

# 인문학 기반 데이터 리터러시 개념에 대한 연구\*

## A Study about the Concept of Data Literacy based on Digital Humanities

한상우 (Sang Woo Han)\*\*

### 초 록

본 연구의 목적은 문헌정보학 중심의 인문학 기반 데이터 리터러시 개념과 세부 역량을 제안하는 것이다. 이를 위해 국내외 데이터, 정보, 디지털 리터러시 등의 관련 연구를 비교·분석하여 데이터 리터러시의 개념을 정의하고, 인문학 관련 개념으로 디지털 인문학 주요 개념을 연계하여 11개 항목의 인문학 기반 데이터 리터러시의 기본 요소를 설계하였다. 본 연구결과는 향후 데이터 리터러시 프레임워크 설계의 기초자료로 이용할 수 있을 것으로 기대된다.

### ABSTRACT

The purpose of this study is to suggest the concept of data literacy and its detailed competencies based on humanities focused on LIS. To do this, we define the concept of data literacy by comparing and analyzing researches related data literacy, data information literacy, and digital literacy in domestic and foreign. And we design the basic 11 elements of data literacy based on digital humanities concepts. The result of this study is expected to be used as a basic data when design henceforward data literacy framework.

키워드: 데이터 리터러시, 데이터 정보 리터러시, 정보 리터러시, 디지털 인문학  
data literacy, data information literacy, information literacy, digital humanities

---

\* 이 연구는 2018년도 광주대학교 대학 연구비의 지원을 받아 수행되었음.

\*\* 광주대학교 인문사회대학 문헌정보학과 조교수(swhan@gwangju.ac.kr)

■ 논문접수일자: 2018년 11월 18일 ■ 최초심사일자: 2018년 12월 28일 ■ 게재확정일자: 2018년 12월 29일

■ 정보관리학회지, 35(4), 223-236, 2018. [http://dx.doi.org/10.3743/KOSIM.2018.35.4.223]

## 1. 서론

바야흐로 데이터의 시대(data-driven society)이다. 정치·경제·사회·문화 및 일상생활 전반에 있어 데이터의 의미와 역할이 확대됨에 따라 데이터 기반 비즈니스(data based business), 데이터 기반 의사 결정(data-driven decision making), 데이터 기반 인사관리(human resource analytics) 등 데이터를 활용하는 새로운 개념이 출현하고 있으며, 특히 학계에서는 데이터 사이언스, 디지털 인문학 등으로 대표되는 데이터 관련 학문이 주목받고 있다. 이러한 배경에는 과거 데이터와 정보의 생산주체가 한정되어 있던 상황과 달리 이제는 누구나 데이터를 생산하고 활용할 수 있는 기술적, 사회적 분위기가 형성되었고, 이에 따라 정부, 기업, 개인 등 여러 주체로부터 다양한 유형의 데이터가 생산되고 있는 것에 기인한다. 아울러, 데이터의 양과 효용성 등이 증가함에 따라 데이터의 가치에 대한 인식이 제고되어 데이터의 본질을 이해하고 수집, 분석, 예측으로 이어지는 일련의 과정에 비판적 시각, 논리의 정립, 맥락의 이해 및 적용 등 전체적인 이해가 필요한 상황이다(한상우, 2017).

리터러시(literacy)는 사전적으로는 ‘글을 읽고 쓸 줄 아는 능력’으로 정의되고 있으나, 최근에는 단지 글을 읽고 쓸 줄 아는 능력으로 제한되지 않고, 복잡한 사회의 커뮤니케이션 과정에서 발생하는 상황과 맥락을 이해하고 대처할 수 있는 개인의 역량으로까지 확대되어 이해되고 있다. 또한, 현재처럼 다양한 디지털 미디어를 이용하고 있는 상황에서 이용자들은 각각의 정보 형태를 이해하고 활용한 통합적인 리터러시 능력을 갖추고 있어야 한다는 것을 의미하

는 것으로 이해할 수 있다(박종구, 2013). 이러한 관점에서 정보 리터러시, 미디어 리터러시, 디지털 리터러시, 컴퓨터 리터러시 등 다양한 주제 분야에서 리터러시 관련 연구가 진행되어 왔으며, 여러 연구를 살펴봤을 때, 당시의 상황에서 가장 중요하게 다루어지는 대상에 대한 본질적인 이해를 위해 리터러시라는 개념을 이용하여 진행되어 온 것을 볼 수 있다. 따라서 현재 데이터가 사회 전반에 걸쳐 영향을 미치고 있는 상황에서 데이터에 대한 본질적인 이해가 요구되는 상황임을 인식할 수 있고, 데이터 리터러시가 필요함은 분명해 보인다.

인문학(humanities)은 자연과학과 상대적인 개념으로 인간과 인간의 근원문제, 인간의 사상과 문화에 관해 탐구하는 학문이다(위키피디아, 2018). 인문학의 주요 연구 영역은 문학, 철학, 역사학, 언어학, 논리학 등이며 모든 학문의 기본이 되는 주제로 이해할 수 있다. 따라서 모든 학문 분야는 상황과 문맥을 이해할 수 있는 능력과 논리적으로 문제를 해결해나갈 수 있는 인문학적 소양을 갖춘 상태에서 주제 분야의 지식을 심화시킬 수 있을 것으로 판단할 수 있다. 여기에 최근 분석적, 사변적인 인문학의 연구방법에 정보기술을 결합한 디지털 인문학이 주목을 받고 있다. 디지털 인문학(digital humanities)은 인문학 지식을 데이터로 기술하고, 그 데이터 속에서 새로운 해석을 얻는 학문으로 정의되고 있어(김현, 2018), 데이터의 특성과 활용에 대한 이해가 전제되어야 할 것으로 예상된다.

이처럼 데이터가 중요시되는 상황에서 인간이 생산하고 소비하는 데이터의 적합한 활용을 위해서는 그 특성과 본질에 대한 이해가 필수적이며, 기술적인 이해와 더불어 인문학적 관점에

서 데이터에 대한 이해가 요구되나 아직 이에 대한 연구와 교육은 매우 미미한 상황이다.

본 연구의 목적은 데이터 리터러시의 필요성이 제기되는 상황에서 문헌정보학 중심의 인문학 기반 데이터 리터러시 개념과 세부 역량을 제안하는 것이다. 이를 위해 국내외 데이터, 정보, 디지털 리터러시 등의 관련 연구를 비교·분석하여 데이터 리터러시의 개념을 정의하고, 인문학 관련 개념으로 디지털 인문학 주요 개념을 연계하여 10개 항목의 인문학 기반 데이터 리터러시의 기본 요소를 설계하였다. 본 연구결과는 향후 데이터 리터러시 프레임워크 설계의 기초자료로 이용할 수 있을 것으로 기대된다. 따라서 이를 위한 연구문제를 정리하면 다음과 같다.

- 연구문제 1. 데이터 리터러시의 영역 및 범위는 무엇인가?
- 연구문제 2. 데이터 리터러시를 구성하는 주요 개념은 무엇인가?
- 연구문제 3. 인문학 기반 데이터 리터러시를 위해 요구되는 인문학적 소양은 무엇인가?
- 연구문제 4. 향후 데이터 리터러시 모델의 개발을 위해 주요하게 다루어야 하는 특성은 무엇인가?

이에 본 연구에서는 아래의 절차로 연구를 진행하고자 한다. 첫째, 데이터와 리터러시에 대한 선행 연구들을 살펴보고 본 연구의 목적에

맞는 데이터 리터러시를 재정의함으로써 데이터 리터러시의 영역 및 범위를 규정해본다. 둘째, 기존의 문헌에서 나타난 데이터 리터러시를 구성하는 역량을 분석하고, 디지털 인문학 연구의 동향을 고찰함으로써 각 분야에서 다루어지는 주요 개념을 분석한다. 마지막으로, 본 연구의 목적인 인문학 기반 데이터 리터러시의 개념을 구성하기 위해 선행연구를 분석한 내용을 종합하여 데이터 리터러시를 구성하는 주요 개념을 도출하는 순서로 연구를 진행하고자 한다.

## 2. 이론적 배경

리터러시에 대한 연구는 국내외를 막론하고 다양한 주제 분야에서 폭넓게 연구되어 온 상황이다. 리터러시 관련 연구를 구분해보면 언론, 미디어 분야에서는 인터넷 리터러시, (소셜)미디어 리터러시, 광고 리터러시, 청소년 리터러시 등이 주된 연구 분야이고, 교육 관련 분야에서는 디지털 리터러시, 컴퓨터 리터러시 등을, 디자인 분야에서는 비주얼 리터러시, 이미지 리터러시 등의 연구가 진행되어 왔다. 문헌정보학 분야에서도 오랜 기간 동안 정보 리터러시(information literacy)<sup>1)</sup>가 주로 연구되었고 최근 데이터 정보 리터러시, 데이터 리터러시 등으로 영역을 확장해가는 추세이다. 본 연구에서는 문헌정보학 분야에서 주로 연구되는 리터러시의 개념을 중심으로 살펴보고자 한다.

1) 국내의 여러 연구에서 '정보 리터러시'는 '정보문해', '정보소양', '정보활용능력', '정보교육' 등의 여러 용어로 사용되고 있다. 본 연구에서 정보 리터러시는 '정보교육의 일반적 구상과 관련된 것으로서 정보에 대한 태도, 정보활용 방법, 정보기술 활용능력 등의 의미를 포괄하는 개념'인 정보문해(고영만, 2010)와 동일한 의미로 사용하나 데이터 리터러시와의 용어 관련성 측면에서 '정보 리터러시'로 사용한다.

## 2.1 데이터 리터러시

해외의 데이터 리터러시 연구는 2000년 이후 꾸준히 진행되어 왔다.

Stephenson과 Caravello(2007)는 “데이터 리터러시는 사회학과 인접 사회과학 분야에서 학생들에게 필요한 통계 리터러시, 양적 리터러시, 산술 능력이라고 알려진 정보 역량의 중요한 요소 중의 하나”라고 정의하고, 사회과학 분야에서 독자적으로 데이터 리터러시 코스를 운영하고 평가하기는 쉽지 않으므로 기존의 정보 리터러시의 일부로 통계적 데이터 처리를 교육하는 데이터 리터러시를 개발하고 통합하는 것이 바람직할 것이라고 제안하였다. 아울러, 사회과학분야 사서는 교수와 학생들이 통계 리터러시를 향상시킬 수 있도록 통계적 사고와 정보 리터러시, 기술의 활용 등을 적절하게 결합시켜 지원할 수 있어야 한다고 하였다(Stephenson & Caravello, 2007).

Qin과 D'ignazio는 Love(2004)가 정의한 “데이터가 갖는 복합적이고 다중적인 척도와 레벨을 측정하고, 연구를 위해 고려하고, 합당한 추론을 낼 수 있는 능력”이라는 데이터 리터러시의 정의를 반영하고 여기에 “데이터의 수집, 변환, 관리, 사용을 통한 과학적인 탐구를 강조한” 과학 데이터 리터러시(Science data literacy)라는 개념을 주장했다. 또한 과학 데이터 리터러시의 내용은 각 과학 주제 분야에 따라 다양하게 나타날 수 있으나, 그 핵심은 데이터의 수집, 처리, 조작, 평가, 이용 등을 지원하는 공통적인 기술로 정의될 수 있다고 하였다(Qin & D'ignazio, 2010).

Otto는 데이터 리터러시를 “양적인 데이터

를 접근하고 관리하고 조작하는 것이 대부분의 학계나 전문분야에서의 성공에 있어 중요하고, 이를 위해 필요한 정보 리터러시의 하나의 형태로서의 연구 능력이며 토대”라고 정의하였다(Otto, 2012).

Carlson, Fosmire, Miller, Nelson은 “데이터 리터러시는 전형적으로 데이터가 의미하는 바를 이해하는 것과 관련된다. 구체적으로는 어떤 방법으로 그래프와 차트를 적절하게 읽어낼 것인지, 데이터로부터 올바른 결론을 도출할 것인지, 데이터가 오도되거나 부적절한 방법으로 사용되는 것을 인식하는지를 포함한다.”라고 설명하였다(Carlson, Fosmire, Miller, & Nelson, 2011).

Schild는 “데이터 리터러시는 데이터에 접근하고, 조작하고, 요약하는 능력”이라고 정의하고, 데이터를 분석하고, 해석하고, 평가하는 통계 리터러시와는 별개의 개념이라고 주장하였다. 또한, 정보 리터러시와 데이터 리터러시는 비판적 사고와 통계 리터러시를 포함할 수 있도록 그 영역이 확장되어야 한다고 하였으며, 데이터 리터러시를 확장하는 것은 사회과학 분야의 학생들이 연계 속에서 인과관계를 추론하는 것에 도움을 줄 수 있을 것이라고 하였다(Schild, 2004).

Ridsdale, Rothwell, Smit, Ali-Hassan, Bliemel, Irvine, Kelley, Matwin과 Wuetherick은 데이터 리터러시를 비판적 관점에서 데이터를 수집, 관리, 평가하고 활용할 수 있는 능력으로 정의하고 5개 영역에 22개의 데이터 리터러시 역량으로 분석하였으며, 데이터 리터러시는 21세기에 전세계적으로 반드시 필요한 기술이라고 주장하였다(Ridsdale, Rothwell, Smit, Ali-Hassan, Bliemel,

Irvine, Kelley, Matwin, & Wuetherick, 2015).

Carlson과 Johnston은 공학 분야의 연구자와 교수를 대상으로 데이터를 활용한 연구 환경에서 정보 리터러시를 응용한 개념으로 데이터 정보 리터러시라는 용어를 사용하였고, 웹 상에서 접근할 수 있는 많은 데이터들을 재사용하고 재창조하여 전문적인 책무를 다하고 협력형 연구의 장점을 가질 수 있는 기회를 갖는데 중요한 역할을 수행해야 한다고 제시하였다(Carlson & Johnston, 2015).

Shorish는 Carlson과 Johnston이 제안한 데이터 정보 리터러시와 같은 용어를 사용하였다. 융합 연구가 증가하고 있는 상황에서 학문적 성과물로서 데이터를 공유하고 재사용하는 것의 가치가 점차 상승하고 있으며, 이러한 상황에서 학생 및 연구자가 자신이 생산한 데이터에 책임을 갖게 하고 자신의 데이터가 더 많은 지식을 생산하는 것에 사용될 수 있다는 인식을 갖게 하는 것이 연구 진행 과정에 있어 매우 중요한 부분이 될 것이라고 전제하였다. 이러한 상황에서 데이터 정보 리터러시는 데이터의 문서화, 보존, 공유 등에 있어 정보 처리 기술과 폭넓은 연계가 필요하고, 궁극적으로 데이터 정보 리터러시는 여타의 리터러시 숙련도와 같은 개념으로 취급되어야 한다고 하였다(Shorish, 2015).

Herzog는 데이터가 넘쳐나는 시대에 데이터를 식별하고, 습득하고, 평가하고, 정제하고, 분석하여 시각화하는 능력이 필요하고 이러한 능력을 정보기술을 이용하여 실제적으로 데이터를 다룰 줄 아는 능력을 데이터 리터러시라고 정리하였으며, MS Excel, Google Fusion Tables 등을 이용한 컴퓨터 처리 기술에 중점을 두었다(Herzog, 2016).

Koltay는 데이터 리터러시에 대한 일련의 문헌 연구들을 통해 데이터 리터러시의 개념과 유사한 의미로 데이터 정보 리터러시, 과학 데이터 리터러시, 연구 데이터 리터러시의 세 유형을 언급하며 세 가지의 개념을 데이터 리터러시라는 용어로 통합할 필요성을 제기하였다. 그간 진행되었던 여러 학자들이 내세운 정의들 부분적으로 채택하여 다음과 같이 정리하였다: “데이터 리터러시는 개인이 데이터에 접근하고, 해석하고, 비판적으로 평가하고, 관리하고, 처리하고, 윤리적으로 사용하는 것을 가능하게 한다. 관리하는 것에는 보존과 큐레이션이 포함된다. 아울러, 데이터 리터러시를 ‘의사결정을 위해 데이터를 효율적으로 이해하고 사용할 수 있는 능력’이라는 정의에 추가적으로 데이터에서 정보로, 궁극적으로 실행 가능한 지식으로 변환하는 것을 가능하게 하는 구체적인 기술의 집합 혹은 지식 베이스라는 개념을 추가한다(Koltay, 2015; Koltay, 2017).”

Maybee, Carlson, Slebodnik, Chapman은 데이터 정보 리터러시는 연구 데이터의 생산과 소비에 초점을 두고 있는 반면, 데이터 리터러시는 데이터의 해석, 분석 및 기타 데이터의 소비 측면에 초점을 두고 있는 점에서 구분된다고 주장하였다. 또한 데이터 정보 리터러시는 학술적 가치를 지닌 개체로서의 데이터의 처리에 초점을 두고, 본래의 목적을 넘어서는 발견과 사용을 위한 적합성을 보장하기 위하여 관리, 확산, 보존 등을 위한 숙고와 행동이 요구된다고 하였다(Maybee et al., 2015).

해외에서 데이터 리터러시에 대한 연구가 활발히 진행된 반면, 국내에서는 매우 미미한 상황이다. 이재윤(2015)은 데이터 리터러시를 연

구데이터 리터러시, 통계데이터 리터러시, 실무 데이터 리터러시로 구분하여 고찰하고, 데이터와 관련된 기술과 역량을 강조하는 용어로 문헌정보학의 정체성을 유지하며 데이터 사이언스를 수용할 수 있는 개념으로 보았다. 한상우(2017)는 데이터의 활용에 있어 기술적인 요소만이 아닌 인문학적 요소의 중요성을 언급하며 인문학 기반의 데이터 리터러시의 필요성을 제안하고 모형 설계를 위한 연구의 필요성을 주장하였다.

이상의 연구를 통해 요약할 수 있는 데이터 리터러시는 데이터의 식별, 접근, 습득, 평가, 조작, 정제, 분석, 표현, 관리, 보존까지 데이터 활용의 전과정에 대한 통합 이해를 요구하는 것으로 볼 수 있다.

## 2.2 디지털 인문학

인문학은 자연과학에 상대되는 개념으로 인간을 둘러싼 제반의 문제를 연구하는 학문으로 정의되고 있다. 인문학의 대표적인 분야로는 문학, 역사, 철학, 언어, 논리 등의 주제와 예술의 영역까지 인문학으로 이해되기도 한다. 이러한 인문학을 연구함에 있어 디지털 시대에 디지털 기술을 이용하여 보다 창조적인 연구를 수행하고자 하는 분야를 디지털 인문학이라고 한다. 디지털 인문학은 “정보통신기술의 도움을 받아 새로운 방식으로 수행하는 인문학 연구와 교육, 그리고 이와 관계된 창조적인 저작 활동”, “전통적인 인문학의 연구과정에 정보통신 기술의 입력, 저장, 분석, 출력의 과정이 융

〈표 1〉 학자별 데이터 리터러시의 개념 요약

연구자	주요 개념	중점분야
Stephenson & Caravello, 2007	사회학과 인접 사회과학 분야에서 학생들에게 필요한 통계 리터러시, 양적 리터러시, 산술 능력이라고 알려진 정보 역량의 중요한 요소 중의 하나	사회과학
Qin & D'ignazio, 2010	데이터가 갖는 복합적이고 다중적인 척도와 레벨을 측정하고, 연구를 위해 고려하고, 합당한 추론을 낼 수 있는 능력	과학
Otto, 2012	양적인 데이터를 접근하고 관리하고 조작하는 것이 대부분의 학계나 전문분야에서의 성공에 있어 중요하고, 이를 위해 필요한 정보 리터러시의 하나의 형태로서의 연구 능력이며 토대	
Carlson et al., 2011	데이터 리터러시는 전형적으로 데이터가 의미하는 바를 이해하는 것과 관련	
Schiled, 2004	데이터에 접근하고, 조작하고, 요약하는 능력	
Ridsdale et al., 2015	비관적 관점에서 데이터를 수집, 관리, 평가하고 활용할 수 있는 능력	
Carlson & Johnston, 2015	웹 상에서 접근할 수 있는 많은 데이터들을 재사용하고 재창조하여 전문적인 책무를 다하고 협력형 연구의 장점을 가질 수 있는 기회를 갖는데 중요한 역할을 수행	공학
Shorish, 2015	데이터의 문서화, 보존, 공유 등에 있어 정보 처리 기술과 폭넓은 연계가 필요	
Herzog, 2016	데이터를 식별하고, 습득하고, 평가하고, 정제하고, 분석하여 시각화하는 능력이 필요하고 이러한 능력을 정보기술을 이용하여 실제적으로 데이터를 다룰 줄 아는 능력	
Koltay, 2015, 2017	데이터에 접근하고, 해석하고, 비판적으로 평가하고, 관리하고, 처리하고, 윤리적으로 사용하는 것	
Maybee et al., 2015	데이터의 해석, 분석 및 기타 데이터의 소비 측면에 초점	
이재윤, 2015	연구데이터 리터러시, 통계데이터 리터러시, 실무데이터 리터러시	
한상우, 2017	기술적인 요소만이 아닌 인문학적 요소의 중요성	

합되어 탄생된 인문학의 새로운 방법론이며 학문 분과”, “초학제적 성격을 지니는 이른바 융합의 새로운 시도를 포함하는 학문분야” 등으로 정의되고 있다. 디지털 인문학에서 디지털은 인문학 연구의 확장을 가져다 줄 수 있는 수단으로 이해하고 인문학적 지식의 토대 위에 디지털적 접근 방법을 이용하여 인문학 연구의 지평을 넓혀나가기 위한 다양한 시도로 평가받고 있다(김현, 김바로, 임영상, 2016; 김바로, 2014; 문상호, 강지훈, 이동열, 2016; 홍정욱, 김기덕, 2014). 하지만 아직 국내에서의 디지털 인문학에 대한 연구는 초기 단계이며, 디지털 인문학이라는 표현보다는 ‘문화콘텐츠’라는 표현이 제안되고 사용되고 있는 것으로 보이고(홍정욱, 김기덕, 2014), 2010년 이후 비교적 최근에는 디지털 인문학이라는 용어가 사용되기 시작하였다. 이런 상황에서 해외의 경우 인문학자들이 다양한 디지털 기술을 습득하여 연구를 진행하는 것에 비해 국내에서는 아직 인문학자들의 기술 습득이 다소 더딘 것으로 보는 견해도 존재한다(홍정욱, 김기덕, 2014).

이처럼 아직 디지털 인문학은 규정하기 어려운 연구분야로 이해된다. 이에 김현(2018)은 디지털 인문학의 세계는 이론으로 설명하고 이해하면 되는 대상이 아니며, 데이터를 만들고 그것과 씨름하는 가운데 유용한 지식을 찾아가는 실천적인 노력이 전제될 때에만 그 세계에 대한 올바른 접근이 이루어질 수 있다고 주장하고 있다. 또한, 미래 세대를 위한 디지털 인문학 교육은 소통과 협업에 의한 인문학 지식 탐구의 방법을 알려 주는 것에서부터 출발해야 한다고 언급하였다(김현, 2018).

최근 수년 사이 ‘디지털 인문학 연구’로 보고

된 여러 형태의 인문 분야 연구에서 가장 기초적인 공통점을 찾는다면, 그것은 종래에 인간의 언어로만 기술되었던 인문 지식의 요소들을 기계가독적(machine readable) 데이터로 전환하여, 컴퓨터가 그 정보의 해석을 도울 수 있게 하는 것이라고 할 수 있다(김현, 안승준, 류인태, 2018). 즉, 연구 주제가 무엇이든 디지털 인문학에서 빼놓을 수 없는 공통의 방법적 요소는 ‘지식을 담고 있는 인간의 언어’를 ‘데이터’로 기술하는 일이고, 디지털 인문학에서 관심을 갖는 데이터는 바로 언어로 표현되어온 다양한 지식의 데이터로 이해할 수 있다(김현, 2018).

결국, 디지털 인문학은 기존의 인문학적인 주제를 연구함에 있어 논리와 사변으로 대표되는 연구방법에 디지털 기술을 결합한 협력적이고 융합적인 학문 분야로 볼 수 있으며, 여기에 인류가 축적해온 오랜 지식을 데이터화하여 인간 사회를 둘러싼 제반 문제를 해결하기 위하여 데이터를 생산, 조직, 분석하고 소통하여 또 다른 인간의 이야기를 만들어 가는 과정으로 이해할 수 있을 것이다.

### 3. 데이터 리터러시의 주요 요소 도출

#### 3.1 디지털 기술적 요소 도출

앞서 선행연구에서 진행된 데이터 리터러시의 개념을 살펴보면 전반적으로 특정 기술이나 역량으로 제한하는 것보다는 데이터의 활용과 관련된 전체적인 과정으로 데이터 리터러시를 이해하려는 시도를 하였고, 본질적인 이해의 중

요성을 강조한 것으로 이해할 수 있다(〈표 1〉 참조). 하지만 일부 연구자의 경우 데이터 리터러시의 범위 및 정의, 세부 역량 등 구체적인 요소를 정의하기보다는 개념적인 수준에서 명제적으로 정의하는 것에 그친 경우도 있다. 본 연구는 여러 선행 연구의 분석을 통해 데이터 리터러시의 개념을 재정의하고 세부 역량을 구성하여 궁극적으로 데이터 리터러시 모형을 설계하는

것을 목표로 하므로 가급적 구체적인 역량에 대한 분석이 요구되는 상황이다. 이에 선행 연구 중 비교적 데이터 리터러시의 역량을 구체적으로 분석한 Ridsdale 외(2015)의 연구와, Carlson과 Johnston(2015)의 연구에서 도출된 성과를 바탕으로 비교·분석하였다(〈표 2〉 참조).

데이터 리터러시의 세부 역량을 비교한 두 연구를 분석하면 Ridsdale 외는 conceptual

〈표 2〉 데이터 리터러시 세부 역량 비교

	Ridsdale et al. (2015)	Carlson & Johnston (2015)	
Conceptual Framework	Introduction to data		
Data Collection	Data Discovery and Collection	Discovery and Aquisition	F: 3.57 S: 4.12
	Evaluating and Ensuring Quality of Data and Sources	Data Quality and Documentation	F: 4.63 S: 4.12
Data Management	Data Organization	Data Management and Organization	F: 4.00
	Data Manipulation		S: 4.47
	Data Conversion (from format to format)	Data Conversion and Interoperability	F: 4.13 S: 4.24
	Metadata Creation and Use	Metadata and Description	F: 4.57 S: 3.88
	Data Curation, Security, and Re-Use	Data Curation and Reuse	F: 4.25 S: 4.06
	Data Preservation	Data Preservation	F: 3.57 S: 3.75
Data Evaluation	Data Tools	Databases and Data Formats	F: 3.71 S: 3.88
	Basic Data Analysis	Data Processing and Analysis	F: 4.63
	Data Interpretation		S: 4.35
	Identifying Problems Using Data		
	Data Visualization	Data Visualization and Representation	F: 4.63
	Presenting Data (Verbally)		S: 4.35
Data Driven Decisions Making			
Data Application	Critical Thinking		
	Data Culture	Cultures of Practice	F: 3.71 S: 3.88
	Data Ethics	Ethics and Attribution	F: 4.38 S: 4.35
	Data Citation		
	Data Sharing		
Evaluating Decisions Based on Data			

□ : Conceptual Competencies    □ : Core Competencies    □ : Advanced Competencies  
 F: faculty average, S: student average

framework, data collection, data management, data evaluation, data application의 5개 영역으로 구분하고, conceptual, core, advanced competencies의 3단계 수준으로 총 22개 세부 역량을 제안하였다. 개념적 수준에서 데이터의 이해, 비판적 사고, 데이터 도구, 데이터 윤리 등을 학습한 후, 실제 데이터를 습득, 처리, 분석, 조직, 평가, 표현 등의 세부 단계를 거치고, 데이터의 변환, 재사용, 보존 등의 최종 단계를 거치도록 단계별, 수준별로 구성한 것을 특징으로 볼 수 있다.

반면, Carlson과 Johnston은 총 12개의 세부 역량을 제시하고 각각의 역량을 데이터 리터러시에 참여한 교수와 학생을 대상으로 세부 역량의 중요성에 대해 5점 척도로 평가를 진행해 역량별 중요성을 평가하였다. 그 결과 가장 높은 평가를 받은 역량은 '데이터 처리 및 분석'과 '데이터 시각화 및 표현'으로 나타났으며, 가장 낮은 평가를 받은 역량은 '데이터 보존'으로 나타났다. 이는 데이터 리터러시 코스에 참여하는 참여자들은 실제 데이터의 처리와 실용적인 측면을 중요시하고 있으나 데이터의 내구성, 재활용성 등에 대한 인식이 아직 부족한 것에서 기인하는 것으로 판단할 수 있다. 또한, 데이터를 활용하는 사고, 문화적 측면에서의 역량인 'Cultures of Practice'가 다른 역량에 비해 상대적으로 낮게 평가되고 있는 점은 데이터의 기술적인 처리를 중요시하고 있으나 그 배경과 사용 문화에 대해 소홀히 하고 있는 것으로 판단되어 보완이 필요하다.

두 연구에서 공통적으로 언급되는 데이터 리터러시의 세부 역량을 맵핑한 결과는 완전히 일치하지는 않았으나 대체로 유사하게 구성되어

있음을 볼 수 있다. 특히, 데이터의 수집과 관리 영역에 관련된 역량은 완전히 일치하는 것으로 보아 데이터 리터러시에서 중요하게 생각하는 기술적인 요인은 실제 데이터의 분석, 처리, 조직, 표현 등의 역량인 것을 알 수 있다. 아울러, Carlson과 Johnston의 연구에서는 논리적 사고, 데이터 문화, 데이터 인용, 공유 등 데이터 활용을 위한 사고, 배경 등에 대한 역량이 낮게 평가되거나 Ridsdale 외의 연구와 달리 아예 언급되지 않았다는 것을 차이점으로 볼 수 있다.

### 3.2 인문학적 요소 도출

디지털 인문학 분야의 연구를 살펴보면 가장 중심이 되는 개념 중 하나가 인문학적 사고, 인문학적 함의를 도출해내는 것이다. 인문학적인 사고는 인간을 기본으로 환경과 존재, 의미, 논리, 인과, 관계 등에 대해 이해하고 하나의 세계를 형성하는 근간을 이해하는 사고라고 할 수 있을 것이다. 이 모든 것을 단시간 내에 단발성 교육으로 형성할 수는 없으나 단계적인 교육과 세상에 대한 이해와 협력을 통해 장기간에 걸쳐 형성될 수 있음은 분명할 것으로 판단된다.

이러한 관점에서 여러 학자들은 디지털 인문학을 형성해가는 과정에서 디지털 기술을 도구로 이해하고 활동 안에 담겨있는 목적을 이해하고자 한다. 김현(2018)은 디지털 인문학 교육은 소통과 협업에 의한 인문학 지식 탐구의 방법을 알려주는 것에서부터 출발해야 한다고 말한다. 이러한 협력과 소통의 방법으로 사용하는 도구로 위키 소프트웨어를 이용하여 스스로 탐구하고 결과를 공유하고 협업을 통해 인문지식 콘텐츠의 생산을 추구할 수 있다고 주

장하였다(김현, 2018). 심혜령(2018)은 언어교육 중심의 디지털 인문학을 연구하며 언어를 학습하고 학습한 내용을 맥락과 목적에 맞는 스토리로 재구성하는 것은 원천적으로 인문학의 영역이라고 하며, 스토리를 만들어내기 위해서는 서로 관련이 없어 보이는 아이디어를 연결하는 인문학적 사고와 디지털의 이해가 융합되는 체계가 요구된다고 하였다. 이를 위한 방법으로 스토리텔링 프리젠테이션 도구인 Prezi를 이용하는 것을 제안하였다. 김현, 안승준, 류인태(2018)는 고문서 아카이브의 고문서에 담겨 있는 내용을 데이터화하고 분석, 정리, 해석을 통해 찾아낸 사실과 이야기를 확장된 데이터로 축적해나감으로써 지속적으로 다른 연구자와 연결되고 의미가 상통하는 지식체계를 구현하는 것이 중요하다고 하였다.

인문학은 결국 인간에 대한 학문이고, 디지털 인문학은 디지털 기술을 이용하여 인간에 대한 연구를 수행하는 것이다. 이 과정에서 수많은 데이터가 발생하고 그 데이터에 의미를 부여하고, 새로운 상황에 부합하는 새로운 스토리를 만들어서 소통하고 협력하여 공유함으로써 인간에 대한 근원적인 이해와 인류의 문

화 발전에 기여할 수 있는 성과물이 만들어지는 과정으로 이해하는 것이 타당할 것이다.

### 3.3 데이터 리터러시의 주요 개념

앞 절에서 분석한 데이터 리터러시의 디지털 기술적 개념과 인문학적 개념을 요약해보면, 디지털 기술적 주요 개념은 접근, 수집, 조작, 처리, 분석, 표현, 시각화 등으로 정리할 수 있고, 인문학적 주요 개념은 논리, 추론, 소통, 협력, 스토리 등으로 정리해 볼 수 있다.

리터러시라는 개념이 단지 언어를 읽고 쓰고 말하는 기술만이 아닌 그 언어가 가지고 있는 문화와 배경, 상황에 대해 전체적으로 이해해야 완전한 리터러시를 갖추게 되는 것과 마찬가지로 데이터 리터러시를 갖추기 위해서는 기술적인 요소는 물론 기술을 적절하게 활용하기 위한 인문학적 사고와 논리, 생산된 데이터를 공유하고 활용할 수 있는 인문학적인 배경 역시 중요하게 고려되어야 함을 이해할 수 있다.

이렇게 정리된 키워드를 기존의 연구의 결과와 연결하여 도출한 인문학 기반 데이터 리터러시의 주요 개념은 <표 3>과 같다.

<표 3> 인문학 기반 데이터 리터러시 주요 개념

구분	세부역량	주요 개념
공통	데이터의 이해	데이터의 속성, 특성, 전반적인 이해
	논리적 추론	데이터를 이용한 논리 및 인과 관계의 구성
인문학	소통과 협력	데이터를 활용하여 성과물을 형성해가는 과정에서의 절차와 방법론
	커뮤니케이션 스킬	
	스토리텔링	데이터를 이용한 의미의 재구성
디지털 기술	데이터의 접근 및 수집	데이터 도구를 이용한 데이터소스의 접근과 발굴, 수집
	데이터의 정제 및 처리	데이터 도구를 이용한 데이터의 전처리
	데이터의 분석 및 해석	전처리된 데이터의 의미있는 해석과 연결
	데이터의 시각화 및 표현	데이터의 성과에 대한 효과적, 효율적 표현
	데이터의 상호운용성	데이터의 표준화와 활용성 제고
	데이터의 공유 및 보존	데이터의 재사용

#### 4. 결론 및 제언

본 연구는 데이터의 활용도와 중요성이 높은 현재 사회에서 데이터의 적절한 활용과 이를 위한 기본 교육인 데이터 리터러시의 필요성이 제기되는 상황에서 문헌정보학 중심의 인문학 기반 데이터 리터러시 개념을 고찰해보고 세부 역량을 제안하는 것을 목적으로 하였다. 이를 위해 국내외 데이터, 정보, 디지털 리터러시 등의 관련 연구를 비교·분석하여 데이터 리터러시의 개념을 정의하고, 인문학 관련 개념으로 디지털 인문학의 주요 개념을 연계하여 11개 항목의 인문학 기반 데이터 리터러시의 기본 요소를 제안하였다.

본 연구의 결과 공통역역으로 '데이터의 이해'를 도출하였고, 디지털 인문학 개념에 근거한 기본 개념으로는 '논리적 추론', '소통과 협력', '커뮤니케이션 스킬', '스토리텔링'의 4개 역량을 도출하였고, 기술적 개념에 근거한 기본 개념으로는 '데이터의 접근 및 수집', '데이터의 정제 및 처리', '데이터의 분석 및 해석', '데이터의 시각화 및 표현', '데이터의 상호운용성', '데이터의 공유 및 보존'의 6개 역량을 도출하였다. 이는 여러 선행 연구에서 공통적으로 나타나는 부분과 특징으로 나타나는 부분을 비교·분석하여 도출한

결과이다. 데이터 리터러시는 아직 그 정의와 범위가 명확하게 규정되지 않은 신생 용어이며 여전히 진행 중이고 발전해 가고 있는 개념이다. 따라서 관련 분야에 대한 포괄적인 연구를 통해 정의와 범위를 형성해가야 하는 상황에서 데이터 리터러시의 기본 개념을 도출한 본 연구의 의미를 찾을 수 있을 것이다. 아울러, 본 연구의 결과에서 도출된 역량을 기반으로 교육현장에서 데이터 리터러시 커리큘럼을 구성할 수 있을 것으로 기대된다. 해당 역량을 강화할 수 있는 프로그램을 구성하고 실제 교육을 통해 데이터 리터러시의 효과성을 평가하는 연구는 후속 연구로 수행되어야 할 필요가 있다.

본 연구는 데이터 리터러시를 다루는 모든 주제 분야를 대상으로 하지 않은 것과 산업계에서의 논의를 본 연구에 포함하지 않은 한계점을 가지고 있다. 지금까지 여러 학문 분야에서 리터러시와 관련한 연구가 진행되어 왔으며 데이터와 관련한 리터러시에 대해 충분히 논의가 될 수 있을 것이며, 산업계에서 역시 데이터에 대한 관심이 높아짐에 따라 현장에서의 데이터 관련 교육에 대한 요구와 논의가 있을 것으로 예상된다. 그럼에도 불구하고 본 연구의 결과는 향후 데이터 리터러시 프레임워크 설계의 기초자료로 이용할 수 있을 것으로 기대된다.

#### 참 고 문 헌

- 고영만 (2010). 정보문해론(개정판). 서울: 한국도서관협회.  
 김바로 (2014). 해외 디지털인문학 동향. 인문콘텐츠, 33, 229-254.  
 김현 (2018). 디지털 인문학 교육의 현장. 인문콘텐츠, 50, 9-34.  
<https://doi.org/10.18658/humancon.2018.09.50.9>

- 김현, 김바로, 임영상 (2016). 디지털 인문학 입문. 서울: 휴복스.
- 김현, 안승준, 류인태 (2018). 데이터 기반 인문학 연구 방법의 모색. *횡단인문학*, 1, 17-61.
- 문상호, 강지훈, 이동열 (2016). 디지털 인문학의 이해. 파주: 이담Books.
- 박종구 (2013). 뉴미디어 채택 이론. 서울: 커뮤니케이션북스.
- 심혜령 (2018). 디지털 인문학의 새 지평을 열며: 제4차 산업혁명 시대의 언어교육을 중심으로. *한국콘텐츠학회지*, 16(3), 14-18.
- 위키피디아 (2018). 인문학. Retrieved from <https://ko.wikipedia.org/wiki/인문학>
- 이재운 (2015). 데이터 사이언스와 데이터 리터러시. *한국정보관리학회 학술대회 논문집*, 11-15.
- 장석권 (2018). 데이터를 철학하다. 서울: 흐름출판.
- 한상우 (2017). 인문학 기반 데이터 리터러시 모형 설계에 관한 연구. *한국정보관리학회 학술대회 논문집*, 45-71.
- 홍정욱, 김기덕 (2014). '2014 세계 디지털인문학' 학술대회 및 한국의 디지털인문학. *인문콘텐츠*, 34, 53-75.
- Carlson, J., & Johnston, L. R. (2015). *Data information literacy: Librarians, data, and the education of a new generations of researchers*. West Lafayette, Indiana: Purdue University Press.
- Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *Portal: Libraries & the Academy*, 11(2), 629-657. DOI:10.1353/pla.2011.0022
- Data information literacy. Retrieved from <http://www.datainfoit.org>
- Herzog, D. (2016). *Data literacy: A user's guide*. Los Angeles: SAGE.
- Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401-415. <http://dx.doi.org/10.1108/JD-02-2014-0026>
- Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, 49(1), 3-14. <https://doi.org/10.1177/0961000615616450>
- Maybee, C., & Zilinski, L. (2015). Data informed learning: A next phase data literacy framework for higher education. *Proceedings of the Association for Information Science & Technology*, 52(1), 1-4.
- Maybee, C., Carlson, J., Slebodnik, M., & Chapman, B. (2015). "It's in the syllabus": Identifying information literacy and data information literacy opportunities using a grounded theory approach. *The Journal of Academic Librarianship*, 41(4), 369-376. <http://dx.doi.org/10.1016/j.acalib.2015.05.009>
- Otto, J. L. (2012). Assessing and improving data literacy: A study with urban and regional planning students. *PNLA Quarterly*, 76(4), 5-23.

- Qin, J., & D'ignazio, J. (2010). The central role of metadata in a science data literacy course. *Journal of Library Metadata*, 10(2-3), 188-204. DOI: 10.1080/19386389.2010.506379
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... & Wuetherick, B. (2015). Strategies and best practices for data literacy education. Dalhousie university.
- Schild, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly*, Summer/Fall, 6-11.
- Shorish, Y. (2015). Data information literacy and undergraduates: A critical competency. *College & Undergraduate Libraries*, 22(1), 97-106.  
<http://dx.doi.org/10.1080/10691316.2015.1001246>
- Stephenson, E., & Caravello, P. S. (2007). Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference Service Review*, 35(4), 525-540. <http://dx.doi.org/10.1108/00907320710838354>

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

- Han, Sang Woo (2017). A research on the design of data literacy model based on digital humanities. *Proceedings of Korea Society for Information Management*, 45-71.
- Hong, Jung-Wook, & Kim, Ki-Duk (2014). 'Digital Humanities 2014' review and digital humanities in Korea. *Humanities Contents*, 34, 53-75.
- Jang, Seok-Kwon (2018). *Deliberating data*. Seoul: Heuleumchoolpan.
- Kim, Ba-Ro (2014). Trend of digital humanities. *Humanities Contents*, 33, 229-254.
- Kim, Hyeon (2018). Practical methods for digital humanities education. *Humanities Contents*, 50, 9-34. <https://doi.org/10.18658/humancon.2018.09.50.9>
- Kim, Hyeon, An, Seungjun, & Ryu, Intae (2018). The pursuit of data-based humanities research methodologies: Where digital humanities and archives of old documents meet. *Transdisciplinary Humanities*, 1, 17-61.
- Kim, Hyeon, Kim, Baro, & Lim, Youngsang (2016). *Digital humanities*. Seoul: Huebooks.
- Ko, Youngman (2010). *Information literacy (Rev. ed.)*. Seoul: Korea Library Association.
- Lee, Jae Yun (2015). Data science and data literacy. *Proceedings of Korea Society for Information Management*, 11-15.
- Moon, Sangho, Kang, Jihoon, & Lee, Dongyeol (2016). *Understanding of digital humanities*. Paju: Edambooks.

Park, Jonggu (2013). New media adoption theories. Seoul: Communicationbooks.

Shim, Hye-Ryeong (2018). Opening the new horizon of digital humanities: Focus on language education in the age of the 4th industrial revolution. The Korea Contents Association Review, 16(3), 14-18.