

학습 분석 기술 활용 가능성 및 전망 - 유즈케이스와 서비스 모델

조용상
한국교육학술정보원

요약

1본고에서는 교육 분야에서 다양한 데이터를 수집 및 분석하여 개인화된 학습 서비스를 제공하려는 학습 분석(Learning Analytics) 서비스의 개념과 앞으로 기대되는 유즈케이스를 소개한다. 국제적으로 주목 받고 있는 학습 분석 기술은 현재 개념화 수준에 머물러 있지만, 글로벌 기업들이 주축이 된 민간 단체에서는 데이터 수집체계와 같은 구체적인 구현 방법에 대한 논의도 추진되고 있어서 관련 현황에 대한 진단도 해본다. 특히 국제 표준화 기구와 단체를 통해 추진되고 있는 내용을 중심으로 소개한다. 다양한 데이터 응용 기술을 융합해서 기대할 수 있는 학습 분석 서비스 모형을 제시하면서 관련 정책과 제품 개발에 기여할 수 있을 것으로 기대한다.

I. 서론

스마트폰과 소셜네트워크서비스가 일반화되면서 전 세계에서 수십억의 사용자들이 생성하는 데이터의 유형과 양은 가히 폭발적인 속도로 증가하고 있으며, 정부와 공공 분야, 민간기업의 깨어있는 리더들은 이러한 다양한 데이터 활용에 주목하고 있다. 박근혜 정부도 정부 3.0정책을 통해 공공 데이터를 개방하고 데이터를 기반으로 한 신산업을 육성하는 것을 새로운 정부 운영 패러다임으로 제시함으로써 글로벌 트렌드와 맥을 같이하고 있다. 특히 빅데이터는 집단을 대표하는 데이터라기보다는 마치 개인의 필요와 행동을 표현한 그림 같은 것이어서 개인화된 맞춤 서비스를 지향하는 모든 분야에서 매우 유용하게 활용될 수 있을 것으로 기대된다[1].

그러나 교육 분야는 공공 정보 개방이나 민간의 비즈니스 인텔리전스(Business Intelligence) 분야 보다는 데이터 활용에 대한 준비가 더디게 진행되는 것 같다. 아마도 관련 이해관계

자들과 교육 서비스 기관에게 데이터 공유에 대한 동기부여가 되지 못했고, 서비스 모델 부재 역시 주요 원인으로 지적된다.

일반적으로 교실과 온라인에서 ICT(Information and Communication Technology)를 활용한 다양한 학습 활동이 수행되면서 의미 있는 데이터들이 생성되고 있지만, 이 데이터들은 활동이나 과정이 종료되면 폐기 되거나 일부만 발췌되어 기록·관리되고 있다. 그러다 보니 학습자의 이해 수준과 과정을 정확하게 이해하지 못한 상황에서 평균 수준에 맞춰진 자료와 후속 활동이 요구되고 있어서 개인화된 학습 환경을 제공하거나 수준별 학습을 지원하는 일들이 쉽지 않다. 예를 들어, 기존의 오프라인 교육에서는 시험 결과와 같은 최종 평가 데이터만 수작업으로 기록하고 있으며, 학습 과정에서의 의미 있는 활동 기록이 축적되지 못하는 실정이라서 수준별 학습 지도 및 개별 학습을 위한 한계에 직면하게 된다. 더욱이 정형화된 시험을 통해서만 학생들의 학업 수준을 측정하므로 학습 습관, 인지 수준 등 개인화된 피드백 제공에 어려움을 겪기도 한다. 따라서 이와 같은 제한적인 상황을 극복하고 개별 학습자를 위한 지원과 동기부여를 위해 데이터를 기반으로 한 새로운 서비스에 대한 고민이 필요한 시점이다.

본고에서는 학습 분석 서비스로 제공될 것으로 기대되는 유즈케이스를 소개하고 이 시나리오를 구체화하기 위해 국제 표준화 기구 및 민간 단체에서 추진 중인 표준화 동향을 분석한다. 특히, 국제 표준화 기구에서는 학습 분석 개념에 대한 확립과

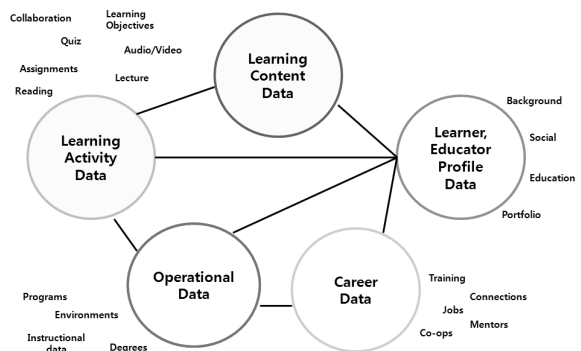


그림 1. 교육 분야의 데이터 모델(Edu Graph) [3]

1 본 연구는 미래창조과학부의 지원을 받는 정보통신 표준화 및 인증지원사업의 연구결과로 수행되었음

이해관계자들을 구체적으로 식별하는 노력을 시작한 반면, 기업들이 참여하는 민간 단체에서는 데이터 수집체계와 같은 실천적인 표준화 노력을 기울이고 있어 대조적인 움직임을 보이는데, 이 두 가지 접근법의 특징도 살펴본다. 결론에서는 학습 분석 기술을 활용하여 개인화된 학습 서비스를 완성할 수 있는 모형을 제시한다.

II. 학습 데이터에 대한 이해

학습 분석은 기본적으로 학습활동으로 인해 생성되는 데이터를 수집 및 가공하여 분석하는 일련의 프로세스 전체를 포함하는 개념이다. 이러한 개념적 이해를 토대로 학습 분석 역시 다양한 형태의 데이터를 수집하여 분석하는 빅데이터 워크플로우와 같은 맥락에서 이해할 수 있다. 대부분의 데이터를 기반으로 한 분석 서비스들은 해당 분야의 데이터 분류와 특성에 대한 정확한 이해를 바탕으로 기획되는데, 학습 분석도 수집 대상이 되는 데이터에 대한 이해가 매우 중요하다. 국제 표준화 단체인 IMS Global Learning Consortium(이하 IMS Global)은 학습 분석을 위해 교육 분야에서 수집·분석할 수 있는 데이터의 유형을 아래 목록처럼 다섯 가지 유형으로 구분하여 학습 분석을 위한 데이터 규명작업을 활발히 추진 중에 있다.

- 디지털 콘텐츠가 발생시키는 데이터
(Learning Content Data)
- 학습 플랫폼을 통해 발생하는 학습 활동 데이터
(Learning Activity Data)
- 교육기관에서 교육 프로그램 운영 중에 발생하는 데이터
(Operational Data)
- 학습자의 경력과 인맥에 대한 데이터(Career Data)
- 학습자 또는 교수자의 프로파일링 중에 발생하는 데이터
(Profile Data)

다섯 가지 유형의 데이터 중에서 학습 분석은 콘텐츠 활용 및 학습 활동과 관련해서 생성되는 데이터들이 학습 분석에서 주목하는 데이터들이다. 아래 <그림 1>은 다섯 가지 데이터 유형 간의 관계와 각 데이터 그룹의 대표적인 예들을 그림으로 정리한 것이다.

학습 분석에서 초점을 맞추는 콘텐츠 활용과 학습 활동 중에 생성되는 데이터를 수집하기 위해서는 이들 데이터의 특성과 형태에 대한 분류와 이해가 선행되어야 한다. 예를 들어, 데이터베이스에 기록된 데이터나 스프레드시트 형태의 데이터는 정형화된 데이터(structured data)로 분류한다. 이러한 정형화된

Learning Activity Metrics				Foundational Metrics
Reading > annotations > page/block use > media use > lookups	Lectures > frameset use > scrub marks > view time > weblink refs	Quiz > scores > attempts > remediation > assoc refs	Projects > deliverables > structure > milestone perf > group profile > patterns	Context > institution > course/section > learner profile > course context > path/sequence > usage context
Homework > scores > attempts > remediation > assoc refs	Media > media type > frameset use > scrub marks > view time > usage context	Tutoring > topics > assoc context > frequency > feedback	Research > searches > patterns > citations > topics	Engagement > activity usage > time on task > session time > last access > activity affinity > context affinity > task patterns > correlation
Assessment > scores > patterns (item) > time utilization > attempts > completion	Collaboration > connections > assoc context > message profile > frequency	Annotation > highlights > notes > marks > tags > attachments	Gaming > progress > cognition > attempts > hints > collaboration	Performance > grades > progress > rubrics -course goals -topic objectives -qualitative evaluation -quantitative scores > patterns > correlations
Social > connections > assoc context > message profile > frequency	Messaging > assoc context > outbound pool > inbound pool > attachments	Scheduling > assoc context > event patterns > event profile > time utilization	Discussions > post mark > frequency > participation > collaboration	

그림 2. 학습 분석을 위한 데이터 측정 프로파일[3]

데이터에 대한 기록, 관리, 분석, 도표화 기법은 오랫동안 활용된 것이어서 친숙한 것이다. 그러나 지금까지 온라인 콘텐츠를 포함한 ICT를 활용한 교육에서 주목하지 않았던 시스템 로그 파일이나, 메시지, 온라인 게시글, 제출된 전자문서 등 정형화되지 않은 데이터(unstructured data)에 대한 활용 가치가 새롭게 조명되고 있다. 정형화된 데이터와 비정형화된 데이터에 대한 구분은 데이터를 기술적인 형태로 구분한 것인데, 형태별로 데이터를 수집하는 방법과 절차상 차이에 기인한 것이라 할 수 있다.

데이터의 기술적인 형태의 차이 외에 수집 대상이 되는 데이터의 내용적인 특성에 대한 분류를 정의해야 한다. 수집 대상이 되는 데이터를 그 자체로 모으지 않고 내용상의 특성에 따라 분류하는 주된 이유는 데이터 수집체계를 효율적으로 구성하고 데이터의 품질과 분석 정확도를 높이기 위한 조치이다. 예를 들어, 학습 활동은 한 곳에서만 발생하지 않고 여러 학습 플랫폼이나 서비스를 통해 수행될 수 있고 학습자가 활용하는 콘텐츠나 소프트웨어도 어느 한 곳이 아니라 여러 곳의 플랫폼이나 단말기에서 활용되는 현실을 감안한다면 체계적인 데이터 수집체계에 대한 필요성에 공감하게 된다. 앞서 언급한 IMS Global은 다양한 이해관계자들과 함께 <그림 2>와 같이 학습 분석을 위한 데이터 매트릭스 원형을 제시한 바 있다. 그림의 좌측은 학습 활동 (또는 학습 과정) 중에 생성되는 데이터 그룹이고, 우측에 있는 데이터들은 온라인 서비스 제공이나 학습 결과(learning outcome)를 측정하기 위해 기존에 사용하던 (대부분 데이터베이스에 기록·관리되던) 데이터 그룹을 나열한 것이다. 학습 분석을 위해서는 <그림 2>의 좌측 영역, 학습 활동

측정을 위한 데이터들이 주목받고 있다. 이들 데이터의 내용적인 특성 중에는 독서 기록, 강의자료 활용 기록, 퀴즈와 시험, 과제, 프로젝트 참여 현황, 소셜 미디어 활용, 온라인 토론 참여 현황 등 매우 다양한 활동 유형이 포함될 것으로 예상된다.

Ⅲ. 학습 분석 서비스 개념

1. 학습 분석 개념과 분석 수준에 대한 이해

2000년대 초반에도 인터넷을 통해 콘텐츠와 정보를 활용하면서 다양한 데이터들이 생성되고 있었지만 스마트폰과 태블릿 등 모바일 디바이스가 보편화되면서 더 다양해진 학습 환경으로부터 생성되는 데이터는 이전 어느 때 보다 가히 폭발적으로 증가하는 추세이다. 이렇게 급증하는 데이터들을 지금까지는 이질적인 데이터로 인식하여 별도로 관리를 하거나 폐기해 왔지만, 대용량의 데이터를 처리하고 분석하는 기술의 발전 덕분에 교육계는 학습 분석을 학습자의 학습 활동을 체계적으로 지원할 수 있는 새로운 기회로 받아들이기 시작했다.

UNESCO의 보고서[2]는 학습 데이터를 수집해서 분석하는 수준을 크게 세 가지 단계로 구분하여 소개한 바 있는데, 아래 목록과 <그림 3>은 각 분석 수준별 효과와 분석 수준 간 융합에 대한 개념을 설명한 것이다.

- 마이크로 수준 분석(Micro-level analytics): 개별 학습자나 그룹의 학습활동 데이터를 추적하고 해석하는데 이용된다. 이 데이터들은 학습 분석에서 가장 미세하고 세밀한 데이터로 분류되는데, 학습자의 이해 수준을 진단하기 위해 온·오프라인에 존재하는 수집 가능한 모든 데이터를 수집한다.
- 메소 수준 분석(Meso-level analytics): 기관 수준의 분석인데, 데이터베이스의 자료와 수집 가능한

비정형 데이터들을 수집 및 분석해서 교육기관의 워크플로우를 재설계하고 교육 서비스 고객의 이탈을 미리 진단하거나 미래 서비스를 준비하는 목적으로 수행된다.

- 매크로 수준 분석(Macro-level analytics): 개인, 기관 수준의 데이터들을 축적해서 지역, 국가, 세계 수준 등 거시적인 분석을 수행하는 것인데, 예를 들면, 지역이나 국가의 교육과정 목표가 달성되고 있는지를 수집된 데이터들을 이용해서 진단할 수 있을 것으로 기대한다.

2. 학습 분석 서비스에 대한 기대와 우려

위에 소개한 학습 분석에 대한 개념과 분석 수준에 대한 정의는 매우 추상적인 수준의 내용이다. 이러한 개념을 보다 구체적으로 실현하기 위해서는 이해관계자를 식별하고 각 이해관계자 별로 기대할 수 있는 서비스에 대한 식별이 이루어져야 한다. 국제 표준화 기구인 ISO의 교육 기술 분과 위원회(ISO/IEC JTC1 SC36)는 학습 분석 기술을 전담할 워킹그룹을 신설하면서 다음과 같은 이해관계자와 각 이해관계자 별 기대를 정리한 바 있다[6].

<학습자 관점>

- 학습 활동과 진도를 측정하여 학습 지원
- 감정, 동기, 준비상태를 측정하여 학습 지원
- (장애 등) 개인의 필요와 (매체) 선호 등을 사전에 파악하여 학습 환경 구성
- 학습 활동과 평가를 통해 개선된 피드백 제공
- 정상적인 학습 과정을 완료할 수 있을 지에 대한 사전 예측을 통해 조기 경고 제공
- 개인화된 학습 경로 생성 및 학습 자원 제공

<교사 관점>

- 학습자와 그룹 활동 및 진도를 측정하여 정확한 이해 도모
- 개별 학습자의 필요와 행동에 대한 맞춤형 대응 및 조언 제공
- 학습 참여 부진 학생과 그룹을 조기에 발견
- 학업 수행 능력을 측정할 수 있는 방법과 기회 확대
- 개인과 그룹에 대한 학습 결과를 직관적으로 표현하는 시각화된 정보 활용
- 사실을 기반으로 한 개선된 학습 경험과 자원을 설계 능력

<교육 기관 관점>

- 학급 및 그룹의 활동 및 결과에 대한 효과성 측정
- 기관에서 제공한 학습 환경의 효과성 및 타당성에 대한 모니터링 및 측정 데이터 제공
- (고등 및 평생교육기관의 경우) 재등록율 제고를 위한 전략

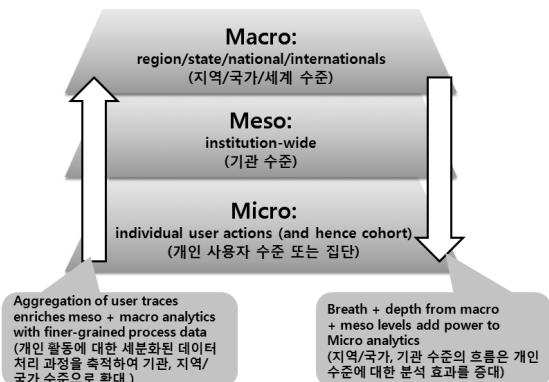


그림 3. 학습 분석 수준 간 융합 개념도[2]

수립용 기초자료 제공

- 교육과정 만족도 제고를 위한 전략 수립용 기초자료 제공
- 교육과정 설계를 지원하기 위한 데이터 제공

이상 소개한 기대에서 쉽게 추론해 볼 수 있듯이 모든 이해관계자들의 기대에는 먼저 개인에 대한 데이터가 수집될 수 있어야 한다는 전제조건이 있다. 이는 매우 민감한 사항인데, 어떤 형태이든 개인정보가 포함된 데이터를 수집 및 분석하는 것이므로 철저한 개인정보 보호 정책과 기술이 적용된 환경과 수집된 데이터를 활용한 패턴을 분석할 경우에는 익명성이 보장되어야 한다는 점이다. 어느 나라를 막론하고 개인정보를 임의로 수집해서 개인의 동의 없이 분석하는 행위가 용납되는 경우는 없을 것이다. 따라서 이해관계자들의 불안과 우려를 해소할 수 있을 정도의 법·제도적인 조치와 기술 개발이 병행되어야 할 것이다.

보통 정보 시스템은 세상에 흩어져 있는 데이터들을 거르거나 분류해서 가치 있는 정보 형태로 가공하는 역할을 수행한다. 이 역할이 정상적으로 잘 수행된다면 우리는 복잡한 현상도 단순하고 명료하게 이해할 수 있다. 반면 시스템이 정보를 잘못 가공한다면 중요한 세부점들을 놓칠 수도 있다. 학습 분석 분야도 비슷한 문제에 직면하는데, 학습 분석 시스템이 어떤 정보들을 표현하거나 표현하지 못하는지에 따라 해석력의 차이가 생길 수 있으므로 실효성과 공정성 논란이 발생할 수 있다[1]. 아래 목록은 최근 쟁점이 되고 있는 몇 가지 사항을 나열한 것이다[2].

- 자동화된 조사와 연구는 지식에 대한 정의를 변화시키고 있다
- 객관성과 정확성에 대한 주장은 오해의 소지가 있다
- 데이터가 많은 것이 항상 나은 것은 아니다
- 모든 데이터가 등가를 지니는 것은 아니다
- 누구나 데이터에 접근 가능했을 때 모두가 윤리적으로 행동하지는 않는다
- 빅데이터에 제한된 접근은 또 다른 정보격차를 야기한다

위에 나열한 쟁점들과 더불어 특히 심각하게 지적되는 데이터 중립성에 대한 논란도 있다. 이 이슈는 데이터 수집, 가공, 분석, 조정 기능까지 이어지는 학습 분석의 생명주기 각 단계에는 사람의 전문성이나 경험에 기초한 판단이 개입되기 때문에 중립적이지 않다는 것이다. 다시 말해서, 모든 상황(맥락)을 식별할 수 있는 간단한 문장으로 표현한다는 것, '데이터는 그 자체로 표현한다(data speaks for itself)'는 말을 믿는 것은 매우 순진한 발상이라는 지적도 있다[1].

그 밖에도 윤리 문제에 대한 우려도 있는데, 앞서 언급한 바와 같이 학습 분석 기술이나 서비스도 기본적으로 개인정보를 공유하고 해석하는 것이어서 정보 윤리 문제와 직결되어 있다. 미

국과 달리 국내 여건은 개인정보를 수집·보관·이용하는 절차가 엄격하게 규제되고 있기 때문에 무분별한 수집은 제도적으로 불가능하다. 그러나 제도적으로 불가능한 것과 기술적으로 불가능한 것에는 차이가 있기 때문에 윤리 문제가 심각하게 고려되는 것이다. 주로 적법한 절차를 거쳐 데이터를 수집하는 것인지, 수집된 데이터는 사전 동의를 구한 범위 내에서만 사용되고 있는지에 대한 모니터링 절차가 뒷받침되어야 하고, 모니터링 결과는 각 개인에게 투명하게 제공될 수 있어야 한다. 이 외에도 수집된 데이터를 가공하는 과정에서 다른 목적으로 데이터를 조작하는 등의 비윤리적인 행동 역시 관찰될 수 있어야 한다는 지적이 윤리 문제에 포함되고 있다. 아래 목록은 학습 분석과 관련하여 주로 대두되는 윤리 문제들을 정리한 것이다[1].

- 누가 로그 데이터의 가치를 판단할 것인가?
- 다른 데이터 세트와 결합 및 축적되면서 개인정보는 깨끗이 정리되었는가?
- 결합 및 축적된 데이터 세트는 객관적으로 비교 가능한 수준인가? (의도된 누락 또는 조작된 데이터는 없는지에 대한 이슈)
- 누가 데이터를 시각적으로 표현하도록 결정할 것인가? 그리고 표현된 데이터는 충분히 해석 가능할 정도로 보이는가?
- 학습자들은 자신과 또래 그룹의 분석 결과를 조회할 수 있는가?
- 교사들은 분석 결과를 토대로 학습 활동에 적절히 개입할 수 있을 것인가?
- 수집된 데이터는 충분히 익명성이 보장되는가? 그리고 데이터에는 합리적인 수준으로 접근할 수 있도록 허용될 것인가?

지금까지 소개한 바와 같이 학습 분석을 바라보는 기대와 우려는 각각 비슷한 수준으로 제기되고 있는 상황이다. 학습 분석이라는 개념과 기술 개발이 비교적 초기 수준이기도 하지만, 특히 우려의 목소리가 제기되는 배경에는 최근 사회적으로 개인 정보 수집 및 활용과 관련된 이슈들이 계속 등장하는 상황 때문인 것으로 보인다. 따라서 당분간은 학습 분석에 대한 기대와 우려가 같이 커져갈 것으로 예상된다.

3. 학습 분석 유즈케이스

지난 20여 년간 국내외에서 ICT를 활용한 다양한 교육 방법을 시도하면서 풍부한 디지털 자료와 학습용 소프트웨어들이 소개되었다. 그 중에는 온라인 콘텐츠와 문제은행, 게임, 소셜 네트워크 등 최신 기술이 적용된 사례들도 다수 포함된다. 그러나 새로운 기술이 적용된 콘텐츠와 서비스도 개인화된 학습환경을 지원하는 것과 학습 동기를 부여하는 면에서 다소 제한적

이었던 경험은 아쉬운 점으로 남는다. 이와 같은 과거의 경험을 토대로 학습 데이터를 기반으로 한 학습 분석 서비스에 대한 주목할 만한 유즈케이스들이 제시되고 있는데, 몇 가지 대표적인 예를 소개한다.

(1) 학습 분석 대시보드(Analytics Dashboard)

학습 플랫폼이나 학습용 소프트웨어에서 발생하는 로그 데이터들은 매우 기술적인 포맷의 데이터로서 일반 사용자들에게 제공되지 않는 유형의 데이터이다. 보통 로그 데이터는 제한적으로 데이터 마이닝을 위한 소프트웨어를 이용해서 분석 작업을 할 때 사용되기는 했지만, 테이블, 그래프 등 시각화된 포맷으로 가공되어 교수자나 학습자 등 최종 사용자에게 제공되는 정보는 아니었다. 학습 분석 대시보드 기능은 로그 데이터를 개인 사용자 또는 집단에게 이해하기 쉽게 시각화하여 제공하는 기능이다.

분석 대시보드는 로그 데이터 외에도 다양한 소스로부터 데이터를 수집하여 표현할 수 있는데, 온라인 학습 지원을 위해 운영되는 헬프데스크에 요청된 로그, 전자문서 형태로 제출된 리포트, 다양하게 사전에 설정된 변수들 간의 상관관계 등을 종합적으로 분석하여 시각화할 수 있다. 분석 대시보드를 통해서 제공된 정보를 통해 학습자들은 동료 집단 중에서 자신의 (시험 성적, 포럼 기여도, 온·오프라인 세미나 참석/참여 정도에 대한) 상대적 위치를 파악할 수도 있다[1].

(2) 예측 분석(Predictive Analytics)

인구 통계적 수치(demographics)나 과거 성취도(achievement) 등과 같은 정적인 데이터와 로그인 패턴, 온라인 토론 참여 정도 등과 같은 동적인 데이터 등 학습자 통계 데이터의 패턴을 이용해서 사용자를 분석하는 방법을 말한다. 예측 분석은 위험 단계나 우수 단계에 위치한 학습자들의 궤적을 그려서 개별적으로 필요한 시점에 처방적 조치를 제안할 수 있는 이점이 있다. 예를 들어 예측 분석을 통해 위험 단계의 궤적에 근접한 학습자에게 경고 메시지를 전달하고 평균이나 우수 단계로 진입하기 위한 활동 궤적을 안내할 수 있다. 지금까지 총괄평가를 통과할 수 있는지 여부를 판단하는 가장 신뢰할 수 있는 방법은 학습을 시작하는 시점에 시험을 치르게 해서 판단하는 방법이었다. 그러나 이 방법은 정적이고 학습 현황을 반영하지 못하는 제한점이 있기 때문에 실시간 학습 현황과 수준을 토대로 판단하기 위해서는 예측 분석 방법이 더욱 효과적이라고 할 수 있다[1].

(3) 적응형 학습 분석(Adaptive Analytics)

적응형 학습 플랫폼은 특정한 (대수학, 광합성, 치과시술 절차 등) 주제에 대한 학습자의 이해 정도를 측정해서 세부적인 피드백을 제공하고, 후속조치로서 관련된 디지털 자원을 학습자에게 제공하는 모델이다. 측정하는 방법 중 한 가지는 교과 과정을 기반으로 한 시험을 통해 진단하는 것인데 현재 기술로 가장 신뢰할 수 있는 방법이기 때문에 선호되고 있다. 이 모델의 특징은 학습자가 선택한 주제에 대해 현재 어떤 수준의 이해에 도달했는지를 진단해서 적절한 콘텐츠를 표현(adaptive presentation of content)해 주는 것이다. 경우에 따라서는 학습자가 이해할 수 없는 난이도의 콘텐츠를 생략할 수도 있기 때문에 동적인 학습자 인지 또는 적응형 콘텐츠 전달 환경에 관한 모델링이 필요하다. 무엇보다 이 모델은 풍부한 자원을 바탕으로 운영 가능한 것이어서 기존의 학습 플랫폼처럼 소프트웨어적인 기능만 제공하는 빈방 같은 모습과는 대조된다. 다시 말해, 플랫폼의 기능이 강조되지만 이 플랫폼은 풍부한 자원들로 채워진 상태에서 출발해야 한다[1].

(4) 소셜네트워크 분석(Social Network Analytics)

소셜네트워크 분석은 이미 잘 알려진 형태의 관계망 분석 모델이다. 이 분석 모델은 사람들이 어떻게 관계망을 형성하고 유지하는지에 대한 이해를 돕기 위해 이용되며 주제어에 대한 빈도나 트렌드를 관계망과 연계해서 분석할 수도 있다. 사람들은 소셜네트워크를 이용해서 서로 다른 유형의 관계를 형성할 수 있는데, 직접적인 영향을 끼치는 관계에서부터 간접적인 관계까지 관계 설정의 정도나 범위가 다양해지고 있다. 교육 분야에서는 학습자의 개인적 관계 설정이나 그룹의 구조가 효과적인 학습에 미치는 효과가 있는지에 대한 분석이 주 관심 대상이 된다.

(5) 담화 분석(Discourse Analytics)

학습 플랫폼이 제공하는 로그인 횟수, 게시글 조회수, 메시지 등록 건수 등과 같은 정량적인 정보는 매우 단순한 것이고, 학습 분석 관점에서는 너무도 당연한 기능이라고 할 수 있다. 학습 분석에서는 정량적인 데이터 이상의 것이 필요하고 교수자에게는 학습 활동의 질적인 기여도를 판단할 수 있는 데이터가 제공될 필요가 있다. 담화 분석이란 활동의 내실을 판단하기 위한 모델이다. 패턴 인식이나 구문 분석 기술 분야에서 오랫동안 연구되던 주제인데, 이제 교육 분야에 본격적으로 적용을 시도할 수 있는 여건이 조성되고 있다. 관련 연구들은 교사들이 에세이나 토론 글에 어떻게 첨삭을 하는지, 쓰기와 말하기가 학습에 어떻게 영향을 주는지, 어떻게 컴퓨터가 우수한 논쟁 글을

인식하는지에 대한 분석이 한창 진행 중이다[1].

(6) ICT를 활용한 평가(Assessment using ICT)

UNESCO의 정책 기고서[4]는 평가의 대상을 학습 경험에 대한 평가(assessing learning experience)와 개인에 대한 평가(assessing individual)로 구분하는데, 이 중에서 학습 경험에 대한 평가는 학습자의 수행에 대한 평가가 아니라 교육이 잘 이뤄지고 있는지에 대한 과정 평가라고 할 수 있다. 이 보고서에서 세부적인 기술 부분을 언급하지는 않았지만, 학습 경험에 대한 평가는 교육과정이나 목표가 잘 달성되고 있는지를 평가하는 것이므로 개별 학습자들의 인지 수준에 대한 평가결과들이 축적되었을 때 추론할 수 있는 문제라고 할 수 있다. 앞서 학습 분석의 수준을 설명할 때 제시한 메소 수준(Meso-level)에 해당되는 분석이 필요한 주제이다[1].

ICT를 이용한 평가는 단순히 선택형 문항이나 진위형 문항 이상의 평가 방식에 활용될 수 있어야 한다. 이를 위해서는 학습자가 얼마나 학습을 했는지, 어느 정도나 문제 해결능력을 향상시켰는지에 대한 판단을 할 수 있도록 과정 데이터를 수집하고 분석할 수 있어야 한다. 현재 고려할 수 있는 방법은 앞서 고려한 바와 같이 로그 데이터와 클릭스트림 데이터를 수집해서 분석하는 것이다. 그러나 이 방법은 수준을 측정하기에 충분하지 않기 때문에 제한점을 극복하기 위해서 컴퓨터 게임에 적용된 기법을 고려해 볼 수 있다. ‘스텔스 평가(stealth assessments)’ 방식이 고려될 수 있는데, 이것은 게임의 중요한 부분으로 평가되는 수행 데이터 수집방식과 유사하게 평가과정을 교육 프로그램에 녹여 넣는 방식이다. 이 방식은 학습자에게 평가 부담감을 없애기 위해 평가라는 절차를 학습 과정 중에 흐릿하게 녹여서 통합하는 방식이다. 기술적인 용어로 다시 해석하면, 활동 데이터를 수집해서 분석하는 학습 분석으로 기존의 시험형식으로 치러지는 평가를 대체하는 것이다[1].

IV. 학습 데이터 수집체계

2장에서 소개한 학습 데이터에 대한 다양한 특성과 유형들을 고려해 보면, 그에 따른 데이터를 수집하는 절차나 방법도 매우 까다로울 수 있다는 것을 예상해 볼 수 있다. 분석을 위한 데이터 가공 절차도 지금까지와는 다른 새로운 인프라와 방법이 필요해 보인다. 무엇보다 학습 데이터 수집본가 분석은 과거처럼 일부 의사결정자에게만 제공되는 데이터 마이닝과는 달리 데이터를 생성하는 학습자와 교사부터 쉽게 데이터의 수집과정을

투명하게 모니터링하고 분석 결과에도 접근할 수 있어야 한다는 특성을 고려한 데이터 수집체계를 고려해야 한다.

이와 같은 새로운 형태의 데이터 기반 분석 서비스를 구체화하기 위해서는 새로운 환경에 적합한 워크플로우 설계가 선행되어야 하는데, <그림 4>는 빅데이터 워크플로우를 적용한 학습 분석 워크플로우를 설명한 것이다. 데이터 수집단계부터 가공, 분석, 시각화 단계까지 대상이 되는 데이터 유형과 각 단계별로 요구되는 기능들이 설명되어야 한다.

특히 모바일 서비스가 보편화되고 디지털 교과서와 온라인 콘

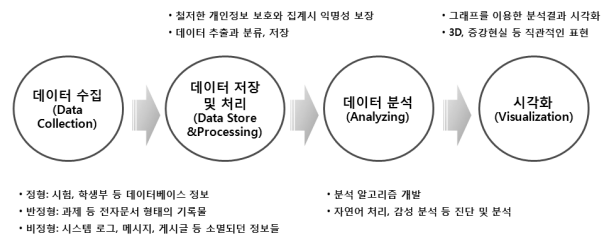


그림 4. 학습 분석 서비스를 위한 주요 워크플로우

텐츠, 문제은행, 학습용 소프트웨어, 소셜네트워크 등 다양한 애플리케이션들이 이용되는 학습 환경을 생각해 본다면, 데이터 수집 단계는 좀 더 구체적으로 설계될 필요가 있다. 아래 <그림 5>는 학습 데이터가 생성되는 부분에 대한 테스트베드와 학습 분석을 위해 데이터가 모여져서 가공, 분석, 시각화 단계의 워크플로우를 수행하는 테스트베드가 어떻게 연계되는지를 설명한 참조모델이다.

학습 데이터 수집체계를 설명하는 참조모델의 특징은 먼저 데

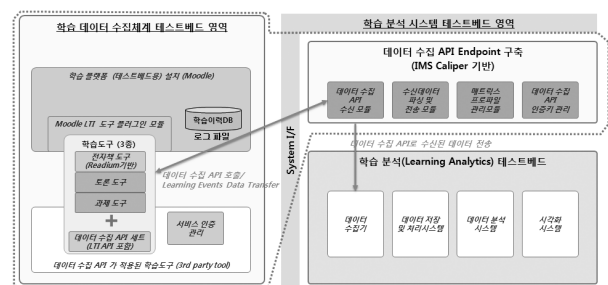


그림 5. 학습 데이터 수집체계에 대한 참조 모델

이터 생성되는 곳이 체계적으로 구성되어야 한다는 점이다. 빅데이터의 경우 비정형데이터들을 특별한 규칙 없이 수집해서 분석 알고리즘을 적용하는 반면, 학습 분석에서는 수집 단계부터 정제된 데이터(qualified data)를 대상으로 함으로써 데이터의 품질과 분석 결과의 정확도를 높일 수 있다고 보는 점에서

차이점을 찾을 수 있다. 예를 들면, 온라인 콘텐츠와 학습 활동을 관리하는 학습 플랫폼에서 생성되는 데이터와 디지털 교과서와 같은 전자책 뷰어에서 생성되는 데이터는 각각 다른 곳에서 생성되는 데이터이지만, 데이터를 수집하는 에이전트 (또는 API)를 이용해서 한 곳으로 모아져야 한다. 일관된 사용자의 식별체계를 유지하면서 로그 파일에 담긴 정보들도 데이터 수집 체계로 통합되어야 하기 때문에 2장의 <그림 5>에서 설명한 것과 같은 데이터 측정 프로파일로 표현될 필요가 있다. <그림 5>의 데이터 수집 API Endpoint에서 볼 수 있는 것처럼, 인증체계가 확립된 데이터 수집환경에서 다양한 데이터 측정 프로파일을 준수하는 데이터 노출과 수집과정이 수행되면 데이터 분석 시스템의 저장소로 모든 데이터가 모아지게 된다. 따라서 표준화된 데이터 노출 및 수집체계는 학습 분석을 위한 데이터 생태계 조성을 위한 핵심 성공요인으로 조명될 필요가 있다. 학습 데이터 수집체계의 중요성을 인식한 국제 표준화 기구와 민간 단체들이 우선 집중하는 주제도 이 부분에 대한 우선순위를 매우 높게 선정하여 현재 표준 개발을 진행 중이다.

V. 결론

본고에서는 학습 분석에 대한 개념 소개와 함께 이해관계자들이 갖고 있는 기대와 우려를 살펴보았다. 아직 개념화 단계에 있는 기술이기 때문에 학습 분석에 대한 대표적인 유즈케이스를 통해 데이터 수집체계를 설계하는 단계도 제시하였다. 유즈케이스에서 소개한 바와 같이 학습 분석이 기대를 모으는 가장 큰 이유는 개인에 대한 정확한 이해를 바탕으로 학습환경을 조성하고 학습에 대한 동기 부여를 하는 것이다. 이러한 기대를 충족하기 위해서 학습 분석은 ‘적응형 학습 분석’을 지향점으로 주변 기술과 디지털 자원들을 연계해 나갈 필요가 있다. <그림

6>은 적응형 학습 분석을 위해 구성할 수 있는 자원과 단위 시스템 간의 연계를 묘사한 것이다.

먼저 학습 분석을 통해 학습자의 취약 부분을 찾아서 보완할 수 있는 기회를 제공하려면, 교육과정에 대한 연계는 필수적이라고 할 수 있다. <그림 6>의 좌측 상단에 있는 것처럼 초중등 교육 분야부터 고등교육 분야까지 각 교육 단계별 교육과정을 의미적으로 탐색할 수 있는 링크드 데이터로 발행한다면 학습 분석 서비스와 연계가 용이할 것이다.

그 다음 과정으로 디지털 자원과 콘텐츠에 대한 접근 정보를 링크드 데이터로 발행하여 교육과정과 연계하는 노력이 필요하다. 이것은 교육과정을 구성하는 각 주제별로 활용 가능한 디지털 자원과 콘텐츠를 탐색할 수 있게 만드는 과정이다.

교육과정과 디지털 콘텐츠를 연결해 놓은 링크드 데이터가 발행되었다면, 이제 학습 데이터를 수집하여 분석한 시스템은 개인별 학습 경로를 산출하여 각 노드를 구성하는 주제별 콘텐츠와 디지털 자원을 학습자에게 제시할 수 있을 것이다.

현재 ICT를 활용한 교육 서비스나 학습 환경은 링크드 데이터를 기반으로 한 자료 탐색 환경도 미흡하고, 무엇보다 학습 데이터를 체계적으로 수집할 수 있는 데이터 생태계 조성도 준비가 안된 상태이다. 따라서 본고에서 소개한 학습 분석 유즈케이스와 데이터 수집체계에 대한 구체적인 논의와 기술 개발이 정부와 민간의 협력 하에 시급히 추진되어야 하며, 특히 데이터 생태계 조성을 위한 개방형 표준체계가 확립되어야 한다. 물론 개인정보가 포함된 민감한 데이터라는 점을 고려하여 안전한 데이터 수집 기술 개발이 역시 균형 있게 검토되어야 하며, 정보격차를 더 심화시키지 않도록 분석 결과에 대한 접근성 제고 방안도 함께 고려되어야 한다. 이와 같은 균형 있는 기술과 서비스 개발을 통해 우리 학습 환경은 모든 학습자들의 잠재력을 더 잘 발전시키고 배우는 즐거움을 갖게 해줄 수 있을 것으로 기대한다.

참고 문헌

- [1] 조용상 “학습분석 활용 가능성 및 전망”, 한국교육학술정보원 (2013)
- [2] Simon Buckingham Shum “Learning Analytics” UNESCO Policy Brief, UNESCO IITE (2012)
- [3] IMS Global “Learning Measurement for Analytics Whitepaper”, IMS Global (2013)
- [4] Kristen Blair and Daniel Schwartz “How Technology Can Change Assessment”, UNESCO Policy Brief, UN-

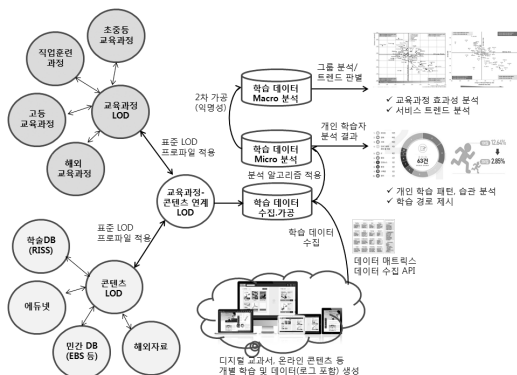


그림 6. 링크드 데이터와 연계한 학습 분석 서비스

ESCO IITE (2012)

- [5] TDWI Research “Big Data Analytics Report”, TDWI (2011)
- [6] 조용상 외 “First draft report of AHG on Learning Analytics Interoperability”, ISO/IEC JTC1 SC36 N2950, ISO (2014)
- [7] World Economic Forum “Big Data, Big Impact: New Possibilities for International Development”, WEF (2012)
- [8] McKinsey & Company “Big data: The next frontier for innovation, competition, and productivity”(2011)
- [9] Crow, M. M., “No More Excuses. EDUCAUSE Review Online”, <http://www.educause.edu/ero/article/no-more-excuses-michael-m-crow-analytics>

약 력



조 용 상

1995년 한림대학교 경영학사
2001년 성균관대학교 공학석사
2009년 성균관대학교 공학박사
2011년~2013년 스마트미디어 분야 국가표준
코디네이터 (국가기술표준원)
2002년~현재 한국교육학술정보원 책임연구원
관심분야: 디지털 콘텐츠 상호운용성, 데이터 수집
및 분석, 접근성 기술, 정보기술 표준