

델파이 방법을 적용한 AI 교육콘텐츠 학생용 사용성 평가도구 개발 연구

김윤영 · 김두규* · 한상준** · 윤혜정*** · 박미현**** · 허 균†

인하대학교(연구교수) · *효림초등학교(교사) · **한국해양대학교(강사) · ***경남정보대학(강사) ·
****부산대학교(강사) · †부경대학교(교수)

A Study on the Development of Usability Evaluation Tools of AI Educational Contents for Students through Delphi Method

Younyoung KIM · Du-Gyu KIM* · Sang-Jun HAN** · Hye-Jung YUN*** · Mi-Hyun PARK**** · Gyun HEO†

Inha University(research professor) · *Hyorim Elementary School(teacher) · **Korea Maritime & Ocean University(lecturer) · ***Kyungnam College of Information & Technology(lecturer) · ****Pusan National University(lecturer) · †Pukyong National University(professor)

Abstract

The purpose of this study is to develop a usability evaluation tool for artificial intelligence educational content that students can use in schools. Literature review were analyzed to derive evaluation elements for usability test. Based on this, a three-round Delphi study was conducted on selected expert panels. As a result of the study, the usability evaluation tool for students consisted of four evaluation areas: system, UI, learning content, learning method, and evaluation. As a sub-area, 19 evaluation categories, 38 evaluation indicators, and 51 evaluation factors were derived for usability evaluation tools for student. The usability evaluation factors were confirmed to be valid and reliable by checking the median, CVR, consensus, and Cronbach's α . The AI education content usability evaluation tool developed through this study is expected to provide a checklist for learner-centered and educationally effective AI education content development.

Key words : AI education contents, Usability evaluation, Delphi technique

I. 서론

인공지능(Artificial Intelligence, AI)의 급속한 발전으로 전 세계 사회, 경제, 문화, 산업 및 인간의 삶 전반에 걸쳐 거대한 변화를 맞이하고 있다. 전 세계 기업의 70% 이상이 2030년까지 AI 시스템을 도입하면서 인공지능 기술이 인간이 하는 업무의 45%를 자동화시킬 것으로 전망되고

(McKinsey and Companany, 2017), 현재 직업의 47%가 20년 이내에 사라질 것으로 예측된다(Frey and Osborne, 2013). 학교 현장에서도 기존 학습 방식과는 다른 정보와 디지털 기술을 활용한 교육의 대변화가 필요한 시점으로, 선진국들은 이미 AI 교육을 강조하고 AI 핵심인재 양성에 총력을 기울이고 있다.

우리나라 정부 역시 과학기술정보통신부와 관

† Corresponding Author: 051-629-5970, gyunheo@pknu.ac.kr

* 이 논문은 2021년도 교육부의 재원으로 한국과학창의재단의 지원을 받아 수행된 성과물임.

제부처 합동으로 ‘인공지능 국가전략’(2019년 12월 17일) 및 ‘전 국민 인공지능(AI)·소프트웨어(SW) 교육 확산 방안’(2020년 8월 7일)을 발표하기에 이른다. 이를 통해 전 국민이 AI 혜택에 소외당하지 않는 교육환경 구축, 맞춤형 비대면 실시간 교육이 가능한 온라인 AI 교육플랫폼을 구축, 수준별 교육콘텐츠 제공과 인공지능 실습이 가능한 교육 환경 제공, 전국에 2,000여 명의 디지털 역량 교육 강사를 양성하여 AI·SW 소양 교육 제공 등의 방안을 발표하였다(Ministry of Science and ICT, 2019; Ministry of Science and ICT, 2020).

이와 관련하여 교육부와 과학기술정보통신부는 초중등 AI·SW 교육 필수화를 대비하고, 현장 안착을 지원하기 위해 AI 교육 선도학교를 지정하여 다양한 지원을 제공하며, 사용자별 수요와 활용 수준에 따른 AI·SW 교육환경구축을 위해 AI 교육 플랫폼과 AI 교육콘텐츠를 개발하기로 하였다.

이미 AI 분야를 선도하는 국가 및 기업들은 학생들이 AI 모델을 만들고 활용할 수 있는 다양한 AI 교육플랫폼을 구축하고 AI 교육 서비스를 제공하고 있다. 대표적인 예로 영국과 IBM이 공동 개발한 Machine Learning for Kids, 미국 구글의 Teachable Machine, 중국의 mblock 등을 들 수 있는데, 우리나라 학교 현장에서도 이러한 해외 AI 교육플랫폼을 대부분 사용하고 있다(Kim et al., 2019; Han, 2021).

하지만 AI 모델은 빅데이터를 기반으로 하는 학습 모델과 서비스를 만들게 되므로 자국민의 데이터를 통해 생성되는 학습 모델의 중요성이 요구된다. 이는 데이터 주권의 문제와도 연결되므로 우리나라 교육 데이터와 교육 현장에 기반한 AI 플랫폼과 AI 교육콘텐츠 개발이 시급하다.

우리나라 학교 현장에서 사용할 AI 교육콘텐츠가 제대로 개발되기 위해서는 우리나라 교육 현장의 특성이 반영된 평가 도구를 개발하는 것이 필요하다.

이에 본 연구는 AI 교육콘텐츠 사용성 평가에

관한 국내외 연구 동향 및 전문가들의 의견을 바탕으로 AI 교육콘텐츠의 사용성 평가를 위한 평가지표를 개발하고자 한다. 이는 학습자의 AI 교육콘텐츠 사용성을 향상하고 교육 효과가 높은 콘텐츠가 개발될 수 있도록 효과적이고 효율적인 체크리스트 역할을 하는 지표를 제공할 수 있을 뿐 아니라 차후 개발 또는 확장되는 콘텐츠들의 개발 방향성을 제시할 수 있을 것으로 기대한다.

이러한 연구 목적을 달성하기 위한 구체적인 연구 문제는 다음과 같다. 첫째, 학생용 AI 교육콘텐츠 사용성 평가지표는 어떻게 구성되는가? 둘째, 학생용 AI 교육콘텐츠 사용성 평가를 위한 구체적인 측정 문항은 무엇인가?

II. 연구 방법

1. 델파이 연구대상

델파이 연구 방법을 적용하기 위하여 전문가 및 전문성을 지닌 집단으로 AI 교육 및 교육공학에 깊은 이해를 지닌 전문가를 패널로 선정하였다. 특히 아래의 선정 준거 중 1개 이상을 만족한 전문가로 설문에 꾸준히 참여할 수 있는 전문가로 패널을 구성하였다.

<Table 1>은 전문가 패널 선정 준거를 나타낸다. 전문가는 총 18명이었다. 구체적으로 18명은 AI 교육 및 연구 경험을 가진 전문가 그룹과 AI 선도학교 현장교사, AI 관련 교재 집필자, 해외 AI 교육 관련 연구자, 교육공학 전문가 등으로 구성되었다. 1차에서 응답한 인원들은 2차, 3차 델파이 조사까지 응답하였다.

2. 연구 방법 및 절차

본 연구는 AI 교육콘텐츠 사용성 평가지표를 개발하기 위해 문헌분석을 기반으로 델파이 연구 분석 절차를 따랐다. 이를 위해 기존의 연구를 바탕으로 평가지표 개발을 위해 전문가 패널을 대상으로 3라운드에 걸쳐 델파이 조사를 진행하

<Table 1> Criteria for Selecting Experts

No	Criteria	Total
1	Professors with various theoretical backgrounds and experiences such as academic, research, and practical experience in relation to AI education and research	8
2	Among the Ph.D.s in pedagogy, a professor with various theoretical backgrounds and experiences such as academic, research, and practical experience in the development and evaluation of learning contents as an educational technology major	10
3	A person with a master's or doctoral degree who is a major in computer, information engineering, or computer education	5
4	Those who have experienced AI education for more than two years (elementary, secondary, and higher education)	6

였다. 그리고 델파이 조사 결과 자료를 분석하여 최종 평가지표를 개발하였다. 지표를 개발하는 과정에서 신뢰도와 타당도를 확인하였다.

가. 1차 델파이 조사 및 분석

3라운드에 걸친 델파이 조사는 전문가 패널을 대상으로 심리적 효과인 편승효과(band-wagon effect)나 간접 효과 또는 후광효과(halo effect) 등을 피하고자 이메일로 진행하였다. 1차 델파이 조사의 질문지는 구안 단계로서, 해결하려는 연구 문제를 개방형 질문으로 제시하였다. 문항 개발의 이해를 돕기 위해 요구분석, 교수설계, 학습 내용, 인터페이스, 상호작용, 평가, 지원 및 운영, 자원관리, 윤리성, 저작권 등에 관한 연구를 동시에 제공하였다. 1차 조사의 응답을 바탕으로 2차 조사에서는 구조화된 폐쇄형 질문을 구성하였다.

1차 델파이 조사는 2021년 11월 첫째 주에 배부되었으며, 2차 설문지는 12월 첫째 주, 3차는 12월 셋째 주에 배부되었다.

나. 2, 3차 델파이 조사 및 분석

2차 델파이 조사는 내적 타당화 단계로, 1차

델파이 조사 결과 분석 내용을 바탕으로 폐쇄형 질문으로 구조화하였다. 2, 3차 델파이 조사에서는 5점 리커트 척도(Likert scale) 문항과 전문가들이 자신의 의견을 수정·추가·삭제하기 쉽도록 개방형 질문란을 함께 제시하였다. 이는 델파이 조사의 절차가 반복되는 동안 전 회차 분석 결과의 통계적 집단반응과 의견 분석 결과를 제공하고 다음 회 조사에 자기 판단을 수정하고 보완할 수 있도록 하기 위함이다(Choi, 2002). 제시된 전문가 의견은 문헌 고찰 및 연구진 협의를 통해 다음 차수 조사 질문지에 반영 여부를 결정하였다.

3차 이상의 델파이 조사는 종결을 위한 단계로, 2차 델파이 조사의 분석 결과를 제공하며, 2차 델파이 조사의 질문지와 같이 폐쇄형으로 구조화하였다.

2차 이상의 델파이 조사 자료 분석은 평균(mean), 표준편차(standard deviation)를 활용한 기술통계치와 중위수(median), 최빈값(mode), 사분위 범위(interquartile range; IQR)를 활용한 집중 경향치를 활용하여 각 항목에 대한 적합도 및 중요도의 동의 정도와 타당성을 평가하였다.

다. 델파이 조사 타당도 검증

내용 타당도를 확보하기 위해서 기본적인 평균, 표준편차, 중앙값, CVR, 합의도를 확인하고 그 결과를 분석하였다. 델파이 연구 결과의 타당도 분석을 하기 위해 내용타당도지수(Content Validity Ratio; CVR)를 활용하였는데, CVR은 아래 수식 (a)와 같이 전문가 패널의 합의 정도를 비율로 나타낸 것이다(Lawshe, 1975).

$$CVR = \{N_e - (N/2)\} / (N/2) \text{ -----(a)}$$

Ne: 문항이 타당하다고 응답한 패널의 수로 5점 리커트 척도의 4점, 5점으로 응답한 전문가 패널의 수

N: 델파이 조사에 참여한 전체 전문가 패널의 수

Lawshe(1975)가 제시한 전문가 패널 수에 따른 내용 타당도지수(CVR) 비율의 최솟값 기준은 아래 <Table 2>와 같다.

<Table 2> Based on the Minimum Value of CVR

Number of Panelists	10	11	12	13	14	15	20	25	30	35	40
CVR minimum value	.62	.59	.56	.54	.51	.49	.42	.37	.33	.31	.29

예를 들어 전문가 패널의 수가 15명일 경우 내용 타당도지수(CVR)가 0.49 값 이상의 값을 가진 항목들만 내용 타당도가 있는 것으로 판단하고 항목으로 선정하는데, CVR 값이 최소값 이상일 경우 문항에 관한 내용 타당도를 확보한 것으로 판단하는 것이다.

본 연구에 참여한 전문가 패널의 수가 18명이므로, 내용타당도(CVR) 값의 기준을 0.45로 설정하여 CVR 값이 0.45에 미치지 못하면 문항을 삭제하였으며, 문항별 합의도 및 전문가 의견을 반영하여 일부 문항을 수정하였다.

합의도 = $(1 - (75\text{백분위수} - 25\text{백분위}) / \text{중위수})$
로 0.75 이상으로 1에 가까워야 의견이 합의된

<Table 3> The results of the first Delphi survey analysis for students: the results of nomadization of evaluation indicators

No.	Evaluation Category	Frequency	Ratio (%)	Factors
1	Interaction	15	5.8	appropriateness of interactive content (1), appropriate interactivity (1), interaction with program (1), social network (1), confidence through interaction (1), satisfaction inducement (1), sharing (1), communication between learners (1), communication with other learners (2), communication with experts/teachers (1), active participation of learners (1)
2	Convenience	12	4.6	ease of operation (2), ease of learning (1), efficiency (1), minimum link (1), link provision (1), ease of use (1), convenience of use (1), ease of operation (1), functionality (1), shortening (1)
3	Aesthetic	14	5.4	attractiveness (4), aesthetic implementation of visual message (1), aesthetic implementation of the interface (1), design according to the message design principle (1), design (1), design consistency (1), screen design (1), visibility (1)
4	Structured	14	5.4	structural clarity (2), Structural consistency (1), structural systemicity (1), structural systemicity (1), step (1), content hierarchy (1), systematization (1), progress (1), modularization configuration (1), cognitive presentation of learning content (1), procedural (1)
5	Appropriateness	13	5.0	suitability (3), purpose (1), suitability for learners' needs and interests (1), degree of conformity with the nature of a specific task (1)

것으로 판단하였다.

델파이 조사연구에 참여한 전문가 패널의 응답에 대한 반응의 일관성을 확인하고자 Cronbach's α 을 분석하여 신뢰성을 확보하였다. Cronbach's α 값은 0~1 사이의 값을 가지며, 1에 가까울수록 높은 신뢰성을 가진다고 판단하였고, Cronbach's α 값이 0.6 이상, 엄격하게 0.7 이상이면 신뢰도가 높다고 판단하였다.

III. 연구 결과

1. 1차 델파이 학생용 평가지표 분석 결과

1차 델파이 조사에서 도출된 259개의 학생용 평가지표를 1차로 30개로 유목화하고, 이를 전문가 패널이 응답한 평가지표의 정의와 이유를 종합적으로 분석하여 평가영역의 평가항목, 평가영역을 유목화하였다.

<Table 4> The results of the 2nd Delphi survey for students

Evaluation Area	Evaluation Category	Evaluation Indicators	Contents of Evaluation Factors	Average	S.D	M.V	CVR	Agreements	Result
UI	Convenience	Convenience of Content Manipulation	Learning content is convenient to use.	4.67	0.49	5.00	1.00	0.80	O
		Ease of Accessing Learning Content	I can easily find the desired learning content and materials.	4.67	0.49	5.00	1.00	0.80	O
	Aesthetic	Aesthetics of the interface	The multi-person of the user environment is beautiful.	4.00	0.77	4.00	0.44	0.81	X
		Design based on the message design principle	The learning content is designed to be easy to understand and memorable.	4.61	0.61	5.00	0.89	0.80	O
	Readability	Readability of the content	The pictures and letters are concise and highly readable.	4.78	0.43	5.00	1.00	1.00	O
			The menu is systematically structured.	4.56	0.62	5.00	0.89	0.80	O
	Structured	Structured menu	If you look at the menu, you can know your current location in the content.	4.61	0.50	5.00	1.00	0.80	O
			The structure of the menu is simple, so you can easily find the content you want.	4.67	0.49	5.00	1.00	0.80	O
			The structure of the menu is visually convenient.	4.50	0.62	5.00	0.89	0.85	O
			The menu is consistently organized across the entire content.	4.44	0.98	5.00	0.78	0.80	O
Intuitive	Intuitive understanding of the menu name	Menu names are intuitively easy to understand.	4.61	0.61	5.00	0.89	0.80	O	

총 4개 평가영역, 22개 평가항목, 48개 평가지표, 68개 평가 요소로 도출되었다. <Table 3>은 학생용 1차 델파이 조사된 평가지표 유목화 결과이다. 상위 5개 평가영역을 나타내었다.

2. 2차 델파이 학생용 평가지표 분석 결과

2차 델파이 조사는 1차 조사에서 도출된 평가 요소에 대한 필요성을 확인하고자 함에 그 목적

이 있다. 따라서 평가요소의 필요도에 따라 ‘매우 필요하다’, ‘필요하다’, ‘보통이다’, ‘필요하지 않다’, ‘전혀 필요하지 않다’의 5점 리커트 척도로 필요성 정도를 체크하도록 하고, 필요시 그 이유를 자유롭게 제시할 수 있도록 하였다.

분석에 있어서 평가항목의 중요도를 파악하기 위하여 평균, 표준편차, 중앙값을 계산하였다. 또한, 내용 타당도(CVR : Content Validity Ratio)를 통해 검증하였고, 전문가별 합의를 계산하여

분석하고 항목의 신뢰도는 Cronbach's α 값을 산출하여 분석하였다. <Table 4>는 학생용 2차 델파이 조사 분석 결과이다. 지면 관계상 본 지에는 UI에 해당하는 결과만 수록하였다.

학생용 2차 델파이 조사 분석 결과 68개의 평가요소 모두 중위값 4 이상으로 나타났으며, 52개의 평가요소가 CVR 0.45 ~ 1.00 범위에 있었다. CVR 0.45 미만인 16개 평가요소는 학생용 3차 델파이 조사 문항에서 제외하였다. 학생용 2차 델파이 분석 결과 제외된 평가 요소는 매뉴얼 제공(CVR 0.44), 보조 학습자료 제공(CVR 0.44), 미리보기 기능(CVR 0.11), 출력 기능(CVR 0.33), 안내 및 도움의 친절(CVR 0.33), 학습 중 오류(CVR 0.33), 학습 분량 조절(CVR 0.44), 인터페이스의 심미성(CVR 0.44), 학습 내용의 논리성(CVR 0.44), AI 학습 전이성(CVR 0.44), AI 학습 내용의 최신성(CVR 0.33), AI 학습 내용 기억 용이성(CVR 0.44), 동료학습자와의 상호작용(CVR 0.44), 북마크 기능(CVR 0.33), 동료학습자 평가(CVR 0.11) 등 15개와 동일한 문항이 반복되어 시스템 영역에서 유지하고, 학습방법 및 평가 영역에서 삭제한 필기 기능을 포함하여 총 16개이다.

합의도는 68개의 항목 모두 0.75 이상 값으로 도출되어 전문가들의 의견이 대부분 일치하는 것으로 나타났다.

학생용 사용성 평가항목의 신뢰도를 나타내는 Cronbach's α 값은 0.93으로 분석되었다.

학생용 2차 델파이 조사 결과 전문가들의 의견과 연구진 회의를 통해 평가 요소의 내용을 일부 수정하였다. 수정된 평가 요소의 수정 이유를 살펴보면, '학습 결과'의 의미가 모호하여 수정하였고, '문의 및 도움'의 내용이 시스템의 기술적인 부분인지 학습의 내용적인 부분인지 구체화하였으며, '오류', '저지르다' 등의 부정적인 단어를 순화하여 표현하였다. 또한, 학습자의 학습이 이루어지는 웹 환경에 관한 내용도 명확하게 기술하는 것으로 수정하였다. 학습 내용의 평등성에 관한 내용은 학습자가 이해하기 어려울 수 있어

수정이 필요하다는 전문가 의견에 따라 동사 부분의 내용만으로 차별적 내용이 있는지를 판단하여 응답할 수 있도록 평가요소를 수정하였다.

3. 3차 델파이 학생용 평가지표 분석

3차 델파이 조사는 2차 델파이 조사 분석 결과를 반영하여 수정한 3차 델파이 조사 평가 요소에 대한 필요성을 확인하고 연구를 종결하기 위한 단계이다. 그러므로 각 평가 요소의 필요도에 따라 '매우 필요하다', '필요하다', '보통이다', '필요하지 않다', '전혀 필요하지 않다' 등 5점 척도로 필요성을 응답하도록 하고, 필요시 그 이유를 제시할 수 있도록 하였다.

분석에 있어서 평가항목의 중요도를 파악하기 위하여 평균, 표준편차, 중앙값을 계산하였다. 또한, 내용 타당도(CVR : Content Validity Ratio)를 통해 검증하였고, 전문가별 합의도를 계산하여 분석하였다. 항목의 신뢰도는 Cronbach's α 값을 산출하여 분석하였다.

3차 델파이 학생용 평가지표 분석 결과 학생용 3차 델파이 조사 분석 결과이다. 그 결과 52개의 평가요소 모든 항목이 중위값 4 이상으로 나타났으며, 52개의 항목 중 51개의 항목이 CVR 0.45 ~ 1.00 범위에 있었다.

CVR 0.45 미만인 평가요소는 학생용 3차 델파이 조사 문항에서 제외하였다. 학생용 3차 델파이 분석 결과 제외된 평가요소는 학습 내용의 직관성(CVR 0.44) 총 1개이다.

합의도는 52개의 항목 모두 0.80 이상 값으로 도출되어 전문가들의 의견이 대부분 일치하는 것으로 나타났다.

<Table 5> The reliability of the 3rd Delphi survey for students

Cronbach's α	항목 수
0.89	52

학생용 사용성 평가항목의 신뢰도를 나타내는 Cronbach's α 값은 0.89로 분석되었다.

학생용 3차 델파이 조사 결과 전문가들의 의견과 연구진 회의를 통해 평가 요소의 내용을 일부 수정하였다. 수정된 평가 요소의 수정 이유를 살펴보면, '시스템 안정성/호환성' 평가지표의 '웹 환경'을 '웹 브라우저 환경'으로 보다 구체화하여 그 내용을 명확하게 제시하였다.

학생용 사용성 평가지표는 아래 <Table 6>과 같다. 51개의 설문 문항으로 구성되어 있으며, 시스템, UI, 학습내용, 학습방법 및 평가의 영역으로 구성되어 있다. 시스템 영역에는 물리적 및 인적 지원체계, 오류지원, 안전성, 개인화 기능으로 구성되었으며, UI 영역에는 편의성, 심미성, 가독성, 구조화, 직관성으로 구성되었다. 학습내용 영역은 구조화, 적절성, 타당성, 실제성, 동기유발, 다양성, 이해성으로 구성되었으며, 학습방법 및 평가 영역은 상호작용과 평가로 이뤄졌다.

4. 최종 사용성 평가지표

1~3차에 걸친 델파이 설문조사를 통해 도출된

<Table 6> Final Usability Evaluation Index for Students

Evaluation Area	Evaluation Category	Evaluation Indicators	Contents of Evaluation Factors
System	Physical Support system	Manual	The guided content is easy to understand.
		Learning Progress and Outcome Management	I can check the progress of learning.
			It has the function of informing if learning is not done or progress is slow.
	Additional Learning Materials	The learner's current progress, assignments, quiz, etc. can be stored and managed.	
		I can find learning contents corresponding to the results of learning activities such as assignments and quiz.	
	Human Support system	Guidance and Help	This provide a link to the relevant or original data.
			There is a contact information for technical inquiries and help.
	Error Support	Error Resilience	This provide an appropriate error message in case of an error. It guides me on how to recover in the event of an error.
		Backward Error Resilient	Learners can cancel the results they have performed during learning activities and do them all over again.
			Learners cannot cancel the results they performed during learning activities, but they can return to their previous state and try again.
	Safety	System Stability/Compatible	Learning is well performed in the web browser environment (chrome, edge, Explorer, etc) currently being studied by learners.
			It can be using PC, laptop, tablet, mobile phone, etc.
Personalized Function	System Responsiveness	It has a search function.	
	Setting Function	It can be set by adjusting the difficulty of learning content.	
	Note Funcion	It has a note function that allows you to write/store/print.	
UI	Convenience	Convenience of Content	Learning content is convenient to use.

	Manipulation	
	Ease of Accessing Learning Content	I can easily find the desired learning content and materials.
Aesthetic	Design based on the Message Design Principle	The learning content is designed to be easy to understand and memorable.
Readability	Readability of the Content	The pictures and letters are concise and highly readable
Structured	Structured Menu	The menu is systematically structured.
		I can know my current location in the content.
		The structure of the menu is simple, so I can easily find the content I want.
		The structure of the menu is visually convenient.
	Consistency of Menu	The menu is consistently organized across the entire content.
Intuitive	Intuitive Understanding of the Menu Name	Menu names are intuitively easy to understand.
	Clarity of AI learning Goals	Learning goals are clearly presented.
Structured	Procedurality of Learning Content	Learning contents are presented in order.
	Relevance of AI Learning Content	There is a link between AI learning contents, so I can understand the context.
Appropriateness	Appropriateness of AI learning Amount	AI learning content consists of an appropriate amount of learning at a given time.
	Appropriateness of AI Learning Content	Learning contents that meet AI learning goals will be conducted.
	Appropriateness of Difficulty in AI Learning Content	The difficulty of AI learning is appropriate.
Validity	Reliability of AI Learning Content	AI learning content is reliable.
Reality	The practicality of AI Learning Activities	Practical AI educational content was helpful for learning.
	The usefulness of AI Learning Content	Learning activities are designed around learners.
Motivation	Causing Interest from Learners	AI learning content and learning activities are also useful in real life.
	Attention	Learning begins with interesting AI content.
		There are interesting learning activities in between AI learning contents, so I can focus on learning contents.
Diversity	Diversity of AI Learning Content	Various AI learning contents are provided.
	Diversity of AI Learning Activities	I can experience various AI learning activities.
	Equality of Learning Content	The learning content is discriminatory according to race, gender, religion, age, and economic level.

Learning Methods and Evaluation	Understanding	Ease of Remembering AI Learning Contents	It is easy to understand AI learning content.	
	Interaction	Interaction with Learning Content	Process feedback is provided according to the understanding of learning content.	
			Feedback on assignments and test results is provided.	
			I can find learning contents corresponding to tasks and learning outcomes.	
	Evaluation	Validity of Evaluation	Practice	I can practice the learned AI content.
			Evaluate what I have learned according to learning goals.	
			The difficulty of evaluation is appropriate.	
			The problem can be solved within a given time.	
			Multi-faceted evaluation	Evaluate my own learning activities.
Learner Satisfaction	Evaluate learner satisfaction with learning.			

IV. 결론 및 제언

본 연구 결과를 바탕으로 결론을 제시하면 다음과 같다.

첫째, 학생용 AI 교육콘텐츠 사용성 평가지표를 3차에 걸친 델파이 연구를 통해 도출하였다. 델파이 연구는 기존 연구가 부족하거나 없을 때 전문가들의 의견을 적용하여 이뤄진다. AI 교육콘텐츠에 대한 사용성 평가 연구가 부재한 상태에서 AI 교육과 관련된 전문가들의 합의를 바탕으로 새로운 AI 교육콘텐츠 사용성 평가지표를 끌어내었다.

둘째, 학생용 AI 교육콘텐츠 사용성 평가지표는 4개 평가 영역하에 51개 세부 문항으로 구성하였다. 즉, 사용성 평가지표는 시스템, UI, 학습내용, 학습 방법 및 평가의 네 가지 대영역으로 구성된 51개의 문항으로 개발되었다.

2022 개정 교육과정의 가장 큰 특징 중 하나는 초·중등 학생의 디지털·AI 리터러시 함양 교육 강화라고 할 수 있다. 이로부터 다양한 AI·SW 연계 수업이 이뤄질 것으로 예상된다. 수업 지원을 위한 다양한 AI 교육콘텐츠들이 개발되고 있다. 본 연구 결과는 이러한 AI 교육콘텐츠의 사용성 평가지표로 활용될 것으로 기대한다.

본 연구의 한계는 AI 교육콘텐츠 사용성 평가

지표 개발까지를 연구 범위로 두었고 그 이후의 단계는 포함하지 않았다. 이에 추후 연구에서는 학습자나 교수자가 실제 개발된 AI 교육콘텐츠를 사용한 뒤 사용성 평가지표를 기반으로 좀 더 깊이 있는 분석이 이루어져야 할 것이다. 또 Rasch 모형을 활용하여 개발된 문항들의 양호도를 확인 하는 연구를 통해 평가지표 개선이 이뤄져야 할 것이다.

References

Big Idea 1 - Progression Chart. V.0.1 - Released May 28, 2020. www.AI4K12.org
https://ai4k12.org/wp-content/uploads/2021/01/AI4K12_Five_Big_Ideas_Poster-1.pdf

Choi YM(2002). A three-round Delphi study on the contents of Business English tests. English Teaching. Unpublished master dissertation, Ewha Womans University.

Frey CB and Osborne M(2013). The future of employment: How susceptible are jobs to computerisation?, 36-42.

Han SG(2021). Educational Contents for Concepts and Algorithms of Artificial Intelligence. Journal of the Korea Society of Computer and Information, 26(1), 37-44.
<https://doi.org/10.9708/jksci.2021.26.01.037>

Joint Ministry of Science and ICT and related ministries(2020). National AI·SW education spread

- plan, 22~24.
- Kim SH, Kim SH and Kim HC(2019). Analysis of International Educational Trends and Learning Tools for Artificial Intelligence Education. In : Proceedings of the Korean Association Of Computer Education, 23(2), 25~28.
- Kim SH, Kim SH, Lee MJ and Kim HC.(2020). Review on Artificial Intelligence Education for K-12 Students and Teachers. The Journal of Korean Association of Computer Education, 23(4), 1~11.
<https://doi.org/10.32431/kace.2020.23.4.001>
- Lawshe, CH(1975). A quantitative approach to content validity. Personnel psychology, 28(4), 563~575.
- McKinsey & Companany(2017). What's now and next in analytics, AI, and automation, 5~6.
- Ministry of Education (2020). Elementary and secondary school curriculum notice (No. 2020-236). 2020.9.11.
- The Korea Foundation for Science and Creativity (2020). Report on the results of AI education platform planning policy research, 29~32.
- The Ministry of Science and ICT (2019). National strategy for artificial intelligence, 5~9.
-
- Received : 04 February, 2022
 - Revised : 03 March, 2022
 - Accepted : 11 March, 2022