

# 교수들의 셀프 아카이빙 저작물의 종류와 저장소에 관한 연구\*

## A Study on Types of Content and Venues for Faculty Self-archiving

김 지 현(Ji-Hyun Kim)\*\*

### 목 차

- |                      |                    |
|----------------------|--------------------|
| 1. 서론                | 3.3 후속 전화 면담       |
| 2. 셀프 아카이빙 방식에 관한 연구 | 4. 교수들의 셀프 아카이빙 방식 |
| 2.1 셀프 아카이빙 저장소      | 4.1 셀프 아카이빙 저장소    |
| 2.2 셀프 아카이빙 저작물      | 4.2 셀프 아카이빙 저작물    |
| 3. 자료의 수집 및 분석       | 4.3 요약             |
| 3.1 표본 추출            | 5. 결론 및 제언         |
| 3.2 설문지 구성 및 배포      |                    |

### 초 록

본 연구는 대학 교수들이 셀프 아카이빙하는 저장소와 저작물의 종류 및 그러한 결정의 배경이 되는 이유에 대하여 조사하였다. 미국 17개 연구 중심 대학의 교수들을 대상으로 설문지와 전화 면담을 수행하였으며 본 연구는 480명의 설문 응답과 41명의 후속 전화 면담 분석을 바탕으로 하였다. 분석 결과 교수들은 개인 홈페이지나 연구 그룹 홈페이지를 통하여 심사를 거친 논문을 셀프 아카이빙하는 경우가 많았다. 이를 통해 교수들이 학술 자료의 질적 수준 관리 기제로서의 논문 심사를 셀프 아카이빙에서도 중요한 요소로 인식한다는 것을 알 수 있다. 기관 레포지토리에서의 셀프 아카이빙 참여도는 저조한 편이었으나 몇몇 교수들은 다양한 종류의 학술 자료들을 디지털 형태로 보존할 수 있다는 점에서 그 발전 가능성을 언급하였다.

### ABSTRACT

This study investigated the self-archiving venues that university faculty members have used, the types of content that they have made publicly accessible on the web, and their rationales for such decisions. The present study was based on the analysis of 480 survey responses and 41 telephone interviews from professors at 17 Carnegie Doctorate-granting universities in the U.S. It was found that faculty members tended to self-archive referred articles on their personal websites or research group websites. This indicated that the faculties perceived peer-review process to be important in self-archiving practices as a quality-control mechanism. The rate of self-archiving in institutional repositories was low, although several interviewees envisioned the potential of the repositories regarding the ability to preserve various types of research works in digital form.

키워드: 셀프 아카이빙, 교수, 기관 레포지토리, 학술 출판, 디지털 보존

Self-archiving, Faculty, Institutional Repositories, Scholarly Publishing, Digital Preservation

\* 본 연구는 저자의 박사 학위 논문(Ji-Hyun Kim, 2008. Faculty Self-archiving: Factors Affecting the Decision to Self-archive, Ph.D. diss., University of Michigan) 일부임을 밝힙니다.

\*\* 오레곤 보건 과학 대학 아카이브 연구 조교(jhkz93@gmail.com)  
논문접수일자: 2010년 1월 30일 최초심사일자: 2010년 2월 3일 게재확정일자: 2010년 2월 12일  
한국문헌정보학회지, 44(1): 53-74, 2010. [DOI:10.4275/KSLIS.2010.44.1.053]

## 1. 서론

전통적으로 학술 커뮤니케이션(scholarly communication)은 연구자들 간 비공식적인 정보의 교류와 공식적인 정보 교류, 즉 심사(peer review)를 통한 논문과 단행본의 출판으로 구분되어 왔다. 인터넷이 광범위하게 수용되기 이전에 이 두 가지 커뮤니케이션 양식은 명확하게 구분되었으나 디지털 학술 자료 생산의 일반화와 인터넷을 통한 정보 교류의 확산은 전통적 관점의 학술 커뮤니케이션 양상에 변화를 일으키고 있다. Borgman(2007)은 이와 관련된 가장 큰 변화로써 “공적 커뮤니케이션과 사적 커뮤니케이션의 조화(the balance between public and private communication)(p.49)”를 언급하였다. 즉, 인터넷이 확산되기 이전에는 주로 사적인 정보 교류 대상이었던 학자들간의 대화나 학회 발표 자료, 혹은 출판 전 논문(preprints) 등은 이메일이나 웹캐스트(Webcasts), 혹은 웹상에 공개적으로 게시하는 방식으로 확산되고 있고, 공식적인 출판 자료의 배포 역시 인터넷을 통해 간소화되고 있는 실정이다. 뿐만 아니라 학술 정보의 배포를 주로 담당하는 출판사들의 학술지 가격 인상은 도서관 재정으로 감당하기 힘든 상황에 이르렀고 이러한 위기에서 학술 자료의 접근성 향상과 신속한 배포, 그리고 이러한 자료의 장기적인 보존을 가능하게 하는 새로운 시도와 관련 연구들이 계속되고 있다.

본 연구에서는 이와 같은 학술 커뮤니케이션, 특히 학술 정보의 배포와 보존에 관련된 변화를 나타내는 현상으로서 미국 연구 중심 대학 교수

들의 셀프 아카이빙 방식을 조사하였다. 셀프 아카이빙의 정의는 넓은 의미로 디지털 문서를 접근 제한이 없는(publicly accessible) 웹 사이트에 저장하는 것으로 특히 셀프 아카이빙을 위한 바람직한 저장소로서 OAI(Open Archives Initiative) 메타데이터 태그를 사용하는 레포지토리(OAI-compliant repositories)가 제안되고 있다.<sup>1)</sup> 이를 통해 레포지토리 간의 상호 운용성과 검색 기능이 향상된다는 장점이 있는데 실제 한 연구에서는 셀프 아카이빙의 장소를 이와 같은 기능을 가진 학문 주제별 레포지토리(disciplinary repositories) 혹은 기관 레포지토리(institutional repositories)로 한정하여 그 개념을 정의하였다(Coleman and Roback 2005). 그러나 또 다른 연구에서는 셀프 아카이빙 장소로서 개인 홈페이지 등 연구자들에 의해 자주 이용되는 웹 공간들을 조사하고 다양한 형태의 연구 저작물을 셀프 아카이빙 대상으로 간주하였다(Swan and Brown 2005). 이와는 달리 셀프 아카이빙의 대상을 심사를 거쳐 출판된 논문(refereed, published articles)과 심사 전 초고(pre-refereed drafts)로만 한정하여 정의하는 경우도 있다(Harnad 2003).

이처럼 셀프 아카이빙의 대상과 장소의 범위가 명확하게 정의되어 있지는 않지만 본 연구에서는 다양한 방식으로 이루어지는 셀프 아카이빙에 초점을 두어 이에 대한 교수들의 경험과 인식을 조사하였다. 셀프 아카이빙의 장소로는 개인 홈페이지, 연구 그룹 웹 사이트, 학과 웹 사이트, 주제별 레포지토리와 기관 레포지토리를 중심으로 조사하였으며, 셀프 아카이빙

1) What is self-archiving? <<http://www.eprints.org/openaccess/self-faq/#self-archiving>>.

되는 저작물은 연구 논문, 단행본 장(chapters) 및 연구 데이터를 중심으로 살펴보았다. 이와 더불어 셀프 아카이빙의 대상과 장소를 결정하는데 바탕이 되는 이유에 대해서도 조사하였다. 본 연구에서 제시하는 연구 문제들은 다음과 같다.

- ① 교수들은 어느 저장소에 셀프 아카이빙하며 왜 그것을 이용하는가?
- ② 교수들은 어떤 종류의 연구 저작물을 셀프 아카이빙하며 그 이유는 무엇인가?

셀프 아카이빙에 대한 선행 연구를 바탕으로 본 연구에서는 미국 17개 연구 중심 대학 교수들을 대상으로 설문지와 전화 면담을 실시하였다. 그 결과 셀프 아카이빙을 하고 있다고 응답한 480명의 교수들의 설문 내용과 41명의 응답자들을 대상으로 한 후속 전화 면담 내용을 분석하였다.

## 2. 셀프 아카이빙 방식에 관한 연구

셀프 아카이빙에 관하여 학술 저자들의 인식을 조사한 기존 연구들은 대부분 특정 셀프 아카이빙 저장소 - 주제별 레포지토리, 기관 레포지토리, 또는 개인 홈페이지에 초점을 맞추고 있다. 주제별 레포지토리, 특히 물리학 분야의 arXiv는 가장 성공적인 셀프 아카이빙 모델로 인정받고 있으며 이와 관련된 연구들이 다수 존재한다. 사례 연구나 질문지를 통해 arXiv에서의 셀프 아카이빙에 대한 저자들의 인식을 조사한 연구들이 있으며(Lawal 2002; Pinfield 2001), arXiv에 저장된 논문들을 분석하여 인

용 패턴과 저자들의 특징을 조사한 연구들도 있다(Brown 2001; Harnad and Brody 2004; Manuel 2001; Meyer and Kling 2002). 기관 레포지토리에서의 셀프 아카이빙에 관하여 교수들의 인식을 조사한 연구들은 주로 교수들의 낮은 참여도와 그 원인을 면담을 통해 심층 조사하거나(Davis & Connolly 2007; Foster & Gibbons 2005), 질문지를 통한 교수들의 인식 조사(Kim 2007; Lercher 2008), 혹은 기관 레포지토리에 저장된 논문들을 분석하여 교수들의 참여도를 조사한 연구들이 있다(Jantz and Wilson 2008; Xia 2007). 주제별 혹은 기관 레포지토리에 비해 개인 홈페이지에서의 셀프 아카이빙에 관한 연구는 소수에 불과하며 주로 홈페이지에 저장된 논문들의 종류 및 주제 분야를 분석한 내용을 바탕으로 셀프 아카이빙 방식을 조사하였다(Andrews 2003; Covey 2009).

이와 같이 선행 연구들의 대다수가 특정 셀프 아카이빙 장소를 중심으로 이루어져 왔으며 교수들의 인식에 관한 연구는 기관 레포지토리에 한정되어 있는 경향을 보이고 있다. 예외적으로 Swan and Brown(2005)은 다양한 셀프 아카이빙 방식에 대한 연구자들의 인식과 경험을 질문지를 통해 조사하였으며 이를 응답자들의 출신 국가와 주제 분야에 따라 비교 분석하였으나, 응답자들의 직위에 따른 인식의 차이는 고려하지 않았다. 본 연구에서는 학술 저자들 중 교수들의 인식에 중점을 두어 그들이 수용하고 있는 다양한 형태의 셀프 아카이빙 방식을 조사했다는 점에서 그 의의가 있다. 이 외에도 본 연구를 위해 수집된 질문지 응답과 면담 자료를 바탕으로 교수들의 셀프 아카이빙 참여에 영향을 미치는 여러 가지 요인들을 회

귀 분석을 통해 제시한 연구가 있다(Kim 2010). 이는 본 연구와 밀접한 관련이 있으나 교수들의 셀프 아카이빙 동기 요인을 중점적으로 분석했다는 점에서 본 연구와는 구별된다.

## 2.1 셀프 아카이빙 저장소

앞서 언급했듯이 arXiv는 가장 널리 수용된 주제별 레포지토리이며 현재 코넬 대학교 도서관에서 운영되고 있는데 물리학, 수학, 컴퓨터 사이언스 분야의 약 583,000건의 논문들을 포함하고 있다.<sup>2)</sup> 이와 같은 arXiv의 성공 요인 중 하나는 저자들이 연구 성과를 하루 속히 배포하고자 하는 요구를 arXiv가 충분히 만족시키고 있다는 것이다. 또한 arXiv에 포함된 대부분의 심사 전 논문들은 이후 학술 저널에 출판되고 있다. 이와 같은 특징은 arXiv에 셀프 아카이빙되는 연구물이 두 단계의 검증 과정을 거친다는 것을 보여준다. 즉 (1) arXiv를 통해 자유롭게 심사 전 논문들을 공개하여 타 연구자들로부터 논문에 관한 피드백을 받고 (2) 이후 학술지에 투고하여 심사를 거친(refereed) 논문을 출판하는 것이다. 이러한 심사 전 논문의 공개가 출판에 제약을 주지 않는다는 점에서 arXiv에서의 셀프 아카이빙이 권위를 가지게 되었고 학술 커뮤니케이션 전반에 영향을 미치게 되었다(Gunnarsdóttir 2005).

arXiv에 셀프 아카이빙하는 저자들은 자신들의 출판 전 논문뿐만 아니라 출판 후 논문(post prints)도 셀프 아카이빙하여 자신의 연구 성과가 타 연구자들에게 널리 알려질 수 있기를 원

했으며, 그와 더불어 권위 있는 학술지에 논문을 출판하는 것 자체도 매우 중요하다고 여기고 있었다. 결국 arXiv는 기존의 학술 논문 출판 시스템과 경쟁하는 것이 아니라 이를 보완하는 역할을 수행한다고 할 수 있다(Pinfield 2001). 또한 저자들이 arXiv에 참여하는 가장 큰 이유로는 신속한 연구물의 배포와 연구 성과의 가시성(visibility)을 향상시킬 수 있다는 점이였다. arXiv에 참여하지 않은 저자들은 학술 출판사들의 정책이 셀프 아카이빙을 지원하지 않는다는 점과 본인들의 정보 기술에 대한 지식 부족을 이유로 들었다(Lawal 2002).

arXiv가 성공적인 셀프 아카이빙 모델로 인식되면서 이와 비슷한 학문 주제별 레포지토리가 다른 분야에도 생기게 되었다. 인지 과학 분야의 CogPrints, 경제학 분야의 RePEc, 그리고 문헌정보학 분야의 DLIST를 예로 들 수 있으나 arXiv에 비해 각 분야 저자들의 셀프 아카이빙 참여도가 높지 않은 실정이다. 이 밖에도 arXiv를 통해 나타난 셀프 아카이빙 현상과 OAI에 의한 레포지토리 간 상호 운용성의 향상은 기관 레포지토리의 개발을 가능하게 한 사회적, 기술적 원동력이 되었다. 기관 레포지토리는 “한 대학에서 대학 구성원들이 생산한 디지털 문서들을 관리하고 배포하기 위해 제공하는 일련의 서비스”라고 정의되었으며 이를 위해 정보 기술의 변화를 관리하고 그 변화에 적합한 기술에 따라 디지털 정보를 이전(migration)하는 것을 기관의 임무로 삼고 있다(Lynch 2003, p.2). 이와 같은 정의는 대학 및 연구 기관에서 생산되는 디지털 학술 정보의 보존과

2) arXiv.org. <<http://arxiv.org>>.

관리가 기관 레포지토리의 핵심적인 역할임을 보여주고 있다.

최근 발표된 미국 내 기관 레포지토리 현황 조사 연구에 따르면 기관 레포지토리를 운영하거나 운영 계획이 있는 미국 내 대학들의 대다수가 카네기 대학 분류(The Carnegie Classification of Institutions of Higher Education)의 박사 학위 수여 대학(Doctorate-granting universities)인 것으로 나타났다(Markey et al., 2007). 이는 미국의 연구 중심 대학에서 기관 레포지토리가 학술 정보 인프라의 구성 요소로서 등장하고 있음을 보여준다. 그러나 대학 구성원들, 특히 교수들이 기관 레포지토리에 셀프 아카이빙 하는 빈도는 낮은 것으로 드러났다. 사례 연구와 교수들과의 면담을 통해 기관 레포지토리를 이용하지 않는 이유를 조사한 연구들에 따르면, 교수들은 자신들의 연구물이 널리 읽히고 인용되기를 원하는데 기관 레포지토리는 그러한 요구를 충족시키지 못하고 있다고 인식하고 있었다(Foster and Gibbons 2005). 자신들의 분야에 있는 연구자들이 기관 레포지토리에 있는 논문을 읽지 않으며 기관 레포지토리 자체가 고립된 저장소로서 일반 검색 엔진을 통해 검색되지 않는다는 의견도 있었다. 교수들은 자신들의 주제 분야에서 통용되는 기준이나 전형에 맞추어 기관 레포지토리를 인식하고 있으며 이렇게 분야별로 다양한 요구와 기대를 어떤 식으로 충족할 것인가가 교수들의 참여를 유도하는 관건이라 할 수 있다(Davis and Connolly 2007).

주제별 혹은 기관 레포지토리가 셀프 아카이빙을 위한 우선적인 장소로 제시되고 있기는 하지만, 일반적으로 연구자들이 자신들의 저작물

을 셀프 아카이빙하는 곳은 개인 홈페이지인 것으로 나타났다. 저자들의 셀프 아카이빙 방식에 대한 한 연구에 따르면 질문지 응답자의 18%가 출판 전 논문을, 27%는 출판 후 논문을 개인 홈페이지에 셀프 아카이빙한다고 하였다. 이는 같은 응답자들의 주제별 레포지토리나 기관 레포지토리에서의 셀프 아카이빙 비율보다 적게는 3%, 많게는 15% 정도 높게 나타난 것이다(Swan and Brown 2005). 프랑스의 수학자들과 컴퓨터 과학자들의 셀프 아카이빙 방식에 대한 연구에서도 응답자 123명 중 66%가 출판 전 논문을, 63%가 심사를 거친 논문을 개인 홈페이지에 셀프 아카이빙 한다고 대답하였다. arXiv에 출판 전 논문을 셀프 아카이빙 한다는 응답자는 34%에 그쳤으며 기관 레포지토리에 셀프 아카이빙하는 비율은 출판 전 논문이 19%, 심사를 거친 논문은 17%에 그쳤다(Wojciechowska 2007). 또 다른 연구에서는 카네기 멜론 대학 교수들의 개인 홈페이지를 조사하여 학술지에 출판된 논문을 어느 정도 셀프 아카이빙 하는지를 분석하였다. 34%의 교수들이 적어도 한 건의 출판된 논문을 셀프 아카이빙하였으며, 학과 홈페이지를 통해서 셀프 아카이빙 한 것까지 합치면 45%의 교수들이 출판된 논문을 셀프 아카이빙 하는 것으로 나타났다. 이를 통해 예상 외로 많은 교수들이 자신의 논문을 오픈 액세스로 웹 상에 공개하고 있음을 알 수 있다(Covey 2009). 개인 홈페이지나 학과 홈페이지 이외에 저자들의 연구 그룹이나 연구 과제 웹 사이트도 셀프 아카이빙의 장소로 언급되고 있다(Gadd, Oppenheim and Proberts 2003). 이렇게 저자들이 이용하는 기존의 셀프 아카이빙 방식을 이해하는 것은 기관 레포지토리나 같은 새로운 서비

스를 계획하고 제공하는 데 있어 중요한 기초가 될 수 있다.

## 2.2 셀프 아카이빙 저작물

전자 형태의 출판 전 논문(electronic preprints, 또는 e-prints)은 주로 arXiv를 통해 셀프 아카이빙되는 대표적인 종류의 연구 저작물이라고 할 수 있다. 그러나 Kling(2004)에 따르면 출판 전 논문의 원래 의미가 아직 출판되지는 않았으나 심사를 거쳐 출판이 허가된(accepted for publication) 논문이기 때문에 심사 중인(under review) 논문을 표현하는데는 부적절하다고 언급하였다. 심사 중인 논문일 경우 전자 형태의 원고(electronic manuscript, 또는 e-scripts)라고 명시하는 것이 정확한 표현임을 주장하였으며 이는 심사 중인 논문과 출판이 허가된 논문 간에 명백한 차이가 있음을 보여주는 것이다. 그러나 다른 연구에서는 심사 중인 논문을 심사를 거치기 전의 논문(pre-refereed article)으로 표현하여 이를 출판 전 논문(preprints)과 동일시 하기도 한다(Guédon 2003). 이렇게 출판 전 논문의 개념이 선행 연구에서 명확하게 정의되지 않았기 때문에 본 연구에서는 출판 전 논문이라는 용어 대신 심사 전 초고(pre-refereed drafts)라는 용어를 사용하였고 이는 학술지에 투고하여 심사 중인 원고를 의미한다. 그리고 출판이 허가된 논문은 모두 심사를 거친 논문(refereed articles)으로 정의하였다.

심사를 거친 논문은 저자들이 가장 많이 셀프 아카이빙 하는 연구 저작물이며 두 종류가 있다. 하나는 저자의 최종본(author's final version)이고 다른 하나는 출판사의 PDF본인데, 출판

사들은 일반적으로 출판된 학술지 논문의 PDF본을 셀프 아카이빙 하지 못하도록 규정하고 있다. 그럼에도 불구하고 흥미로운 것은 많은 저자들이 출판사의 PDF본을 셀프 아카이빙 하고 있다는 사실이다. Antelman(2006)에 따르면 사회 과학 분야의 22개 학술지에서 표본 추출되었고 저자들에 의해 셀프 아카이빙 되어 있는 575개의 논문들 중 절반 정도가 출판사의 PDF본이었다. 이는 출판사의 PDF본이 저자의 최종본에 비해 그 연구의 가치가 심사를 통해 검증되어 출판된 논문이라는 인상을 더 강하게 줄 수 있고, 읽기 쉬우며 인용하기도 편하다는 장점이 있기 때문이다. 이러한 결과를 통해 학술 저자들은 출판사의 정책보다는 자신들의 분야에서 통용되는 기준, 즉 연구물의 질적 수준이나 인용의 용이성을 강조하는 방향으로 셀프 아카이빙하는 경향이 있다는 것을 알 수 있다.

학술지의 논문 심사(peer review)가 연구물의 질적 수준을 평가하는 가장 일반적이고 핵심적인 평가 방법이기도 하지만, 몇몇 주제 분야, 특히 컴퓨터 사이언스, 수학, 혹은 경제학 분야에서는 심사 없는 논문(unrefereed articles)이 널리 읽히기도 한다. 기술 보고서(technical reports)나 워킹 페이퍼(working papers) 등이 그것인데 이러한 논문들이 생산되는 학과나 연구 기관에서 이들을 셀프 아카이빙하여 웹 상에 공개하기도 한다. 이와 같은 형태의 셀프 아카이빙에서 저장된 논문의 질적 수준은 대개 그 학과나 연구 기관의 명성에 따라 평가된다고 할 수 있다(Kling, Spector and McKim 2002).

논문 이외에도 저자들은 자신들이 저술한 단행본이나 교과서를 장 단위로 셀프 아카이빙하

기도 한다. 이는 논문 다음으로 저자들이 자주 셀프 아카이빙 하는 저작물의 종류인 것으로 나타났다. 그 밖에 저자들의 학위 논문이나 수업 자료, 소프트웨어나 오디오/비디오 파일 등도 셀프 아카이빙 하는 것으로 나타났다(Swan and Brown 2005). 이러한 연구 결과는 저자들이 셀프 아카이빙 하는 저작물들이 학술지 논문뿐만 아니라 다양한 형태의 연구 결과물들을 포함하고 있다는 것을 보여 준다.

이 밖에도 연구 데이터 세트(research data set)는 재사용(re-use)을 통해 새로운 연구를 가능하게 한다는 점에서 장기적인 보존의 가치가 있는 중요한 연구 저작물이며 셀프 아카이빙의 주요 대상이기도 하다. 미국이나 영국, 호주 등지에서는 연구 데이터의 공유와 보존을 위한 인프라를 구축하기 위해 국가 단위 프로젝트들을 진행하고 있는데 미국 국가 과학 재단(National Science Foundation)의 Cyber-infrastructure Initiative와 영국의 UK Research Data Service(UKRDS)를 예로 들 수 있다. 이들은 연구 데이터 세트의 셀프 아카이빙을 장려하고 국가나 공공 기관에서 재정 지원을 받아 생산된 데이터 세트를 오픈 액세스로 제공하며 포털이나 네트워크를 개발하여 분산된 관련 자료들에 효율적으로 접근할 수 있도록 하는데 역점을 두고 있다(Doorn and Tjalsma 2007). 연구 데이터 세트의 생산자인 연구자들 역시 데이터 공유와 보존의 중요성에 대하여 인식하고 있지만 실제 데이터를 셀프 아카이빙 하는 것에 대해서는 여러 가지 제약을 느끼고 있는 것으로 나타났다. 연구자들은 데이터 세트를 웹 상에 공개하기 전에 그에 대한 자세한 설명과 메타데이터 등을 준비하는데 시간이 많이 소요된다고

응답했으며 이에 부담을 느끼고 있었다. 그 외에도 어느 장소에 자신들의 데이터를 셀프 아카이빙 할 것인지, 누가 그것을 이용할 것인지에 대한 불확실성, 데이터가 다른 연구자들에 의해 부정확하게 분석될 지도 모른다는 염려, 그리고 다른 연구자들이 자신들의 데이터를 재이용 하더라도 그 출처를 인용하지 않을 수 있기 때문에 셀프 아카이빙하기를 주저한다는 의견도 있었다(Griffiths 2008).

선행 연구를 통해 셀프 아카이빙 되는 저작물의 종류와 셀프 아카이빙에 이용되는 저장소에 대하여 살펴보았으며 이를 바탕으로 본 연구에서는 미국 17개 연구 중심 대학 교수들의 셀프 아카이빙 방식에 대하여 설문지와 전화 면담을 통하여 조사하였다. 다음 장에서 조사 대상 교수들의 표본 추출과 자료의 수집 및 분석 방법에 대하여 기술하고자 한다.

### 3. 자료의 수집 및 분석

#### 3.1 표본 추출

본 연구는 다양한 방식으로 이루어지는 교수들의 셀프 아카이빙에 대하여 조사하고자 하였으며 이를 위해 미국 내 DSpace 기관 레포지토리를 운영하는 대학의 교수들로 표본 추출 범위를 한정하였다. DSpace는 MIT(Massachusetts Institute of Technology)에서 개발된 오픈 소스(open-source) 소프트웨어이며 미국 내 기관 레포지토리를 운영하는 대학들 간에 가장 많이 사용되는 시스템이기도 하다(Markey et al., 2007). 동일한 기관 레포지토리 시스템을 사

용하는 대학들로 범위를 제한함으로써, 특히 기관 레포지토리에 대한 인식을 조사할 때 시스템이 다르기 때문에 발생할 수 있는 인식의 차이를 최소화하고자 하였다. 또한 이들 대학 교수들을 통해 기관 레포지토리 이외의 셀프 아카이빙 방식에 대해서도 조사할 수 있었다.

표본 추출은 2005년 9월 DSpace 웹사이트<sup>3)</sup>에서 제공되는 기관 레포지토리 목록을 통해 이루어졌다. 그 당시 목록에 포함된 22개국 110개의 기관 레포지토리 중 미국 내 34개 대학 기관 레포지토리가 있었으며 이 중 18개 대학이 카네기 대학 분류의 박사 학위 수여 대학이었다. 이 중 1개 대학 교수들을 대상으로 필자가 파일럿 연구를 시행하였으며, 본 연구에서는 나머지 17개 대학에서 표본 추출된 교수들을 대상으로 하였다. 모집단의 수를 알아보기 위해 각 17개 대학들의 참고 사서들에게 조교수, 부교수, 정교수들의 수에 대하여 문의하였고, 응답이 없을 경우 대학 홈페이지에 나와 있는 통계 자료를 바탕으로 전체 교수 수를 추정하였다. 그 결과 모집단에 포함된 교수 수는 28,287명으로 추정되었고 이 가운데 약 5%에 해당하는 1,500명의 교수들을 표본 추출하였다.

본 연구에서 추출된 표본은 크게 두 종류로 나뉜다. 하나는 각 기관 레포지토리에 본인의 저작물이 셀프 아카이빙되어 있는 교수들인데, 이를 찾기 위해 각 기관 레포지토리에 저장된 자료들의 저자 목록에서 저자 성명을 각 대학 홈페이지에 있는 온라인 디렉토리를 통해 검색하여 조교수, 부교수, 정교수, 혹은 석좌 교수인 저자들을 가려내었다. 그 결과 621명의 교수들

이 검색되었고 이들 모두를 표본에 포함시켰다. 이들 교수들의 저작물이 기관 레포지토리에 저장되어 있기는 하지만, 실제 미국의 기관 레포지토리에 저장된 자료들은 대학 사서들을 통해 학과에서 레포지토리로 옮겨진 경우가 대부분이다(Lynch and Lippincott 2005). 이러한 대리 제출 방식(proxy submission)을 염두에 두고 본 연구에서는 검색된 교수들 모두가 자신의 저작물이 기관 레포지토리에 저장되어 있음을 인식하지는 못할 것이라고 가정했다. 하지만 이러한 표본 추출 방법을 통해 기관 레포지토리에 직접 셀프 아카이빙하는 교수들을 파악하여 조사할 수 있다는 점에서 의미가 있다고 할 수 있다.

이들 621명의 교수들을 제외한 나머지는 기관 레포지토리에 셀프 아카이빙하지는 않지만 다른 방식의 셀프 아카이빙에 참여할 가능성이 있는 교수들로 간주하였다. 여기서의 표본 추출은 자연 과학, 공학, 사회 과학, 인문학 영역에서 각각 주제 분야 셋을 선정하여 모두 12개 주제 분야의 교수들을 대상으로 이루어졌다. <표 1>은 이 주제 분야들을 영역 별로 정리한 것이다. 이들은 주제별 레포지토리가 존재하여 셀프 아카이빙에 전반적으로 익숙한 분야(수학, 물리학, 컴퓨터 사이언스, 경제학)와 그렇지 않은 분야를 포함하고 있어 셀프 아카이빙에 대한 교수들의 다양한 인식을 개진하는데 도움이 되었다. 무작위 추출 방식으로 모두 879명의 교수들의 표본이 완성되었으며 이로써 합계 1,500명의 교수 표본 추출이 이루어졌다. 비록 기관 레포지토리와 관련된 621명은 다양한 주제 분야의 교수들로 구성되어 있지만 대다수

3) <<http://www.dspace.org>>.

〈표 1〉 표본 집단의 12개 주제 분야

영역	주제 분야
자연 과학	수학, 물리학, 분자 생물학
공학	컴퓨터 사이언스, 전자 공학, 기계 공학
사회 과학	경제학, 심리학, 사회학
인문학	미술(Art), 영문학, 역사학

가 〈표 1〉의 주제 분야에 속해 있으므로 이들 주제 분야가 표본 집단 전체를 대표한다고 볼 수 있다. 이 주제 분야들 외에 다수의 교수들이 포함된 분야로는 경영학과 언어학이 있으며 그 외에 화학 공학, 항공우주 공학, 행정학, 정치학 등이 있다.

### 3.2 설문지 구성 및 배포

설문지에 포함된 질문들은 선행 연구를 바탕으로 개발되었으며 몇몇 질문들은 기존 셀프 아카이빙 연구에서 사용된 설문지의 질문들을 본 연구에 맞게 바꾸어 사용하였다(Allen 2005; Gadd, Oppenheim and Probets 2003; Kankanhalli, Tan and Wei., 2005; Ober 2005; Rowlands, Nicholas and Huntington., 2004; Swan and Brown 2004, 2005). 설문지는 네 부분인 (1) 셀프 아카이빙 방식; (2) 셀프 아카이빙에 관한 인식; (3) 향후 셀프 아카이빙에 대한 계획; (4) 개인별 특성(demographics)으로 구성되었다. 본 연구는 설문지의 첫 번째 부분인 셀프 아카이빙 방식에 관해 수집된 데이터의 분석을 바탕으로 하였다. 이 부분은 셀프 아카이빙의 참여 여부, 대학에서 운영하는 DSpace 기관 레포지토리에 대한 인식 여부에 관한 질문, 13개의 선다형(multiple-choice) 질문들과 개방형 질문 하나로 구성되었다. 셀프 아카이빙

인식에 관해서는 36개의 리커트 스케일(Likert scale) 질문들을 이용하였으며, 셀프 아카이빙의 향후 계획에 관해서는 8개의 리커트 스케일 질문을 통해 기관 레포지토리에서의 셀프 아카이빙을 위한 동기 부여 요인을 조사하였다.

설문지는 온라인 모드로 작성되었으며 설문지 작성 요구는 일반 우편을 통하여 이루어졌다. 표본 추출된 교수들에게 연구의 목적과 설문지의 URL, 그리고 응답자의 고유 번호를 적은 편지와 2달러 지폐를 사례금으로 동봉하였다. 배포의 편의를 위해 응답자를 세 집단으로 나누어 2006년 10월 13일부터 27일 사이에 일주일 간격으로 편지를 발송하였으며 편지를 최초 발송한 일자에서 1주 후에 연구의 목적과 설문지 URL을 포함한 이메일을 발송하였다. 다시 2주 후 이메일을 보내 설문문에 응해 줄 것을 요청하였다. 이를 통해 설문을 작성한 교수들은 736명이었으나 이 중 684명(45.6%)의 응답지가 분석 가능하였으며 이를 활용하였다. 수집된 데이터는 Stata통계 패키지를 이용하여 분석하였다.

### 3.3 후속 전화 면담

본 연구에서 사용된 설문지에는 후속 전화 면담에 참여 의사를 묻는 질문이 있었으며 이에 대답한 응답자들을 대상으로 이메일을 보내

참여 의사를 다시 확인하였다. 이 과정에서 41명의 응답자들이 전화 면담 요청을 수락하였다. 전화 면담은 2007년 3월부터 5월 사이에 이루어졌으며 평균 20분 정도 소요되었다. 이들 41명 중 32명은 셀프 아카이빙을 하는 교수들이었으며 9명은 하지 않는 교수들이었다. 셀프 아카이빙에 참여하는 교수들에게는 가장 최근의 셀프 아카이빙 경험과 어떤 종류의 저작물을 어느 곳에 셀프 아카이빙하며 그러한 결정의 배경이 되는 이유, 그리고 셀프 아카이빙에 영향을 미치는 요인을 중심으로 질문하였다. 셀프 아카이빙을 하지 않는 교수들에게는 그 이유와 셀프 아카이빙에 대한 인식, 향후 셀프 아카이빙에 대한 계획을 중심으로 질문하였다. 전화 면담 내용은 디지털 녹음기를 이용해 녹취하였으며 전문 전사가(transcriber)가 녹음 내용을 전사하였다. 전사된 면담 내용을 선행 연구에서 도출한 24개 코드로 구성된 코딩 체계(coding scheme)를 바탕으로 코딩하였으며 이 과정에서 질적 연구 자료 분석을 위한 소프트웨어인 Nvivo를 사용하였다.

#### 4. 교수들의 셀프 아카이빙 방식

응답자 684명 가운데 480명(70.2%)이 연구 저작물을 저장소의 종류에 관계 없이 셀프 아

카이빙하고 있다고 응답했다. 이는 응답자 가운데 대다수의 교수들이 웹 상에 자신들의 연구물을 공개하고 있다는 것을 나타내며, 본 연구에서는 이들 480명의 셀프 아카이빙 방식 - 어떤 종류의 저작물들을 어느 곳에 셀프 아카이빙하는가 - 에 중점을 두어 분석하고자 한다.

분석에 앞서 본 연구에서 사용된 표본 집단과 응답자 집단, 그리고 자신의 저작물을 셀프 아카이빙하고 있다고 응답한 집단을 <표 2>에서 요약하였다. 표본 추출시 기관 레포지토리에 저작물이 셀프 아카이빙 되어 있었던 교수들 가운데 질문지에 응답한 수는 269명이며 이들 중 셀프 아카이빙에 참여하고 있다고 응답한 수는 221명인 것으로 나타났다. 하지만 이는 기관 레포지토리 뿐만 아니라 다른 저장소에서의 셀프 아카이빙도 포함하는 것이기 때문에 이들 221명이 모두 기관 레포지토리에 셀프 아카이빙한 것을 인식하고 있다고 단정지을 수 없으며, 이에 대한 자세한 분석은 '4.1 셀프 아카이빙 저장소'에서 다루고자 한다(표 2 참조).

셀프 아카이빙에 참여하고 있는 교수들의 직위를 보면 정교수의 비율(36.7%)이 가장 높았으며 조교수의 비율이 그 다음을 차지했다(표 3 참조). 표본 추출시 정교수 비율이 45.6%, 조교수 비율이 30.2%이었고 부교수와 석좌 교수가 그 뒤를 따랐으며 전체 응답자에서의 직위별 교수 비율의 순위 양상도 이와 같았다.

<표 2> 표본 집단, 응답자 집단과 셀프 아카이빙 참여 집단

	기관 레포지토리에 저작물이 셀프 아카이빙 되어 있는 교수	기관 레포지토리에 저작물이 셀프 아카이빙 되어 있지 않은 교수	합계
표본 집단	621(41.4%)	879(58.6%)	1,500(100.0%)
응답자 집단	269(39.3%)	415(60.8%)	684(100.0%)
셀프 아카이빙 참여자	221(46.0%)	259(54.0%)	480(100.0%)

〈표 3〉 셀프 아카이빙하는 교수들의 직위

	빈도	퍼센트
조교수	145	30.2
부교수	121	25.2
정교수	176	36.7
석좌 교수	38	7.9
합계	480	100.0

주제 영역을 보면 공학 교수의 비율(33.3%)이 가장 높았고 사회 과학 교수의 비율(30.4%)이 그 다음이었다(표 4 참조). 실제 표본 추출시 공학 교수 비율(31.0%)이 가장 높았고 사회 과학 교수들의 비율(29.1%)이 그 뒤를 이었으며 전체 응답자도 이와 같은 양상을 보였다.

다만 자연 과학 교수 비율은 표본에서의 비율(21.3%)보다 4% 가량 높으며 인문학 교수 비율은 표본 비율(18.7%)보다 8% 정도 낮은 것으로 나타났다(표 4 참조). 본 연구에서의 자연 과학 영역에는 수학과 물리학이 포함되어 있고 이들 분야 교수들의 셀프 아카이빙은 특히 arXiv를 통해 활발히 이루어지며, 인문학 교수들의 셀프 아카이빙 참여도는 낮다는 것을 보여준다.

〈표 4〉 셀프 아카이빙하는 교수들의 주제 영역

	빈도	퍼센트
자연 과학	123	25.6
공학	160	33.3
사회 과학	146	30.4
인문학	51	10.6
합계	480	100.0

셀프 아카이빙 기간에 대한 질문에서는 과반 수 이상이 6년 이상 셀프 아카이빙을 해 왔다고 응답하였으며 이는 셀프 아카이빙에 참여하는 교수들의 상당수가 오랜 기간 자신들의 저작물

을 웹 상에 공개해 왔으며 이러한 배포 방식에 익숙하다는 것을 보여 준다(표 5 참조).

〈표 5〉 셀프 아카이빙 기간

	빈도	퍼센트
1년 이하	17	3.5
1년에서 3년 이하	113	23.5
4년에서 5년 이하	89	18.5
6년 이상	249	51.9
무응답	12	2.5
합계	480	100.0

셀프 아카이빙을 실제로 수행하는 사람들에게 대한 질문에서는 복수 응답이 가능했으며 대다수(83.1%)가 교수 본인이 셀프 아카이빙한다고 응답하였다(표 6 참조). 학생 조교가 대신한다는 응답(29.3%)이 그 다음이었으며 학과 직원이나 협동 연구자(collaborators)들이 대신한다는 응답도 있었다. 도서관 직원이 대신한다는 응답은 7.9%에 그쳤는데 이는 기관 레포지토리에서의 대리 제출 방식과 관련이 있으며 소수의 교수들만이 이를 인식하고 있는 것으로 나타났다. 기타 응답으로는 연구 센터 직원, 응답자가 고용한 웹 디자이너, 출판사 직원, 학회 간사 등이 있었다.

〈표 6〉 셀프 아카이빙을 실제 수행하는 사람들

	빈도	퍼센트
본인	399	83.1
학생 조교	143	29.8
학과 직원	117	24.4
협동 연구자 (collaborators)	100	20.8
도서관 직원	38	7.9
기타	21	4.4

#### 4.1 셀프 아카이빙 저장소

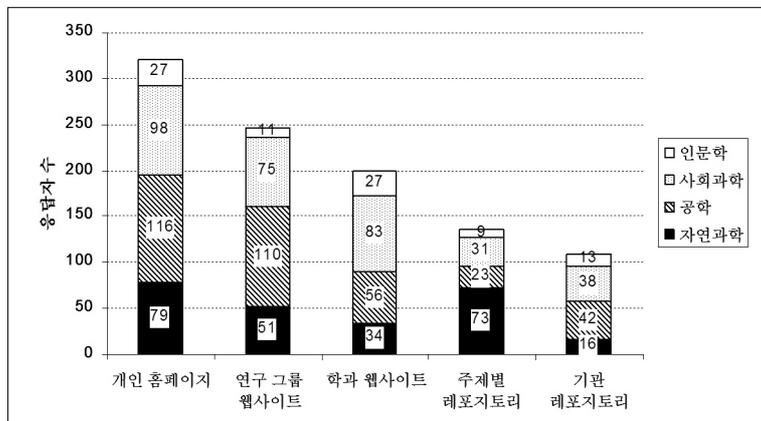
본 연구에서는 어느 저장소에 셀프 아카이빙 하는지에 관한 질문에서 복수 응답을 허용하였다. <표 7>에 제시된 셀프 아카이빙 저장소들 중 한 종류만을 이용하는 응답자는 480명 중 122명(25.4%)이었으며 나머지 358명(74.6%)의 응답자들은 두 가지 이상의 저장소에 셀프 아카이빙하고 있었다. 선행 연구와 마찬가지로 교수들의 셀프 아카이빙은 개인 홈페이지에서 가장 많이 이루어지는 것으로 나타났다. 개인 홈페이지 이외의 웹 공간, 특히 연구 그룹 웹사이트를 셀프 아카이빙의 장소로 활용하는 교수들이 많았으며(51.5%), 학과 웹사이트도 셀프 아카이빙의 장소로 자주 활용되고 있음을 알 수 있다. 셀프 아카이빙의 바람직한 장소로 제안되는 주제별 혹은 기관 레포지토리에서의 셀프 아카이빙은 상대적으로 저조한 것으로 나타났다. 앞서 <표 2>에서 나타난 바와 같이 기관 레포지토리에 자신의 저작물이 셀프 아카이빙 되어 있는 교수들 621명의 표본 집단 중 269

명이 질문에 응답했고 이 중 221명이 셀프 아카이빙을 하고 있다고 응답했다. 그러나 실제 본인이 기관 레포지토리에 셀프 아카이빙한다고 응답한 수는 109명에 불과했다(표 7 참조). 이를 통해 미국 기관 레포지토리의 자료 수집이 저조, 특히 교수들의 자발적인 셀프 아카이빙에 의해서라기보다는 사서들을 통한 대리 제출 방식으로 이루어진다는 것을 보여 준다.

<표 7> 셀프 아카이빙 저장소

	빈도	퍼센트
개인 홈페이지	320	66.7
연구 그룹 웹 사이트	247	51.5
학과 웹 사이트	200	41.7
주제별 레포지토리	136	28.3
기관 레포지토리	109	22.7

<그림 1>에서는 이와 같은 셀프 아카이빙 저장소의 이용을 학문 영역별로 구분하여 나타내었다. 주제별 레포지토리를 이용한다고 응답한 136명 가운데 73명(53.7%)이 자연 과학 영역의 교수들이므로 드러났으며 이들 중 물리



<그림 1> 학문 영역별 셀프 아카이빙 저장소의 이용

학 분야 교수 36명과 수학 분야 교수 34명이 포함되어 있었다. 이를 통해 분석에 포함된 자연 과학 영역 교수 대다수가 셀프 아카이빙하는 주제별 레포지토리는 arXiv임을 알 수 있다. 사회 과학이나 공학 영역에서는 arXiv와 같이 널리 수용된 주제별 레포지토리가 없으므로 이들 분야 교수들은 개인 홈페이지나 연구 그룹 웹 사이트, 혹은 학과 웹사이트를 셀프 아카이빙의 장소로 활용하고 있었다. 기관 레포지토리 역시 사회 과학 교수들과 공학 교수들의 참여율이 높았다. 인문학 교수들의 셀프 아카이빙 참여율은 모든 종류의 저장소에서 가장 낮은 것으로 나타났다.

전화 면담에서 교수들은 개인 홈페이지에 셀프 아카이빙 하는 이유로 관리의 용이성을 가장 많이 언급하였다. 즉 자유롭게 저작물을 셀프 아카이빙하고 필요할 경우 수정하거나 삭제할 수 있으며 이러한 작업을 원하는 시간에 할 수 있다는 것이다. 면담에 응한 사회학 분야 조교수의 말을 인용하면 다음과 같다.

My personal web site is what I use more often because it's something that I have control over. I can update it and it's easy for me if I want to post something at 2:00 in the morning and that's when I want it done. I don't have to worry about someone else's schedule so the amount of control that I have, what's there, how it's presented is useful (ID 1702).

이러한 개인 홈페이지 이용의 장점은 종종 다른 종류의 셀프 아카이빙 저장소, 특히 기관

레포지토리를 이용하지 않는 이유로 자주 언급되었다. 교수들은 기관 레포지토리에 일단 제출한 논문은 삭제할 수 없다고 믿고 있었으며, 기관 레포지토리의 사서를 통해 자신의 저작물이 관리된다는 것에 불편함을 느끼고 있었다. 이처럼 교수들은 본인의 저작물을 기관 레포지토리에 셀프 아카이빙하면 그 저작물에 대한 모든 통제를 상실하는 것이라고 인식하고 있었고 이는 교수들이 기관 레포지토리 수용을 꺼리는 한 요인임을 알 수 있다.

연구 그룹 웹 사이트에서의 셀프 아카이빙에 대하여 사회학 분야의 교수(ID 1803)는 이를 통해 비슷한 연구 관심사를 가진 다른 연구자들과 교류할 수 있기 때문에 중요하다고 하였다. 또한 물리학 분야의 부교수(ID 2031)는 연구 그룹 웹 사이트가 지금까지의 자신의 연구 성과를 총망라하는 저장소이며 인터넷 접속이 있는 어느 곳에 가든지 접근 가능하기 때문에 본인에게는 매우 유용한 정보원임을 강조하였다. 또 다른 물리학 교수(ID 1925)는 대규모 협동 프로젝트에 참여하고 있기 때문에 자신의 연구 그룹 웹 사이트는 다양한 특성을 지닌 웹 페이지들로 구성된다고 하였다. 즉 자신의 학과 서버에 속해 있고 오픈 액세스로 제공되는 연구 그룹 페이지와 협동 연구자들 서버에 속한 위키(wiki), 그리고 국립 연구소 서버에 속한 접근 제한 페이지 등이 그것이다. 이는 셀프 아카이빙 공간이 다른 형태의 웹 자원들과 연결되어 있을 경우 장르의 구분이 모호할 수 있음을 보여주는 예라고 할 수 있다.

학과 웹 사이트의 경우, 컴퓨터 사이언스 분야 교수(ID 1984)는 학과 교수나 학생들이 작성한 기술 보고서들을 학과 웹 사이트에 셀프

아카이빙한다고 하였으며 이는 심사 없는 논문의 셀프 아카이빙에 대한 예를 보여 주는 것이다. 그러나 면담에 응한 교수들 중 학과 웹사이트를 학술 자료의 셀프 아카이빙에 활발히 이용하는 사람들은 거의 없었으며 학과 웹사이트는 주로 수업 자료나 뉴스레터 등 비학술 자료들을 위한 공간이라는 의견이 많았다.

주제별 레포지토리의 경우 면담에 응한 교수들 중 8명이 arXiv에 셀프 아카이빙함으로써 얻을 수 있는 장점들을 제시하였다. 예를 들어 물리학 분야의 한 교수(ID 2452)는 본인의 주제 분야에 있는 거의 모든 연구자들이 arXiv에 셀프 아카이빙된 페이지를 읽기 때문에 arXiv를 통한 연구 성과의 배포는 필수적이라고 하였다. 또한 수학 분야의 조교수(ID 2109)는 학술지 논문 심사를 통과하는데 걸리는 시간이 1년 여 임을 감안할 때 arXiv를 통하여 빠르게 연구 성과를 배포할 수 있다는 것이 장점을 언급하였다. 뿐만 아니라 물리학 분야의 다른 조교수(ID 1839)는 새로운 연구 결과를 발견하고 자신과 연구 관심사가 비슷한 연구자들을 파악할 수 있다는 점에서 arXiv는 매우 유용한 자원임을 강조하였다.

설문 결과 교수들의 기관 레포지토리 수용도는 저조한 것으로 나타났으나, 기관 레포지토리에 실제 셀프 아카이빙하는 교수들은 그 가능성에 대해 긍정적으로 평가하였다. 이들은 기관 레포지토리의 중요성으로 다양한 형태의 디지털 학술 저작물에 대한 접근과 보존을 언급하였으며 한 사회학 교수(ID 1803)의 말을 인용하면 다음과 같다.

Institutional repositories are really impor-

tant because that's a place where there's easy access to your work over a longer period of time(ID 1803).

보건 의료 정보학(health informatics) 분야의 조교수(ID 1958)는 기관 레포지토리를 통해 텍스트 자료 이외의 저작물, 예를 들어 포스터 발표 자료나 연구 논문에 속한 도표 등을 배포할 수 있으며 이를 통해 자신의 연구 성과를 알릴 수 있다고 하였다. 그는 기관 레포지토리에 셀프 아카이빙한 자신의 포스터를 통해 미국 국립 보건원(National Institute of Health)의 한 연구 부서(study section)로부터 발표자로 초청받을 수 있었던 예를 들면서 기관 레포지토리의 뛰어난 접근성을 강조하였다. 또한 역사학 분야의 부교수(ID 1756)는 심사 없는 논문과 같은 저작물의 경우 배포할 수 있는 경로가 제한적이기 때문에 기관 레포지토리는 이와 같은 종류의 저작물들을 배포할 수 있는 새로운 통로가 될 수 있다고 하였다. 그러나 이렇게 심사를 거치지 않은 저작물들은 배포 이전에 그 질적 수준에 대한 평가가 선행되어야 함을 강조하였으며 이를 인용하면 다음과 같다.

Even though I've circumvented peer review just by publishing my own things [unrefereed articles] [in the institutional repository], I myself would say they don't really count as knowledge unless some mechanism is put in place to have someone else review it and say that's knowledge. So peer review is still the unsolved problem in online publication(ID 1756).

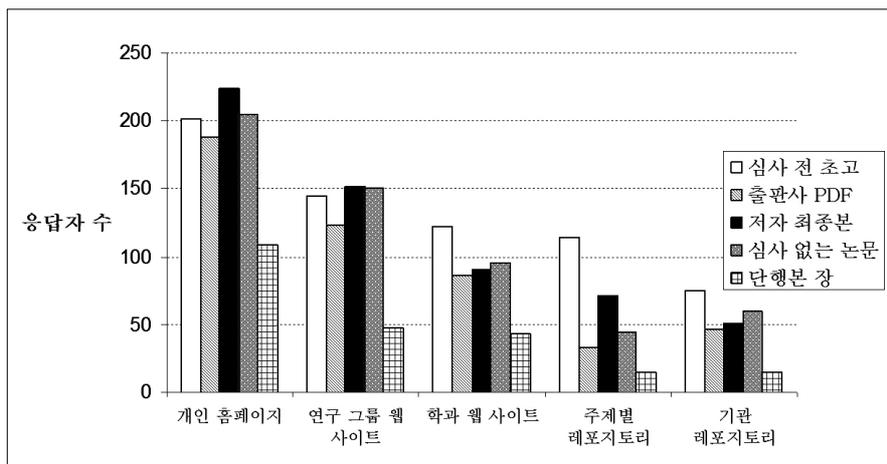
### 4.2 셀프 아카이빙 저작물

선행 연구에서도 나타났듯이 셀프 아카이빙의 대상이 되는 저작물은 주로 학술지 논문이며 본 연구에서는 논문 심사의 여부와 진행 과정에 따라 (1) 심사 전 초고; (2) 심사를 거친 저자의 논문 최종본; (3) 출판사의 PDF 본; (4) 심사 없는 논문; (5) 단행본의 장으로 구분하여 조사하였다. 특히 심사를 거친 논문의 경우 저자 최종본과 출판사 PDF본으로 나누어 그 차이에 대한 교수들의 인식과 셀프 아카이빙 양식을 살펴보았다.

〈그림 2〉는 이들 저작물들이 앞 장에서 기술한 셀프 아카이빙 저장소에 어떤 식으로 분포되어 있는지를 보여주며 응답자들이 설문에 응할 시점을 기준으로 지난 3년 간의 셀프 아카이빙 방식에 대한 데이터를 바탕으로 하였다. 셀프 아카이빙 저장소 별로 그 분포에 차이를 보이는 하지만 심사를 거친 논문들, 특히 저자 최종본의 셀프 아카이빙 빈도가 높은 것으로

나타났다. 출판사의 PDF본도 자주 셀프 아카이빙되고 있는 것으로 나타났는데 면담에 응한 교수들의 상당수가 자신의 최종본보다는 PDF본이 진정한 의미의 최종본이며 읽기 쉽다는 점에서 이를 선호하였다. 이는 교수들이 출판사의 PDF본을 VoR(Version of Record)로 믿고 있다는 것을 나타낸다. 그러나 출판사에서 PDF본의 셀프 아카이빙을 허용하는지의 여부가 이를 셀프 아카이빙하는데 영향을 미치는지에 대해서는 교수들마다 상이한 반응을 보였다. 물리학 분야의 한 교수(ID 2337)는 출판사의 PDF본을 셀프 아카이빙 하는 것이 출판사의 정책에 위배된다는 생각은 들지만 이를 셀프 아카이빙하면서 직접 확인한 적은 없다고 답했으나, 기계 공학 분야의 조교수(ID 1435)는 출판사의 PDF를 선호하지만 출판사의 허가 없이 셀프 아카이빙할 의사는 없다고 밝혔다.

어떤 교수들은 심사를 거치지 않은 논문을 웹 상에 공개적으로 배포하기에는 질적 수준의 측면상 부적합하다는 의견을 보였으며 심사를



〈그림 2〉 교수들의 셀프 아카이빙 저장소 별 저작물의 종류와 셀프 아카이빙 빈도

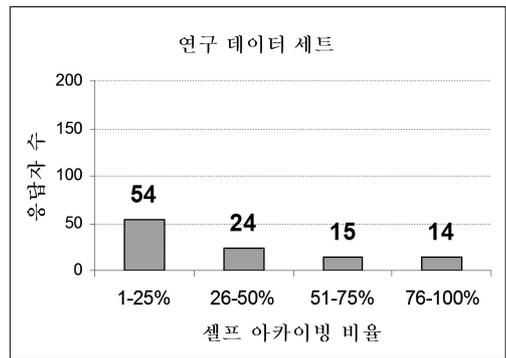
거쳐 출판된 논문을 셀프 아카이빙함으로써 자신들만의 연구 아이디어가 보호될 수 있다고 믿었다. 또한 무기명 심사(blind review)를 통해 심사 중인 논문은 심사가 마무리되어 출판이 허가될 때까지 공개될 수 없기 때문에 셀프 아카이빙 할 수 없다는 의견도 있었다. 이와 같이 교수들은 학술지 논문 심사가 학술 커뮤니케이션의 질적 수준을 유지하는 핵심 기제이며 이는 셀프 아카이빙에서도 예외가 될 수 없음을 강조하였다.

그러나 심사 전 초고의 경우 주제별 레포지토리, 기관 레포지토리, 그리고 학과 홈페이지를 통해 심사를 거친 논문들보다 많이 셀프 아카이빙되는 것으로 나타났다. 주제별 레포지토리, 특히 arXiv의 경우 심사 전 초고를 셀프 아카이빙하는 것이 일반적이므로 그에 대한 빈도가 다른 종류의 저작물에 비해 높았다. 심사 없는 논문 역시 개인 홈페이지나 연구 그룹 웹사이트를 통해 자주 셀프 아카이빙되는 것으로 드러났다. 심리학 분야의 한 교수(ID 1457)는 기술 보고서들을 개인 홈페이지에 셀프 아카이빙 하는데 이들이 학술지에 출판할 만한 수준의 내용은 아니지만 그 주제 분야에서 중요한 기술적인 정보를 담고 있기 때문이라고 했다.

단행본의 장을 셀프 아카이빙하는 빈도는 대체적으로 낮았으며 한 설문 응답자(ID 1327)는 자신이 저술한 단행본은 판매되는 것이기 때문에 셀프 아카이빙을 통해 무상으로 웹 상에 공개하기를 원하지 않는다고 했다. 그러나 면담에 응한 한 수학과 교수(ID 2112)는 미간행 교과서(unpublished textbook)를 개인 홈페이지에 셀프 아카이빙 하고 있었으며 단행본으로 이를 출판할 계획이 없기 때문에 셀프 아

카이빙을 통하여 자신의 교과서가 널리 이용될 수 있다는 것에 만족하고 있었다.

그 밖에 중요한 셀프 아카이빙 대상으로 연구 데이터 세트가 있다. 본 연구에서는 셀프 아카이빙에 참여하는 교수들에게 설문에 응할 당시를 기점으로 지난 5년 간 자신이 가지고 있는 연구 데이터 세트를 어느 정도 비율로 셀프 아카이빙하였는지 질문하였다. 연구 데이터를 셀프 아카이빙하는 교수들은 모두 107명이었으며 이 중 자신이 가진 데이터 세트 중 25% 이하를 셀프 아카이빙한다는 교수들이 54명(50.5%)으로 가장 많았다(그림 3 참조).



〈그림 3〉 연구 데이터 세트의 셀프 아카이빙 비율

일반적으로 생태학이나 분자 생물학 분야 교수들 간에 연구 데이터 공유가 활발하며 특히 해당 분야 학회나 미국 과학 재단 등 연구비 지원 기관에서 연구 데이터의 셀프 아카이빙을 장려하고 있다. 이와 관련하여 생태학 분야의 한 교수(ID 1507)은 미국 생태학회(Ecological Society of America)에서 제공하는 데이터 등록 프로그램(data registry program)을 언급하였다. 이를 통해 해당 학회 저널의 논문 저자

들은 논문에 이용된 연구 데이터 세트를 메타 데이터와 함께 제출해야 하며 제출된 데이터들은 데이터 등록 웹 사이트<sup>4)</sup>를 통해 오픈 액세스로 검색 가능하다. 뿐만 아니라 이 교수는 기관 레포지토리와 자신이 속해 있는 연구 실험실 웹 사이트에도 데이터 세트를 셀프 아카이빙하고 있었는데 이렇게 복수의 저장소에 셀프 아카이빙함으로써 데이터의 유실을 막고 접근을 향상시킬 수 있다고 설명하였다.

교수들은 셀프 아카이빙을 통한 연구 데이터의 공유가 과학의 발전을 가속시킬 것이라는 데 동의하였지만, 연구 데이터가 오픈 액세스로 제공됨으로써 데이터가 오용될 가능성이 높아진다는 것에 우려를 표시하였다. 또한 어떤 교수들은 인간을 대상으로 하는(human subjects) 연구를 통해 얻어진 데이터를 주로 사용하는데 이러한 데이터 세트에는 여러 가지 개인 정보가 포함될 수 있고 이러한 정보가 셀프 아카이빙을 통해 유출되지 않도록 주의해야 하기 때문에 데이터 관리가 쉽지 않다는 점을 언급하기도 했다.

논문, 단행본 장 및 연구 데이터 세트 이외에도 교수들은 자신들의 연구 과정에서 생산된 다양한 형태의 디지털 자료들을 셀프 아카이빙하고 있었다. <표 8>에 따르면 학회 발표 자료를 셀프 아카이빙하는 비율이 가장 높았으며 (48.5%), 이미지 파일과 학위 논문이 그 뒤를 이었다.

학회 발표 자료, 특히 파워포인트 슬라이드는 이용자가 연구의 내용을 빨리 이해하고자 할 때 유용하기 때문에 셀프 아카이빙한다는 의견이 있었으며, 실무자들에게 많이 읽히기

때문에 이를 셀프 아카이빙한다는 응답도 있었다. 셀프 아카이빙되는 기타 자료로는 실험실 프로토콜, 연구에 이용된 설문지, 측정 도구 (measurement scales) 등이 있었다.

<표 8> 교수들이 셀프 아카이빙하는 기타 연구 관련 자료

	빈도	퍼센트
학회 발표 자료	233	48.5
이미지	190	39.6
학위 논문	148	30.8
소프트웨어	126	26.3
소프트웨어 관련 문서	94	19.6
오디오/비디오	76	15.8
기타	24	5.0

### 4.3 요약

본 연구에서는 셀프 아카이빙의 의미를 주체별 레포지토리나 기관 레포지토리에 한정하지 않고 개인 홈페이지 등 다양한 종류의 저장소를 통해 연구 저작물을 웹 상에 오픈 액세스로 공개하는 것으로 정의하였다. 셀프 아카이빙에 참여한다고 응답한 480명의 교수 중 66.7%에 해당하는 320명이 개인 홈페이지에 셀프 아카이빙하고 있었으며 연구 그룹 웹사이트를 이용하는 교수들이 51.5%로 그 뒤를 이었다. 복수 응답을 허용하였는데 두 종류 이상의 셀프 아카이빙 저장소를 활용하는 교수들이 74.6%를 차지해 대다수의 교수들이 복수의 저장소에 셀프 아카이빙하고 있음을 알 수 있다. 셀프 아카이빙의 대상이 되는 연구 저작물 중 심사를 거친 논문들을 개인 홈페이지나 연구 그룹 홈페이지

4) ESA(Ecological Society of America) Data Registry. <<http://data.esa.org/esa/style/skins/esa/index.jsp>>.

이지에 셀프 아카이빙하는 빈도가 가장 높았다. 관리의 용이성이 개인 홈페이지를 이용하는 가장 큰 이유였으며 이는 기관 레포지토리를 이용하지 않는 이유로도 자주 언급되었다. 주제별 레포지토리의 이용자들은 주로 물리학, 수학 분야 교수들인 것으로 나타났는데 이들 분야에서의 셀프 아카이빙은 arXiv를 통해 활발히 이루어짐을 알 수 있었다. 기관 레포지토리에 셀프 아카이빙하는 교수들의 수는 다른 저장소에 비해 적었지만 이들은 기관 레포지토리가 다양한 형태의 디지털 연구 저작물, 예를 들어 연구 데이터 세트나 심사 없는 논문 등을 인터넷을 통해 신속히 접근 가능하게 하고 보존한다는 측면에서 그 가능성을 제시하였다.

## 5. 결론 및 제언

본 연구의 참여자인 미국 연구 중심 대학 교수들은 주로 개인 홈페이지나 연구 그룹 웹사이트를 통하여 심사를 거친 논문을 셀프 아카이빙하는 것을 선호하였다. 즉, 학술지 논문 심사를 통해 연구 논문의 질적 수준을 평가하고 유지하되 이를 웹 상에 오픈 액세스로 공개함으로써 검색 엔진을 이용하여 누구나 신속하게 학술 논문을 열람할 수 있게 하는 것이다. 이는 학술지의 논문 심사 과정이 셀프 아카이빙에서도 여전히 중요한 요소임을 보여 주고 있다.

그러나 셀프 아카이빙에 참여한다고 응답한 480명의 교수들 중 273명(56.9%)이 출판사의 PDF본을 셀프 아카이빙하고 있는 것으로 드러

났으며 이는 분석 대상자 중 과반수 이상의 교수들이 출판사의 정책과는 무관하게 셀프 아카이빙을 수행하고 있다는 것을 보여 준다. 실제 연구자들의 대부분이 학술 논문을 저널에 출판할 때 저작권을 배타적 권리(exclusive right)로서 출판사에 양도하고 있으며 복사 및 배포에 관한 권리도 이에 포함되기 때문에 출판된 논문의 셀프 아카이빙은 저작권 문제와 필연적으로 연결된다. 그러므로 기관 레포지토리와 같이 셀프 아카이빙에 바탕을 둔 서비스 제공자들은 출판사의 셀프 아카이빙에 대한 정책을 이해하고 기관 레포지토리에 셀프 아카이빙하는 이용자들에게도 이를 알리고 교육할 필요가 있다. 최근 조사된 영미권과 유럽 일부 국가 523개 출판사의 셀프 아카이빙 정책에 따르면 63.2%가 출판 후 논문의 셀프 아카이빙을 허용하고 있으며<sup>5)</sup> 이는 출판사들이 저자들의 셀프 아카이빙을 인정하는 추세로 변화하고 있음을 보여준다.

또한 교수들은 심사를 거친 논문 이외에도 다양한 디지털 형태의 연구 결과물들을 셀프 아카이빙 하고 있었다. 심사 전 초고나 심사 없는 논문, 학술 논문의 바탕이 되는 연구 데이터나 논문의 일부인 도표, 그리고 논문의 축약형인 학회 발표 자료나 포스터 등이 이에 속하는데 이들은 기존의 출판 체제에서는 배포 경로를 찾기 어려운 자료들이다. 기관 레포지토리는 이와 같은 자료들을 접근 가능하게 하고 장기적으로 보존하는 저장소로서 그 역할을 수행할 수 있으며 본 연구에서 기관 레포지토리에 셀프 아카이빙하는 교수들도 이와 같은 측

5) Journal Policies - Summary Statistics So Far. <<http://romeo.eprints.org/stats.php>>.

면에서 기관 레포지토리의 가능성을 높이 평가하였다. 위에서 열거한 자료들은 특정 학술 논문과 연관되어 그 논문의 일부분으로서, 혹은 논문의 바탕이 되는 연구 데이터로서 이용될 수 있는데, Lynch(2007)는 이와 관련하여 이용자들이 각각의 자료를 하나의 관련된 정보로 열람할 수 있게 하는 사이버 인프라(Cyberinfrastructure)가 필요하다는 것을 강조하였으며 이것은 학술 정보를 관리하고 이용하는 행동 및 관련 정책을 변화시키는 새로운 패러다임이 될 것이라 전망했다. 미국의 기관 레포지토리 역시 넓은 의미에서 이와 같은 학술 정보 교류 패러다임 변화의 일환으로 운영되고 있으며 이렇게 다양한 형태의 학술 자료들을 어떻게 효과적으로 관리하고 보존할 것인지에 관심이 모아지고 있다.

디지털 보존은 교수들이 인식하는 기관 레포지토리의 가장 핵심적인 기능이지만 실제 미국의 기관 레포지토리들은 디지털 보존보다는 자료의 수집(content recruitment)에 더 치중하는 경향을 보였다(Markey et al., 2007). 이는 기관 레포지토리 서비스와 이용자의 요구 사이의 괴리를 단적으로 보여주는 것이며, 이를 개

선하기 위해 장기적인 관점에서 기관 레포지토리의 디지털 보존 서비스를 정립하는 것이 필요하다. 미국 연구 도서관 그룹(Research Library Group)과 OCLC에서 개발한 신뢰할 수 있는 디지털 보존소(Trusted Digital Repository)의 요건과 이를 위한 인증 시스템은 기관 레포지토리에 디지털 보존 기능을 통합하는 바탕이 될 수 있다(OCLC-CRL 2007). 이와 더불어 기관 레포지토리는 디지털 보존에 필요한 도구와 전문성 개발을 위한 실험적인 장소로 활용될 수 있으며, 이와 같은 노력을 통해 대학 내 다양한 학술 정보의 보존소로서 거듭날 수 있을 것이다. 실무적인 관점에서 기관 레포지토리는 디지털 보존에 대한 임무와 한계를 명확히 하고 이를 이용자들에게 알리며, 저작물의 수집시 디지털 보존에 필요한 메타데이터를 갖출 수 있게 하는 도구와 지침을 개발하는 것이 필요하다(McGovern and McKay 2008). 디지털 보존 기능은 기관 레포지토리의 가치를 높이고 대학 내 연구자들의 셀프 아카이빙을 유도하는 동기 요인이 될 수 있으며 이에 대한 연구와 모범 사례(best practices)의 개발이 계속되어야 할 것이다.

## 참 고 문 헌

- [1] Allen, J. 2005. *Interdisciplinary differences in attitudes towards deposit in institutional repositories*, M. A. thesis, Manchester Metropolitan University. [online]. [cited 2010.02.17]. <<http://eprints.rclis.org/archive/00005180/01/FULLTEXT.pdf>>.
- [2] Andrew, T. 2003. "Trends in self-posting of research material online by academic staff." *Ariadne*, 37. [online]. [cited 2010.02.17]. <<http://www.ariadne.ac.uk/issue37/andrew/>>.
- [3] Antelman, K. 2006. "Self-archiving practice and the influence of publisher policies in the

- social sciences.” *Learned Publishing*, 192: 85-95.
- [4] Borgman, C. L. 2007. *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press, 47-74.
- [5] Brown, C. 2001. “The E-volution of preprints in the scholarly communication of physicists and astronomers.” *Journal of the American Society for Information Science and Technology*, 523: 187-200.
- [6] Coleman, A., & J. Roback. 2005. “Open Access Federation for Library and Information Science: dLIST and DL-Harvest.” *D-Lib Magazine*, 1112. [online]. [cited 2010.02.17]. <<http://www.dlib.org/dlib/december05/coleman/12coleman.html>>.
- [7] Covey, D. T. 2009. “Self-archiving journal articles: A case study of faculty practice and missed opportunity.” *portal: Libraries and the Academy*, 92: 223-251.
- [8] Davis, P. M., & M. J. L. Connolly. 2007. Institutional Repositories: Evaluating the reasons for non-use of Cornell University’s installation of DSpace. *D-Lib Magazine*, 133/4. [online]. [cited 2010.02.17]. <<http://www.dlib.org/dlib/march07/davis/03davis.html>>.
- [9] Doorn P., & H. Tjalsma 2007. “Introduction: archiving research data.” *Archival Science*, 7: 1-20.
- [10] Foster, N. F., & S. Gibbons. 2005. “Understanding faculty to improve content recruitment for institutional repositories.” *D-Lib Magazine*, 111. [online]. [cited 2010.02.17]. <<http://www.dlib.org/dlib/january05/foster/01foster.html>>.
- [11] Gadd, E., C. Oppenheim, & S. Probeta. 2003. “RoMEO studies 2: How academics want to protect their open-access research papers.” *Journal of Information Science*, 293: 333-356.
- [12] Griffiths, A. 2008. “The publication of research data: Researcher attitudes and behaviour.” Paper presented at the 4th International Digital Curation Conference, Scotland: Edinburgh.
- [13] Guédon, J. C. 2003. “Open access archives: From scientific plutocracy to the republic of science.” *IFLA Journal*, 292: 129-139.
- [14] Gunnarsdóttir, K. 2005. “Scientific journal publication: On the role of electronic preprint exchange in the distribution of scientific literature.” *Social Studies of Science*, 354: 549 - 579.
- [15] Harnad, S. 2003. “Open Access to peer-reviewed research through author/institution self-archiving: Maximizing research impact by maximizing online access.” *Journal of Postgraduate Medicine*, 494: 337-342.
- [16] Harnad, S., & Brody, T. 2004. “Comparing the impact of Open Access OA vs. Non-OA articles in the same journals.” *D-Lib Magazine*, 106. [online]. [cited 2010.02.17].

- <<http://www.dlib.org/dlib/june04/harnad/06harnad.html>>.
- [17] Jantz, R. C., & M. C. Wilson 2008. "Institutional repositories: Faculty deposits, marketing, and the reform of scholarly communication." *Journal of Academic Librarianship*, 343: 186-195.
- [18] Kankanhalli, A., B. C. Y. Tan, & K.-K. Wei. 2005. "Contributing knowledge to electronic knowledge repositories: an empirical investigation." *MIS Quarterly*, 291: 113-143.
- [19] J. Kim, 2007. "Motivating and impeding factors affecting faculty contribution to institutional repositories." *Journal of Digital Information*, 82. [online]. [cited 2010.02.17]. <<http://journals.tdl.org/jodi/article/view/193/177>>.
- [20] J. Kim. 2010. "Faculty self-archiving: Motivations and barriers." *Journal of the American Society for Information Science and Technology*, forthcoming.
- [21] Kling, R. 2004. "The Internet and unrefereed scholarly publishing." *Annual Review of Information Science and Technology*, 39: 591-631.
- [22] Kling, R., L. Spector, & G. McKim. 2002. "Locally controlled scholarly publishing via the Internet: The Guild Model." *The Journal of Electronic Publishing*, 8. [online]. [cited 2010.02.17]. <<http://www.press.umich.edu/jep/08-01/kling.html>>.
- [23] Lawal, I. 2002. "Scholarly communication: the use and non-use of e-print archives for the dissemination of scientific information." *Issues in Science & Technology Librarianship*, 36. [online]. [cited 2010.02.17]. <<http://www.istl.org/02-fall/article3.html>>.
- [24] Lercher, A. 2008. "A survey of attitudes about digital repositories among faculty at Louisiana State University at Baton Rouge." *Journal of Academic Librarianship*, 345: 408-415.
- [25] Lynch, C. A. 2003. "Institutional repositories: essential infrastructure for scholarship in the digital age." *ARL*, 226: 1-7.
- [26] Lynch, C. A., & J. K. Lippincott. 2005. "Institutional repository deployment in the United States as of early 2005." *D-Lib Magazine*, 119. [online]. [cited 2010.02.17]. <<http://www.dlib.org/dlib/september05/lynch/09lynch.html>>.
- [27] Lynch, C. 2007. "The shape of the scientific article in the developing cyberinfrastructure." *CTWatch Quarterly*, 33: 5-10.
- [28] Manuel, K. 2001. "The place of e-prints in the publication patterns of physical scientists." *Science & Technology Libraries*, 201: 59-85.
- [29] Markey, K, Rieh, S.Y., St. Jean, B., Kim, J., & Yakel, E. 2007. "Census of Institutional Repositories in the United States MIRACLE project research findings." *CLIR publication*, 140: 167. [online]. [cited 20100217]. <<http://www.clir.org/pubs/reports/pub140/contents.html>>.

- [30] McGovern, N.Y., & A.C. McKay. 2008. "Leveraging short-term opportunities to address long-term obligations: A perspective on institutional repositories and digital preservation programs." *Library Trends*, 572: 262-279.
- [31] Meyer, E. T., & Kling, R. 2002. *Leveling the playing field, or expanding the bleachers? Socio-Technical Interaction Networks and arXiv.org Working paper*. Bloomington: Indiana University Center for Social Informatics. No. WP-02-10
- [32] Ober, J. L. 2005. *Postprint repository services: context and feasibility at the University of California*. California: The University of California's Office of Scholarly Communication. [online]. [cited 2010.02.17].  
<[http://osc.universityofcalifornia.edu/responses/materials/UC\\_postprintstudy\\_final.pdf](http://osc.universityofcalifornia.edu/responses/materials/UC_postprintstudy_final.pdf)>.
- [33] OCLC-CRL 2007. "Trustworthy Repositories Audit & Certification TRAC: Criteria and checklist." [online]. [cited 2010.02.17].  
<[http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)>.
- [34] Pinfield, S. 2001. "How do physicists use an e-print archive? Implications for institutional e-print services." *D-Lib Magazine*, 712. [online]. [cited 2010.02.17].  
<<http://www.dlib.org/dlib/december01/pinfield/12pinfield.html>>.
- [35] Rowlands, I., D. Nicholas., & P. Huntington. 2004. "Scholarly communication in the digital environment: What do authors want?" *Learned publishing*, 17: 261-273.
- [36] Swan, A., & S. Brown. 2004. "ISC/OSI Journal author survey report:" JISC, HEFCE. [online]. [cited 2010.02.17]. <<http://eprints.ecs.soton.ac.uk/11002/01/JISCOAreport1.pdf>>.
- [37] Swan, A., & S. Brown. 2005. *Open access self-archiving: an author study*. Key Perspective Ltd. [online]. [cited 2010.02.17]. <<http://cogprints.org/4385/1/jisc2.pdf>>.
- [38] Wojciechowska, A. 2007. "Analysis of the use of Open Archives in the fields of Mathematics and Computer Science." *OCLC Systems & Services*, 231: 54-69.
- [39] Xia, J. 2007. "Assessment of self-archiving in institutional repositories: Across disciplines." *Journal of Academic Librarianship*, 336: 647-654.