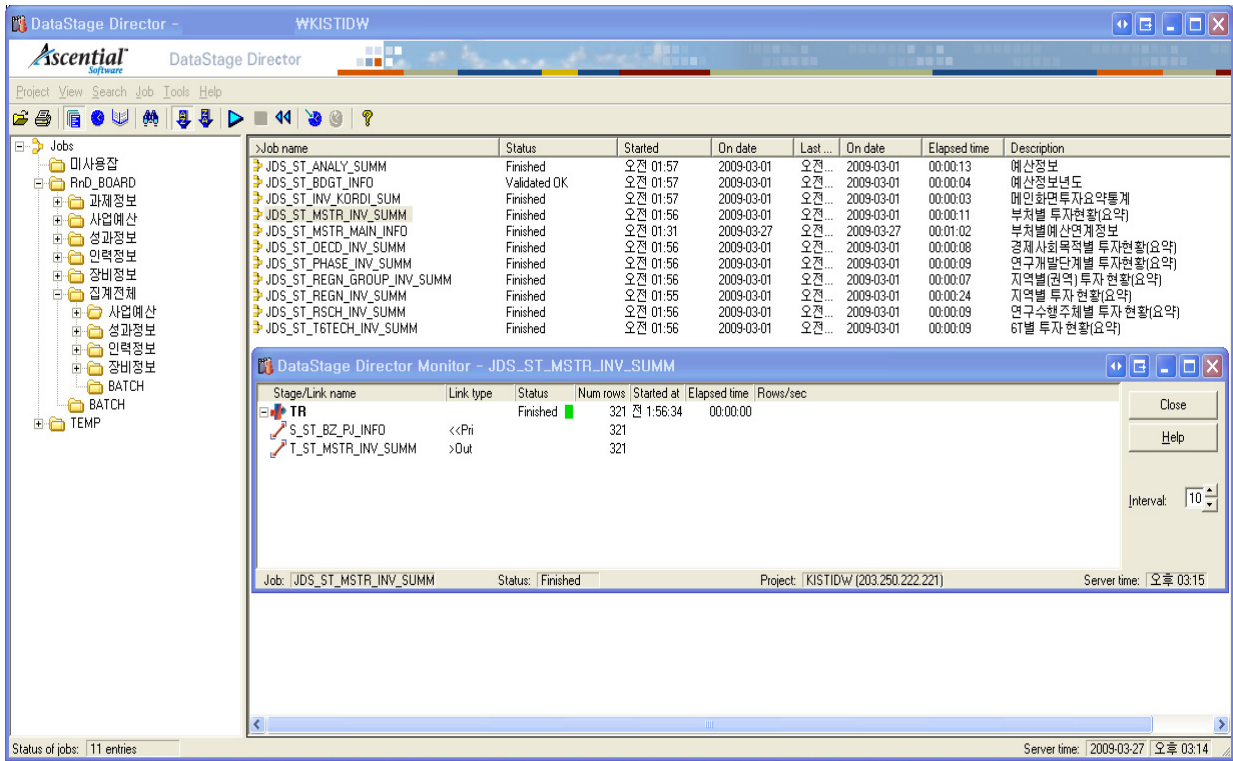
국가R&D 종합모니터링 시스템 구축을 위해 <표 3>에서 보는 것처럼 전체 91개 Job을 설계하여 구현하였으며, 이 가운데 60%가 사실테이블의 데이터생성과 관련



<그림 5> Data Stage Designer의 Job설계(예시)

<표 5> 국가R&D Board 메뉴 구성

메뉴		설 명	
메인	국가R&D Board 메인	· 정부 R&D 예산, 범부처연계현황, 국가R&D 사업추이를 요약하여 제공	
범부처 연계현황	연계현황	부처별 연계현황	· 범부처 연계기관현황과 대표연구관리전문기관별 R&D 투자금액 관리현황을 제공
	사업현황	부처별 사업현황	· NTIS를 통해 연계된 15개 부처의 투자, 사업/과제, 참여인력, 장비현황, 성과물현황을 부처별/연구수행주체별로 제공
국가R&D 사업추이	투자현황	부처별 투자현황	· 2002년부터 2007년까지 수행된 국가R&D사업의 투자현황을 부처별/연구개발단계별/연구수행주체별/경제사회목적별/6T별/지역별 현황과 추이를 제공
		연구개발단계별 투자현황	
		연구수행주체별 투자현황	
		경제사회목적별 투자현황	
		6T별 투자현황	
	지역별 투자현황		
	인력현황	연구수행주체별인력현황	· 2002년부터 2007년까지 국가R&D 사업에 참여한 연구인력의 연구수행주체별/학위별/전공별 현황과 추이를 제공
		학위별 인력현황	
		전공별 인력현황	
	공동활용 필수장비	금액별 장비현황	· 2002년부터 2007년까지 51개 장비구축사업을 통해 등록된 장비의 금액별/지역별/기관유형별 장비현황과 추이를 제공
		지역별 장비현황	
		기관유형별 장비현황	
	기타 R&D 장비	금액별 장비현황	· 2002년부터 2007년까지 51개 장비구축사업을 통해 등록된 장비의 금액별/지역별/기관유형별 장비현황, 등록율과 추이를 제공
		지역별 장비현황	
기관유형별 장비현황			
기관유형별 등록률 현황			
성과현황	부처별 성과현황	· 2002년부터 2007년까지 국가R&D 사업을 통해 창출된 부처별/연구개발단계별/연구수행주체별/경제사회목적별/6T별/지역별 성과현황과 추이를 제공	
	연구개발단계별 성과현황		
	연구수행주체별 성과현황		
	경제사회목적별 성과현황		
	6T별 성과현황		
지역별 성과현황			
관리 서비스	분석	추세분석	· 정부 R&D 투자, 사업수, 과제수, 참여인력수 등의 과거 데이터를 이용하여 회귀분석과 연평균증가율을 적용하여 미래 특정기간동안의 예측치를 제공
		코드관리	· NTIS 공통코드와 국가R&D Board 자체코드를 등록, 수정, 삭제 등의 관리기능을 제공
		로그관리	· DataStage에 등록된 Job별 소요시간, Job Status, 추출건수, 적재건수 모니터링 기능을 제공



<그림 6> DataStage Director의 Job Monitor(예시)

된 Job이고 나머지 40%가 차원테이블의 데이터 생성과 관련된 Job으로 구성되어 있다. Datastage Designer는 실행되어야 하는 서버의 순서나 병렬작업을 명시할 수 있게 해주는 그래픽기반 작업순서 설정기능을 제공한다. 순서설정 기능은 특정작업의 성공 혹은 실패여부에 따라 수행해야 할 서로 다른 과정을 명시할 수 있으므로 관리 정보를 포함시킬 수 있다. 본 연구에서는 25개 Sequence를 정의하였으며 이들 가운데 1일 주기로 갱신되고 있는 NTIS의 범부처 연계데이터를 고려하여 일별 Schedule을 설계하였으며, 월별 연계현황 집계작업을 위해 월별 Schedule을 설계에 반영하여 구현하였다.

<그림 5>는 부처별 투자현황 정보를 제공하기 위하여 설계된 Job의 사례를 예시한 것으로 Source로부터 Target으로 관련정보를 적재하기 위한 작업을 수행하고 있다. 예산과제 사실테이블로부터 조사분석 대상 데이터를 추출하여 부처별로 사업연도, 사업수, 과제수, 정부연구투자비를 집계하여 부처별 투자현황 요약 테이블에 관련 데이터를 적재하는 작업을 수행한다. 그림에서 TR은 변환단계(Transformer Stage)를 의미하며 Source의 컬럼과 Target의 컬럼 매핑정보와 내장함수를 이용하여 정제작업을 정의할 수 있는 기능을 제공한다. Source(S_ST_BZ_INFO)로부터 생성되는 결과는 변환단계를 거쳐 Target(T_ST_MSTR_SUMM)으로 적재되는데 <그림 5>에서 Source로

부터 생성되는 결과는 <표 4>와 같은 SQL로 정의되어 있다.

DataStage Director는 개발된 Job에 대해서 실행, 중지, 성능 모니터링 기능을 제공한다. 또한 Job에 대한 상세한 작업 로그 및 스케줄링(특정 시간, 일자에 자동으로 실행하는 기능) 기능을 제공한다. <그림 6>은 부처별투자현황 Job에 대한 Monitor기능을 예시한 것으로 Source로부터 321개 행을 추출하여 변환·정제작업을 거친 후 Target으로 321개 행을 적재하였다는 것을 의미한다.

3.4 국가R&D 종합모니터링시스템의 메뉴구성 및 주요 기능

본 연구를 통해 구축된 국가R&D Board는 <표 4>에서 보는바와 같이 4개 메인메뉴와 9개 하위 메뉴로 구성되어 있으며, 전체 28개 세부메뉴로 구성되어 있다. 관리 서비스를 위한 분석메뉴를 제외한 나머지 세부 메뉴들은 국가R&D 정책결정자와 핵심 사용자의 의견수렴을 통해 국가R&D 현황과 추이를 한눈에 파악할 수 있는 25개 핵심지표를 도출한 후 이를 반영하였다.

국가R&D Board 세부 메뉴들의 주요 기능과 특징들을 구현된 화면을 중심으로 설명하면 다음과 같다.

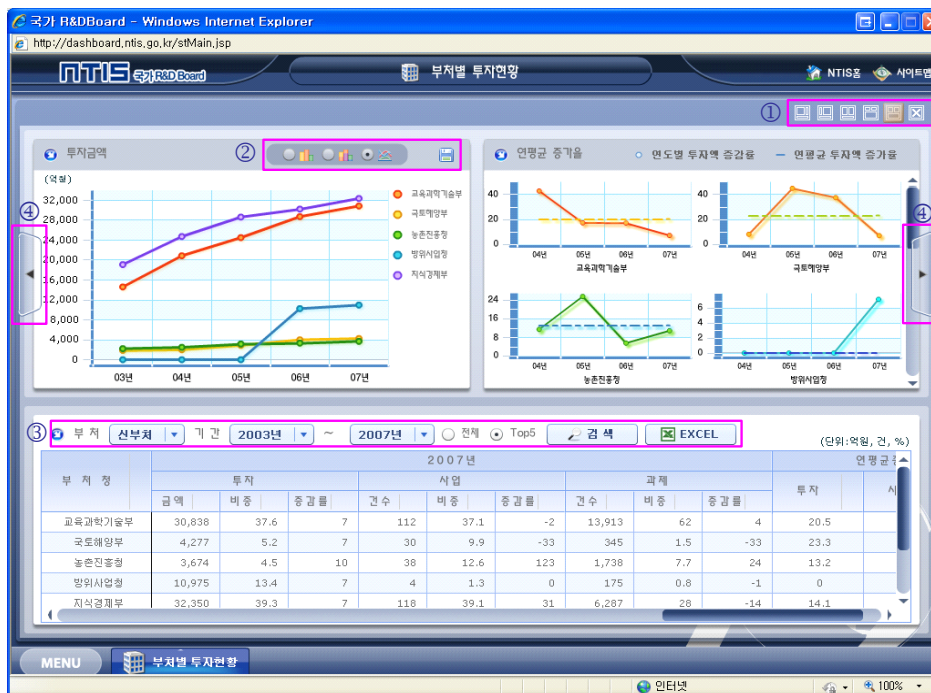
국가R&D Board의 메인화면은 정부 R&D 예산, 범부



<그림 7> 국가R&D Board 메인화면

치연계현황, 투자, 참여인력, 성과, 공동활용필수장비, 기타R&D장비로 구성되어 있으며, 정부 R&D 예산은 최근 5년 동안의 정부R&D 투자를 예산과 기금으로 구분하여 총투자금액 현황과 추이를 그래프와 테이블 형태로 <그림 7>

과 같이 제공되고 있다. 제공된 서비스결과에 따르면 정부R&D예산은 2005년부터 12.2%의 연평균증가율을 나타내고 있으며 2009년 현재 12조 3,437억 원의 투자가 이루어지고 있다. 사용자 인터페이스는 핵심지표들을 아이콘



<그림 8> 부처별 투자현황



<그림 9> 정부 R&D 투자액 추세분석

화하여 식별성을 제고하였으며 Windows XP형식의 메뉴 구성과 인터페이스를 제공하고 있다.

부처별 투자현황은 <그림 8>과 같이 최근 5년 간 투자금액 기준으로 상위 5개 부처의 투자금액과 사업수, 과제수와 비중 및 연평균증가율이 그래프와 표로 제공된다. 조회되는 정보는 2개의 그래프와 1개의 표로 구성되어 있으며 각각의 그래프와 표는 창크기 조절기능(①)을 이용하여 특정 그래프와 표를 최대화할 수 있다. 제공되는 그래프는 그래프 선택기능(②)을 이용하여 Bar, Stack, Line차트로 변경할 수 있으며 그래프는 JPG형식의 그림 파일로 저장된다. 투자금액, 사업수, 과제수 정보항목에

따른 서비스화면은 화면 스크롤기능(④)을 이용하여 변경할 수 있다. 최근 5년 간 상위 5개 부처의 투자금액 현황은 검색기준 선택기능(③)을 이용하여 신부처, 구부처를 선택할 수 있으며 2002년부터 2007년까지 데이터를 분석대상으로 선정할 수 있고, 전체를 선택하면 전체 부처에 대한 투자현황을 분석할 수 있으며 테이블에 제공되는 데이터는 엑셀파일로 저장이 가능하다.

추세분석은 과거 데이터를 이용하여 미래 특정기간 동안의 데이터를 예측하여 그 결과를 그래프와 표로 제공하는 서비스이다. 분석대상 데이터는 정부R&D 총투자금액, 조사분석대상 투자금액, 사업수, 과제수, 참여인력

수이고 분석방법은 회귀분석과 연평균증가율을 지원하며 분석기법에 따른 통계량을 제공한다. 본 연구를 통해 구현된 추세분석 서비스의 예시화면은 <그림 9>와 같다.

4. 결 론

본 연구에서는 사용자 요구분석과 의견수렴 결과를 바탕으로 국가R&D 현황을 종합적이고 체계적으로 모니터링할 수 있는 주요 정보항목 및 요소와 25개 핵심지표를 선정하고, 핵심지표 서비스를 위해 데이터웨어하우스를 구축하고 국가R&D 종합모니터링시스템인 국가R&D Board를 구축하였다.

비용대비 효율이라는 측면에서 국가R&D 종합모니터링 시스템 구축의 목적은 최소한의 비용으로 최대의 효율을 얻는 것이다. 그러나 모든 정보시스템은 장점과 더불어 한계점 또한 존재한다. 국가R&D 종합모니터링 시스템 구축을 통해 얻은 대표적인 이슈와 쟁점들을 정리하면 다음과 같다.

첫째, 일반적으로 정보시스템 구축 사업의 성공여부는 최고경영자 혹은 사업추진 주체의 강력한 의지와 관심에 의하여 결정된다고 보고되고 있다. NTIS는 다양한

이해관계자가 참여하고 있는 범부처 연계사업이므로 개별 부처의 이해와 협력을 이끌어내는 게 가장 중요한 사업 성공요인이었다고 할 수 있다. 사업초기에 이러한 이해와 협력체제 부재로 사업진행이 다소 부진하였으나, 교육과학기술부의 강력한 의지와 사업추진위원회와 실무추진위원회를 통한 역할분담 및 조정을 통해 다양한 이해관계자의 협력과 참여를 유도할 수 있었다.

둘째, 분산되어 있는 정보를 통합함에 있어서 정보의 일관성과 품질을 유지하기 위해서는 표준화가 선결되어야 한다. 개별 대표연구관리전문기관의 데이터 관리 및 품질수준이 상이하여 한순간에 통합하는 것은 불가능하다. 따라서 NTIS에서는 필요최소한의 정보를 공동활용 정보라는 표준안을 마련하고 단계적 확대방안에 따라 연계를 추진중에 있다.

셋째, 데이터웨어하우스와 데이터마트는 기존의 운영계 시스템으로부터 데이터를 추출하여 사용하게 된다. 운영계 시스템에서 추출되는 데이터의 품질수준이 낮으면 전체 데이터웨어하우스의 품질수준이 저하될 수 밖에 없다. 따라서 NTIS는 이러한 품질저하를 막기 위하여 개별 부처로부터 제공되는 정보의 정제작업에 많은 노력을 기울임과 동시에 NTIS 품질상 제정 등을 통해 대표 연구관리전문기관에서 관리중인 데이터의 품질수준을 제고하고 있다.

본 연구에서 구축한 국가R&D Board를 통해 얻을 수 있는 대표적인 정량적 기대효과로는 평균 검색시간 절감이라고 할 수 있다. 2008년 3월 서비스를 시작한 이후 월별 평균 페이지뷰는 2,500회이고 시스템 구축을 통한 조회시간 절감은 DBMS상의 쿼리수행시간 기준으로 평균 10초 정도의 조회속도가 개선된 것으로 분석되었다. 이를 검색 절감시간으로 환산하면 연간 83시간(=2,500회/월×12월×10초/페이지) 정도의 절감효과가 있는 것으로 분석되었다. 이외의 기대효과로는 핵심지표를 바탕으로 사용자에게 종합현황 및 추세데이터 제공이 가능하고, 이를 바탕으로 국가R&D 정책의사결정이 효율적이고 효과적으로 이루어질 수 있는 환경이 마련되었다는 점이다. 또한, 시스템 통해 핵심정보가 제공되므로 종합현황 분석시간과 정형 보고서 작성시간이 절감되었으며 결과적으로 EUC(End-User Computing)환경을 구축하였다는 점이다. 무형의 효과로는 국가R&D 분석데이터를 제공함으로써 국가R&D 투자전략 수립의 기초자료로 활용이 가능하며 2002년 이후 데이터에 대한 분석정보를 제공함으로써 시계열분석을 위한 기반을 제공하였다는데 있다.

본 연구를 통해 구축한 시스템에서는 현황정보 제공을 1차적인 목적으로 하여 구축되었으므로 고급 분석정보나 고부가가치 정보제공 측면에서는 상당히 미흡하고 독자적인 solution이나 알고리즘 적용하지 못한 관계로

학문적인 측면에서 많은 부분 미흡하다고 할 수 있다. 학문적으로 미흡한 부분들은 향후 개발할 고부가가치 정보제공을 위한 분석정보서비스를 통해 보완해 나갈 계획이다. 유럽의 IST world[9]는 CORDIS 정보를 바탕으로 키워드 매칭방식에 기반한 네트워크분석을 적용하여 협업 다이어그램, 경쟁 다이어그램, 컨소시엄 예측서비스 등을 제공하고 있다. 본 연구를 통해 구축한 데이터웨어하우스를 활용하여 데이터마이닝 분석기법과 네트워크분석 등을 적용하여 국가R&D의 주요 특성, 연구자 네트워크, 연구 방향성 등의 고부가가치 정보제공을 통해 정책 의사결정을 지원할 수 있는 시스템으로 개선·발전시켜나갈 계획이다. 이를 위해 다음과 같은 콘텐츠 확충 및 분석서비스를 통해 국가R&D 정보의 활용도와 정보가치를 제고하고자 한다.

첫째, NTIS에서 구축된 국가R&D 과제, 참여인력 정보를 바탕으로 네트워크분석을 적용하여 주요 경쟁관계에 있는 과제나 전문가가 속한 주제 클러스터를 시각화하여 제공하는 경쟁관계 다이어그램, 미래에 협업할 확률이 가장 높은 클러스터들을 시각화하여 제공하는 컨소시엄 예측서비스를 제공할 계획이다.

둘째, 국가R&D 성과(SCI논문, 특허 등)에 영향을 미치는 주요 특성변수(사업금액, 사업기간, 과제수, 사업분류, 연구수행주체, 참여인력 등)간의 관계를 의사결정나무분석 기법을 적용하여 분류하고 가시화함으로써 국가R&D 정보에 대한 이해도를 제고할 계획이다.

참고문헌

- [1] 교육과학기술부; “연구개발투자의 경제성장기여도 국제비교”, 2007.
- [2] 국가과학기술위원회; “국가연구개발사업 조사·분석 보고서”, 2007.
- [3] 한국과학기술정보연구원; “국가과학기술종합정보서비스”, 지식정보인프라, 30 : 30-34, 2008.
- [4] Harris, S. and Englert, S.; “A flexible data integration architecture using WebSphere DataStage and WebSphere Federation Server,” IBM, 2007.
- [5] Inmon, W.H., and Hackathorn, R.D.; “Using the data warehouse,” John Wiley and Sons, Inc., New York, 1994.
- [6] OECD; “A new Economy?; The Changing Role of Innovation and Information Technology in Growth,” 2000.
- [7] CORDIS(http://cordis.europa.eu/home_en.html).
- [8] e-RAD(<http://www.e-rad.go.jp>).
- [9] IST World(<http://www.ist-world.org>).
- [10] JST(<http://www.jst.go.jp>).
- [11] NTIS(<http://www.ntis.gov>).