

ORCID 기반의 학술 연구 결과물 저자명 식별 시스템 구축 방안에 관한 연구*

A Study on the Construction Methods for Author Identification System of Research Outcome based on ORCID

조재인 (Jane Cho)**

초록

전통적인 전거 통제는 단위 시스템에 한정될 뿐 아니라, 그 대상도 도서관 소장 자원으로 국한되어 있다. 그러나 학술 논문은 인터넷 환경에서 상업 출판사, 기관 레포지토리 등으로 자유롭게 유통되고 있어, 시스템을 넘어선 글로벌한 개념의 이름 식별 체계가 필요하다. 최근 연구자 이름 식별 체계가 도서관계, 출판계를 망라한 글로벌 링크를 지향하게 되면서 ORCID가 출범하였다. ORCID는 학술지 출판사들이 그동안 기계적 클러스터를 통해 운영해 왔던 각종 ID 시스템, 대학의 연구 업적 시스템, 학술지 논문 투고 시스템 등을 파트너로 기존의 ID를 투명하게 링크할 뿐 아니라, 저자가 셀프 클레임을 통해 자신의 실적물을 식별할 수 있도록 하는 방식이다. 그러나 해외 출판물을 중심으로 운영되고 있는 ORCID를 국내 저자 이름 식별에 바로 적용하기는 어렵다. 따라서 본 연구에서는 저자명 식별의 글로벌한 동향을 밝히고, ORCID를 심층 분석한 후, 이를 응용한 국내 학술 연구 결과물 저자명 식별 시스템 구축 방향을 제안하였다.

ABSTRACT

Since research articles aren't be fixed in particular system, author name of articles have been handled not by authority control but by unique identifier. While unique identifier has a tendency that moves toward open-transparent global link system, lately ORCID has been initiated. ORCID is not only link diverse existing ID system as a partner, for example, publisher's ID systems, university's research assessment systems, manuscript tracking systems but also let researchers self-claim by themselves to identify their own articles. Though, since ORCID run based on overseas' publications, it is not appropriate to adapt ORCID shortly to Korean researcher's identification. So this study suggests the direction of domestic researcher's identification system by applying ORCID.

키워드: 학술 연구 결과물, 저자, 학술 연구자, 이름 식별, ORCID

Research Outcome, Researcher, Name Disambiguation, ORCID

* 본 연구는 2012 인천대학교 교내학술연구지원비로 수행됨.

** 인천대학교 문헌정보학과 조교수(chojane123@naver.com)

논문접수일자 : 2013년 2월 13일 논문심사일자 : 2013년 2월 26일 게재확정일자 : 2013년 3월 11일

1. 서론

1.1 연구의 배경

학술 논문은 연구자들의 커뮤니케이션을 촉진하고 연구 결과의 인용을 통해 지식을 선순환하는 기본 기능 이외에도 특정 분야의 지식 형성에 기여한 공헌자나 공헌 기관을 파악하고 평가하는데 있어 주요 도구로 활용되고 있다. 그러나 학술 논문의 이름 표기에는 동일 인물의 다양한 명칭 문제, 그리고 같은 이름 표기를 갖는 동명이인의 문제가 존재할 수 있다. 또한 내국인이 국외 저널에 논문을 출판할 때 발생하는 로마자, 약자 표기 방식에 의해서도 다양한 이름 식별의 문제가 발생할 수 있다. 이 문제는 연구 성과를 출판하거나 서비스하는 기관, 그리고 이를 관리하거나 평가하는 기관뿐 아니라, 연구자 자신과 이용자에게 혼란을 유발할 수 있게 된다.

따라서 영미권 출판계, 도서관계 등에서는 모호한 이름 식별이라는 의미를 갖는 “Name Disambiguation”을 매우 중요한 과제로 인식하고 있다. 일본에서도 “저자 명기(著者の名寄)”라는 용어로 명명하면서, 다양한 각도에서 논의하고 있다(藏川圭 2011). 국내에서는 관련 개념과 기법(김현정 2011)이 소개된 바 있으나, 이 과제를 제대로 명명할 수 있는 용어가 아직 통용되지 않고 있으며, 이 분야의 연구도 저자 식별 알고리즘 개발(강인수 2008)이나 전거 통계의 연장선상(이용호 2000; 이석형, 박승진 2010)에서 이루어지고 있다.

한편, 학술지 출판사들은 모호한 저자 식별 문제에 확률적으로 동일 저자를 판별하는 클러

스터링 알고리즘과 저자 자신의 직접 선언을 유도하는 셀프 클레임 방식을 병용해 왔다. 그러나 최근 다양한 학술 결과물이 웹을 통해 자유롭게 유통되고 레포지토리 등을 통해 상호 연계되기 시작하면서, 단위 시스템에 제한된 이름 통제방식은 큰 의미가 없음을 공감대를 형성하게 되었다. 따라서 시스템, 유틸리티, 도메인을 벗어난 글로벌한 이름 관리 체계의 필요성을 절감하게 되었고, 이러한 필요성은 기존 ID 시스템들을 투명하게 링크할 수 있는 글로벌 이름 식별 체계의 출범으로 이끌게 되었다.

본 연구는 최근 시스템을 초월한 글로벌 연구자 식별 체계로 제안된 ORCID(Open Researcher and Contributor ID) 개념을 심층 분석하고 이런 조류에 있어, 국내 학술 연구 결과물 저자명 식별 체계가 어떤 지평을 향해 나아가야 할지에 대한 새로운 측면의 담론을 제시한다.

1.2 연구의 목적과 내용

이름 식별은 도서관 소장 자원을 위주로 한 전거 통계, 출판계의 콘텐츠 창조자 식별, 그리고 학술 연구 결과물의 저자 식별을 망라하는 개념이다. 본 연구는 그 중 학술 연구 결과물의 저자에 집중하여 ORCID를 응용한 국내 저자명 식별 시스템 제안을 목적으로 한다. 본 연구의 목적과 내용을 구체적으로 기술하면 다음과 같다.

첫 번째, 도서관계, 학술지 출판계, 비영리 정보 유통 기관의 기존 저자 식별 체계에 대해 살펴본다. 전통적 전거 통계 방식과 전거 통계 유사 활동을 연계하는 협력적 허브, 그리고 상업적 학술지 출판사와 비영리 정보 유통 기관을 중심으로 개발된 연구자 식별 도구를 분석하여,

이들 기존 식별 체계가 가진 문제점을 종합해 본다.

두 번째, 현존하는 복수의 저자 ID 체계 간 링크 매카니즘을 구축하는 ORCID의 원칙과 이름 식별을 위한 기본 매카니즘을 분석해 본다.

세 번째, 국내 학술연구자를 위한 저자명 식별 체계가 이러한 조류에 있어 어떠한 방향성을 가질 수 있는지 논의해 보고, ORCID를 기반으로 한 국내 연구 결과물 저자명 식별 시스템 구축 방향을 제안해 보도록 한다.

2. 연구자 이름 식별의 복잡성과 중요성

영국의 ETOC(Electronic Table of Contents) 데이터베이스를 중심으로 저자 표기 현황을 분석한 사례를 살펴보면, 다음과 같이 저자명 식별 문제의 복잡성을 엿볼 수 있다. The Names Project에서 ETOC 데이터베이스를 대상으로 "Birtwhistle"이 저자명으로 등록된 아티클 140건을 추출하고 그 중 독립된 실제 저자를 판별하였다. 그 결과 140건의 아티클 중에 41명의 각기 다른 저자가 존재하였고, "Birtwhistle, G"로 저자명을 축소한다고 해도 총 7개의 각기 다른 저자가 존재한다고 밝혔다(Danskin, Hill, & Needham 2011). 요컨대, 학술 논문 데이터베이스에는 같은 표기의 이름을 사용하는 다수의 다른 연구자들이 존재하고 있어, 이름만으로 이들을 명확히 식별하고 이들과 이들의 연구 실적물을 정확하게 연계하기 어렵다는 것이다.

한편, Strotmann과 Zhao(2012)는 연구의 글로벌화가 저자명 식별 문제를 복잡하게 만드

는 주요 원인 중 하나가 되고 있다고 논평하였다. 특히 한국인은 동일한 성씨를 사용하는 경우가 빈번함에도 불구하고, "Kim, H."와 같이 이름의 이니셜 표기가 일반화 되어 있어, 동일 성명으로 표기되는 다수의 다른 연구자가 발생하고 있다는 것이다(김현정 2011). 국내 연구자의 해외 학술지 논문 게재 편수가 매해 증가하는 상황에서 김씨, 이씨, 박씨의 3개 성씨가 전체 인구의 44.9%를 이루는 국내 저자의 성명 문제는 위에서 지적한 바와 같이 단지, 해외 논문에 게재된 국내 연구자 식별의 복잡성뿐만 아니라, "Name Disambiguation"이라는 글로벌 과제의 부담감까지 가중시키고 있다.

이름 표기에는 같은 이름 표기를 갖는 동명이인의 문제뿐 아니라, 동일 인물의 다양한 명칭 문제가 존재할 수 있다. 그렇기 때문에 동명의 다른 사람을 명확하게 다른 것으로 식별하고, 동일인의 다양한 이름 표기 형태를 하나로 연결시키는 것은 연구자와 그들의 저작물들을 명확하게 식별하기 위하여 반드시 필요하다. 그 밖에도 여러 가지 의미를 가질 수 있는데, 그 중요성을 다음과 같이 정리해 볼 수 있겠다. 첫 번째, 특정 분야 학술 성과의 기여자를 정확하게 판별하고 영향력을 파악하는데 필요하다. 두 번째, 출판사가 논문심사 및 출판 프로세스에서 연구자를 효과적으로 관리하고, 연구지원 기관이 연구비 응모 과정과 수혜 결과에 대한 평가를 효율화하는데 필요하다. 세 번째, 정보 이용자가 특정 분야에 성과를 낸 연구자의 연구 경향을 파악하거나 연구자가 공저자를 모색할 때에도 필요하겠다.

최근 전자출판물이 급증하고 기관 레포지토리가 확대되면서, 학술논문 저자의 다양한 이름

변형이 발생하게 되었다. 연구자의 셀프 아카이빙 도구인 기관 레포지토리는 사서가 전거 통제를 통해 저자명을 표준적으로 입력하는 전통적 프로세스의 적용 대상이 아니기 때문이다. 더구나 웹상에서 메타데이터가 하베스팅되어 유통되고 상호간 교환·통합되면서, 이름 식별의 혼란이 가중될 수밖에 없게 되었다. 이러한 조류에 의해 도서관계는 기존의 전거 통제를 넘어선 글로벌한 식별 체계에 대한 필요성을 절감하게 되었고, 이는 출판계를 비롯한 정보 유통 분야에 있어, 매우 중요한 공통의 이슈가 되고 있다.

영국의 경우, 도서관의 전거 통제 대상이 되는 신간 도서가 연간 13만건인데 반해, ETOC에 추가되는 저널의 아티클 수는 250만건이나 된다고 한다(Danskin, Hill, & Needham 2011). 요컨대 이는 저자 식별이 필요한 연구자의 숫자가 20배 가까이 많다는 것으로, 정보조직 분야에 있어 전거 통제 방식이 손에 닿지 않는 영역이 얼마나 더 크고 넓은 지를 함의하고 있다.

3. 각종 이름 식별 도구의 분석

지금까지 이름 식별에 대한 노력은 국가나 지역, 주제 분야, 상업 출판사, 도서관 등에서 독자적으로 이루어져 왔다. 본 장에서는 먼저 도서관계에서 그동안 이루어진 전통적 방식의 전거 통제와 도서관 이외의 다른 커뮤니티에서 이루어진 전거 통제 유사 활동을 연계하는 협력적 노력에 대하여 살펴본다. 더불어 범위를 구체화시켜, 학술지 출판사, 국가 단위 비영리 학술정보 유통 기관의 연구자 식별 도구에 대해 살펴보고 이들의 문제점을 종합해 보도록 한다.

3.1 도서관의 전거 통제와 유사 활동을 연계하는 협력적 노력

전거 통제는 도서관에서 소장자원의 식별기능, 배열기능, 집중기능을 실현하기 위하여 존재해 왔다. 이는 전거가 되는 이름, 주제, 서명 등을 서지자료의 표목으로 일관성 있게 채택하도록 한 방식을 의미한다. 이론상으로는 통일표목을 연결 고리로 하지만, 실제적으로는 시스템상에서 전거데이터가 서지데이터와 링크되는 방식으로 운영되고 있다. Watanabe(2010)는 전통적인 도서관 전거 통제의 한계를 다음과 같은 측면에서 통찰하고 있어 주목된다.

첫 번째는 통일 표목의 상호 운영성 문제이다. 통일 표목을 일개 기관에서 신중하게 결정하여 채택하여도 단위 시스템을 넘어서면 상호 운영이 보장되지 않는다. 두 번째는 비용 효과성 문제이다. 전거를 유지하는 데에는 막대한 비용이 소요되지만, OPAC 환경에서 이용자가 제목, 저자 등을 고루 검색하면, 어느 정도 전거 통제와 비슷한 수준의 검색 효과를 달성할 수 있게 된다. 세 번째는 대상으로 하는 자료의 실체 문제이다. 디지털 환경의 진전에 따라 콘텐츠, 케리어의 양면에서 동일 저작의 다양한 변형이 증가하고 있다. 따라서 전거 통제가 목표로 하는 이름 식별의 실현뿐 아니라, 동일 콘텐츠가 담기는 다양한 매체를 어떻게 식별할 것인가도 중요한 의미를 가지게 되었다. 네 번째로는 대상으로 하는 자료의 범위 문제이다. 네트워크 자원이 급증하면서 도서관은 소장 자원뿐 아니라, 구축 자원, 구독 자원 등 도서관 내외부에 있는 다양한 자원에 대한 일원화된 접근점 제시를 요구받게 되었다. 그러나 전통적 전거

통제는 단지 도서관 소장 자원만을 대상으로 하고 있어 매우 제한적인 범위를 가진다는 것이다.

이러한 이유에서 최근 도서관계는 전통적인 전거 통제 시스템의 한계에 대해 성찰하고 새로운 전거 통제 모델을 만들어 가기 위해 많은 노력을 기울이고 있다. 특히 RDA(Resource Description and Access)와 FRBR(Functional Requirements for Bibliographic Records) 개념을 통해, 전거 통제 방식에 커다란 변화를 시도하고 있다. IFLA(International Federation of Library Associations)에서 제안한 전거데이터 기능상의 요건(FRAD: Functional Requirements for Authority Data)은 FRBR의 주요 개념들을 전거 데이터의 작성과 활용을 위해 적용시킨 개념이다. 여기에서는 전거데이터 참조를 위한 구조화된 프레임워크를 제공하고 있는데, 구현 형태와 개별 자료 이외의 모든 실체가 전거 통제의 대상이 되어야 한다고 정의하고 있다. 따라서 개체인 개인, 단체, 저작, 표현형, 구현형, 개별자료, 개념, 대상, 사건, 장소, 가족이 모두 전거 통제의 대상이 되며, 서지데이터에 통일표목을 부여하는 기존의 연계 방식과 달리, 특정 개념의 관계 설정에 의해 서지데이터와 전거데이터가 연계된다.

그 밖에 도서관계는 전통적인 전거 시스템을 현재 기술 환경에 맞게 개선하기 위하여 많은 노력을 기울이고 있다. 특히, 다른 커뮤니티와의 상호 운영성을 제고하기 위한 데이터 모델 개발과 협력적 허브 구축 방안을 통해 새로운 패러다임을 찾아가고 있다. 그 내용을 요약하면 다음과 같다. 첫 번째, 미국 의회 도서관(LC: Library of Congress)은 사람, 기관, 이벤트, 장소, 통일표목에 대한 800만 건 이상의 이름 전거 데이터를 수록한 “Library of Congress Name Authority File”을 Linked Data 형식으로 다운로드할 수 있는 “Authorities and Vocabularies” 서비스를 공개하였다. 두 번째, OCLC(Online Computer Library Center)는 NETWORK NAMES¹⁾라는 협력적 허브를 통해 동명이인의 문제, 동일인의 표기가 언어별, 국가별로 다양하게 전개되는 문제를 해결하고, 도서관, 기록관, 박물관, 기관 레포지토리 등에서 범용적으로 활용할 수 있는 협력적 전거 통제 시스템을 제안하였다. 세 번째, LC와 OCLC는 전거데이터의 국제적 교환을 위해 VIAF(Virtual International Authority File)²⁾를 추진하고, 콘텐츠 산업계의 저자 표준인 ISNI(The International Standard Name Identifier)에도 기초 데이터로 제공하였다(Library of Congress 2011).

- 1) 각 기관이 보유한 전거 정보를 연결하고 수집하는 총체적 프레임 워크라고 말할 수 있겠다. 이 허브는 도서관, 박물관, 기록관 등의 숨겨진 모든 전거 관련 엔티티를 결합하며, 우선 표목 없이 사용할 수 있는 게이트웨이를 제공한다. 또한 소셜 네트워크 모델로 확장하여 저자가 스스로 자신의 정보를 생성하거나 삭제할 수 있으며, 오픈소스를 기반으로 다양한 소프트웨어 어플리케이션과 연동될 수 있도록 제안되었다. 이렇게 함으로써 ① 언어나 분야에 상관없이 동일 엔티티에 대한 식별을 가능하도록 하며, ② 고유한 맥락에 따라 우선 표목을 결정할 수 있고, ③ 참여 기관이 그들의 고유한 네임 데이터 리소스를 확충할 수 있도록 지원할 수 있다. 이 프로젝트에서 제시된 내용은 영국의 Names 프로젝트나, NISO Institutional Identifiers의 레퍼런스 모델이 되고 있다.
- 2) 도서관계는 전거데이터의 국제적 교환을 촉진하고 전거 생성과 유지를 단순화하며, 이용자가 선호하는 언어, 문자 형식으로 정보를 접근시키기 위해 이 프로젝트를 추진하였다. 각국에서 대표 전거를 만들고 그것을 교환하거나 통합하는 방식으로 추진된 국제 가상 전거 파일 프로젝트에서는 총 15개 기관의 저자이름 파일을 결합한 후, 통일표목을 정하지 않고 각국에서 만들어진 통일 표목을 남겨 상호 링크시키는 방식을 지향하고 있다.

3.2 상업적 학술지 출판사의 이름 식별 도구

이름 식별 영역에서 학술 논문의 저자 식별은 앞에서 언급한 바와 같이, 부분적 이슈이긴 하지만, 숫자적으로 매우 방대하다고 말할 수 있다. <표 1>과 같이 상업적 학술지 출판계, 비영리 학술정보 유통 기관을 중심으로 그 움직임을 보이고 있다.

먼저, 상업적 학술지 출판계의 연구자 식별 체계를 살펴보도록 한다. 기존 영리 출판사들은 저자 식별 문제에 기계적 클러스터링과 저자의 셀프 클레임 방식을 병용해 왔다.

기계적 클러스터링 방식의 대략을 살펴보면 다음과 같다. 먼저, 논문의 서지 집합에 알고리즘을 기반으로 한 클러스터링 기술을 이용하여 저자별로 분류한다. 그 다음 성명, 소속, 공저자, 분야 등 다양한 조건을 통하여 확률적으로 논문 서지에 명시된 저자와 다른 논문 서지에 명시된 저자가 동일 저자임을 판정하는 것이다. Thomson Reuters Web of Science의 “Distinct Author Identification System”, Elsevier Scopus

의 “Scopus Author Identifiers”가 대표적인 저자 식별 체계로 손꼽힌다. 그러나 기계적인 처리에 의존하는 이와 같은 방식으로는 신뢰할 만한 수준의 정확도를 제공하지 못한다는 평가가 있다. 따라서 이용자의 예러 통보나 다음과 같은 보조 도구들에 의해 보완시켜 나가고 있다. Web of Science는 특정 연구자의 성과물들을 모두 트랙하는 보조 도구로 “Author Finder”를 제공하고 있는데, 다음과 같은 방식으로 검색된 저자를 축소시켜 나가고 있다. 먼저, 이용자로부터 성, 명, 미들 이니셜을 입력하도록 유도한다. 그러면 ‘Weinstein B’, ‘Weinstein B*’, ‘Weinstein BA*’과 같이 다양한 수준의 레코드에 대한 검색 결과를 제시해 준다. 그 다음 이용자가 적절한 셋을 선택하면, 주제 분야 리스트를 이용하여 순차적으로 그 결과를 축소시켜 나간다. 마지막으로 소속 기관 리스트를 제공함으로써 이용자가 최종적으로 원하는 연구자를 찾을 수 있도록 지원하는 것이다(Wagner 2009).

한편, 위와 같은 방법 이외에도 연구자가 자신의 성과물을 스스로 선언하는 셀프 클레임 방

<표 1> 대표적인 학술연구자 식별 시스템(Fenner 2011a)

이름	조직	특징	
Scopus Author ID	엘스비어	출판사 스코퍼스 데이터 기반 식별 체계	
Research ID	톰슨 로이터		웹어브사이언스 기반 식별 체계
LATTES	National Council for Scientific and Technological Development	국가적 비영리 학술 정보 유통 기관 브라질의 연구자 데이터베이스 기반 식별 체계	
Digital Author Identifier	Royal Netherlands Academy of Arts and Science		네덜란드 출판물 및 연구 프로젝트 데이터베이스 기반 식별 체계
ArXiv Author ID	코넬 대학		ePRINT 아카이브 기반 식별 체계
PubMed Author ID	NLM		의학 데이터베이스 기반 식별 체계
Names Project	Mimas, British Library		영국 연구자, 기관에 대한 식별 체계
연구자 resolver	NII		일본 기관 레포지토리 기반 식별 체계

식이 존재한다. 이는 연구자가 자신의 ID를 등록하고, 자신의 실적물을 직접 찾아 선언하는 방식으로 2008년 1월 시작된 Thomson Reuters의 연구자 디렉토리 “ResearcherID”가 대표적이라고 말할 수 있겠다. 구체적으로 설명하면 연구자에게 개개의 고유 식별자와 영구 URL을 부여한 후, “Web of Knowledge” 등을 통해 자신의 실적물을 선별하여, “이것은 내 것이다”라고 선언하도록 유도하는 방식이다. 따라서 정확하게 연구자와 그들의 실적물을 연계시킬 수 있으나, 충분한 정도의 참여 규모를 얻지 못하고 있다.

3.3 국가 단위 비영리 학술정보 유통 기관의 이름 식별 도구

출판사들은 위와 같이 각기 자신들의 영리 시스템에 적용할 독자적인 연구자 식별 도구 마련에 주력하는 반면, 비영리 학술정보 유통 기관에서는 국가적 차원의 표준화된 이름 식별 체계 마련에 관심을 기울이고 있다.

언급한 바와 같이 레포지토리에 등록된 콘텐츠에는 이니셜 표기나, 성명의 순서 등 이름과 기관명 표기에 일관성이 부재한데, 기관 레포지토리에 등록되는 콘텐츠 수가 증가하고 레포지토리간 메타데이터 공유가 활성화되면서, 더욱 심각한 문제가 되고 있기 때문이다. 따라서 국가 단위 비영리 학술정보 유통 기관에서는 고유 ID 기반의 식별 체계를 구축하고 이를 레포지토리 간에 표준화시킴으로서 이 문제를 해결해 나가고 있다.

이러한 배경으로 Mimas는 영국국립도서관(BL)과 공동으로 레포지토리 전용 명칭 전거 프

젝트 ‘Names(<http://names.mimas.ac.uk>)’를 추진하고 있다(Danskin, Hill, & Needham 2011). 이 프로젝트는 기본 식별키로 전거형을 사용하는 것이 아니라 번호 기반의 식별 체계를 사용함으로써, 전거 통제가 아닌 접근 제어를 지향하고 있음을 강조하고 있다. 데이터 셋을 구축하기 위해서는 다음과 같은 두 가지 방법을 채택하고 있다. 첫 번째는 개인이나 기관과 관련된 외부의 데이터 소스를 반입하여 자동적으로 동일인을 클러스터링하는 방법이며, 두 번째는 저자의 셀프 클레임과 같이 연구자가 스스로 자신의 정보를 제공하도록 하는 방식이다. 그러나 ETOC에 추가되는 수많은 저널 아티클량에 의해 기계적 클러스터링에 의존하는 첫 번째 방식이 주로 사용될 수밖에 없다고 말하고 있다. 데이터 소스로는 ZETOC(British Library’s table of contents)의 3,800만건, JISC(Joint Information Systems Committee)가 펀드한 연구 결과물인 MERIT(Making Excellent Research Influential and Transferable)의 45,000건이 활용되고 있다.

일본 역시 학술 논문 저자 식별을 위하여 연구자 리졸버(研究者リゾルバー, <http://rns.nii.ac.jp/html/help.html>)를 구축하고 있다(藏川圭 2011). 이는 일본국립정보학연구소(NII: National Institute of Informatics)가 과학 연구 비 보조금 데이터베이스 KAKEN(Database of Grants-in-Aid for Scientific Research)과 ReaD(JST ReaD Researcher code)의 연구자 번호를 기반으로 자국내 연구자들에게 고유한 ID와 URI를 부여하는 프로젝트이다. 2011년 현재 192,717명의 연구자가 등록되어 있으며, 226개 대학을 대상으로 부여된 ID는 개별 대학의

연구 업적 관리시스템, 그리고 기관 레포지토리 시스템과도 싱크(Sync)를 유지하도록 유도되고 있다.

그 밖에 네덜란드 SURF(www.surf.nl) 재단도 저자 고유 ID인 DAI(Digital Author Identifier)를 연구자들에게 할당하는 프로젝트를 운영하고 있다.

이렇게 국가 단위 학술정보 유통 기관은 대용량 데이터 소스를 기반으로 기계적인 클러스터링을 수행하고 여기에 고유 ID 기반의 식별 기호를 부여함으로써, 국가적 차원의 표준화를 지향하고 있다. 그러나 클러스터링 자체의 정확성, 글로벌한 상호 연계성, 상업적 출판사의 기존 ID 시스템간 호환성 등 여러 가지 측면의 숙제가 아직 남겨져 있다.

3.4 분석 내용의 종합 및 문제의 발견

도서관, 상업출판사, 그리고 국가 단위 비영리 정보 유통 기관은 이름 식별 문제에 있어 각각의 탐구 방식을 찾아가고 있지만, 아직 공통의 기반을 둔 짜임 관계 속에 녹아들지 않고 있다. 각각의 탐구 방식에 대한 주요 내용을 다음과 같

이 요약하고 문제점을 발견해 볼 수 있겠다.

첫 번째, 도서관 소장 자원을 중심으로 한 전통적 전거 통제는 대상의 제한성, 상호 운용성의 부재, 비용 효과성 등에 의해 지속적으로 문제가 제기되어 왔다. 최근에는 네트워크 자원이 급증하면서, 통제의 범위를 확장하거나 다른 도메인과 상호 연계할 수 있는 글로벌 식별 체계의 모색이 시급한 요구 사항이 되고 있다.

두 번째, 학술 논문 저자의 이름 식별 체계는 상업적 학술지 출판사에 의해 독자적인 도구로 개발되고 있다. 상업적 학술지 출판사들은 클러스터링 알고리즘에 의해 유일 저자를 식별하고 셀프 클레임 방식을 도입하여 이를 보완하고 있으나, 단위 시스템 내 유일 저자를 식별하는 이 방식은 웹 환경에서 망라성을 보장하지 못하고 있다.

세 번째, 국가 단위 비영리 정보 유통 기관은 연구인력 DB, 연구비 수혜 결과물 DB 등의 소스를 활용한 연구자 식별 방식을 채택하고 있으며, 이를 기반으로 한 국가 표준 식별 번호 체계 마련에 관심을 기울이고 있다. 그러나 이는 상업적 출판사의 기존 ID 시스템과의 호환, 글로벌한 상호 연계성 측면에서 아직 완전하지 못하다.

〈표 2〉 주요 이름 식별 체계 종합 및 제한점

주체	주요 대상	이름 식별 방식	제한점
도서관	도서관 소장자원	- 전거통제	- 대상의 제한성 - 상호 운용성 - 비용
영리 학술지 출판계	학술 논문	- 기계적 클러스터링 및 보조 도구 - 셀프 클레임	- 부정확성 - 상호 운용성
국가 단위 비영리 정보 유통 기관	연구비수혜과제, 연구인력 DB, 학술 논문	- 기계적 클러스터링 - 고유 식별 기호	- 부정확성 - 상업출판사 기존 ID 시스템과의 호환성 - 글로벌 상호 운용성

4. ORCID 기반의 국내 학술 논문 저자명 식별 체계 구축 방향

앞에서 통찰한 바와 같이 기존 학술 연구 결과물 저자명 식별 도구는 기계적 처리 방식에 의존한 부정확성, 글로벌 상호 연계성 부재가 주요 문제로 지적될 수 있겠다. 본 장에서는 새로운 연구자 식별 모형으로 최근 제안된 ORCID를 심층 분석해 보고, 이를 기반으로 국내 연구자에게 적용 가능한 ORCID 기반의 학술 연구 결과물 저자명 식별 시스템 구축 방향을 제시해 보도록 한다.

4.1 ORCID 분석

2012년 말 학술지 출판계와 정보 유통 기관은 그간의 모든 노력과 경험을 망라하여, 글로벌 연구자 식별 체계인 ORCID 레지스트리를 출범하게 되었다. 본 장에서는 그 설립 취지와 의의, 구현 방식을 자세히 분석해 본다.

4.1.1 설립 취지 및 의의

ORCID(Open Researcher and Contributor ID)는 학술 정보 유통에 관련된 다양한 이해 관계자가 모여 연구자 식별의 문제를 해결하기 위하여 조성된 국제적인 조직이다. 홈페이지(ORCID Homepage 2013)에 게시된 설립 취지는 아래와 같이 요약할 수 있다.

ORCID는 학술 커뮤니케이션에서 저자 이름 모호성 문제를 해결하기 위하여 개별 연구자에게

고유 식별자를 부여하는 체계이다. 중앙 레지스트리, ORCID, 현존하는 다른 저자 ID 체계간 투명한 링크 매커니즘을 구축하여, 과학적 발견 과정을 확대하고 연구 커뮤니티 협력을 촉진한다.

광범위한 도메인을 대상으로 하는 유니크한 식별자인 ORCID는 다음과 같은 의의를 가진다. 첫 번째, 연구자는 본인의 연구 업적 및 프로필을 편리하게 관리할 수 있으며, 공동 연구자를 효과적으로 탐색해 낼 수 있다. 두 번째, 출판사는 이용자에게 특정 저자의 다른 논문들을 편리하게 추적하여 열람할 수 있는 인터페이스를 제공할 수 있다. 세 번째, 연구 기관은 소속 연구자들의 업적물 관리 및 평가 체계를 효율화할 수 있다. 네 번째, 학회나 출판사는 논문 게재를 신청하는 연구자의 레포팅 부담을 줄일 수 있을 뿐 아니라, 프로세스 전반의 관리 효율성을 제고할 수 있다.

4.1.2 원칙

ORCID는 학술 커뮤니케이션에 관심이 있는 모든 조직에게 열린, 어디에도 적합한 전방위 시스템을 지향한다. 다시 말해 오픈되어 있으며, 무료로 프로필을 생성하고, 편집, 유지할 수 있다. ORCID에 입력된 데이터는 CCO Public Domain Dedication³⁾에 의해 무료로 다운로드 가능하며, 오픈소스에 의해 운영되는 것을 원칙으로 한다.

공개된 기본 원칙에는 다음과 같은 10가지 사항이 명시되어 있다(ORCID Homepage 2013).

3) CC0 1.0 Universal(CC0 1.0) Public Domain Dedication Creative Commons. [cited 2012.1.14].
 <<http://creativecommons.org/publicdomain/zero/1.0/>>.

첫 번째, ORCID는 신뢰성 있게 저자와 기여자를 찾을 수 있도록 지원함으로써 학술 커뮤니케이션을 활성화시킨다. 두 번째, 분야, 지리, 국적, 기관의 경계를 넘어선 투명함을 지향한다. 세 번째, 참여자는 학술 커뮤니케이션에 관심이 있는 어떠한 조직에게도 개방된다. 네 번째, ORCID 서비스는 웹을 통해 투명하고 차별 없는 용어로 제공된다. 다섯 번째, 연구자는 ORCID를 통해 자유롭게 ID와 프로필을 등록하며, 개인정보는 충분히 보호된다. 여섯 번째, 연구자의 프로필 데이터는 프라이버시 설정을 통해 제어될 수 있다. 일곱 번째, 정보는 카피레프트의 개념을 가지는 CCO 라이선스 형태로 공개된다. 여덟 번째, ORCID가 개발한 소프트웨어는 오픈소스로 공중에 배포된다. 아홉 번째, 비영리이면서 지속 가능한 비즈니스 모델을 추구하기 때문에 최소한의 수익을 얻는 것을 목적으로 한다. 열 번째, 조직 내부의 구성은 최대한 투명성을 확보한다.

4.1.3 참여 기관

ORCID는 출판사, 학회, 연구지원기관, 대학, 연구소 등이 참여하게 된다. 출범 이전 단계부터 300개 이상의 단체가 지원을 약속하였고 그 중 50개 이상이 재정 지원을 하고 있다. 학술 연구와 관련된 다양한 이해 관계자로 구성되어 있지만, 대학과 학술저널 출판사가 주류를 이루며, 미국과 영국이 압도적이고, 아시아 국가의 참가는 아직 미미하다. 2012년 9월 시작된 파트너 시스템 테스트에는 “CrossRef”, “Elsevier”, “Thomson Reuters”, “미국물리학회”, “보스톤 의과대학”, “웰컴 트러스트(Wellcome Trust)” 등 주요 출판사와 연구 기관이 참여하였으며,

최근 2013년에는 “미국 국립보건원(NIH)”이 합류하였다.

4.1.4 구성 요소

ORCID의 핵심 구성 요소는 최종사용자, 파트너시스템, ORCID의 Identity System이다. 최종 사용자는 저자, 공저자, 연구 관리자 등이 될 수 있으며, 파트너 시스템은 출판사의 전자 저널 시스템이나 논문 투고 시스템, 대학의 연구 업적 시스템 및 각종 학술연구정보 포털의 논문 검색 시스템 등이 될 수 있다.

중앙 레지스트리라고 할 수 있는 ORCID Identity System은 연구자가 자신의 계정을 등록하고 자신의 연구 업적을 선언할 수 있도록 지원한다. 뿐만 아니라, 현존하는 다른 ID 체계와 연계되어 상호간 연구자와 그들의 실적 정보를 공유할 수 있도록 한다. 이를 위해 여러 파트너 시스템에서 생산된 프로필을 수집하여 통합하고 중복을 제거하는 역할을 수행해야 한다. 한편, 파트너 시스템은 ORCID Identity System에 프로필을 업로드하거나 다운로드함으로써 데이터를 공유하는 역할을 수행한다. ORCID Identifier 레졸루션 API를 파트너 시스템이 탑재하면, ORCID의 Identity System과 효과적인 프로필 교환이 가능하다(Ratner 2011b).

4.1.5 이름 식별 매카니즘

ORCID 이름 식별 매카니즘의 핵심은 중앙 레지스트리를 통해 연구자가 스스로 자신의 실적물을 선언하고 이를 현존하는 ID 체계간에 투명하게 링크시키는 것이다. 따라서 ORCID 시스템에서 가장 중요한 콘텐츠는 저자의 ORCID

와 연구 실적정보라고 말할 수 있겠다. ORCID 식별자는 체크섬을 포함하여 4개의 그룹으로 구분된 16자리 숫자로 구성되며, 다음과 같이 HTTP URI(<http://orcid.org/0137-1963-7699-2319>)로 표현된다. 또한 ORCID는 ISNI ISO 표준과 호환이 가능한 형식을 사용하고 있다. 한편, 연구 실적의 등록은 저자가 직접 수행할 뿐 아니라, 저자의 소속 기관이 대행하는 하이브리드 방식을 병용하게 된다. 다시 말해, 연구자 자신이 자유롭게 셀프 클레임에 참여할 수 있으나, 이 방식만으로는 많은 수의 연구자 확보가 어렵기 때문에, 저자가 소속된 기관 단위로 외부 클레임 방식이 일괄 수행될 수 있다는 것이다(Fenner 2011b).

클레임 방식에 대하여 조금 더 구체적으로 설명하면 다음과 같다. 첫 번째, 연구자들은 ORCID 서비스를 통해 자유롭게 ID와 프로필을 등록한다. 두 번째, 연구자의 성과 목록은 ORCID의 바깥 세상에 분산되어 있기 때문에 연구자는 스스로 그것들을 검색하고 확인한 후, “이것은 내 것이다” 또는 “이 저작물의 저자가 바로 나다”라는 선언을 통해 자기 명기(self disambiguation)하게 된다. 이용자는 ORCID의 Identity System과 연계된 Elsevier의 “Scopus” 서지 데이터를 검색하여 자신의 업적 정보를 찾아올 수 있을 뿐 아니라, 개인 블로그 등 다양한 콘텐츠에 대한 신고와 외부에 생성되어 있는 프로필 데이터를 반입할 수 있다. 언급했다시피 대학이나 학회, 및 연구지원기관이 연구 업적 관리, 연구 논문 심사, 연구 보조금 지급을 구실로 구성원을 ORCID에 일괄 등록할 수도 있다. 특히 대학은 연구 업적 시스템을 통해 구성원의 프로필을 단체로 등록시킬 수 있는데, ORCID

는 많은 수의 연구자를 확보하기 위하여 이 방식을 권유하고 있다.

4.2 ORCID 기반의 국내 학술 연구 결과물 저자명 식별 시스템 구축 방향 제시

현재 국내에서는 영리 전자저널 출판사와 연구과제 관리 기관에서 구축한 몇 개의 연구자 이름 식별 사례를 찾아볼 수 있다. 전자저널 출판사인 누리미디어의 DBPIA, 국가과학기술지식정보서비스(NTIS)의 국가 R&D 참여인력정보 DB와 한국연구재단의 KRI(Korean Researchers Information) 및 KCI(Korea Citation Index) 시스템에서 연구자 식별 서비스가 제공되고 있다. 대표적인 국내 전자저널 서비스인 DBPIA는 자체 서비스를 위한 동일 저자 판별 시스템을 운영하고 있다. 그러나 이메일, 공저 관계 등 기계적 판별에 의존하고 있어, 소속 변경 등이 이루어진 연구자의 아이덴티티를 정확하게 식별해 내지 못하고 있다. 한편, 국가 R&D 참여인력정보 DB는 국가연구개발 사업 참여 인력 중심으로 프로필과 연구 성과가 수작업으로 관리되고 있으며(이석형, 박승진 2012), KRI는 대학의 연구 업적 시스템과 연계 운영되고 있어, 비교적 높은 정확도를 보이고 있다. 그러나 이들은 연구자의 업적 정보를 국가적 차원에서 관리하기 위한 목적으로 개발된 시스템이기 때문에, ORCID와 같이 고유식별번호를 위한 중앙 레지스트리 기능은 제공하지 않는다.

아울러 해외 학술지에 게재된 국내 연구자명의 이름 식별 문제도 관심을 가져 볼 필요가 있다. 내국인 연구자의 해외 학술지 논문 게재율이 지속적으로 증가하고 있어, 교육과학기술부가

집계한 대학 정보 공시 내용에 따르면 2011년 대학 전임교원 1인당 국외 학술지 게재 논문 수는 전년보다 7% 증가한 0.3편으로 집계되었다. 그러나 앞에서 지적한 바와 같이, 해외 학술지에 투고한 국내 연구자는 “Kim”, “Lee”, “Park” 등의 동일성을 사용하는 경우가 다수 존재하며, 이름의 약자 표기가 일반화되어 있어, 아무리 정교한 알고리즘으로 처리된다 할지라도 명확한 이름 식별에 한계가 존재할 수밖에 없다.

따라서 이러한 문제에 대한 통찰과 앞에서 살펴본 글로벌 조류를 종합해 볼 때, 국내 학술 연구 결과물 저자명 식별 시스템 구축 방향에 대해 다음과 같은 공감대를 형성할 수 있을 것이다.

첫 번째, 국내 연구자에게 고유 식별자를 부여하고 이를 관리하기 위한 중앙 레지스트리를 구축해야 한다.

두 번째, 고유 ID와 업적 정보를 학술정보 포털, 상용 전자저널 시스템과 같은 파트너 시스템과 효과적으로 공유하여, 파트너 시스템의 자체 식별 매카니즘과 시스템을 넘어선 링크를 지원해야 한다.

세 번째, 국내 학술연구자를 위한 고유 식별 체계를 ORCID와 효과적으로 연계시켜 해외 학술 논문에 기재된 국내 저자의 명확한 이름 식별을 도모해야 한다.

4.2.1 핵심 구성 요소 및 기능

국내 학술 연구 결과물 저자명 식별 시스템의 핵심 구성 요소는 ORCID와 같이 오픈 소스 기반의 중앙 레지스트리로 개발된 핵심 ID 시스템과 파트너 시스템으로 구성될 수 있다. 핵심 ID 시스템에 연구자의 고유 ID와 연구 업적정보를

구축하고 이 정보를 다양한 파트너 시스템이 갱신하거나 다운로드 받아갈 수 있도록 공개한다. 한편, 핵심 ID 시스템은 <그림 1>에서 보는 바와 같이 ORCID Identity System의 파트너 시스템 중 하나가 되어, 국내 연구자의 실적 정보를 업로드함으로써 글로벌 식별 체계의 일부가 되어야 할 것이다. ORCID는 글로벌 링크 매카니즘으로 제안되었으나, 영미권 학술지 출판사를 중심으로 시작되고 있어, 아직은 아시아 국가의 참여가 어렵다. 따라서 일단 국내에서 핵심 ID 시스템을 기반으로 고유 식별 체계를 운영하고 이 시스템이 추후 ORCID 파트너가 되는 방향으로 추진하는 것이 바람직할 것이다. 핵심 구성 요소와 기능에 대해 구체적으로 설명하면 다음과 같다.

1) 핵심 ID 시스템

핵심 ID 시스템은 연구자가 자신의 프로필을 등록하고 셀프 클레임을 수행하는 기능을 지원해야 한다. 다시 말해 국내 주요 학술논문 메타데이터 DB와 연동하여 연구 업적물을 직접 검색한 후, 이것이 자신의 논문이라고 선언할 수 있는 기능이 제공되어야 한다는 것이다. 또한 파트너 시스템으로부터 프로필을 수집하고 이를 통합하는 역할을 수행할 뿐 아니라, 파트너 시스템이 데이터를 다운로드 받아 자체 식별 시스템에서 활용할 수 있도록 API를 배포해야 할 것이다. 핵심 ID 시스템은 한국연구재단이 전국 대학에 기 배포한 KRI(한국연구 업적통합정보)와 가장 유사한 기능을 가지고 있으므로 이를 보완 발전시켜 구현하는 것이 효과적일 것이다. 한편, 핵심 ID 시스템은 언급한 바와 같이 국내 중앙 레지스트리이자 ORCID의 파트너 시스템

이 되어 추후, 연구 실적 정보를 공유할 수 있어야 할 것이다.

2) 파트너 시스템

파트너 시스템은 핵심 ID 시스템에 프로필을 업로드하거나 다운로드 받음으로써 연구자 정보를 상호 공유하게 된다. 대학의 연구 업적 시스템, JAMS(Journal & Article Management System), ACOMS(Article Contribution Management System)와 같은 논문투고 시스템, RISS(www.riss4u.net), 학회마을(society.kisti.re.kr) 등의 학술논문 서비스, DBPIA(www.dbpia.co.kr)와 같은 상업 전자저널 출판사 등이 파트너가 될 수 있을 것이다. 파트너 시스템은 연구자 고유 ID, 논문리스트, 프로필 정보를 다운로드 받아 자체 식별 매커니즘에 활용하는 한편, 파트너 시스템에 먼저 등록된 연구자의 실적 정보를 핵심 ID 시스템에 업로드함으로써 데이터 갱신에도 기여할 수 있을 것이다.

4.2.2 운영 전략

연구자에게 고유 ID를 부여하고 그들의 실적 정보를 효과적으로 구축하기 위한 방법은 외부 소스를 활용하는 일괄 구축 방식과 연구자 자신의 직접 선언을 유도하는 셀프 클레임 방식이 병용될 수 있을 것이다. 여기에서는 먼저, 이 두 가지 전략에 대하여 설명한다. 그 다음, 핵심 ID 시스템을 중심으로 국내 파트너 시스템간 투명한 링크 체계를 구현하기 위한 몇 가지 전략과 해외 출판 논문의 내국인 저자명 식별을 도모할 수 있는 방안에 대하여 설명하도록 한다.

1) 연구자 고유 ID 부여 및 실적 정보 구축
연구자 고유 ID를 효과적으로 부여하고 관련 데이터를 편리하게 구축하기 위한 전략으로 다음과 같은 2가지 방식을 제안해 볼 수 있겠다.

첫 번째는 대학의 연구 업적 시스템이나 외부의 각종 소스를 활용하여 일괄 구축하는 방법이다. 연구 업적 시스템은 대학마다 존재하므로 이를 이용하여 연구자의 고유 ID를 효과적으로 생성시킬 수 있을 것이다. 대학에서 핵심 ID 시스템에 일괄적으로 프로필과 연구 업적 정보를 업로드함으로써 초기 DB를 효과적으로 구축할 수 있으며, 주기적 업로드를 통해, 효과적으로 갱신·유지할 수 있을 것이다. 이를 위해 대학의 연구 업적 시스템은 핵심 ID 시스템의 API를 설치해야 하며, 연구자의 고유 ID와 대학 내부 ID간 싱크(Sync) 유지가 필요하겠다. 또한 파트너 시스템의 기구축 학술정보 소스를 활용하는 방법도 있을 것이다. 파트너 시스템인 KCI, RISS, 학회 마을과 같은 학술논문 포털을 대상으로 유니크한 연구자를 추출해 내고, 추출된 연구자에게 고유 ID를 부여하는 것이다. 서지로 부터 동일저자 클러스터를 형성하고 육안식별 과정을 통해 고유 ID를 생성시키면(Ikki 2010), 대학의 업적 시스템에 포함되지 않은 연구자들까지 고유 ID를 생성시킬 수 있으며, 소스가 된 파트너 데이터베이스에 구축된 논문들과 링크가 생성될 수 있을 것이다.

두 번째는 연구자 자신의 직접 선언을 유도하는 것이다. 개별 저자의 참여를 유도하는 것은 쉬운 일이 아니지만, 저자 스스로가 자신의 저작물임을 선언하는 셀프 클레임은 빠르고 정확하게 모호한 저자명을 식별할 수 있는 방법이다. 또한 누락된 정보를 갱신시켜 나가고 복수의 아

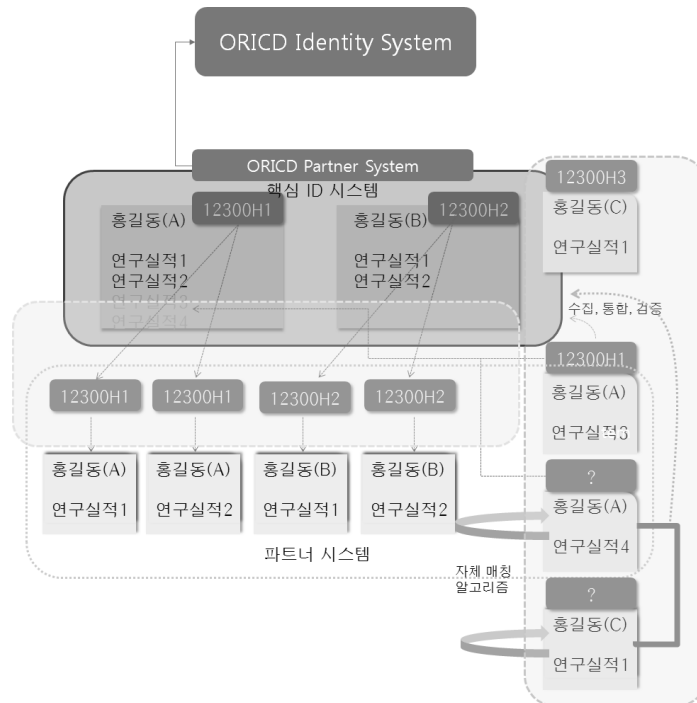
이덴티티로 분리된 개체를 통합시킴으로써 핵심 ID 시스템의 효과적인 유지를 가능하도록 한다. 따라서 ORCID에서는 저자의 셀프 클레임을 유도하기 위한 방식 중 하나로 논문 투고 시스템의 활용을 제안하고 있다(Shillum 2012). 저자가 학술지에 논문을 투고할 때, 본인에게 부여된 연구자 고유 ORCID를 밝히고 동시에, "CrossRef", "Scopus"등을 통해 본인의 실적을 선별하도록 하는 방식이다. 우리도 논문투고 시스템인 JAMS나 ACOMS 등을 통해 이러한 방식을 적용시킬 수 있을 것이다. 물론 투고 시스템은 핵심 ID 시스템과 API로 연계되어 있어야 하고, KCI나 학회마을 DB등과도 연동되어 저작물을 검색할 수 있도록 지원되어야 할 것이다.

2) 시스템을 넘어선 링크 체계 구현

핵심 ID 시스템을 기반으로 파트너 시스템간 투명한 링크 체계를 구현하기 위하여 다음과 같은 몇 가지 전략을 제안할 수 있을 것이다.

첫 번째는 <그림 1>과 같이 핵심 ID 시스템에 등록된 연구자 고유 ID를 파트너 시스템이 다운로드 받아가 각자의 시스템에 구축된 저자 정보와 일치시키는 방법이다. 파트너 시스템은 연구자 고유 ID, 논문 리스트, 프로필 정보 등을 다운로드 받아 자체 식별 매커니즘에 활용하고 해당 시스템에 구축된 학술 논문과 링크를 형성시킬 수 있을 것이다.

두 번째는 논문 투고 단계에서부터 신규 생성될 논문에 연구자의 고유 ID가 부여될 수 있도록 유도하는 것이다. 학회의 논문 투고 프로세



<그림 1> 핵심 ID 시스템과 파트너 시스템(학술정보포털, 전자저널)

스에서 연구자가 자신의 고유 ID를 명시한다면, 심사가 종료되어 출판이 완료된 논문이 상업적으로 유통되거나 인터넷을 통해 자유롭게 유통되더라도 저자의 명확한 이름 식별을 보장할 수 있게 된다.

세 번째는 기관 레포지토리를 활용해 투명한 링크 체계를 유지하는 것이다. 기관 레포지토리는 기관의 연구 성과를 장기 보존하지만, 허베스팅 프로토콜에 의해 세상 밖으로 전파시키는 역할도 수행한다. 따라서 학내 연구 업적 시스템에 등록된 연구 결과물이 기관 레포지토리에 아카이빙될 때, 연구자에게 부여된 고유 ID가 명확하게 유지된다면 연구자의 학내 업적 정보 시스템, 기관 레포지토리, 기관 레포지토리 포털 시스템과 핵심 ID 시스템을 연구자 고유 ID로 투명하게 연계시킬 수 있을 것이다.

3) 해외 출판 논문의 내국인 저자명 식별 도모
 내국인 연구자의 ORCID 등록은 연구자의 명확한 식별뿐 아니라, 해외에 국내 연구자들의 연구 영역과 범위를 이해시키고 국내 학술연구 업적의 글로벌한 전파에도 많은 도움을 줄 수 있을 것이다. 가장 효과적인 방법은 앞에서 설명한 바와 같이 한국의 핵심 ID 시스템이 ORCID의 파트너가 되는 것이다. 연구자가 개별적으로 ORCID에 가입하거나, ORCID Identity 시스템을 통해 셀프 크레임을 수행하지 않아도 일괄적으로 국내 연구자의 ID와 그들의 정확한 실적 정보가 등록되어 글로벌하게 공유될 수 있을 것이다. 이렇게 된다면, “Scopus”, “Web of Knowledge”와 같은 파트너 시스템이 이 데이터를 활용하여, 기존에 모호했던 내국인 저자의 아이덴티티를 명확하게 식별해 낼 수 있게 될

것이다.

5. 결론 및 요약

기계적 처리 방식에 의존한 부정확성, 글로벌 상호 연계성 부재가 지적되어 온 학술 연구자 식별 문제를 해결하기 위하여 최근 ORCID가 출범하였다. ORCID는 현존하는 다른 저자 ID 체계간 투명한 링크 매커니즘을 구축하기 위한 글로벌 식별 체계로 연구자를 대상으로 한 DOI (Digital Object Identifier)라고도 표현할 수 있겠다. ORCID의 이상은 전 세계 연구자 커뮤니티의 패스포트가 되는 것이다. 개인에게 ID를 부여하는 것에는 프라이버시와 유용성이라는 이율배반적 관계가 존재하지만, 연구자의 ORCID는 개인의 ID임에도 불구하고 다분히 공공적 성격을 띠고 있다. 또한 최근 오픈 액세스 운동에서 보여지듯이 연구 결과의 공공성이 계속 높아지고 있어, 여러 가지 우려에도 불구하고 ORCID의 확대 가능성은 긍정적으로 전망해 볼 수 있을 것이다(藏川 圭, 武田 英明 2011).

본 연구는 ORCID 체계를 심층 분석하고, 이를 응용하여 핵심 ID 시스템, 파트너 시스템으로 구성된 국내 학술연구자 식별 시스템 구축 방향과 운영 전략을 제시하였다. 그 내용을 아래와 같이 요약해 볼 수 있겠다.

첫 번째, 핵심 ID 시스템은 국내 연구자의 고유 ID와 업적 정보를 효과적으로 학술정보 포털, 상용 전자저널 시스템과 같은 파트너 시스템과 공유하여, 시스템을 넘어선 투명한 링크를 지원할 수 있어야 할 것이다.

두 번째, 핵심 ID 시스템은 글로벌 식별 매카

니즘인 ORCID Identity 시스템의 파트너가 되어, 해외 학술 논문에 기재된 국내 연구자의 명확한 이름 식별을 도모해야 할 것이다.

세 번째, 핵심 ID 시스템은 이미 전국 대학의 연구 업적 데이터베이스와 연동된 KRI를 확대 발전시키는 것이 가장 효과적일 것이다. 데이터 구축과 갱신은 대학의 연구 업적 시스템과 파트너 시스템의 기구축 학술정보 소스를 활용할 수 있으며, 저자의 셀프 클레임 유도하여 보완해 나갈 수 있을 것이다.

네 번째, 연구자가 학술논문 투고시, 개인에게 기부여된 연구자 고유 ID를 명시하도록 유도하거나 미등록자의 신규 등록을 유도한다면, 신규 생성될 논문에 연구자 고유 ID가 자연스럽게 포함되어, 추후 논문이 상업적으로 출판되거나 인터넷상에서 자유롭게 유통되어도 명확한 식별을 보장할 수 있게 된다.

다섯 번째, 파트너인 전자저널 출판사, 학술

정보 유통기관은 핵심 ID 시스템을 통해 연구자 고유 ID와 매칭된 학술 논문 정보를 다운로드 받아, 자체 식별 매커니즘에 활용할 수 있고, 이를 기반으로 연구자 페이지 같은 연구자 기반 서비스를 효과적으로 제공할 수 있을 것이다.

ORCID는 2012년 말에 출범하여 이제 막 걸음마를 시작한 국제 조직이므로, 크고 작은 운영상의 변화는 불가피하다고 본다. 그러나 이는 “Thomson Reuters”, “Elsevier”, “웰컴 트러스트(Wellcome Trust)”, “미국 국립보건원(NIH)” 등 전 세계적으로 가장 영향력 있는 학술지 출판사와 연구 지원 기관을 파트너로 추진하고 있는 글로벌 규모의 사업이다. 따라서 아시아 국가의 참여를 비롯한 앞으로의 행보를 주목하고 이러한 조류에 부응할 수 있는 국내 학술 연구 결과물 저자명 식별 시스템을 고민할 필요가 있을 것이다.

참 고 문 헌

- 강인수 외. 2008. 저자 식별을 위한 자질 비교. 『한국콘텐츠학회논문지』, 8(2): 41-47.
- 이용효. 2000. 『학술정보 전거 DB 시스템이 개발과 구축에 관한 연구』. 서울: 한국교육학술정보원.
- 이석형, 권승진. 2012. 전거데이터 기반 학술논문 검색시스템의 구축 및 평가에 관한 연구. 『한국문헌정보학회지』, 46(2): 133-156.
- 김현정. 2012. 인용분석에서의 모호한 저자명 식별을 위한 방법들에 관한 고찰. 『한국비블리아학회지』, 23(3): 5-17.
- 藏川 圭, 武田 英明. 2011. 研究者識別子ORCIDの取り組み. 『情報管理』, 54(10): 622-631.
- 藏川 圭. 2011. 動向レビュー：著者の名寄せと研究者識別子ORCID, カレントアウェアネス, CA 1740, [online]. [cited 2011.7.7]. <<http://current.ndl.go.jp/ca1740>>.
- 谷藤 幹子. 2010. ORCID 活用と(日本學術

- 誌への)提案 — 著者名の同定という
共通課題を克服するために, 『著者ID
の動向』, 第7回 SPARC Japanセミ
ナー2010, [online]. [cited 2012.9.7].
〈<http://pubman.mpg.de/pubman/faces/viewItemFullPage.jsp?itemId=escidoc:639846:2>〉.
- 武田 英明. 2010. ORCID とは何か, 第 64回
著者 HC の動向, [online]. [cited
2012.5.7].
〈http://www.nii.ac.jp/sparc/event/2010/pdf/7-2/doc2_takeda_20110114.pdf〉.
- 守本 瞬. 2011. 著者の識別に向けて: オープ
ンアクセス環境下における同定機能
導入のための恒久識別子実証実験,
[online]. [cited 2012.5.7].
〈<http://hdl.handle.net/2297/28168>〉.
- Danskin, A., A. Hill, and D. Needham. 2011.
The Names Project: A New Approach
to Name Authority. *ISQ Information
Standards Quarterly*, 23(3): 14-17.
- Fenner, M. 2011a. Author Identifier Overview.
LIBREAS. *Library Ideas*, 18: 24-29
[online]. [cited 2012.5.18].
〈<http://edoc.hu-berlin.de/docviews/abstract.php?lang=&id=37867>〉.
- Fenner, M. 2011b. ORCID: Unique Identifiers
for Autors and Contributers. *ISQ In-
formation Standards Quaterly*, 23(3):
10-13.
- ORCID Homepage. [cited 2013.3.8].
〈<http://orcid.org>〉.
- Ratne, H. 2011a. ORCID Brief Introduction
[online]. [cited 2012.5.20].
〈<http://orcid.org/sites/default/files/orcid-introduction-vivo-march-2010.ppt>〉.
- Ratne, H. 2011b. ORCID: An Open Registry
of Scholarly IDs, AAP-PSP Annual
Meeting Feburary, 2011 [online].
[cited 2012.5.25].
〈<http://www.slideshare.net/hratner/orcid-update-aap-bsp-annual-meeting-february-2011>〉.
- Ikki, O. 2010. Metadata design on academic
information services, *The journal of
Information Science and Technology
Association*, 60(12): 495-500 [online].
[cited 2012.11.25].
〈<http://ci.nii.ac.jp/naid/110007989449>〉.
- Library of Congress. 2011. Library of Congress
Name Authority File(LC / NAF).
[online]. [cited 2012.11.25].
〈<http://id.loc.gov/authorities/names.html>〉.
- Shillum, C. 2012. *ORCID Update*, CNI CNI
Scholarly Identity Workshop Baltimore,
Apr [online]. [cited 2013.2.5].
〈<http://www.cni.org/wp-content/uploads/2012/07/Shillum-CNI-Indentify-Workshop-ORCID-Update.ppt>〉.
- Smith-Yoshimura, K. 2009. Networking Names
[online]. [cited 2012.12.25].
〈<http://ci.nii.ac.jp/naid/110007989449>〉.
〈<http://www.oclc.org/research/publications/library/2009/2009-05.pdf>〉.

- Strotmann, Andreas and Danzhi Zhao. 2012. "Author name disambiguation: What difference does it make in author-based citation analysis?" *Journal of the American Society for Information Science and Technology*, 63(9): 1820-1833.
- Wagner, A.B. 2009. Author Identification Systems. *Issues in Science and Technology Librarianship*, 59 [online]. [cited 2013. 2.10]. <DOI:10.5062/F40K26HX>.
- WATANABE, T. 2010. Authority control: Today and the future, *The journal of Information Science and Technology Association*, 60(9): 371-377 [online]. [cited 2012.12.29]. <http://ci.nii.ac.jp/els/110007701194.pdf?id=ART0009507369&type=pdf&lang=en&host=cinii&order_no=&ppv_type=0&lang_sw=&no=1335401864&cp=>.