

데이터 리터러시 교육 방향성 수립을 위한 제언: CO-DATA Station 해커톤 사례를 중심으로

Suggestions for Establishing Directions for Data Literacy Education: Focused on the CO-DATA Station Hackathon Case

김형진 (Hyungjin Lukas Kim) 명지대학교 미래융합경영학과 조교수

요약

4차 산업혁명으로 촉발된 정보화에 대한 필요성으로 지난 10년간 관련 기술에 대한 교육적 투자와 사회적 지출이 상당하였다. 그럼에도 불구하고 우리나라 성인과 청소년의 정보화 분야의 리터러시 수준은 OECD 평균을 밑돌고 있으며, 이는 데이터와 관련된 종합적 역량인 데이터 리터러시(Data Literacy)와 관련된 교육이 Python 등의 분석 도구 교육에 편중되었기 때문에 발생한 결과이다. 본 연구는 이러한 한계를 극복하고 데이터 관련 교육을 위한 방향성을 제시하고자 한다. 이를 위해 해외의 우수 사례를 바탕으로 데이터 리터러시 교육에 필요한 조건(기준)들을 제시하였으며, 한국연구재단 첨단분야 혁신융합대학(Convergence and Open Sharing System) 사업에서 주최한 데이터 리터러시 해커톤 대회 CO-DATA Station의 사례를 분석하여 교육적 방향에 대한 제언을 도출하였다.

키워드 : 데이터 리터러시, 데이터리터러시 교육, 해커톤, 사례연구, CO-DATA Station, 첨단 분야 혁신융합대학 사업

I. 서론

정부는 2022년부터 2026년까지 총 100만 명의 디지털 인재를 양성한다는 ‘디지털 인재양성 종합방안’을 2022년 8월 19일 발표하였다. 모바일로 인해 디지털기술에 대한 교육의 붐이 일었던 2010년 이후 약 10여 년간의 IT기술 교육을 살펴보면, 교육의 방향성에 있어 근본적인 문제가 있었다. 이를 단적으로 보여주는 것이, 우리나라 성인과 청소년의 정보화 영역에서의 리터러시 수준은 OECD 평균을 밑돈다는 사실이다(OECD, 2021). 특히, 여러 디지털기술 중에 실질적인 역량을 강

화하기 위한 새로운 접근이 필요한 것은 데이터(data)와 관련된 분야이다. 이는 데이터와 관련된 역량이 코딩교육이라 불리는 Hard Skill 외에 실제 현상을 분석하고 문제를 제기하고 해결해가는지 생각하는 힘인 Soft Skill이 동시에 개발되어야 하기 때문이지만, 지금까지의 교육 대부분이 Python이나 R의 라이브러리 구현 등 Hard Skill에만 집중되었기에, 실무에서 데이터를 이용한 문제해결 과정에 적용이 어렵다는 한계가 지속해서 제기되어 왔다.

4차 산업혁명으로 촉발된 데이터의 무수한 증가는 다양한 분야에서 데이터를 통해 가치가 증대

될 것으로 기대되었다. 하지만, 데이터는 그 자체로 가치를 발하는 것이 아니라 데이터를 통해 문제를 발견하고, 패턴과 관계를 파악하고, 의미를 통찰하고 예측하는 등의 의사결정 과정에 적절하게 활용될 수 있을 때 그 목적에 맞는 가치부여가 가능하다. 따라서 데이터를 수집, 분석, 해석하고 이를 토대로 의사결정을 내릴 수 있는 종합적인 능력의 개발이 필요한데, 이를 데이터 리터러시(Data Literacy) 역량이라고 한다. 데이터 리터러시 역량을 갖춘 사람은 데이터에 대한 신뢰성을 판단하고 다양한 형태의 데이터를 읽고 이해할 수 있고, 시각화 및 분석 도구를 사용하여 정보를 추출하고 효과적으로 전달할 수 있다.

데이터 리터러시 역량의 여부는 한 개인의 경쟁력과도 깊은 연관성이 있다. IT기반의 산업구조로의 변화에 따라 데이터 기술과 관련된 역량은 필수로 되었지만, 이를 실제 상황에서 원하는 수준으로 활용할 수 있는 인재는 희소하다. 따라서, 데이터와 관련된 Soft/Hard Skill을 모두 포함하는 데이터 리터러시 역량을 기를 수 있는 방향성의 제시가 시급하다.

대학에서의 데이터 리터러시 역량을 기르기 위한 교육은 점차 증가하고 있으나, 그 실효성을 갖기에는 아직 개선해야 할 점들이 많다. 본 연구는 한국연구재단에서 진행한 데이터 리터러시 해커톤 CO-DATA Station의 사례를 분석하여 앞으로 데이터 리터러시 교육이 갖춰야 할 점들을 제안한다.

II. 문헌연구

2.1 데이터 리터러시의 정의

데이터 리터러시는 관련된 연구자의 관점에 따라 약간의 정의가 조금씩 다르며 포함하는 세부요소에 대한 범위에서도 차이가 존재한다. Carlson et al.(2011)은 데이터 리터러시를 데이터가 무엇을 의미하는지 이해하는 것을 바탕으로 적절하게 그래프와 도표를 읽고 올바른 결론을 도출하고 데이

터가 오해의 소지가 있거나 부적절하게 사용되고 있는지 인식하는 것을 포함하는 것으로 정의하였다. Mandinach and Gummer(2013)는 의사결정에 정확한 정보를 제공하기 위해 데이터를 효과적으로 이해하고 사용하는 능력으로 정의하였으며, 데이터를 식별, 수집, 정리, 분석, 요약 및 우선순위를 매기는 방법을 알고 있는 것을 포함한다. 또한, 이 과정에서 가설을 개발하고 문제를 식별하고 데이터를 해석하며 행동 계획을 결정하고 계획하고 실행하고 모니터링하는 방법을 포함한다. D'Ignazio and Bhargava(2015)는 데이터 리터러시를 데이터 읽기(reading data), 데이터 다루기(working with data), 데이터 분석하기(analyzing data), 데이터로 논쟁하기(arguing with data)의 기본적인 요소에 더해 빅데이터 시대에 필요한 세 가지 추가적인 리터러시 역량을 정의하여 부여하였다. 이는 (1) 자신의 행동과 상호작용에 대해 수동적으로 데이터가 수집되는 시기와 장소를 파악하는 것, (2) 대규모 데이터 집합에서 패턴을 식별하기 위해 수행되는 알고리즘 조작을 이해하는 것, (3) 개인과 사회에 대한 데이터 주도적 결정의 실제적 및 잠재적 윤리적 영향을 고려하는 것을 포함한다. 배화순(2019)은 데이터 리터러시를 개인이 정보 시대에 적응하고 학습하며, 신속하게 변화하는 환경에서 의사결정을 내릴 때 필요한 데이터를 이해하고 분석하며, 적절하게 사용할 수 있는 능력으로 정의하였다. 한미영 등(2022)은 효과적인 의사결정을 위해 데이터에 접근하고 이해하며 목적에 맞게 데이터를 활용하고 평가, 조작, 요약하는 능력으로 정의하였다. 김주섭, 김선태(2023)는 이러한 정의들을 바탕으로 데이터 리터러시는 데이터의 의미를 이해하는 것부터 출발하여 데이터 작업과 관련된 역량을 포함한다고 정리하였다. 이에 필요한 역량은 데이터 수집, 정리, 관리, 정제, 정렬, 비교, 분석, 시각화 등을 포함한다.

앞서 내린 연구자들의 정의를 종합해보면 ‘데이터 리터러시는 데이터를 이해하고 활용할 수 있는 종합적인 능력으로, 데이터를 수집, 분석, 해석

하고 이를 토대로 의사결정을 내릴 수 있는 능력'을 포함한다. 따라서, 데이터 리터러시 역량을 갖춘 사람은 데이터에 대한 신뢰성을 판단하고 다양한 형태의 데이터를 읽고 이해할 수 있고, 시각화 및 분석 도구를 사용하여 정보를 추출하고 효과적으로 전달할 수 있다.

2.2 데이터 리터러시의 기반이 되는 양적 추론 교육

데이터 리터러시 역량은 데이터를 기계적으로 분석하는 능력과는 다르다. 따라서, 데이터 분석 도구 위주의 교육만으로는 데이터 리터러시 역량을 종합적으로 기르는 데에 한계가 있을 수밖에 없다. 데이터 리터러시를 포함하고 있는 모든 하위 요소의 역량을 종합적으로 길러야만 제대로 된 역량을 발휘할 수 있는데, 이를 위해 접근할 수 있는 방법은 양적 추론(quantitative reasoning)이다. 양적 추론은 수량적인 정보나 데이터를 활용하여 추론하고 문제를 해결하는 능력으로, 데이터에 근

거한 비판적 사고력의 활용이 중요하다. 미국을 비롯한 해외 대학에서는 이러한 양적 추론과 데이터 과학 교과목을 대학생의 교양 기초교육으로 운영하고 있다(김혜영, 2020). 대표적인 예로 UC Berkeley는 2015년 가을학기부터 4학점의 “데이터 과학의 기초(foundations of data science)” 과목을 개설하였으며, 본 과목은 전공기초가 아니라 전공생을 대상으로 개설한 교양과목이다. 본 교과목의 2023년 가을학기의 세부 주제는 크게 다음의 표와 같다.

표의 내용을 자세히 살펴보면, 한국 대학의 데이터 관련 교육의 세부 주제와 큰 차이가 없다. 다루는 세부 주제에 대해서는 큰 차이가 없지만, 본 교육을 대하는 목적과 관점에서의 차이가 교육적 효과의 차이로 이어지고 있었다. UC Berkeley는 본 과목을 현실 세계의 이슈와 문제들을 탐구할 수 있는 새로운 렌즈를 제공하는 과정이라 설명하고 있다. 단순히 수학적, 통계적 지식을 기르고 이를 바탕으로 분석역량을 키우는 것이 아닌 추론적 사고와 컴퓨팅 사고를 바탕으로 현실 세계

〈표 1〉 데이터 과학의 기초 커리큘럼(UC Berkeley)

주차	주제	비고
1	Introduction, Cause and Effect	
2	Tables, Data Types, Building Tables	
3	Census, Charts	
4	Histograms, Functions, Groups	Project1: World Population and Poverty
5	Pivots and Joins, Table Examples, Iteration	
6	Chance, Sampling, Models	
7	Comparing Distributions, Decisions and Uncertainty, A/B Testing	
8	Causality, Examples	Midterm Review and Exam
9	Confidence Intervals, Interpreting Confidence, Center and Spread	
10	Normal Distribution, Sample Means, Designing Experiments	Project2: Climate
11	Correlation, Linear Regression, Least Squares	
12	Residuals, Regression Inference	
13	Classification	Project3: Classifying Movies
14	Updating Predictions	
15	Case Study	
16	Final Exam	

를 이해하고 표현할 수 있는 역량에 초점을 맞추고 있다. 이를 가능하게 하려고 주당 3시간의 강의와 2시간의 실습으로 진행되며 개발 과제와 팀 프로젝트가 지속해서 진행된다. 다양한 전공 교수들이 팀티칭으로 운영하는 것도 이러한 목적을 이루기 위함이다.

한국대학교육협의회와 한국교양기초교육이 2021년 12월 발간한 <모든 학생을 위한 양적 추론(Quantitative Reasoning) 프로그램 개발 방안 연구(1차 연도)에서는 기초수학 및 통계학을 기반으로 하여 교육용으로 가공된 인위적 데이터를 분석하는 반복 훈련이 아닌 일상생활에서 발생한 실제 데이터를 갖고 문제를 해결하는 과정으로 인식시킬 수 있는 교육과정을 개발하고 다양한 학문 분야에 적용하고자 하여 UC Berkeley의 접근과 유사하게 앞으로의 데이터 리터러시 교육의 방향성을 제시하였다.

미국 마이애미대학교의 양적 추론 교과목 구성은 인문, 예술, 사회, 과학, 공학 등 여러 학문 분야의 통합적 통찰력 제시는라는 측면에서 살펴볼 만한 점이 있다. 마이애미대학교는 1990년 양적 추론을 학제적 과목으로 개발하였는데, 데이터로부터의 학습, 양적 표현, 증거와 주장, 그리고 양적 직관 등 4가지 요소의 상호 관련성을 고려하였다(Wolfe, 1993). 데이터로부터의 학습은 자연과학과 사회과학 분야의 데이터를 수집하고 분석하는 능력으로, 일반적인 데이터 관련 기초 교과목에서 배우는 통계적 개념과 실습을 통한 분석방법을 배운다. 양적 표현은 '양적 언어'를 사용하고 이해하는 능력으로 양적 개념을 시각적으로 표현하는 능력이다. 증거와 주장은 여러 증거로부터 도출된 결론이 합리적인지 판단하는 능력이다. 양적 직관은 숫자 등 정량적 개념에 대한 감(feel)을 개발하는 것에 해당한다. 마이애미대학교는 생물학, 화학, 수학, 심리학, 정치학, 경제학, 지질학, 인류학, 물리학, 통계학, 지리학, 융합 등 여러 다양한 전공에서 양적 추론과 관련된 교과목들을 개발하여 운영하고 있다. 이와 유사하게 예일대학교도 범학문

적으로 양적 추론 교과목들을 개설하고 있다. 상당수의 교과목은 수학, 통계, 데이터 과학, 컴퓨터 과학, 경제학, 공학 등의 분야에서 담당하고 있으나 이외의 자연과학 및 인문사회과학 분야에서도 적절한 교과목을 개발하여 개설하고 있으며, 모든 학생은 관련 영역에서 2개의 교과목을 필수로 이수해야 한다.

2.3 실제 활용 가능한 정량적 사고력의 필요성

앞서 살펴본 다양한 데이터 관련 교육 프로그램들을 분석하고 비교하는 과정에서, 이러한 교육이 효과적으로 이루어지기 위해서는 몇 가지 핵심적인 조건들이 충족되어야 한다는 것을 확인할 수 있다. 첫째, 데이터 리터러시 교육은 공학부터 인문학까지 다양한 학문영역에 걸쳐 범용적으로 적용 가능한 능력을 개발하는 것을 목표로 해야 한다. 이는 데이터 분석과 활용이 점차 모든 분야에 필수적인 기술이 되고 있음을 반영하는 것으로, 기초적인 데이터 소양과 교양 교육의 형식으로 접근하여, 모든 전공 학생들이 데이터에 대한 기본적인 이해와 활용 능력을 갖추도록 해야 함을 의미한다.

둘째, 실효성 있는 데이터 교육을 위해서는 이론에만 국한되지 않고, 실제 데이터를 활용한 프로젝트 기반 학습이 중요하다. 교육용으로 간소화되거나 가공된 데이터가 아닌, 실제 현장에서 마주할 수 있는 복잡하고 다양한 형태의 데이터를 학습 과정에 도입함으로써, 학습자들이 실생활 문제해결에 필요한 데이터 확보와 분석 능력을 실제로 연습하고 발전시킬 수 있도록 해야 한다.

셋째, 데이터 리터러시 교육은 문제해결 지향적이어야 한다. 학습자들이 스스로 문제를 정의하고, 필요한 데이터를 찾아 분석하여 해결책을 도출할 수 있는 프로세스를 경험하는 것은 중요한 핵심이다. 이 과정에서 학습자는 데이터 분석의 목적성과 적용성을 명확히 이해하고, 실제 문제에 대해 비판적이고 창의적인 해결 방안을 모색하게 된다.

넷째, 데이터 분석 도구의 원활한 사용은 데이터 리터러시 교육에서 빼놓을 수 없는 요소다. 이를 위해 수리적, 통계적 지식뿐만 아니라 해당 도구를 사용할 때 필요한 실습이 연계되어야 한다. 학습자가 데이터 분석 소프트웨어나 플랫폼을 효과적으로 활용할 수 있는 능력을 갖추는 것은 복잡한 데이터 집합을 분석하고 해석하는 과정에서 필수적이다.

마지막으로, 데이터 리터러시 교육은 다원화된 해결 방안을 수용하는 교육의 방식을 지향하는 것이 효과적이다. 데이터와 관련된 문제해결 과정은 매우 다양하고 복잡하며, 때로는 예상치 못한 결과를 도출할 수도 있다. 따라서 학습 과정에서는 다양한 해석이 가능하고, 여러 가지 해결 방안을 모색할 수 있는 유연성을 포함해야 한다.

이상의 조건들을 충족하는 교육 모델을 개발하고 실행하기 위하여, 본 연구에서는 데이터 리터러시 해커톤 사례를 상세히 분석하려 한다. 해커톤은 본 조건들을 실천할 수 있는 이상적인 플랫폼으로, 참가자들이 실시간으로 데이터 문제를 정의하고, 해결책을 모색하며, 다양한 분석 도구를 사용해 봄으로써 데이터 리터러시 교육의 핵심 요소들을 체험할 기회를 제공한다. 이 사례를 통해 현재 대학에서 진행되고 있는 데이터 리터러시 교육의 실제적 효과를 분석하고, 앞으로의 교육 방향성을 제시하는 데에 중요한 시사점을 제공할 것이다.

III. 데이터 리터러시 해커톤: CO-DATA Station

본 연구에서 분석하고자 하는 해커톤 사례는 한국연구재단의 첨단분야 혁신융합대학(Convergence and Open, Sharing System, 이하 COSS) 사업의 Pop-up 캠퍼스인 제2회 CO-WEEK ACADEMY (2023.07.03.~2023.07.07.)의 데이터 리터러시 분야 해커톤 행사 <CO-DATA Station>으로 2023.07.06.~2023.07.07.의 1박 2일 일정으로 건국대학교 X-스

페이스에서 본 연구자가 직접 기획하여 운영하고 그 결과를 분석하였다. 참가대상은 COSS 사업을 수행하고 있는 대학의 재학생(휴학생 포함)으로 총 11개 팀 51명(4~5인/팀 구성)이 참여하였다.¹⁾

2021년부터 운영된 COSS 사업은 인공지능, 빅데이터, 차세대반도체, 미래자동차, 바이오헬스, 실감미디어, 지능형로봇, 에너지신산업 등 8개 신기술 분야에 대해 대학들이 컨소시엄을 이뤄 공동으로 교육과정을 개발 및 운영하고 인정, 물적 자원을 공유, 활용하는 사업으로 학과(전공) 간, 대학 간 경계를 허물고 기업과 협업해 첨단분야 인재양성을 지원하는 데 방점을 둔다.²⁾

해당 해커톤의 개념과 구성은 다음과 같은 고려사항에 기반을 두었다. 첫째, 이 프로그램은 무엇보다 접근성을 최우선으로 취급했다. 개발이나 분석 코드 작성에 익숙하지 않은 인문사회계 학생 및 데이터 관련성이 상대적으로 낮은 다양한 학문 분야에서 온 연구생들이 참여할 수 있도록 장벽을 허물기 위해 문제의 난이도를 신중하게 조절했다. 이는 탄탄한 기초 교육과 함께, 참가자들이 기술에 대한 두려움 없이 데이터를 접하고, 조작하며, 분석할 수 있는 실질적인 자신감을 심어주는 것을 목표로 했다.

둘째, 해커톤은 실세계의 산업 데이터를 활용하여 문제를 설정함으로써 참가자들이 다양한 관점에서 문제해결 과정을 경험할 수 있도록 했다. 이를 통해 참가자들은 실제 산업 데이터를 분석하고, 그 안에서 패턴을 발견하며, 해결책을 제안하는 과정에서 데이터 리터러시 능력을 실질적으로 개발할 수 있다. 문제는 단순히 학문적인 배경 지식을 요구하는 것이 아닌, 실제 상황에서 발생할 수 있는 복잡한 문제들을 모델링하여, 참가자들이 실

- 1) 인공지능, 빅데이터, 차세대반도체, 미래자동차, 실감미디어, 지능형로봇, 에너지신산업 컨소시엄의 재학생들이 팀을 이뤄 참가하였다.
- 2) 2023년부터는 항공·드론, 반도체소부장, 이차전지, 차세대통신, 예코업 등의 5개 분야의 컨소시엄이 추가되어 총 13개의 컨소시엄으로 운영되고 있다.

세계 문제에 대한 솔루션을 모색하도록 유도했다.

셋째, 본 해커톤은 문제해결에 있어 단일 정답을 제시하는 전통적인 접근 방식을 배제하고, 다양한 해결책이 존재할 수 있는 개방형(Open-ended) 문제들을 제시함으로써 참가자들의 창의적 사고와 문제해결 능력을 강화했다. 이는 참가자들이 자신만의 독창적인 해결 방안을 모색하고, 다른 참가자들과의 협업을 통해 다양한 관점에서 문제를 바라보고 해결 방안을 찾아 나가는 과정을 중요시했다.

전체 1박 2일 일정 중에 첫 번째 날은 해커톤 대회의 운영 취지 및 방식에 대해 설명하는 오리엔테이션을 시작으로 데이터 리터러시의 개념과 문제 해결법 그리고 데이터의 방향성에 대한 특강을 통해 학생들이 갖는 데이터 역량에 대한 수준의 편차를 완화할 수 있도록 하였다. 이후, 문제를 공개함과 동시에 해커톤이 시작되었으며 결과물 제출 및 발표는 이튿날 오전 10시부터 시작되었다. 본 대회에서 제시된 문제는 실제 소비재 기업 데이터를 바탕으로 추후 해당 기업의 신규점포 개설과 매출 추정에 관한 전략적 의사결정을 도출하기 위한 양적 추론을 하는 것이었으며, 참가자들에게 모두 공평하게 20시간이 작업을 할 수 있는 시간으로 주어졌다. 참가자들의 원활한 문제해결을 위해 데이터 전문가 3명을 퍼실리테이터(facilitator)로 현장에 배치하였다. 평가는 논리력, 창의력, 데이터 분석 능력, 전달 및 표현력을 기준으로 이뤄졌다.

IV. 데이터 리터러시 교육 방향성 도출

CO-DATA Station 진행 결과 참가팀 11팀 중 1팀이 중도 포기하여 총 10개의 팀을 대상으로 평가하였으며, 각 팀별 발표 10분 질의응답 5분으로 진행되었다. 심사위원은 총 7명으로 대학에서 관련 교육을 진행하고 있는 교수 5인, 데이터 분석 전문가 1인, 교육 전문가 1인으로 구성하였다. 최고점 84.8점, 최저점 74.8점, 평균점수: 79.3, 표준

편차: 2.9였다. 참가자들이 제출하고 발표한 최종 결과물에 대한 분석과 평가, 참가자 및 평가위원의 의견들을 바탕으로 하기의 교육 방향성을 제안하고자 한다.

제안 1. 프로젝트 기반 교육의 필요성

본 연구자가 가장 시급하고 필요하다고 생각하는 것은 데이터 리터러시 교육을 위한 프로젝트 기반의 접근이다. 오늘날의 기술 및 사회의 변화 속도는 상상을 초월하고 이로 인한 복잡성도 상당히 커졌다. 이러한 이유로 우리가 직면하는 대부분의 문제는 여러 분야의 지식을 융합해야만 풀 수 있는 융합적 성격이 강하고 실제 경험하지 않으면 곧바로 풀기 어려운 경우가 대부분이다. 기술과 지식 측면에서 융합적이고 복잡하기 때문에 혼자서 모든 임무(task)를 완료하기도 어렵다. 결국, 산업의 종류와 관계없이 자신이 직면한 문제 혹은 이슈에 대해 학문 내 또는 학문 간 지식 및 기능을 융합하고 데이터를 활용해 가치를 창출할 수 있는 융합역량이 필수적인데 이를 기를 수 있는 접근은 프로젝트 기반 교육이다(OECD, 2018).

물론 지금까지의 대학 교육에서 프로젝트기반 데이터 교육이 없었던 것은 아니다. 15~16주의 한 학기 교과를 진행하면서 수업 내에 조별과제 형태로 프로젝트 수행이 대부분 구성되어 있다. 하지만, 많은 수업이 프로젝트로 평가를 진행하지만, 강의(lecture)가 중심인 경우가 상당수다(박영주, 박진희, 2023). 기초 혹은 입문 단계의 수업에서는 이러한 방식이 효율적일 수는 있으나 중급 혹은 고급 단계의 수업에서는 비효율적일 수 있다. 오히려 학생이 스스로 풀고 싶은 문제를 갖고 수업에 들어와서 교수자와 함께 이를 풀어나가고 교수자는 필요한 경우 수업을 진행하는 형태가 효과적일 수 있다.

이러한 프로젝트 형태의 교육이 진행되기 위해서는 교수자의 자질과 강의 기획도 중요하지만, 교육 수혜자인 학생이 학습 행동에 대한 스스로의 높은 주체성을 지니고 있어야만 한다.

**제안 2. 여러 관점에서의 다양한 답을 내는
훈련의 필요성**

일부 참가자들은 조금 더 어려운 분석기법을 쓴 자신들보다 더 기초적인 분석기법을 활용한 다른 팀이 더 좋은 평가를 받는지에 관한 불만을 제기하기도 하였다. 이러한 불만을 제기한다는 것은 리터러시 역량이 낮다는 것을 의미한다. 더 어렵고 고급화된 분석기법이 더 나은 해결책을 의미하지 않는다. 만약, 두 기법의 결과가 같다면, 실무적 관점에서는 더 어렵고 고급화된 분석기법은 더 많은 컴퓨팅 파워(자원)를 낭비한 것이 되고 조직 내에서의 커뮤니케이션에 있어서도 어려움이 발생할 수 있다. 동일한 결과의 도출에 대해 다양한 접근이 가능함을 이해하고 상황에 맞춰 활용할 수 있어야 한다.

이러한 이슈가 발생하게 된 것은 대부분의 데이터 관련 교육이 분석스킬에만 초점을 맞추고 하나의 답을 찾는 것을 보여주기 때문이다. 또한, 교육 과정에서 활용되는 데이터가 실제 데이터가 아니라 특정 분석에 적합하게 가공된 데이터인 경우가 대부분인 점도 이러한 문제를 부추기고 있다. 따라서, 관련 교과목을 가르치는 교수자는 하나의 문제에 대해 여러 종류의 답을 낼 수 있는 예제와 접근 프로세스의 개발이 필요하다.

제안 3. 질문(question) 잘하기의 필요성

앞서 언급한 다양한 접근에 대한 필요성과 연결되는 내용으로 답을 내는 훈련이 아닌 질문하는 훈련이 더 필요하다. 문제를 만든 사람은 답을 이미 알고 있으며, 문제를 만드는 과정에서 출제자는 특정한 의도를 갖고 있다. 따라서, 더 나은 답을 도출하기 위해서는 질문을 잘하는 것이 상당히 중요할 수밖에 없다.

본 해커톤에서 제시된 문제의 최종 요구사항은 특정 시점에서의 매출 규모에 대한 추정이었다. 올바른 매출을 추정하기 위해서는 매출을 추정하기 위한 공식과 이를 뒷받침하기 위한 전략적 접근에 대한 논리가 필요하다. 사실상 매출을 추정

하라고 한 것은 점포를 어떠한 전략을 바탕으로 확장할 것인가를 물어본 것과 동일하다. 하지만, 정답 찾기에 익숙한 상당수의 참가자가 매출 추정과 신규점포 확장 전략 간의 논리적 연결 관계를 만드는 데에 어려움을 겪었다.

**제안 4. 데이터 분석 도구 활용 능력 강화의
필요성**

일부 팀들의 경우 데이터 분석 과정에서의 어려움을 보였다. 이로 인해 한 팀의 경우 중도에 포기하였고 일부 팀들의 경우 훌륭한 논리적 접근과 추론에도 불구하고 데이터 분석에서의 어려움으로 올바른 결론을 도출하지 못하기도 하였다. 코딩교육 중심의 데이터 관련 교육의 문제점이 있다고 하더라도 결국에 분석역량이 뒷받침되지 못한다면 실전에서의 활용도는 떨어질 수밖에 없다. 본 대회에서 상위권 팀의 대부분은 데이터 분석 도구를 활용하는 능력이 우수하였다.

최근 들어, ChatGPT 등 생성형 인공지능(Generative Artificial Intelligence, 이하 GenAI)의 등장으로 인해 분석 코드를 짜는 행위의 노력이 감소할 것으로 예측되고 있다. 하지만, 아직까지는 시기상조일 뿐만 아니라 GenAI를 분석에 활용하기 위해서는 기본적으로 분석에 대한 역량과 스킬이 체득되어야만 한다.

제안 5. 비판적 사고력 및 추론력 강화의 필요성

본 해커톤 참가팀의 결과물을 분석해본 결과 종합적으로는 데이터를 활용한 비판적 사고력과 추론력의 강화가 필요하다. 특히, 일부 팀들의 경우 문제를 정의하고 해결책의 방향성을 만들어가는 과정에서의 퍼실리테이터에 대한 의존성을 보여주었다. 당연히 도움을 주고자 배치하였지만, 대회가 아닌 실제 문제를 직면했을 때에는 이러한 도움이 불가능할 것이므로, 스스로의 비판적 사고력과 추론력을 강화할 수 있어야 한다.

현재 대학의 비판적 사고력 및 추론력을 강화할 수 있는 관련 교육들은 많이 개설되어 있다. 다만,

〈표 2〉 데이터 리터러시 교육 방향성 도출

제안점	세부사항
프로젝트 기반 교육의 필요성	<ul style="list-style-type: none"> 기술 및 사회 변화 속도에 맞는 융합적 능력을 기를 수 있는 프로젝트 기반 교육방법 필요 교육 수혜자의 수준에 따른 교육 콘텐츠 난이도 및 프로젝트 구성 비율 조절 필요 교수자의 강의 기획력과 학생의 주체성 동시충족 필요
다양한 관점에서의 여러 답을 내는 훈련의 필요성	<ul style="list-style-type: none"> 동일한 결과에 대한 다양한 접근법 제시 필요 교육용으로 가공된 데이터 외에 실제 사용되고 있는 데이터의 교육적 활용 방안 제시 필요
질문(question) 잘하기의 필요성	<ul style="list-style-type: none"> 답(answer)을 내기 위한 질문(question)하기 훈련 접근 문제에 숨어있는 행간을 읽어낼 수 있는 역량 강화 논리적 연결 관계를 만들 수 있는 역량 필요
데이터 분석 도구 활용 능력 강화의 필요성	<ul style="list-style-type: none"> 실전 활용도를 높이기 위한 분석 도구 활용역량 강화 필수
비판적 사고력 및 추론력 강화의 필요성	<ul style="list-style-type: none"> 상황을 분석하고 해결책을 만들어낼 수 있는 논리적 사고력과 추론력 훈련 필요 관련 교육 진행 시, 실제 데이터의 다양한 활용 필요하며 이에 맞춘 교과목 구성 방향 수정 필요

이와 관련된 교육과정에 데이터를 활용하는 부분이 배제되거나 축소된 경우들이 많다. 따라서, 데이터에 대한 활용도를 높일 수 있는 형태의 교과목 세부내용의 수정이 필요하다.

V. 결 론

본 연구는 데이터 리터러시에 대한 정의 및 국내외 사례를 바탕으로 데이터 리터러시 교육이 제대로 된 효과를 나타내는 데 필요한 조건들을 도출하였으며, 해커톤 결과를 바탕으로 관련 교육이 나아가야 할 방향을 제안하였다. 본 연구에서 제안한 부분들을 실제 교육에 적용한다면 데이터 리터러시 역량의 실질적 강화와 국가 경쟁력 강화에 도움을 줄 수 있을 것으로 기대한다.

다만, 본 연구가 비교분석을 위한 여러 사례를 분석하였지만, 직접적으로는 CO-DATA Station 해커톤의 단일 케이스만을 다뤘다는 점에서 추가적인 연구가 필요할 것이다. 특히, 관련 교육 혹은 해커톤 등의 과정에서 교육 수혜자의 데이터 리터러시 역량에 어떠한 변화가 있는지 데이터를 수집

하여 정량적으로 접근할 필요가 있다.

참 고 문 헌

- [1] 김주섭, 김선태, “데이터 리터러시 역량 강화 방안에 관한 연구: 연구데이터 관리 교육을 중심으로”, *한국문헌정보학회지*, 제57권, 제1호, 2023, pp. 115-137.
- [2] 김혜영, “21세기 핵심 역량 강화를 위한 대학 교양기초교육으로서의 데이터 리터러시 (Data Literacy) 분석”, *교양교육연구*, 제14권, 제6호), 2020, pp. 147-159.
- [3] 박영주, 박진희, “프로젝트 학습에서의 대학생 경험에 대한 관찰연구”, *교양교육연구*, 제17권, 제1호, 2023, pp. 227-243.
- [4] 배화순, “데이터 리터러시의 사회과 교육적 함의”, *시민교육연구*, 제51권, 제1호, 2019, pp. 95-120.
- [5] 최병문, 김혜영, 박희문, 배영희, 안현효, 조연정, “모든 학생을 위한 양적추론(Quantitative Reasoning) 프로그램 개발 방안 연구(1차 연

- 도)”, 한국대학교육협의회, 한국교양기초교육원, 2021.
- [6] 한미영, 정대홍, 김현정, “데이터 기반 과학 탐구 사례 연구”, *현장과학교육*, 제16권, 제2호, 2022, pp. 165-178.
- [7] Carlson, J., M. Fosmire, C. C. Miller, and M. S. Nelson, “Determining data information literacy needs: A study of students and research faculty”, *Portal: Libraries and the Academy*, Vol.11, No.2, 2011, pp. 629-657.
- [8] Christopher, R. W., “Quantitative reasoning across a college curriculum”, *College Teaching*, Vol.41, No.1, 1993, pp. 3-9.
- [9] D’Ignazio, C. and R. Bhargava, *Approaches to building big data literacy*, Bloomberg data for good exchange, 2015.
- [10] Mandinach, E. B. and E. S. Gummer, “A systemic view of implementing data literacy in educator preparation”, *Educational Researcher*, Vol.42, No.1, 2013, pp. 30-37.
- [11] OECD, *21st-Century Readers: Developing Literacy Skills in a Digital World*, PISA, OECD Publishing, Paris, 2021.
- [12] Organisation for Economic Co-operation and Development (OECD), *The future of education and skills: Education 2030*, OECD Education Working Papers, 2018.

Suggestions for Establishing Directions for Data Literacy Education: Focused on the CO-DATA Station Hackathon Case

Hyungjin Lukas Kim*

Abstract

The necessity of informatization triggered by the Fourth Industrial Revolution has led to significant educational investments and social expenditures in related technologies over the past decade. However, despite this, the levels of literacy in the field of informatization among adults and adolescents in Korea are below the OECD average. This is due to a bias towards education in analysis tools such as Python, rather than a comprehensive competency related to data literacy. This study aims to overcome these limitations and propose directions for data-related education. To this end, we present the necessary conditions (criteria) for data literacy education based on excellent overseas cases and derive suggestions for educational directions by analyzing the case of the CO-DATA Station data literacy hackathon competition organized by the Korea Foundation for the Advancement of Science and Creativity (Convergence and Open Sharing System) project.

Keywords: *Data Literacy, Data Literacy Education, Hackathon, Case Study, CO-DATA Station, Convergence and Open Sharing System*

* Assistant Professor, Department of Future and Convergence Business Administration, Myongji University

○ 저 자 소 개 ○



김 형 진 (kimhj@mju.ac.kr)

고려대학교에서 경영학 학사와 석사 학위를 받았으며, 한국과학기술원(KAIST) 경영공학부에서 IT경영 전공으로 공학 박사학위를 취득하였다. 현재 명지대학교 미래융합경영학과 조교수로 재직 중이며, Journal of Business Research, Journal of Intellectual Capital, Information Technolgy & People, Telematics & Informatics, Computers & Security 등에 논문을 발표하였다. 주요 관심분야는 Customer Relationship Management, Digital Marketing, Digital Strategy in ICT and New Media, Information Security 등이다.

논문접수일 : 2024년 02월 12일

게재확정일 : 2024년 02월 21일

1차 수정일 : 2024년 02월 19일