

# 데이터 인용의 현황과 제언

## The Current State and Recommendations for Data Citation

김지현 (Jihyun Kim)\*

정은경 (EunKyung Chung)\*\*

윤정원 (JungWon Yoon)\*\*\*

이재윤 (Jae Yun Lee)\*\*\*\*

### 초 록

학술 커뮤니티 내에서 논문의 인용은 보편적인 규범으로 자리 잡은 데 비해 데이터의 인용은 아직 초보적인 단계에 머물러 있다. 이를 개선하기 위해 제기되고 있는 데이터 인용의 필요성 및 원칙과 가이드라인에 대해서 살펴보았다. 또한 데이터 인용체계 구축 사례에서는 데이터 인용 요소들을 정의하고 서비스를 제공하는 DataCite, Dataverse Network, Data Citation Index 사례를 중심으로 살펴보았다. 마지막으로 한국종합사회조사 데이터 인용 분석을 통해 국내 데이터세트 인용/이용 정보 제공 실태를 조사하였다.

### ABSTRACT

Data citation remains in its infancy, although providing the citation to a journal article is a typical norm in an academic community. This study examines the need for data citation, its principles and guidelines for improving the issue. In addition, the study investigates cases that established data citation mechanism, including DataCite, Dataverse Network and Data Citation Index that define elements of data citation and provide relevant services. At the end, it explores the current state of data citation in Korea through the analysis of citations to dataset from Korean General Social Survey.

키워드: 데이터 인용, 인용지침, 데이터세트, 연구 데이터, 한국종합사회조사  
data citation, citation guideline, dataset, research data, Korean General Social Survey

\* 이화여자대학교 사회과학대학 문헌정보학전공 부교수(kim.jh@ewha.ac.kr) (제1저자)

\*\* 이화여자대학교 사회과학대학 문헌정보학전공 부교수(echung@ewha.ac.kr) (공동저자)

\*\*\* School of Information, University of South Florida, Tampa, Florida, U.S.A.  
(jyoon@usf.edu) (공동저자)

\*\*\*\* 명지대학교 인문대학 문헌정보학과 부교수(memexlee@mju.ac.kr) (교신저자)

■ 논문접수일자: 2016년 12월 5일 ■ 최초심사일자: 2016년 12월 19일 ■ 게재확정일자: 2017년 3월 15일

■ 정보관리학회지, 34(1), 7-29, 2017. [http://dx.doi.org/10.3743/KOSIM.2017.34.1.007]

## 1. 서론

과학적 지식(scientific knowledge)의 일환으로 동료 심사를 통해 출판된 논문은 인용을 통해 집합적으로 연결되어 있으며 이들 논문은 방대한 양의 데이터에 기반을 두고 있다. 논문에서는 데이터의 분석 결과를 제시함으로써 주장에 대한 근거를 보여주는 것이 일반적이지만, 실제로 논문을 읽은 연구자가 해당 데이터에 접근할 수 있도록 하는 규범이나 실천은 미흡한 실정이다. 데이터가 과학적 지식의 영역에서 논문만큼이나 중요한 부분임에도 불구하고 데이터를 논문 내에서 공식적으로 인용하는 경우가 드물고 데이터 인용에 대한 가이드라인 역시 부족하다는 문제점이 제기되고 있다(Parsons, Duerr, & Minster, 2010).

이러한 문제점을 해결하기 위한 방안으로 해외에서는 데이터 인용을 통해 얻을 수 있는 학술적 가치와 연구 성과의 인정 등 긍정적인 효과에 주목하고 데이터 인용의 원칙과 인용 체계에 관한 논의를 활발히 진행하고 있다. 이에 비해 국내에서는 데이터 인용에 관한 논의가 매우 제한적으로 이루어지고 있는 상황이다. 따라서 본 연구는 해외에서 제안된 데이터 인용 원칙과 체계에 대한 동향을 살피고 국내의 데이터 인용 현황 분석을 토대로 개선방안을 제시하는 것을 목적으로 한다.

이를 위해 본 연구에서는 학술 커뮤니케이션 및 연구데이터 관리 분야의 대표적인 기관 또는 커뮤니티에서 제시하고 있는 데이터 인용의 원칙과 가이드라인을 통해 데이터 인용의 개념과 형식에 대한 논의를 살펴보았다. 또한 데이터 인용체계를 구축하고 있는 대표적인 사례인

DataCite, Dataverse Network와 Data Citation Index를 중심으로 데이터 인용의 구성요소와 데이터 인용이 활용되는 방식을 조사하였다.

데이터 인용에 대한 해외 동향을 살펴본 다음에는 실제 국내 연구 논문에서 데이터 인용이 어떻게 이루어지고 있는지를 조사해보았다. 해외 데이터 인용색인을 분석한 선행연구(조재인, 2016)에서 사회, 경제 등 사회과학 분야의 서베이 데이터가 가장 많이 인용되고 있는 것으로 나타났으므로, 국내 사회과학 분야에서 가장 널리 이용되고 있는 한국종합사회조사(Korean General Social Survey) 데이터를 인용한 참고문헌 사례를 살펴보았다. 한국학술지인용색인의 참고문헌 검색 기능을 통해 추출된 사례들을 바탕으로 한국종합사회조사 데이터가 인용되고 있는 학문 분야와 해당 데이터의 참고문헌 표기 형식을 분석하였다. 이를 통해 데이터 인용에 포함되어야 하는 핵심적인 요소들의 제공 여부를 조사 분석하였고 데이터 인용 활성화를 위해 관련 기관들의 역할을 강화할 것을 제안하였다.

## 2. 데이터 인용의 필요성 및 원칙

### 2.1 데이터 인용에 따른 긍정적 효과

데이터 인용을 수행함으로써 발생할 수 있는 긍정적인 효과는 다양한 연구에서 논의되어 왔으며 이는 학술지 논문의 인용을 통해 얻을 수 있는 이익과 유사하다. 즉 학술지 논문 인용을 통해 저자는 해당 연구에 대한 인정을 받을 수

있으며, 인용분석을 통해 연구의 영향력을 측정하여 그 결과가 저자와 저자가 속한 기관에 보상체제로 작용할 수 있다. 뿐만 아니라 인용은 논문의 이용도를 나타내므로 해당 논문의 지속적인 서비스 제공 여부를 판단할 수 있는 근거가 된다(Ball & Duke, 2015). 이러한 긍정적 효과로 인해 학술 커뮤니티 내에서 논문의 인용은 보편적인 규범으로 자리 잡은 데 비해 데이터의 인용은 아직 초보적인 단계에 머물러 있다. 이를 개선하기 위해 제시되고 있는 데이터 인용의 필요성은 다음의 네 가지 측면에서 설명할 수 있다.

첫째, 데이터 인용을 통해 연구자에게 데이터 생산과 공유에 대한 인정 및 보상을 제공할 수 있다(Mayernik, 2012; Mooney & Newton, 2012). 데이터 공유는 최근 미국이나 영국의 연구비 지원기관을 중심으로 연구자들에게 의무 또는 권고사항으로 제시되고 있다. 데이터를 공유하기 위해서는 데이터를 정비하고 메타데이터를 제공하는 등 연구자들의 시간과 노력이 투자되어야 하는 반면 이러한 노력에 대한 명확한 보상체계가 없다는 것이 데이터 공유를 저해하는 요인으로 지적되고 있다. 또한 연구자들이 데이터 공유를 꺼리는 대표적인 요인으로 데이터의 오용에 대한 우려와 공식적인 인정체계의 부족을 들 수 있다. 이와 같은 문제점을 완화하고 데이터의 공유와 재이용을 촉진할 수 있는 방안으로 데이터의 인용이 강조되고 있다(Helbig, Hausstein, & Toepfer, 2015; Kotarski, Reilly, Schrimpf, Smit, & Walshe, 2012).

둘째, 데이터 인용은 연구 결과의 재현과 검증을 가능하게 하며 이를 통해 연구의 투명성

을 향상시킬 수 있다(Callaghan, 2014; Mayernik, 2012). 연구에 대한 신뢰는 전적으로 그 연구의 방법론과 데이터의 투명성 및 데이터에 대한 명확한 인용을 바탕으로 하기 때문에 연구의 재현과 검증을 위해서는 데이터의 공개와 더불어 일관되고 엄격한 데이터 인용 체계가 수립되어야 한다(Parsons, Duerr, & Minster, 2010). 이러한 논의는 빅데이터를 주로 다루는 계산과학(computational science) 분야에서도 제시되고 있는데 예전에는 논문만으로도 가능했던 연구결과에 관한 검증이 빅데이터를 활용하는 연구에서는 불가능함을 지적하고 있다. 결국 다른 연구자들도 재현 가능한 연구(reproducible research)를 제시하기 위해서는 데이터와 데이터의 속성을 기호화한 코드의 공유가 필수적이며 표준화된 인용 체계 내에서 데이터와 코드에 대한 인용이 필수적임을 제안하고 있다(Stodden, 2010).

셋째, 데이터 인용을 통해 데이터의 영향력과 이용을 파악할 수 있다(Ball & Duke, 2015; Mayernik, 2012). 데이터의 영향력을 측정하기 위해서는 일관된 데이터 식별 방법을 사용하는 것이 필요하므로 최근 DOI(Digital Object Identifier)와 같은 디지털 객체 식별 시스템이 공통적으로 데이터 인용에 적용되고 있는 추세이다. 또한 톰슨-로이터(Thomson-Reuters)사에서 2013년 최초의 데이터 인용 색인 시스템인 Data Citation Index(DCI)를 개발하였으며 이를 기반으로 하는 데이터 인용 분석 연구들이 최근 수행되고 있다(Robinson-García, Jiménez-Contreras, & Torres-Salinas, 2015). 아울러 데이터의 재이용 패턴을 분석하기 위한 목적으로 데이터 인용 분석이 수행되고 있으며 데이

터 재이용 분석을 위한 방법론의 개발과 정립이 강조되고 있는 실정이다(Chao, 2011; Piwowar & Vision, 2013).

넷째, 출판된 연구논문과 해당 논문의 기초 자료가 되는 데이터 간의 연계를 강화할 수 있다(Callaghan, 2014; Mayernik, 2012). 데이터 인용을 통해 획득된 데이터에 관한 인용 정보와 기존의 논문 인용 정보를 연계하여 분석에 활용함으로써 출판물 배포와 데이터의 공유를 포함하는 학술 유통 절차를 보다 면밀히 파악하고 궁극적으로는 연구의 재현 가능성을 높일 수 있다(Dragan, Luczak-Rösch, Berendt, Simperl, Packer, & Moureau, 2015). 이와 같이 데이터 인용을 바탕으로 출판물과 데이터 간에 의미 있는 링크를 제공하는 것은 학술 커뮤니케이션의 질적인 수준과 투명성을 제고하는 데 핵심적인 요소라고 할 수 있다(Hoogerwerf, Löscher, Schirwagen, Callaghan, Manghi, Iatropoulou, Keramida, & Rettberg, 2013).

## 2.2 데이터 인용 원칙과 가이드라인

데이터 인용 원칙을 논의한 대표적인 사례들을 살펴보면 우선 미국 과학·공학·의학 한림원(The National Academies of Sciences, Engineering, and Medicine)에서 2012년 주최한 국제 워크숍에서 제안된 데이터 인용 원칙을 들 수 있다. 이 원칙은 다섯 가지 범주인 저작자표시(Attribution), 지속성(Persistence), 접근(Access), 검색(Discovery), 출처(Provenance)로 구성된다. 저작자표시는 데이터 인용이 데이터에 대한 적법하고 학술적인 저작자표시를 가능하게 하는데 활용된다는 것을 의미한다. 지

속성은 데이터 인용이 해당 데이터를 지속적으로 참고하고 관리하는데 활용됨을 나타낸다. 접근은 데이터 인용이 데이터의 장단기적인 접근을 용이하게 하는데 활용된다는 것을 의미하며, 검색은 데이터 인용이 데이터 세트와 데이터로부터 파생되거나 연관된 결과물을 찾아내는데 활용됨을 가리킨다. 출처는 데이터 인용은 출판된 결과물과 이를 지원하는 증거를 연결하고 그 증거가 변경되지 않았음을 검증하는데 활용된다는 것을 나타낸다(Uhlir, 2012).

또한 데이터 인용의 제 1원칙으로 “데이터 인용이 출판에 있어 중요한 대상으로 다루어져야 한다(data citations should be treated as first-class objects for publication)”는 것을 제안하였다. 이는 데이터의 인용이 논문 등 자료에 대한 인용과 마찬가지로 참고문헌에 포함되어야 하며, 데이터를 인용하는 방식 역시 다른 자료를 인용할 때와 비슷한 수준으로 용이해야 함을 의미한다. 데이터 인용에 추가적인 요구 사항을 부여하거나 다른 자료를 인용할 때 포함되는 핵심적인 사항이 누락되는 것은 둘 다 바람직하지 않다는 것을 명시하고 있다.

이러한 원칙은 CODATA(Committee on Data for Science and Technology)와 ICSTI(International Council for Scientific and Technical Information)의 데이터 인용 표준과 실무에 관한 태스크포스(Task Group on Data Citation Standards and Practices)에서 보다 구체적으로 논의되었다. CODATA와 ICSTI는 모두 국제과학연맹이사회(International Council for Science: ICSU)라는 비정부기구에 속해 있으며 데이터 인용 원칙, 데이터 인용 현황, 데이터 인용을 지원하는 기술적 인프라와 사

회 문화적 측면, 향후 연구과제를 제안하였다 (CODATA-ICSTI, 2013).

CODATA-ICSTI에서 제안하고 있는 데이터 인용 원칙은 모두 10가지로 1) 데이터의 위상(status of data); 2) 저작자표시(attribution); 3) 지속성(persistence); 4) 접근(access); 5) 검색(discovery); 6) 출처(provenance); 7) 세분성(granularity); 8) 검증가능성(verifiability); 9) 메타데이터 표준(metadata standards); 10) 유연성(flexibility)이다. 이 중에서 미국 한림원 워크숍에서 논의된 5가지 범주 외의 내용을 살펴보면 데이터의 위상은 데이터 인용이 다른 자료에 대한 인용과 동일한 중요성을 가진다는 것을 나타낸다. 이는 앞서 제시된 데이터 인용의 제 1원칙과 유사함을 알 수 있다. 세분성은

데이터 인용이 데이터를 식별하는데 필요한 세부적인 기술(finest grained description)을 지원해야 함을 의미한다. 검증가능성은 데이터 인용이 데이터를 분명하게 식별할 수 있을 만큼 충분한 정보를 포함해야 한다는 것을 말한다. 메타데이터 표준은 데이터 인용에 널리 적용되는 메타데이터 표준을 활용해야 함을 의미한다. 유연성은 데이터 인용이 커뮤니티 간 상호운용성을 저해하지 않는 수준에서 여러 커뮤니티의 다양한 인용 방식을 포괄할 수 있어야 함을 나타낸다.

2014년 FORCE11 Data Citation Synthesis Group에서는 이러한 기존의 논의들을 종합하여 데이터 인용의 8가지 원칙을 제안하였다(〈표 1〉 참조). FORCE11은 the Future Of Research

〈표 1〉 FORCE11 데이터 인용 원칙

원칙	정의
중요성(Importance)	데이터는 정당하고 인용 가능한 연구의 산물로 간주되어야 한다. 출판물과 같은 다른 자료의 인용과 마찬가지로 데이터 인용에도 동일한 중요성이 부여되어야 한다.
인정 및 저작자표시(Credit and Attribution)	데이터 인용을 통해 데이터의 모든 기여자에게 학술적인 인정과 더불어 규범에 부합하면서 적법한 저작자표시를 제공할 수 있어야 하며, 단일한 형식의 저작자 표시가 모든 데이터에 적용될 수는 없음을 인지해야 한다.
증거(Evidence)	학술 문헌에서 제시된 주장은 언제 어디서든지 데이터에 근거를 두고 있으며 해당 데이터는 반드시 인용되어야 한다.
고유 식별(Unique Identification)	데이터 인용은 데이터 식별을 가능하게 하는 지속적인 방법을 포함해야 하며 그 방법은 기계에 의해 수행될 수 있고 전세계적으로 고유하면서 커뮤니티에서 널리 활용되는 것이어야 한다.
접근(Access)	데이터 인용은 데이터 그 자체와 관련된 메타데이터, 도큐멘테이션, 코드 및 그 밖의 자료에 접근하는 것을 가능하게 해야 하며, 이는 인용된 데이터를 인간과 시스템이 의미 있게 사용하는데 필요한 수준이어야 한다.
지속성(Persistence)	고유식별기호와 데이터의 메타데이터 및 disposition은 지정된 데이터의 수명을 넘어서는 기간에도 지속 가능해야 한다.
구체성 및 검증가능성(Specificity and Verifiability)	데이터 인용은 어떠한 주장을 지원하는 구체적인 데이터를 식별, 접근, 검증하는 것을 가능하게 해야 한다. 인용 및 인용 메타데이터는 데이터의 특정 시분할, 버전 및 세분된 영역을 검증하기에 충분한 정도로 출처와 무결성에 대한 정보를 제공해야 한다.
상호운용성 및 유연성(Interoperability and Flexibility)	데이터 인용은 커뮤니티 간 상호운용성을 저해하지 않는 수준에서 여러 커뮤니티의 다양한 데이터 인용 방식을 충분히 포괄할 수 있을 정도로 유연해야 한다.

출처: Data Citation Synthesis Group (2014). Joint declaration of data citation principles. In Martone, M. (Ed.). San Diego, CA: FORCE11. <https://www.force11.org/datacitation/>

Communication and E-scholarship의 약어인 FORCE와 이 단체가 설립된 연도인 2011년을 나타내는 11을 합친 명칭이다. 연구자, 사서, 아카이스트, 출판사, 연구비지원기관을 포함하는 커뮤니티로 정보기술의 효과적인 활용을 통해 학술 커뮤니케이션의 변화를 주도하는 것을 목적으로 한다(FORCE11, 2016). 이들이 제안한 8가지 원칙은 앞서 제시된 데이터 원칙에 대한 논의와 선행연구들에 기반을 두고 있으며 널리 수용되고 있다(Ball & Duke, 2015).

데이터 인용에 대한 요구가 높아지면서 데이터 아카이브 또는 기존의 인용 양식 가이드를 중심으로 데이터 인용 형식에 대한 가이드라인이 제시되고 있다. 영국의 Data Curation Center (DCC)에서 발간된 데이터 인용 가이드에서는 선행연구에서 공통적으로 제안된 데이터 인용의

요소를 <표 2>와 같이 제시하였다. 이들 구성요소 중에서 가장 핵심적인 요소는 저자, 제목, 날짜, 소장위치 및 출판기관이며, 이를 바탕으로 저자는 데이터 공유 및 생산에 대한 인정을 받을 수 있고 이용자들은 데이터의 적절성을 판단하고 데이터에 접근할 수 있으며 데이터의 품질을 재확인할 수 있다(Ball & Duke, 2015).

기존의 인용 양식 가이드 중에서 APA (American Psychological Association) 양식에서는 데이터에 대한 인용 형식을 정의하고 있는데 인용의 최소 요구사항으로 ‘저자(연도), 제목(버전)[자원 유형].소장위치’를 제시하고 있다(McAdoo, 2013). 이에 추가될 수 있는 인용 항목으로 미시건주립대학교(Michigan State University, 이하 MSU) 도서관 가이드에서는 자원 유형과 소장위치 사이에 데이터세트가 생

<표 2> 데이터 인용 구성요소

구성요소	의미
저자(Author)	데이터세트의 생산자
일자(Publication date)	다음 중 가장 최근의 것으로 1) 데이터를 이용 가능하게 한 일자; 2) 데이터의 품질 확인 절차가 완료된 일자; 3) 데이터의 엠바고가 종료된 일자
제목(Title)	인용된 자원의 제목뿐만 아니라 시설(facility)의 명칭과 데이터세트를 구성하고 있는 최상위 컬렉션 및 주요 하위 컬렉션의 제목도 포함될 수 있음
판(Edition)	데이터가 처리된 수준 혹은 단계를 말하며 데이터가 정제된 정도를 나타냄
버전(Version)	데이터가 변경될 때, 예를 들어 데이터 요소가 추가되거나 데이터 파생 과정을 반복 실행할 경우 그 결과로서 발생된 데이터의 버전 번호가 증가함
하위 데이터세트 표시 (Feature name and URI)	데이터의 하위세트를 기술하기 위한 것으로 feature는 ISO 19101인 지리정보 참조모형에서 제안된 개념이며 feature name은 GridSeries나 ProfileSeries와 같은 데이터세트의 하위 컬렉션을 나타내는 용어들을 지칭함. URI는 feature를 식별하는 기능을 수행
자원 유형(Resource type)	‘데이터베이스’ 또는 ‘데이터세트’ 등의 유형을 나타냄
출판기관(Publisher)	데이터를 호스트하거나 데이터에 대한 품질 확인을 수행하는 기관
Unique Numeric Fingerprint (UNF)	데이터가 생산된 이후 변경되지 않았음을 확인할 수 있는 데이터에 대한 암호화 해시(cryptographic hash)
식별기호(Identifier)	지속적으로 데이터를 식별 가능하게 함
소장위치(location)	영구적인 URL(persistent URL)을 통해 데이터를 이용 가능하게 함

출처: Ball, A., & Duke, M. (2015). How to cite datasets and link to publications. Digital Curation Centre, 1-15. [http://www.dcc.ac.uk/webfm\\_send/525](http://www.dcc.ac.uk/webfm_send/525)

산된 기관(producer)과 데이터를 배포한 기관(distributor)을 표시하고 있다. 데이터세트가 생산된 기관은 저자의 연구를 지원한 기관 또는 데이터의 코드화나 디지털화 등 데이터세트의 생산을 가능하게 한 기관을 의미한다. 데이터를 배포한 기관은 데이터를 다운로드하여 활용할 수 있게 서비스를 제공하는 기관을 말한다(Michigan State University Libraries, 2016).

데이터와 관련된 정부기관에서도 데이터 인용 형식에 관한 가이드를 제공하고 있다. 예를 들어 미국 지질조사국(U.S. Geological Survey, USGS)에서는 데이터 관리에 관한 웹사이트를 별도로 운영하면서 데이터 인용 관련 정보를 제공하고 있다. 이들은 데이터 인용의 핵심 요소로 1) 저자 혹은 연구책임자; 2) 배포 일자/연도; 3) 제목; 4) 버전/판 정보; 5) 아카이브/배포기관; 6) 식별기호; 7) 접근일자과 시간을 제시하고 있다. 또한 데이터 인용을 효과적으로 수행하기 위해서는 데이터세트에 영구식별기호를 부여하고 데이터에 대한 메타데이터를 생성할 수 있는 소프트웨어를 사용하며 데이터 기술 시 표준화된 키워드를 사용할 것을 제안하고 있다. 아울러 논문에서 데이터를 인용할 때 해당 학술지나 출판사에서 요구하는 인용 양식을 준수하고 연구자들에게 데이터 인용 및 공유를 장려할 것을 언급하고 있다(USGS, 2016).

다음 장의 데이터 인용체계 구축 사례에서는 데이터 인용 요소들을 정의하고 서비스를 제공하는 DataCite, Dataverse Network, Data Citation Index 사례를 중심으로 살펴보고자 한다. 또한 해외 대학도서관에서 연구자에게 제공하는 서비스의 하나로 데이터인용을 각 대학도서관의 LibGuide 등에서 제공하는 사례를 찾아볼 수 있다.

### 3. 데이터 인용체계 구축 사례

#### 3.1 DataCite

DataCite는 2005년 이래로 독일 국립과학기술도서관(German National Library of Science and Technology, TIB)은 연구 데이터에 DOI를 부여하는 서비스를 제공하고 있으며, 현재는 국제적인 컨소시엄으로써 4백만 건 이상의 데이터세트에 DOI를 부여하고 있다. 2009년에는 TIB와 영국도서관, 스위스 ETH Zurich 도서관, 프랑스 과학기술정보연구원(Institute for Scientific and Technical Information, INIST), 덴마크 기술정보센터(Technical Information Center), 캐나다 과학기술정보연구원(Institute for Scientific and Technical Information, CISTI), 호주 국립데이터서비스(Australian National Data Service, ANDS), 네덜란드의 TU Delft Library는 협력하여 인터넷 상의 연구 데이터의 접근 향상 증진을 위한 MOU를 체결하였다. 이러한 협력의 목적은 비영리기관으로 DataCite를 설립하여 연구 데이터 셋을 등록하여 영구적인 identifier를 제공하고자 하는 것이다. 그 결과로 연구 데이터세트는 독립적이고, 인용가능하며, 고유한 과학적인 객체로서 존재하게 된다(Brase, 2009; Brase, Lautenschlager, & Sens, 2015).

Starr와 Gastl(2011)은 DataCite의 메타데이터 스키마를 제시하고 있는데, 5개의 필수항목(identifier, creator, title, publisher, and publication year)과 12개의 선택항목(subject, contributor, date, language, resourcetype, alternateidentifier, relatedidentifier, size, format, version, rights,

description)으로 구성되었다. 2016년 5월 14일 발표된 DataCite 메타데이터 스키마는 4.0으로 변화와 진화를 거쳐 초기의 메타데이터 스키마와의 차이를 찾아볼 수 있다.<sup>1)</sup> 우선 필수항목, 추천항목, 선택항목 세 그룹으로 나누었으며, 필수항목은 반드시 입력해야 하며, 추천항목은 상호운용성 증진을 위해서 강력하게 추천된 항목이며, 선택항목은 풍부한 기술을 위한 항목으로 구성되었다. 초기 버전과 비교하면, resourcetype 항목이 선택항목에서 필수항목으로 변경되어 필수항목은 6개로 늘어났다. 선택항목 7개와 추천항목 6개로 총 13개 항목으로 증가하였다. 추천항목은 subject, contributor, date, relatedidentifier, description, geolocation 6개이며, 선택항목 7개는 language, alternateidentifier, size, format, version, rights, fundingreference이다.

실제 다양한 학문 분야의 데이터 리포지터리가 DataCite에 데이터를 등록하여, 접근, 이용, 인용이 가능한 상황이다. Arend, Junker, Scholz, Schüler, Wylie, Lange(2016)는 식물의 발현형체학과 유전체학 분야의 데이터 리포지터리인 Plant Genomics and Phenomics Research Data Repository(PGP)를 소개하였다. PGP 리포지터리는 Leibniz Institute of Plant Genetics and Crop Plant Research에서 호스트하고 있으며, e!DAL이라는 소프트웨어를 사용하여 데이터를 수집하고 있다. 내부 평가 과정을 거친 후에 데이터는 DOI와 메타데이터가 부여되어 DataCite를 통해 출판된다. 또한 기후 분야의 데이터는 Stockhouse, Höck, Toussaint, Lautenschlager(2012)에 의해서 DataCite 활

용이 제시되었다. 기후 분야의 기후와 기후 모델 데이터는 일반적으로 데이터의 규모가 크고, 여러 서버에 분산되어 보관되어 있는 상황이다. 따라서 기후 분야의 데이터는 적절하게 아카이브되고, DOI가 부여되며, DataCite 규정에 따라서 World Data Center for Climate(WDCC) 적용되는 것이 중요하다고 여겨진다. Arend, Lange, Chen, Colmsee, Flemming, Hecht, Scholz(2014)는 생명과학 분야의 데이터 관리 측면의 어려움을 해결하는 방안으로써 e!DAL 소프트웨어를 제안하고 있으며, 그 주요한 기능으로 버전 추적, 메타데이터 관리, 정보검색, DOI등록, 공공 데이터 접근을 위한 임베디드 HTTP서버, 네트워크 파일 시스템 접근, 확장 가능한 저장 기능을 살펴볼 수 있으며, DataCite 컨소시엄의 중요한 구성요소로 역할을 수행하고 있다.

### 3.2 Dataverse Network

Dataverse Network(<http://dataverse.org/>)는 Harvard 대학의 계량사회과학연구소(Institute for Quantitative Social Science)에서 오픈소스 웹 기반 어플리케이션으로 연구 데이터의 공유, 보존, 인용, 탐구, 분석을 목적으로 개발하였다. King(2007)이 제시한 바와 같이 Dataverse Network라는 기술적인 수단을 통해 데이터 공유의 정치적 사회적 문제를 해결하여 궁극적으로 연구 커뮤니티와 개인 연구자에게 혜택이 주어지도록 하는 것이다. 또한 Dataverse Network가 전자적인 데이터가 만연한 환경에서 발생하는 여

1) [https://schema.labs.datacite.org/meta/kernel-4.0/doc/DataCite-MetadataKernel\\_v4.0.pdf](https://schema.labs.datacite.org/meta/kernel-4.0/doc/DataCite-MetadataKernel_v4.0.pdf)

리 가지 문제점에 대응할 수 있는 해결책이라고 제시하였다. 이와 함께, Dataverse Network는 데이터의 이용자들이 공통적인 데이터 이용과 인용 표준을 수용할 수 있도록 연구자 커뮤니티의 인정, 가시화, 데이터 소유권 등과 같은 장려책을 포함하고 있다(Appleton, 2008; Crosas, 2011).

dataverse.org(2016)에서 제시한 바에 따르면, Dataverse Network의 메타데이터에서 인용에 사용되는 요소는 7개로 구성되어 있으며, 저자, 연도, 제목, DOI/Handle, 출판사, 버전, UNF(테이블형식의 데이터)이다. 다른 메타데이터와의 상호운용성을 살펴보면, 우선 Document, Discover and Interoperate(DDI) 스키마 버전 2와 호환이 가능하며, 더블링크어와도 호환이 가능하다. 메타데이터 요소는 인용정보, 초록과 범위, 데이터수집과 방법론, 데이터 이용가능성, 이용조건, 기타 정보, 문서 기술의 총 7개 그룹으로 구성되었다. 각 그룹을 살펴보면, 인용정보는 40개의 하위 요소로 구성되어 있으며, 초록과 범위는 23개, 데이터수집과 방법론은 20개, 데이터 이용가능성은 5개, 이용조건은 8개, 기타 정보는 3개, 문서 기술은 2개로 각각 구성되었다. DataCite와의 비교하여 살펴보면, 상당히 광범위하게 나열된 메타데이터 요소로 볼 수 있으며, 타 메타데이터와의 상호운용성을 위한 매핑 테이블도 갖추고 있다.

한편 Crosas(2013)는 Dataverse Network의 사례를 제시하고 있다. 첫 번째는 선거 데이터(Harvard Election Data Archive)의 Dataverse Network를 이용한 공유 사례이다. Dataverse Network를 통해서 공유되기 전에는 인쇄매체 등을 통해서 한정된 접근만이 가능했으나, Dataverse

Network을 통해 미국 43개 주에서 5,000건 이상의 다운로드가 이루어진 것으로 나타났다. 두 번째 사례는 개별 연구자가 대규모 연구비 지원을 신청하는 경우에 Dataverse Network은 흔히 요구되는 연구 데이터 관리계획 수립에 유용하다. 일례로 NSF(National Science Foundation) 연구비 지원 과정에서 연구지원 요구사항인 데이터관리계획 작성에 있어서 Dataverse Network을 데이터의 지속적인 접근, 저장, 이용에 대한 보장을 가능하게 하였다. 세 번째는 저널의 Dataverse Network 이용 사례이다. Review of Economics and Statistics 저널은 발표되는 논문의 데이터를 Dataverse Network에 공유하도록 하였다. 그 결과 데이터를 공유한 논문은 더 많은 인용을 받았으며, 연구 커뮤니티로부터 데이터 자체에 대한 직접적인 인정을 받은 것으로 나타났다.

Dataverse Network의 사례 이외에도 실제로 사회과학 분야에서 Dataverse Network을 통해서 데이터를 다운 받아 해당 연구를 복제하여 연구를 수행할 수 있도록 제시한 사례를 찾아볼 수 있다. Selway(2011)는 경제학과 정치학 분야에서 흔히 사용되는 인종의 다양성을 구분하는 연구들에서 인종의 구분만이 유일하게 분석의 대상이 되어왔으나, 세 가지 다른 측면의 구분을 통해서 논의가 다시 촉발될 수 있는지를 밝히고자 하였다. 이와 함께 연구에서 사용된 데이터셋이 Dataverse Network을 통해서 다시 이용될 수 있음을 밝혔다. 이와 유사하게 Menaldo(2016)는 정치학 분야의 연구 결과를 제시하면서 관련된 모든 데이터, 코딩, 분석이 Dataverse Network에서 이용가능하다고 명시하였다. 한편 분석의 편리함을 추구하기 위

해서 통계분석도구인 R은 dvn패키지를 사용하여 Dataverse Network의 데이터셋, 분석 파일, 코드북, 관련 메타데이터를 이용할 수 있도록 연계하였다(Leeper, 2014).

### 3.3 Data Citation Index

Data Citation Index는 톰슨로이터에 의해서 구축되었으며, 2012년부터 연구데이터의 색인과 인용정보를 제공하기 시작하였으며(Torres-Salinas & Martín-Martín, 2013) 데이터셋(dataset), 데이터연구(data studies), 리포지터리(repositories)의 세 종류를 포함하고 있다. DCI는 세 종류의 document를 포함하고 있는데, 데이터셋은 단일의 데이터 집합이며, 데이터 스투디는 데이터와 연관된 실험에 대한 기술(description)이다. 마지막으로 리포지터리는 데이터셋과 데이터 스투디를 저장하여 접근을 제공하는 데이터베이스라고 정의할 수 있다(Robinson-García, Jiménez-Contreras, & Torres-Salinas, 2015). 2016년 9월 현재 2백 만 건 이상의 데이터 연구와 데이터셋을 포함하고 있는 데이터 리포지터리가 색인되었으며, 저자명, 기관명, 기관명, 인용 정보 등의 메타데이터가 사용되었다. 또한 Web of Science에 색인된 저널의 논문이 연관이 있는 데이터와 연결되었다(Clarivate Analytics, [n.d.]).

최근에는 데이터 인용이 학문 분야에 미치는 영향에 관한 연구들이 나타나기 시작했다. 리포지터리를 분야별로 살펴보면, 생명과학 분야가 가장 많은 48%를 차지하고, 물리학이 23%, 사회과학이 20%로 주된 분야로 구성되어 있다. 이 외에도 예술/인문학 분야의 리포지터리

가 7%이고 다학제 분야의 리포지터리가 2%를 구성하고 있다(Clarivate Analytics, [n.d.]). Robinson-García, Jiménez-Contreras, Torres-Salinas(2015)가 제시하는 바와 같이 DCI는 학문의 전 영역인 과학, 사회과학, 인문학 분야의 데이터 검색을 위해서 설계되었다. 또한 데이터 인용 관행이 연구 분야에 따라 극명한 차이를 보여준다고 밝혔다. 일반적으로 대부분의 학문분야에서 데이터 인용은 거의 이루어지지 않고 있으나, 결정학(crystallography)과 유전체학(genomics) 분야에서 급격하게 증가하고 있는 추세이다. 전체적인 분야별로 살펴보면, 과학, 공학/기술 분야가 가장 많이 인용되었으나, 사회과학, 예술, 인문학 분야에서는 데이터 스투디가 두각을 나타낸 것으로 제시되었다. Peters, Kraker, Lex, Gumpenberger, Gorraiz(2016)이 제시한 바와 같이 연구 데이터의 대부분(85%)은 인용되지 않고 있지만, 연구 데이터의 인용은 증가추세이다. 연구데이터가 소셜 웹에 미치는 영향을 altmetrics(PlumX, ImpactStory, Altmetrics.com)로 측정해 보면, 상관관계가 없는 것으로 나타났으나 특정 데이터 형태(서베이, aggregate 데이터, 연속데이터)의 경우에는 자주 인용되며, 높은 altmetrics 지수를 보여준다. 이와 유사한 연구로서 조재인(2016)은 앞선 연구인 Robinson-García, Jiménez-Contreras, Torres-Salinas(2015)와 유사하게 연구데이터의 규모는 유전학과 생명공학 분야가 크게 차지하는 것으로 밝혔다. 그러나 실제로 인용된 분야는 인구, 고용 등 경제 등 사회과학분야인 것으로 나타났다. 리포지터리별로 살펴보면, UK Data Archive, ICPSR(Inter-University Consortium For Political And Social Research)

에 구축된 연구데이터가 가장 많이 인용되고 있는 것으로 분석되었다. 피인용된 연구 데이터의 특성을 살펴보면, 데이터세트보다는 조사방법과 연구방법론이 포함된 데이터스터디가 높은 피인용도를 보이는 것으로 나타났으며, 연구데이터의 알트메트릭스 분석 결과에서도 사회과학분야의 데이터스터디가 상대적으로 많이 참조되고 있는 것으로 나타났다.

### 3.4 데이터 인용요소 공통사항

2장 2절에서 조사한 인용 가이드라인과 상기한 논의한 사례에서 제시된 공통된 데이터 인용요소들을 정리하면 <표 3>과 같다. 인용 가이드라인의 경우 DCC에서 제시한 인용요소(<표 2>

참조)와 APA 양식, 미시건주립대학교(MSU) 도서관가이드, 미국지질조사국(USGS) 데이터 관리 웹사이트에서 제시하는 인용요소를 조사하였다. 또한 DataCite, Dataverse Network, DCI 등 각각의 사례에서 제시한 메타데이터 중 인용에 사용되는 요소들을 살펴보았다.

인용 가이드라인 4개와 사례 3개에서 모두 제시되고 있는 데이터 인용요소는 저자, 배포일자, 제목, 출판기관(배포기관)인 것으로 나타났다. 저자(Author)는 데이터를 생산한 개인, 단체 또는 기관을 의미한다. 일자(Publication date)는 여러 측면에서 정의되고 있었으나 조사된 가이드라인과 사례에서는 공통적으로 데이터를 배포한 일자 또는 이용가능하게 한 일자로 정의되었다. 이 외에도 데이터의 엠바고가 종료된

<표 3> 인용 가이드라인과 사례에서의 데이터 인용요소

	가이드라인				사례			빈도
	DCC	APA 양식	MSU 도서관가이드	USGS	Data Cite	Dataverse network	DCI	
저자(Author)	●	●	●	●	●	●	●	7
일자 (Publication date)	이용가능하게 한 일자 또는 배포일자	●	●	●	●	●	●	7
	엠바고가 종료된 일자	●			●			2
	품질확인이 종료된 일자	●						1
	생산된 일자						●	1
제목(Title)	●	●	●	●	●	●	●	7
출판기관 (Publisher)	배포기관(Distributor)	●	●	●	●	●	●	7
	제작기관(Producer)			●	●	●		3
판(Edition)	●		●	●				3
버전(Version)	●		●	●		●		4
자원유형(Resource type)	●	●	●		●			4
식별기호(Identifier)	●			●	●	●		4
소장위치(Location)	●	●	●				●	4
Unique Numeric Fingerprint(UNF)	●					●		2
하위 데이터세트 표시(Feature name and URI)	●							1
편집자(Editor)			●					1
접근일자(Access date and time)				●				1

일자가 DCC 가이드라인과 DataCite 메타데이터에서 제시되었으며 데이터의 품질확인이 종료된 일자 또는 데이터가 생산된 일자로 정의되는 경우도 있었다. DCC 가이드라인에서는 세 가지로 정의된 일자 가운데 가장 최신의 것을 사용할 것을 제안하였다.

데이터의 제목(Title) 역시 조사된 가이드라인과 사례에서 공통된 인용요소로 제시되었다. 출판기관(Publisher)의 경우 모든 경우에서 배포기관(Distributor)으로 정의되고 있었으며, 이러한 역할과 함께 MSU 도서관가이드와 DataCite, Dataverse network에서는 출판기관을 데이터 생성을 지원하는 제작기관(Producer)으로도 정의하고 있었다.

절반 이상인 4개 가이드라인 또는 사례에서 제시된 인용요소로는 버전, 자원유형, 식별기호, 소장위치가 있었다. 판(Edition)과 버전(Version)의 경우 DCC 가이드라인에서는 구별하여 정의하였으나 MSU 도서관가이드와 USGS에서는 상호 호환적인 의미로 사용하고 있었다. Dataverse network에서는 버전만을 인용요소로서 제시하고 있었다. 자원유형(Resource type)은 다양한 종류의 데이터 유형을 의미하는데 예를 들어 DataCite에서는 통제된 자원유형 리스트로서 시청각(Audiovisual), 컬렉션(Collection), 데이터세트(Dataset), 이벤트(Event), 이미지(Image), 인터랙티브 자원(InteractiveResource) 등을 제시하고 있었다. 식별기호(Identifier)의 경우 DOI나 Handle과 같은 표준 식별기호가 주된 예시로 제시되었으며 소장위치(Location)는 데이터에 접근할 수 있는 URL을 지칭하는 요소로 언급되었다.

## 4. 국내 논문의 데이터 인용 현황 분석

### 4.1 한국종합사회조사 데이터 정보 제공 실태

이 장에서는 실제 국내 연구 논문에서 데이터 인용이 어떻게 이루어지고 있는지를 조사해보았다. 해외 데이터 인용색인을 분석한 선행연구(조재인, 2016)에서 사회, 경제 등 사회과학 분야의 서베이 데이터가 가장 많이 인용되고 있는 것으로 나타난 바 있다. 국내의 서베이 데이터로는 한국종합사회조사(Korean General Social Survey, 이하 KGSS라고 칭함)가 대표적이다. 한국사회과학데이터센터에서는 자주 활용되는 양적 및 질적 데이터를 소개하고 분석 방법을 시연하거나 강연하는 행사인 'KOSSDA 데이터페어'를 2016년 9월부터 시작했는데, KGSS 데이터가 국내에서 가장 광범위하게 이용되고 있다는 이유로 첫 행사인 제1회 데이터페어의 대상으로 채택하였다. 2016년 9월 현재 KGSS를 이용하여 한국어/영어/일본어/중국어 등으로 발표된 국내외 연구 문헌이 279건, 국내외 단행본에서 활용된 사례가 91건, 국내외 학술회의 발표가 219건, 국내외 학위논문이 66건으로 보고된바 있다(김솔이, 김하영, 이연희, 한수진, 이가현, 현리정, 김민주, 박은선, 김혁우, 김지범, 2016).

이처럼 국내에서 생산되어 국내뿐만 아니라 해외에서도 널리 활용되고 있는 데이터세트인 KGSS는 한국 사회의 시계열적 연구 및 국제비교 연구를 위한 기초 통계 자료의 생산 및 확산을 위해 성균관대 동아시아학술원 서베이리서

치센터가 시행하고 있는 연차적 대단위 전국표본 조사사업의 결과물이다. 이 조사사업은 미국 시카고대학교 NORC(National Opinion Research Center)의 GSS(General Social Survey)를 기본 모형으로 하여 설계되었다(김상욱, 고동우, 고지영, 김석호, 김재은, 남은영, 박병진, 양중희, 이경미, 이장영, 이정진, 정병은, 최유정, 2007). 한국학술진흥재단(현 한국연구재단)의 지원을 받아 2002년에 시범 사업이 실시된 이후 2003년부터 연차 조사가 실시되어 2014년 실시된 12차 조사 결과까지의 누적 데이터가 한국사회과학자료원에 기탁되어 배포되고 있다. 한편 이 사업을 지원하는 한국연구재단의 기초학문자료센터에는 2002년의 예비연구와 2005년부터 2007년까지의 3차, 4차, 5차 조사 데이터만 제공되고 있다.

KGSS 데이터를 이용한 연구성과에 대한 정보도 기초학문자료센터에서는 거의 제공되지 않고 있지만 한국사회과학자료원에서는 다수의 이용 성과물 정보가 제공되고 있다. 기초학문자료센터에서는 해당 연구팀이 과제결과물로 발표한 논문만 포함되어 있으며 다른 연구자가 이용한 논문은 전혀 없다. 한국사회과학자료원의 KGSS 검색 결과에서 '이용논문'이라는 메뉴를 클릭하면 해당 버전의 데이터셋을 이용한 논문 목록이 제공되고 있다. 한국사회과학자료원이 제공하는 데이터셋을 이용할 경우에는 출판된 논문을 다시 한국사회과학자료원에 제출하는 것이 의무 사항이므로 이와 같은 데이터셋 이용논문 정보 수집이 가능하였다. 그러나 KGSS를 이용한 것으로 등록된 논문 중에서도 KGSS를 참고문헌에 누락한 경우가 흔하다. 예를 들어 KGSS 1차 조사 데이터셋을 이용한

논문은 34건이 등록되어 있는데, 이중에서 학술지에 발표된 첫 논문인 2005년 <한국사회학>에 발표된 글에서는 KGSS와 관련된 참고문헌은 전혀 포함되어 있지 않다. 심지어 해당 논문의 제1저자가 KGSS 조사를 수행한 성균관대 동아시아학술원 서베이리서치센터의 연구원임에도 불구하고 데이터셋의 인용 표기가 누락된 것이다.

현재까지 지속적으로 연구비를 지원하고 있는 한국연구재단의 기초학문자료센터에서조차도 2008년 이후의 KGSS 데이터셋이 누락되어 있는 것은 기초학문자료센터의 데이터 수집 및 활용 정책에 심각한 문제가 있음을 반증한다. 또한 해당 데이터를 이용한 연구 성과 정보도 거의 수집되어 있지 않아서, 기초학문자료센터가 데이터 이용과 인용을 통한 연구 활성화에 정책적으로 기여하기에는 개선할 부분이 많다고 판단된다.

한국사회과학자료원의 경우는 기탁된 KGSS 데이터셋을 충실히 관리하고 이용성과물에 대한 정보도 연구자들이 제공한 사례를 수집하여 목록을 공개함으로써 데이터셋의 이용과 인용 활성화에 큰 도움이 되고 있다. 그러나 KGSS 데이터셋을 이용한 것으로 등록된 논문에서조차도 KGSS 데이터셋의 참고문헌 표기가 누락된 경우가 있어서 데이터 인용에 있어서는 아직 개선점이 많다고 판단된다.

#### 4.2 한국종합사회조사 데이터 참고문헌 표기 분석

KGSS를 이용한 논문에서 참고문헌에 'KGSS'나 '한국종합사회조사'가 표기된 사례를 수집하

여 참고문헌 표기 현황을 분석해보고자 하였다. 실제 참고문헌에 표기된 경우를 찾기 위해서 한국학술지인용색인(Korea Citation Index: 이하 KCI로 표기)의 참고문헌 검색 기능을 통해 '한국종합사회조사'와 'KGSS'로 각각 검색한 후 결과를 통합하였다. KCI 인용 데이터의 구축 범위인 2004년부터 2015년 사이에 발간된 논문 중 '한국종합사회조사'나 'KGSS'가 포함된 참고문헌으로 파악된 건수는 153건이었으며, 이들을 인용한 논문의 수는 107건이었다. 인용한 논문 107건 중에서 데이터 자체가 아닌 관련 논문만을 인용한 경우에 해당하는 13건을 제외한 결과 94건의 논문이 한국종합사회조사

데이터를 인용한 것으로 파악되었다.

KCI를 이용한 조사결과인 <표 4>에서 한국종합사회조사 데이터를 한 번이라도 인용한 학문분야는 총 21개 분야로 나타났다. 대분류별로 살펴보면 사회과학 분야에서 사회학을 비롯한 13개 학문분야, 인문학 분야에서 종교학을 비롯한 3개 학문분야가 포함되었으며, 이외에는 복합학과 자연과학 분야에서 2개 학문분야, 그리고 예술체육의 체육 분야에서 인용되었다. 중분류 단위에서 한국종합사회조사 데이터를 가장 많이 인용한 학문분야는 사회학으로서 28건이 파악되었으며, 그 다음으로 정치외교학 10건, 행정학 9건, 학제간연구 8건, 사회복지학 7

<표 4> KCI에서 학문분야별로 2004-2015까지 KGSS 데이터세트를 인용한 논문 수

대분류	중분류	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15	합계
사회과학	사회학			3	2	1	1	4	1	3	5	3	5	28
	정치외교학					1	1	1	1	1	2		3	10
	행정학					1		1	1		2	3	1	9
	사회복지학	1							3	1	1		1	7
	기타사회과학									2		1	2	5
	정책학								2		2	1		5
	교육학						1	1				1		3
	경영학											1		1
	사회과학일반										1			1
	관광학								1					1
	법학									1				1
	심리과학							1						1
지역개발										1			1	
인문학	종교학							1			2		1	4
	기타인문학											1	1	2
	철학		1											1
복합학	학제간연구						1	1			4	2		8
	여성학								1					1
예술체육	체육							1					1	2
자연과학	통계학					1	1							2
	생활과학						1							1
합계		1	1	3	2	4	7	10	10	8	20	13	15	94

건의 순이었다.

한 논문에 여러 개의 KGSS 참고문헌이 수록되기도 하므로 94건의 논문에 등장한 KGSS 관련 참고문헌 표기는 101건이었다. 참고문헌에 표기된 주요 항목을 조사해본 결과 저자(기관), 저자(개인), 연도, 버전, 배포기관, 식별기호로 나타났다. 이 7가지 항목 각각의 표기 여부에 따라 <표 5>와 같은 23가지 참고문헌 표기

유형이 조사되었다.

<표 6>에는 각 유형에 해당하는 실제 참고문헌 표기 사례를 제시하되, 사례가 4건 이하로 적게 파악된 유형은 지면 관계상 포함하지 않았다. 다만 예외적으로 가장 많은 정보를 표기한 유형1과 극단적으로 적은 정보만 표기한 유형22, 유형23에 해당하는 사례는 <표 6>에 포함하였다.

<표 5> KGSS를 이용한 논문의 참고문헌에 표기한 유형별 표기 항목

유형 번호	사례 수	데이터세트 인용 여부	표기 항목						
			저자(기관)	저자(개인)	출판년도	제목	버전(조사년도)	배포기관	식별기호
1	2	●	●	●	●	●	●	●	●
2	1	●		●	●	●		●	●
3	1	●		●	●	●	●		●
4	1	●	●		●	●	●	●	
5	3	●		●	●	●	●	●	
6	2	●	●	●	●	●	●		
7	5	●			●	●	●	●	●
8	5	●			●	●	●	●	
9	4	●	●		●	●	●		
10	37	●		●	●	●	●		
11	1	●	△		●	●	●		
12	1	●	●			●		●	
13	1	●			●	●		●	
14	10	●	●		●	●			
15	4	●			●	●	●		
16	6	●		●	●	●			
17	2	●	△		●	●			
18	1	●	●		●				
19	1	●		●		●			
20	6	●			●	●			
21	1	●				●			
22	2	●							
23	4								
건수	101	97	22.5	53	92	94	65	19	9
비율	100.0%	96.0%	22.3%	52.5%	91.1%	93.1%	64.4%	18.8%	8.9%

(●는 해당 항목을 온전히 표기한 경우, △는 '성균관대학교'와 같이 '서베이리서치센터'를 제외하고 불완전하게 기관명을 표기한 경우임, '데이터세트 인용 여부'는 데이터세트가 아닌 관련 발표논문집을 참고문헌에 표기한 유형23만 제외표시함)

〈표 6〉 KGSS를 인용한 참고문헌 표기 주요 유형별 사례

유형번호	KGSS 데이터세트의 참고문헌 표기 사례(실제로 논문에서 표기된 형태)		
1	김상욱 (2012). 『2012 한국종합사회조사』. 연구수행기관: 성균관대학교 동아시아학술원 서베이리서치센터. 자료서비스기관: 한국사회과학자료원. 자료공개년도: 2014년. 자료번호: A1-2012-0156.		
7	한국사회과학자료원 조사자료		
	자료번호	조사기관	조사명
	A-2003-0006	성균관대학교 서베이리서치센터	한국종합사회조사, 2003
8	한국사회과학자료원 (2008). 한국종합사회조사 2003-2007 누적자료.		
10	김상욱, 김지범, 문용갑, 신승배 (2012). 한국종합사회조사 2012. 성균관대학교출판부.		
14	성균관대학교 서베이리서치센터 (2010) 《한국종합사회조사(KGSS)》 성균관대학교.		
16	김상욱 외 (2010). 한국종합사회조사. 성균관대학교 동아시아학술원 서베이리서치 센터.		
20	한국종합사회조사(KGSS) (2004). 성균관대학교 출판부.		
22	한국종합사회조사 인터넷 웹사이트( <a href="https://kgss.skku.edu/index_ssl.htm">https://kgss.skku.edu/index_ssl.htm</a> )		
23	삼성경제연구소 (2008). 제5차 한국종합사회조사(KGSS) 심포지엄 자료집-문화, 여가, 삶의 질. 삼성경제연구소 (www.seri.org).		

23개 표기 유형 중에서 가장 자세한 경우는 7개 항목을 모두 제시한 2건이었다. 2건 모두 2015년에 〈한국정당학회보〉에 발표된 논문에서 인용되었으며 인용한 저자는 두 논문이 서로 다른 연구자였다. 7개 항목이 모두 제시된 참고문헌의 예는 〈표 6〉의 유형1과 같다.

반면에 가장 적은 정보가 제시된 참고문헌의 예는 2013년과 2015년에 〈사회이론〉에 발표된 동일 저자 논문 두 편에서 〈표 6〉의 유형22와 같이 한국종합사회조사 웹사이트 정보만 제시한 경우였다. 또한 KGSS 데이터를 활용했음에도 불구하고 유형23과 같이 데이터세트 인용을 하지 않고 관련 학술행사 자료집만을 인용한 경우도 4건에 달했다.

KGSS 데이터세트를 참고문헌에 표기한 가장 흔한 유형은 유형10과 같이 단행본으로 출판된 KGSS 코드북을 인용한 경우로서, 총 37건의 문헌에서 이와 같은 방식으로 참고문헌 표기가 이루어졌다. 이 코드북에는 KGSS 문항과 각 문항에 대한 응답자의 사회인구학적 특성별 응답

분포만 수록되어 있다. 두 번째로 흔한 유형은 유형14와 같이 제작기관인 성균관대학교 서베이리서치센터를 밝히고 데이터세트 이름과 제작년도를 출판년도처럼 표기한 경우로서 10건이 있었다. 세 번째로 흔한 유형은 유형16과 유형20으로서, 유형 16에서는 개인저자, 제작년도, 제목이 표기되었으며, 유형20에서는 제작년도와 제목만 표기되어 있다.

유형7과 유형8은 제작기관과 개인저자는 표기되지 않고 배포기관 위주의 정보가 표기된 경우로서 두 유형 모두 5건의 사례가 파악되었다. 유형7의 사례 5건은 모두 강원대학교 사회과학연구원에서 발간하는 〈사회과학연구〉에 실린 논문 한 편에 수록된 것으로서, 특히하게 배포기관인 한국사회과학자료원의 조사자료를 참고문헌의 서두에 별도의 표로 제시하였다. 유형8은 배포기관인 한국사회과학자료원을 저자 표기 위치에 적고 연도와 자료명 및 조사년도(버전)를 표기한 사례이다.

전체 101건 중에서 90% 이상의 참고문헌 사

례에서 표기된 항목은 제목(93.1%)과 출판년형식으로 표기된 제작년도(91.1%)이었다. 출판년도 표기 형식이 아니라 '한국종합사회조사 2007'과 같이 조사년도에 해당하는 버전이 제목과 함께 표기된 경우는 64.4%였고, 연구진에 해당하는 개인저자가 표기된 경우는 52.5%로서 전체의 절반을 약간 초과하였다. 개인저자 표기는 책자로 출판된 KGSS 코드북에 개인저자가 명기된 것에 기인한 것이라고 판단된다.

미시간주립대학교(MSU) 도서관 가이드에서 권고되고 있는 데이터세트 제작기관(producer)과 배포기관(distributor)이 표기된 사례는 전체의 1/4에도 미치지 못하였는데, 각각 22.3%와 18.8%의 사례에서만 표기되었다. 가장 적게 표기된 항목은 데이터세트 식별기호로서 전체의 1/10에 못미치는 8.9%의 사례에서만 표기되었다. 데이터세트 식별기호가 표기된 사례는 2011년 이전에는 전무하다가 2012년 1건, 2013년 1건, 2014년 5건(모두 동일한 논문 1편으로부터 인용됨), 2015년 2건이 나타났다. 데이터세트 식별기호의 사용이 국내에서 2012년 이후 시작되었지만 아직은 사용 사례가 연 1~2회에 불과한 수준으로 드문 것으로 판단된다.

### 4.3 데이터인용 권고 항목과 KGSS 참고문헌 표기 항목 비교

국내 데이터 인용 표기 관행이 인용 지침 및 사례에 어느 정도 부합하는가를 파악하기 위해서 KGSS 참고문헌 표기 조사결과와 앞 장의 인용 지침 및 사례의 주요 권고 항목(절반 이상에서 공통됨)을 비교해보았다. <표 7>을 보면 해외 지침 및 사례에서 100% 포함된 4가지 필수 항목 중에서 제목과 출판년도의 두 가지 항목만 90% 이상의 KGSS 참고문헌에 표기된 것으로 나타났다. 기관저자나 개인저자를 포함한 저자는 70.8%의 참고문헌 사례에서 표기되어 어느 정도 충족되었으나, 배포기관을 표기한 경우는 18.8%에 머물러서 매우 저조하게 나타났다. 특히 배포기관은 표기 사례가 1/5에도 미치지 못할 뿐만 아니라 앞의 <표 6>의 유형8 사례처럼 저자를 표시할 자리에 잘못 표기된 경우도 있었다. 배포기관의 표기가 드물었던 것에는 연구자들이 데이터세트의 제작기관과 배포기관 항목을 혼동한 것에도 원인이 있는 것으로 보인다.

저자, 연도, 제목, 출판/배포기관을 제외한 나머지 과반 이상의 해외 지침 및 사례에서 포

<표 7> 데이터 인용 지침 및 사례의 권고 항목과 KGSS 참고문헌 표기에 실제로 제시된 항목 비교

인용 지침 및 사례의 권고 항목	권고 비율	KGSS 참고문헌 표기에 제시된 항목	표기 비율
저자	100%	저자(기관/개인)	70.8%
배포일자(연도)	100%	출판년도	91.1%
제목	100%	제목	93.1%
출판기관(배포기관)	100%	배포기관	18.8%
버전	57%	버전	64.4%
식별기호	57%	식별기호	8.9%
자원유형	57%		
소장위치	57%		

함된 항목 중에서는 버전(조사년도)이 64.4%의 사례에서 표기되었다. 실제 사례에서는 이전에 해당하는 조사년도를 <표 6>의 유형14나 유형16 사례처럼 출판년도만 표기하고 연도가 동일한 조사년도를 제외한 경우가 상당하였다. 한편 식별기호를 표기한 경우는 8.9%에 불과했고 자원유형('[데이터세트]' 등의 표기)이나 소장위치를 표기한 경우는 KGSS 참고문헌 사례에서 전무하였다.

이상의 비교 결과를 통해 국내 데이터 인용 참고문헌 표기에서는 출판년도와 제목을 제외한 나머지 항목 표기가 미비하거나 혼동되고 있는 것으로 나타났다. 이를 개선하기 위한 두 가지 중점 개선 사항은 다음과 같다. 첫째, 데이터세트 핵심 항목의 명확한 구분이 필요하다. 특히 기관저자(제작기관)와 배포기관, 그리고 조사년도와 출판년도를 각각 혼동하지 말고 구분해서 표기해야 한다. 둘째, 식별기호, 자원유형, 소장위치와 같은 데이터세트 상세 항목의 표기를 누락하지 말아야 한다. 특히 국내에서도 Data Citation Index와 같은 데이터인용색인을 원활하게 구축하기 위해서는 '[데이터세트]'와 같이 일반 문헌이 아닌 데이터를 인용하였음을 명시적으로 밝히는 자원유형표기가 반드시 포함되어야 한다.

## 5. 결론

해외에서는 데이터의 인용 원칙과 형식 및 이를 적용하기 위한 실무 차원의 논의가 활발히 이루어지고 있음을 확인할 수 있었다. DataCite, Dataverse Network와 Data Citation Index와

같은 사례에서 보듯이 학술 커뮤니케이션 및 연구데이터 관리 분야의 주요 기관 및 커뮤니티에서 실제적인 지침을 개발하고 있을 뿐만 아니라, 대학도서관에서 LibGuide 등을 통해 데이터 인용과 관련된 정보서비스에 적극적으로 나서고 있음을 확인할 수 있었다.

국내에서 활용되고 있는 대표적인 데이터세트인 KGSS를 사례로 하여 데이터 정보 제공 현황 및 데이터인용표기 현황을 파악해보았다. 그 결과 데이터세트의 입수 및 이용을 위해 필수적인 정보인 제작기관과 배포기관이 혼동되고 있으며 식별기호 표기가 매우 미미한 실정임으로 이를 개선하기 위한 노력이 데이터 관리 기관과 연구 진흥 기관, 그리고 학술단체에서 다각적으로 이루어져야 할 것으로 판단된다. 특히 연구비 지원 기관인 한국연구재단이 운영하는 기초학문자료센터의 경우 데이터세트와 이용 논문 등 관련 정보가 집약될 수 있는 핵심적인 기관임에도 불구하고 이를 위한 제도적인 장치가 미흡하다는 점에서 속히 개선이 이루어져야 할 것이다. 인문사회분야 연구자들을 대상으로 면담을 수행한 신영란과 정연경(2012)은 기초학문자료센터가 데이터 제공 방식 면에서 개선이 절실하다는 결론을 내린바 있으나, 본 연구 결과 근본적으로 데이터 및 관련 정보 수집 단계부터 기초학문자료센터의 활동이 매우 미흡하다는 것이 KGSS의 사례에서 확인되고 있다.

심원식, 안혜연, 변제연(2015)은 연구 데이터의 수집 및 활용성 증진을 위해서 데이터 인용의 활성화가 필요하다고 제안한 바 있다. 본 연구에서 살펴본 데이터 인용 관련 국내 현황은 데이터 인용 지침의 미비, 데이터 인용을 위

한 정보 제공 부족, 불안정한 데이터 인용 실태로 요약할 수 있다. 이의 개선을 통해 연구 데이터 공유 및 활용이 활성화되기 위해서는 연구진흥 기관 및 학계에서뿐만 아니라 국내 도서관계에서도 데이터 인용에 관한 다양하고 심도 있는 연구와 실무 개발이 필요하다. 해외의 대학도서관을 중심으로 데이터 인용 가이드라인을 개발하고 이에 대한 교육을 강화하고 있는 현실을 반영하여 국내의 대학도서관에서도 데

이터 인용을 활성화할 수 있는 정보서비스 방안을 모색해야 할 것이다. 데이터 인용이 학술 커뮤니티의 규범으로 자리 잡기 위해서는 연구자들이 데이터 인용의 중요성을 인식하고 이를 실행하는 것이 필수적이다. 따라서 데이터 인용의 긍정적인 영향에 대한 연구자들의 인식을 높이고 연구자들의 데이터 인용을 지원하는 방안 및 유인책을 개발할 필요가 있다.

## 참 고 문 헌

- 기초학문자료센터 홈페이지 (n.d.). Retrieved from <http://www.krm.or.kr/>
- 김상욱, 고동우, 고지영, 김석호, 김재은, 남은영, ... 최유정 (2007). 한국종합사회조사 2007. 서울: 성균관대학교 출판부.
- 김솔이, 김하영, 이연희, 한수진, 이가현, 현리정, ... 김지범 (공편) (2016). 한국종합사회조사 이용 문헌자료집. Retrieved from <http://kgss.skku.edu/wp-content/uploads/2016/06/Bib-최종290616.pdf>
- 신영란, 정연경 (2012). 국내 인문사회 연구데이터 아카이브의 개선방안에 관한 연구. 한국기록관리학회지, 12(3), 93-115.
- 심원식, 안혜연, 변제연 (2015). 인문학 분야 연구데이터의 수집 및 활용성 증진을 위한 전략 연구: 기초학문자료센터를 중심으로. 한국문헌정보학회지, 49(3), 155-183. <http://doi.org/10.4275/KSLIS.2015.49.3.155>
- 조재인 (2016). Data Citation Index를 기반으로 한 연구데이터 인용에 관한 연구. 한국문헌정보학회지, 50(1), 189-207. <http://doi.org/10.4275/KSLIS.2016.50.1.189>
- 한국사회과학자료원 홈페이지 (n.d.). Retrieved from <http://www.kosssda.or.kr/>
- 한국학술지인용색인 홈페이지 (n.d.). Retrieved from <http://www.kci.go.kr/>
- Appleton, A. (2008). The Dataverse Network Project: A new architecture for data citation and data-sharing. French Politics, 6(2), 178-186. <http://link.springer.com/article/10.1057/fp.2008.7>
- Arend, D., Junker, A., Scholz, U., Schüler, D., Wylie, J., & Lange, M. (2016). PGP repository:

- A plant phenomics and genomics data publication infrastructure. Database, baw033.  
<https://doi.org/10.1093/database/baw033>
- Arend, D., Lange, M., Chen, J., Colmsee, C., Flemming, S., Hecht, D., & Scholz, U. (2014). e!DAL: A framework to store, share and publish research data. *BMC bioinformatics*, 15(1), 214. <http://doi.org/10.1186/1471-2105-15-214>
- Ball, A., & Duke, M. (2015). How to cite datasets and link to publications. Digital Curation Centre, 1-15. Retrieved from [http://www.dcc.ac.uk/webfm\\_send/525](http://www.dcc.ac.uk/webfm_send/525)
- Brase, J. (2009). Datacite: A global registration agency for research data. *Proceedings of the 4th Cooperation and Promotion of Information Resources in Science and Technology*, 257-261. <https://doi.org/10.1109/coinfo.2009.66>
- Brase, J., Lautenschlager, M., & Sens, I. (2015). The tenth anniversary of assigning DOI names to scientific data and a five year history of DataCite. *D-Lib magazine*, 21(1/2), 10. <http://doi.org/10.1045/january2015-brase>
- Callaghan, S. (2014). Preserving the integrity of the scientific record: Data citation and linking. *Learned Publishing*, 27(5), 15-24. <http://doi.org/10.1087/20140504>
- Chao, T. C. (2011). Disciplinary reach: Investigating the impact of dataset reuse in the earth sciences. *Proceedings of the American Society for Information Science and Technology*, 48(1), 1-8. <http://doi.org/10.1002/meet.2011.14504801125>
- Clarivate Analytics (n.d.). The Data Citation Index: Connecting the data to the research it informs. Retrieved from [http://wokinfo.com/products\\_tools/multidisciplinary/dci/](http://wokinfo.com/products_tools/multidisciplinary/dci/)
- CODATA-ICSTI Task Group on Data Citation Standards and Practices (2013). Out of cite, out of mind: The current state of practice, policy, and technology for the citation of data. *Data Science Journal*, 12, CIDCR1-CIDCR75. Retrieved from [https://www.jstage.jst.go.jp/article/dsj/12/0/12\\_OSOM13-043/\\_article](https://www.jstage.jst.go.jp/article/dsj/12/0/12_OSOM13-043/_article)
- Crosas, M. (2011). The dataverse network®: An open-source application for sharing, discovering and preserving data. *D-lib Magazine*, 17(1/2). <http://doi.org/10.1045/january2011-crosas>
- Crosas, M. (2013). A data sharing story. *Journal of eScience Librarianship*, 1(3), e1020. <http://doi.org/10.7191/jeslib.2012.1020>
- Data Citation Synthesis Group (2014). Joint declaration of data citation principles. In Martone, M. (Ed.). San Diego, CA: FORCE11. Retrieved from <https://www.force11.org/datacitation/>
- Dataverse.org (2016). The dataverse project. Retrieved from <http://dataverse.org/>
- Dragan, L., Luczak-Rösch, M., Berendt, B., Simperl, E., Packer, H., & Moureau, L. (2015).

- A-posteriori provenance-enabled linking of publications and datasets via crowdsourcing. *D-Lib Magazine*, 21(1/2). <http://doi.org/10.1045/january2015-dragan>
- FORCE11 (2016). About FORCE 11. Retrieved from <https://www.force11.org/about>
- Helbig, K., Hausstein, B., & Toepfer, R. (2015). Supporting data citation: Experiences and best practices of a DOI allocation agency for social sciences. *Journal of Librarianship and Scholarly Communication*, 3(2), eP1220. <http://doi.org/10.7710/2162-3309.1220>
- Hoogerwerf, M., Lösch, M., Schirrwagen, J., Callaghan, S., Manghi, P., Iatropoulou, K., ... & Rettberg, N. (2013). Linking data and publications: Towards a cross-disciplinary approach. *International Journal of Digital Curation*, 8(1), 244-254. <http://doi.org/10.2218/ijdc.v8i1.257>
- King, G. (2007). An introduction to the Dataverse Network as an infrastructure for data sharing. *Sociological Methods & Research*, 36(2), 173-199. <https://doi.org/10.1177/0049124107306660>
- Kotarski, R., Reilly, S., Schrimpf, S., Smit, E., & Walshe, K. (2012). Best practices for citability of data and evolving roles in scholarly communication. Geneva: Opportunities for Data Exchange. Retrieved from <https://core.ac.uk/download/files/324/30437756.pdf>
- Leeper, T. J. (2014). Archiving reproducible research with R and Dataverse. *The R Journal*, 6(1), 151-158. Retrieved from <http://journal.r-project.org/archive/2014-1/leeper.pdf>
- Mayernik, M. S. (2012). Data citation initiatives and issues. *Bulletin of the American Society for Information Science and Technology*, 38(5), 23-28. <https://doi.org/10.1002/bult.2012.1720380508>
- McAdoo, T. (2013). How to cite a data set in APA style. Retrieved from <http://blog.apastyle.org/apastyle/2013/12/how-to-cite-a-data-set-in-apa-style.html>
- Menaldo, V. (2016). The fiscal roots of financial underdevelopment. *American Journal of Political Science*, 60(2), 456-471. <http://doi.org/10.1111/ajps.12240>
- Michigan State University Libraries (2016). How to cite data: General info. Retrieved from <http://libguides.lib.msu.edu/citedata>
- Mooney, H., & Newton, M. P. (2012). The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship and Scholarly Communication*, 1(1), eP1035. <http://doi.org/10.7710/2162-3309.1035>
- Parsons, M. A., Duerr, R., & Minster, J. B. (2010). Data citation and peer review. *Eos, Transactions American Geophysical Union*, 91(34), 297-298. <https://doi.org/10.1029/2010eo340001>
- Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. (2016). Research data explored:

- An extended analysis of citations and altmetrics. *Scientometrics*, 107(2), 723-744.  
<http://doi.org/10.1007/s11192-016-1887-4>
- Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. <https://doi.org/10.7717/peerj.175>
- Robinson-García, N., Jiménez-Contreras, E., & Torres-Salinas, D. (2015). Analyzing data citation practices using the Data Citation Index. *Journal of the Association for Information Science and Technology*, 67(12), 2964-2975. <http://doi.org/10.1002/asi.23529>
- Selway, J. S. (2011). The measurement of cross-cutting cleavages and other multidimensional cleavage structures. *Political Analysis*, 19(1), 48-65.  
<https://doi.org/10.1093/pan/mpq036>
- Starr, J., & Gastl, A. (2011). isCitedBy: A metadata scheme for DataCite. *D-lib Magazine*, 17(1/2), 9. <http://doi.org/10.1045/january2011-starr>
- Stockhause, M., Höck, H., Toussaint, F., & Lautenschlager, M. (2012). Quality assessment concept of the World Data Center for climate and its application to CMIP5 data. *Geoscientific Model Development*, 5(4), 1023-1032. <https://doi.org/10.5194/gmd-5-1023-2012>
- Stodden, V. C. (2010). Reproducible research: Addressing the need for data and code sharing in computational science. *Computing in Science & Engineering*, 12(5), 8-12.
- Torres-Salinas, D., & Martín-Martín, A. (2013). An introduction to the coverage of the Data Citation Index (Thomson-Reuters): Disciplines, document types and repositories. arXiv preprint arXiv:1306.6584.
- U.S. Geological Survey (2016). Data citation. Retrieved from  
<https://www2.usgs.gov/datamanagement/describe/citation.php>
- Uhlir, P. F. (Ed.). (2012). For attribution—developing data attribution and citation practices and standards: Summary of an international workshop. Washington, DC: National Academies Press. Retrieved from  
<http://www.nap.edu/catalog/13564/for-attribution-developing-data-attribution-and-citation-practices-and-standards>

• 국문 참고문헌에 대한 영문 표기  
(English translation of references written in Korean)

- Cho, Jane (2016). Study about research data citation based on DCI (Data Citation Index). *Journal of the Korea Society for Library and Information Science*, 50(1), 189-207.

- <http://doi.org/10.4275/KSLIS.2016.50.1.189>
- Kim, S., Kim, H., Lee, Y., Han, S., Lee, G., Hyun, L., ... Kim, J. (Ed.) (2016). Literature list using KGSS dataset. Retrieved from <http://kgss.skku.edu/wp-content/uploads/2016/06/Bib-최종290616.pdf>
- Kim, Sangwook, Ko, Dong-Woo, Ko, Jee-Young, Kim, Seokho, Kim, Jae Eun, Nam, Eun Young, ... Choi, Yu-Jung (2007). Korean General Social Survey 2007. Seoul: Sungkyunkwan University Press.
- Korea Citation Index Homepage (n.d.). Retrieved from <http://www.kci.go.kr/>
- Korea Social Science Data Archive Homepage (n.d.). Retrieved from <http://www.kosssda.or.kr/>
- Korean Research Memory Homepage (n.d.). Retrieved from <https://www.krm.or.kr/>
- Shim, Wonsik, Ahn, Hye-yeon, & Byun, Jeayeon (2015). Strategies for improving the collection and use of research data in the humanities. *Journal of the Korean Society for Library and Information Science*, 49(3), 155-183.  
<http://dx.doi.org/10.4275/KSLIS.2015.49.3.155>
- Shin, Young-Ran, & Chung, Yeon-Kyoung (2012). A study on the improvement plans of the humanities and social sciences research data archives in Korea. *Journal of Records Management & Archives Society of Korea*, 12(3), 93-115.

