

# 국가 R&D정보 식별메타데이터에 관한 연구

## A Study on the Identification Metadata of National R&D Information

권이남, 김재수\*

한국과학기술정보연구원 선임연구원,  
한국과학기술정보연구원 기반시스템개발팀장\*

Lee-nam Kwon, Jae-soo Kim\*

KISTI Senior Researcher,  
KISTI Infrastructure System Development Team Leader\*

### 요약

국가 R&D정보에 필요한 모든 정보를 정보의 소유에 관계없이 범 부처적으로 공동 활용하고자 하는 필요성이 대두되고 있다. 그러나, 국가 R&D정보는 부처별 각 연구관리전문기관별로 별도로 관리하고 있으며 상호 연계를 위한 식별체계의 부재로 인해 유사·중복과제 검증, 정보의 추적관리 등 통합적인 정보관리가 어려운 실정이다. 전 부처의 국가 R&D정보를 통합 관리하고 공동활용하기 위해서는 모든 정보에 항구적으로 유일한 URN기반 식별체계의 부여가 필요하다. 본 논문에서는 식별체계를 부여하기 위해 필수적인 식별메타데이터의 개념과 각 연구관리전문기관별 국가 R&D 과제관리번호의 현황을 살펴보고, 범부처 과제고유번호 및 정보식별자 부여를 위한 국가 R&D정보의 식별메타데이터 요소를 제안하고자 한다.

### Abstract

A need for sharing the R&D information at a national level regardless of the ministry in charge of it is being aroused. However, this information is being managed by different institutes under different ministries without a standard identifier which makes it difficult to manage the information such as to look up duplicate projects. A URN-based identifier is required for the integrated management and utilization of the national R&D information. In this paper, the concept of the identification metadata for applying the identification system and the status of the information management by each institute will be the subject for discussion. Based on this, the elements required for issuing a unique identification number and identifier to the national R&D projects will be suggested in this paper.

## I. 서론

국가 R&D정보는 국가 R&D의 전주기에 걸쳐 발생하는 정보로서 국가 R&D 사업·과제, 참여인력, 연구개발 결과로 산출되는 성과, 연구 장비·기자재 등에 관한 정보이다. 현재는 각 부처별, 기관별로 과제관리번호 체계가 상이하여 통합적인 과제관리가 어려우며, 국가 R&D정보의 식별체계 부재로 정보의 상호연계 및 공동활용이 불가능하며, 국가연구개발사업으로 수행한 연구성과인 특허, 논문 등의 관련 기술정보의 관리나 추적도 불가능하다.

전 부처의 국가 R&D사업 과제를 통합 관리하고 공동활용하기 위해서는 모든 연구과제에 항구적으로 유일한 고유번호를 부여하고 이를 관리할 수 있는 표준화된 과제고유번호 체계의 부여가 필요하다. 또한, 국가 R&D사업 전주기에 발생하는 연구보고서, 논문, 특허원문 등 각종 원문정보를 체계적으로 관리하고, 지속적인 성과관리를 위한 정보간 연계를 위해 국가R&D사업 관련 원문정보에 대하여 유일하게 식별할 수 있는 표준적인 정보식별자가 필요하다[1].

국가 R&D정보에서의 식별체계는 국가 R&D관련 정보객체

에 유일한 번호를 부여하고, 이를 활용하여 정보객체를 식별하고 접근할 수 있도록 하는 표준적인 관리체계로서 과제고유번호와 정보식별자로 구성되어 상호간 연계성을 가지도록 설계되었다.

국가 R&D정보의 식별메타데이터는 이러한 국가 R&D 관련 정보자원을 유일하게 식별할 수 있는 메타데이터로서 과제를 식별하기 위한 과제고유번호 식별메타데이터와 정보식별자 식별메타데이터로 구성된다.

본 논문에서는 식별체계를 부여하기 위해 필수적인 식별메타데이터의 개념과 각 부처의 연구관리전문기관별 과제관리번호 체계 현황과 범부처 과제고유번호 구분구조 및 등록방식 현황을 조사한다. 또한, 국가 R&D정보의 식별메타데이터 요소 수립을 위한 기본 원칙과 범부처 과제고유번호 및 정보식별자의 식별메타데이터 요소를 제시한다.

## II. 식별메타데이터와 식별체계

### 1. 식별메타데이터

국내외에서 분야와 목적에 따라서 메타데이터는 다양하게

존재한다. 범용적인 메타데이터 표준인 Dublin Core(이하 DC)는 인터넷 자원 증가에 따른 네트워크 기반 정보자원의 기술요소라고 할 수 있다. 그러나, URN기반 식별체계의 관점에서 보면 식별을 위한 요소로 포함하지 않아도 되는 요소(예를 들면, 주제(Subject), 설명(Description), 언어(Language), 관계(Relation) 등)를 가지고 있다.

URN(Uniform Resource Names)은 디지털콘텐츠에 대한 영구적인 식별자로서 1994년 IETF(Internet Engineering Task Force)가 디지털콘텐츠의 위치와는 상관없이 콘텐츠의 이름을 부여하는 영구 식별체제로 제안한 것이다[7]. 즉, URL이 위치정보가 바뀌면 식별자로서 역할을 제대로 수행하지 못하는 문제점을 해결하기 위해 URN은 정보자원의 위치와는 상관없이 정보자원에 영구적인 고유의 식별기호를 부여하여 해당 정보자원의 위치로 접근할 수 있게 하는 관리체계로서 해외의 DOI, OpenURL, 국내의 KOI, UCI 식별체계 등이 이에 해당된다. 이러한 URN기반 식별체계 활용을 위한 메타데이터의 개발 및 표준화는 필수적이다.

식별체계의 메타데이터는 크게 식별메타데이터와 응용메타데이터로 구분하는데 식별메타데이터는 식별체계 부여 대상 정보를 유일하게 식별할 수 있도록 핵심적으로 정의된 요소들의 셋(Set)이다.

응용메타데이터는 식별메타데이터를 확장하여 응용분야로 추가되는 메타데이터로서 해당 콘텐츠의 다양한 특성을 반영한 메타데이터이다. 본 연구에서는 국가 R&D정보의 식별메타데이터로 제한하기로 한다.

## 2. 메타데이터 수립 원칙

식별메타데이터 요소를 개발하는데 있어서 고려되어야 할 요구사항을 수립하기 위하여 DC와 UCI의 메타데이터 수립을 위한 기본 원칙을 살펴보면 다음 [표 1], [표 2]와 같이 정리할 수 있다.

[표 1] DC의 메타데이터 수립 원칙

수립 원칙	설 명
고유성	자원의 본질적인 특성이 있어야 함
확장성	규정된 요소외에 부차적인 내용이나 특성을 기술요소로 사용가능해야 함
구문 독립성	응용분야나 표현기법을 규정하지 않아야 함
선택성	각 요소의 수록여부를 강제하지 않아야 함
반복성	모든 요소는 반복사용 가능해야 함
수정가능성	한정어를 사용하여 세부사항을 수정할수 있어야 함

[표 2] UCI의 식별메타데이터 수립 원칙[4]

수립 원칙	설 명
식별성	유일하고 식별가능 해야함
단순성	최소한의 공통적인 핵심요소로 구성됨
확장성	분야에 관계없이 적용될 수 있는 기본 프로파일과 특정 분야의 특성에 맞게 구성된 응용 프로파일로 구성됨
공개성	공개 가능한 요소로 구성되어야 함
상호운용성	다른 메타데이터 표준과의 연계를 위한 상호운용성을 보장해야 함

## 3. 국가 R&D정보의 식별체계

국가 R&D정보에서의 식별체계는 각 연구관리전문기관별로 관리하고 있는 과제의 관리번호에 대하여 추가적으로 유일하고 고유하게 부여하는 표준화된 과제고유번호를 부여하여 활용하고, 국가 R&D사업의 성과물인 각각의 정보에 URN기반 정보식별자를 부여하는 것이 효율적이다. 국가 R&D과제에 범부처 과제고유번호를 부여하고, 이와 관련된 산출물에 대하여 정보식별자를 부여하면 과제의 수행단계 뿐만 아니라 사후 관리 단계(평가, 기술이전, 사업화 등)에 있어서도 일관성 있는 종합관리가 가능하며, 국가 R&D정보의 참조연계 및 포워드연계 서비스 기반을 마련할 수 있다[2].



▶▶ 그림 1. 국가R&D정보 식별체계 적용대상도

즉, 이용자가 특정 과제 참여연구원의 성과물에 대한 정보식별자로 검색을 해도 관련 과제정보 및 과제계획서까지 연계되어 볼 수 있고 과제 관련 장비 정보 및 과제 수행후의 관련 성과(논문, 특허, S/W, 생물소재 등)까지 모두 연계해서 볼 수 있게 된다. 즉, 어느 방면에서 조회를 하더라도 국가 R&D과제와 관련된 모든 정보를 검색없이 한꺼번에 연계해서 볼 수 있게 된다[1].

## III. 현황 조사

### 1. 과제관리번호 현황

현재 각 부처별 연구관리전문기관은 국가R&D 과제관리번호

호를 별도로 부여하고 있으며 동일 부처내에서도 연구관리전문기관별로 상이한 형식으로 과제관리번호를 부여하고 있다.

[표 3] 기관별 국가연구개발사업 과제관리번호

연구관리전문기관	과제번호 형식
한국과학재단	A99-9999-999-99999-9 (예: R82-2002-000-00010-0)
한국학술진흥재단	KRF-YYYY-999-AA9999 (예: KRF-2002-074-HS1002)
정보통신연구진흥원	A9999-YYMM-9999 (예: A1100-0401-0001)
한국산업기술평가원	99999999 (예: 10000026)
환경기술진흥원	YYYY9999999999 (예: 20050100200010)
한국문화콘텐츠진흥원	9-YY-AAAA-999-9999-99-9999 (예: 1-05-5003-002-1074-00-0001)
한국보건산업진흥원	YY-AAAA-AAA-AAAAAA-9999 (예: 01-PJ10-PG6-01GN16-0003)
중소기업기술정보진흥원	99999999 (예: 10000031)
농림기술관리센터	9YY999-999-AA999 (예: 205001-031-WT111)
한국건설교통기술평가원	AA-YYYY-A-999-9-999-99 (예: 03-2001-A-001-3-005-01)

[표 3]의 10개 기관 과제관리번호는 8자리에서 30자리까지 다양하며, 사업구분, 과제종류, 기관코드, 연도, 일련번호 등의 정보를 포함하고 있다. 단순히 과제관리번호를 코드화 하지 않고 일련번호(8자리)를 사용하는 경우도 2개 기관이 있어서 기존 번호를 그대로 사용할 경우 중복이 발생하게 된다.

## 2. 범부처 과제고유번호

부처별·기관별로 독자적인 과제관리번호를 사용하고 있어 범부처적인 국가R&D정보 연계 시에 과제 중복번호 발생에 대한 문제를 해결하기 위해 “국가연구개발사업의 관리 등에 관한 규정 일부개정령”에는 과제의 최초협약연도, 사업구분, 국가과학기술표준분류 등을 알 수 있도록 과제별 고유번호를 부여하도록 규정하고 있다[5].

부처별로 과제식별과 인력, 성과, 장비·기자재 등 국가 R&D 사업정보간 연계를 위한 키(Key)값으로 과제고유번호를 관리하면, 국가 R&D사업·과제의 부처별, 연도별 통계 산출에 활용가능하며 연구성과의 전주기적인 관리체계 확립이 가능해진다.

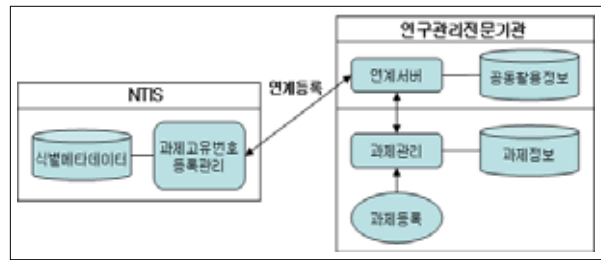
범부처 과제고유번호 구문구조는 다음 [표 4]와 같이 부처코드 3자리와 과제일련번호 7자리로 구성되며 총 10자리로 구성된다.

[표 4] 범부처 과제고유번호 구문구조

①부처코드	②과제일련번호
3자리(A99) : 행정표준코드 기준	7자리(9999999) : 과제일련번호

연구관리전문기관의 시스템 변경을 최소화하면서, 연구성과의 전주기적 관리 및 중복번호 방지가 가능한 방안으로는 국가연구개발사업종합관리시스템(이하 ‘NTIS’) 내에서만 범부처 과제고유번호를 새로 부여하여 관리 및 제공하고, 기관별 과제번호는 그대로 유지하는 것이다.

즉, 과제의 신규 등록시에 자체 과제관리번호를 발급하여 사용하고, 연계서버에 과제 기본데이터와 자체 과제관리번호를 제공하면 된다. [그림 2]와 같이 NTIS에서 배치작업으로 과제고유번호를 부여하여 연계서버에 저장하고, 이를 연구관리전문기관에서 사용하는 방식이다. 기존의 과제관리번호체계를 유지하면서 과제고유번호와 자체 과제관리번호를 병행하여 사용할 수 있게 된다.



▶▶ 그림 2. 연계방식의 과제고유번호 등록

## IV. 국가 R&D정보 식별메타데이터

### 1. 개발 원칙

국가 R&D정보의 URN기반 식별체계를 부여하기 위한 식별메타데이터를 개발하는데 있어서 고려한 기본 원칙은 다음 네가지이다.

첫째, 유일한 식별성이다. 부처별 연구관리전문기관별로 다양한 과제나 정보를 유일하게 식별하기 위한 최소한의 공통된 핵심요소로 구성되어야 한다.

둘째, 공개성이다. 국가 R&D정보를 공동으로 활용하기 위한 목적이므로 모든 요소는 공개 가능한 요소로 구성되어야 한다.

셋째, 상호운용성이다. 기존에 존재하는 다른 메타데이터 표준과의 연계를 고려하고 다양한 콘텍스트 환경을 지원할 수 있도록 상호운용성을 확보해야 한다.

넷째, 적절한 접근성이다. 메타데이터 요소에 대해 적절한 접근(Access)이 가능해야 한다.

## 2. 국가 R&D정보 식별메타데이터

### 2.1 과제고유번호 식별메타데이터

앞에서 제시한 원칙에 따라 국가 R&D정보의 범부처 과제 고유번호를 식별하기 위한 식별메타데이터 항목은 다음과 같다.

[표 5] 범부처 과제고유번호 식별메타데이터

NO	요소명	설명
1	과제고유번호	과제의 유일한 과제식별번호
2	과제명	총괄 또는 세부·협동 과제명을 한글로 기입
3	발주기관명	과제를 발주한 연구관리전문기관의 명칭
4	주관연구기관명	과제의 주관 연구기관명
5	연구책임자	해당 과제의 총괄(주관) 연구책임자 또는 세부·협동 연구책임자
6	과제번호	과제를 구분하기 위해 연구관리전문기관에서 부여한 자체관리번호
7	연차	다년도 과제인 경우 당해연도의 연차
8	분류코드	국가과학기술표준분류 또는 기술분류 등의 분류 코드값
9	최초협약연도	과제를 최초로 협약한 연도
10	종료연도	과제 종료연도
11	사업코드	해당 과제의 사업에 대한 예산사업코드
12	이전과제번호	다년도 과제일 경우, 이전 과제의 과제번호

### 2.2 URN기반 정보식별자의 식별메타데이터

국가 R&D정보의 URN기반 정보식별자 부여 대상 정보는 매우 다양하다. 국가 R&D사업 전주기에 발생하는 모든 관리 대상 정보로서 인력정보, 장비정보, 논문, 특허, 연구보고서, S/W, 사업화, 기술거래, 생물소재, 생물정보, 화합물, 과제관련정보(과제계획서, 과제연구보고서), 사업정보 등의 상세정보 및 원문정보에 부여한다.

정보식별자의 식별메타데이터를 도출하기 위하여 [표 6]과 같이 국가 R&D정보의 주요 데이터 항목들에 대하여 공통적 의미를 가진 항목들로 구분하여 정리하였다.

[표 6] 국가 R&D정보 주요데이터 항목 구분

구분	인력	장비·기자재	성과			
			특허	논문	SW	생물소재
식별번호	인물ID	장비ID	정보식별자	정보식별자	정보식별자	정보식별자
제목	성명	장비명	특허명	논문명	프로그램명	생물소재명
저자명	성명	등록자	출원인	저자	저작자	기탁자
저작일	등록일	취득일	출원일/등록일	게재일	창작일	기탁일/증부여일
기존식별자	주민등록번호	장비관리번호	출원번호/등록번호	ISSN,DOI,KI등	SW등록번호	생물소재번호
설명	-	설명	요약	요약	설명	설명
분류	국가과학기술표준분류	국가과학기술표준분류	국제특허분류	주제분류	적용분야	생물소재분류

이들의 공통적이면서 정보의 식별에 필수적인 항목을 중심으로 다음 [표 7]과 같이 11개의 국가 R&D정보 식별메타데이터 구성요소를 정의하였다.

[표 7] 국가 R&D정보 정보식별자 식별메타데이터

NO	요소명	정의	빈도
1	정보식별자	국가 R&D정보를 식별하기 위한 URN기반 식별번호, 객체번호(인물에 대한 식별번호인 인물ID, 장비 객체에 대한 식별번호인 장비ID)를 포함함	1..1
2	과제고유번호	국가 R&D 과제고유번호	1..1
3	정보등록기관코드	정보식별자 대상정보의 메타데이터를 등록하는 기관(연구관리전문기관 또는 직접수행기관 등)의 기관코드	1..1
4	제목	알려져 있는 자원의 명칭	1..n
5	저자명	정보의 저작자	1..n
6	저작일	정보의 저작 관련 날짜	1..1
7	유형	정보식별자 대상정보에 대한 자료유형 코드	1..1
8	분류	정보의 분류	0..n
9	기존식별자	정보에 부여된 기존에 존재하는 식별자	0..1
10	파일형식	정보의 파일 표현형식	1..1
11	URL	정보에 접근하기 위한 Access URL	1..1

## V. 결론

본 연구에서는 국가 R&D정보의 각 연구관리전문기관별 국가 R&D 과제관리번호의 관리 현황 및 범부처 과제고유번호의 구분구조를 살펴보고, 식별메타데이터 개발시에 고려해야 할 기본 원칙을 조사하였다.

또한, 국가 R&D정보의 공동 활용 및 정보 연계를 위해 필수적인 식별체계인 범부처 과제고유번호 및 정보식별자를 위한 식별메타데이터를 개발하기 위한 기본 원칙을 수립하고, 그에 따른 국가 R&D정보 식별메타데이터 요소를 제안하였다.

국가 R&D정보의 정보식별자 부여 대상 정보는 매우 다양한 속성을 가지고 있으며 향후에는 기술·산업정보 및 사실정보에 대해서도 확대 적용될 것으로 전망한다. 따라서, 식별체계 부여를 위한 식별메타데이터 요소 뿐만 아니라 다양한 R&D정보의 특성을 반영한 응용메타데이터의 국제표준과 동향을 연구하고 응용메타데이터 요소에 대한 개발과 연구도 필요하다.

### ■ 참고 문헌 ■

- [1] 권이남, 김재수, 신동구, 정택영 “식별체계기반 국가R&D 정보의 참조연계방안에 관한 연구”, 2006 한국전자거래학회 추계학술대회, pp.135-142, 2006
- [2] 김재수, 권이남, 김영철, “메타데이터 기반 과학기술정보 참조연계에 관한 연구”, 2006 모바일학회 추계학술대회, pp.279-284, 2006

- [3] 김재수, 권이남 외, "디지털콘텐츠 식별체계 응용기술," 한국과학기술정보연구원, 2003.12.
- [4] 한국정보사회진흥원(NIA), "UCI 명세서 V2.1", 2006
- [5] 국가연구개발사업의 관리 등에 관한 규정 [일부개정 2007.6.29 대통령령 제20137호] 제7조 4항
- [6] 국가과학기술종합정보시스템(NTIS) "<http://www.ntis.go.kr>" [cited 2006.12.20]
- [7] R. Moats, "URN Syntax", May 1997.  
<http://www.ietf.org/rfc/rfc2141.txt>
- [8] Dublin Core Metadata Initiative "<http://www.dublincore.org>" [cited 2007.10.10]