AIとラーニングアナリティクス

Nehal Hasnine
情報メディア教育研究センター
2024年2月20日

Agenda

- ラーニングアナリティクス(LA)のご紹介
- 教育データにAIを使用：LAコミュニティでTalk of the Town
- 自分のLA研究にAIの可能性

ラーニングアナリティクス(LA)

- LAというのは“the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens & Long, 2011, p. 34).”
- ラーニング・アナリティクス(LA)の定義は様々あ るが、「情報通信技術を用いて、教員や学生からどのような情報を獲得して、どのように分析・フィードバックすれば、どのように学習・教育が促進されるか、を 研究する分野」と考えている(緒方広明 , 2017)

ラーニングアナリティクス(LA)

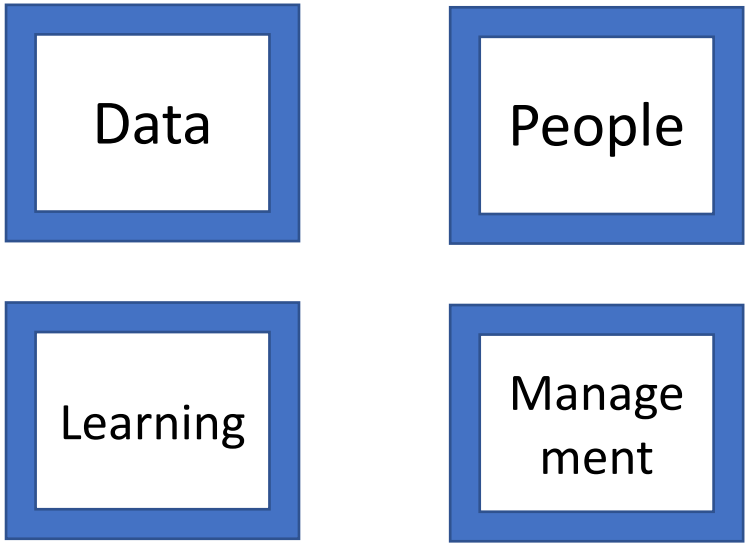
- 最初の LAK (Learning Analytics and Knowledge) カンファレンスは、2011年にカナダのアルバータ州バンフで開催された
- LAKカンファレンスはACMとACM SIGCHIおよびSIGWEBと連携して開催されている

LAK24
Kyoto, Japan
March 18-22, 2024

SOLAR
SOCIETY for LEARNING ANALYTICS RESEARCH

ラーニングアナリティクス(LA)

- 4 key elements
  - データ
  - 人々(学生・教員)
  - 学び
  - マネジメント
- 一般的なデータソース
  - ログデータ(LMS)
  - 学生のプロフィールと特徴(SIS)
  - 学歴
  - 学業成績(academic performance)



ラーニングアナリティクス(LA)

- 学習者がオンラインおよびオフラインの学習環境とやり取りする際に生成されたデータを収集および分析する
- LAでは、ラーニングデザイン、ラーニングセオリー、teaching methodsを組み合わせて、actionable insightsを生成している
- LAは、教育と学習をより深く理解し、情報を提供するためにデータをどのように使用できるかに焦点を当てている

教育研究コミュニティでLA

- 教育研究(EdTech)やHigher Education分野のhot topicになっている
- LAKだけではなくAIED(Artificial Intelligence in Education)やEDM(Educational Data Mining)研究分野でもコアになっている
- Pre-AI LAテクニック
  - Knowledge tracing
  - Predictive modelling
  - User modelling
  - Social network analysis
  - Discourse analysis
  - Text mining

Learning Analytics as a prediction model [edit]

One earlier definition discussed by the community suggested that Learning Analytics is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections for predicting and advising people's learning.[8] But this definition has been criticised by George Siemens[9] [*non-primary source needed*] and Mike Sharkey.[10][*non-primary source needed*]

Learning Analytics as a generic design framework [edit]

Dr. Wolfgang Greller and Dr. Hendrik Drachsler defined learning analytics holistically as a framework. They proposed that it is a generic design framework that can act as a useful guide for setting up analytics services in support of educational practice and learner guidance, in quality assurance, curriculum development, and in improving teacher effectiveness and efficiency. It uses a general morphological analysis (GMA) to divide the domain into six "critical dimensions".[11]

Learning Analytics as data-driven decision making [edit]

The broader term "Analytics" has been defined as the science of examining data to draw conclusions and, when used in decision-making, to present paths or courses of action.[12] From this perspective, Learning Analytics has been defined as a particular case of Analytics, in which decision-making aims to improve learning and education.[13] During the 2010s, this definition of analytics has gone further to incorporate elements of operations research such as decision trees and strategy maps to establish predictive models and to determine probabilities for certain courses of action.[12]

George Siemens is a writer, theorist, speaker, and researcher on learning, networks, technology, analytics and visualization, openness, and organizational effectiveness in digital environments. He is the originator of Connectivism theory and author of the article *Connectivism: A Learning Theory for the Digital Age* and the book *Knowing Knowledge* – an exploration of the impact of the changed context and characteristics of knowledge.[6][7] He is the founding President of the Society for Learning Analytics Research (SoLAR).

Learning Analytics as an application of analytics [edit]

Another approach for defining Learning Analytics is based on the concept of Analytics interpreted as the *process* of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data.[14][15] From this point of view, Learning Analytics emerges as a type of Analytics (as a *process*), in which the data, the problem definition and the insights are learning-related.

In 2016, a research jointly conducted by the New Media Consortium (NMC) and the EDUCAUSE Learning Initiative (ELI) -an EDUCAUSE Program- describes six areas of emerging technology that will have had significant impact on higher education and creative expression by the end of 2020. As a result of this research, Learning analytics was defined as an educational application of web analytics aimed at learner profiling, a process of gathering and analyzing details of individual student interactions in online learning activities.[16]

Learning analytics as an application of data science [edit]

In 2017, Gašević, Kovanović, and Joksimović proposed a consolidated model of learning analytics.[17] The model posits that learning analytics is defined at the intersection of three disciplines: data science, theory, and design. Data science offers computational methods and techniques for data collection, pre-processing, analysis, and presentation. Theory is typically drawn from the literature in the learning sciences, education,

Source: Learning analytics Wikipedia

教育データにAIを使用

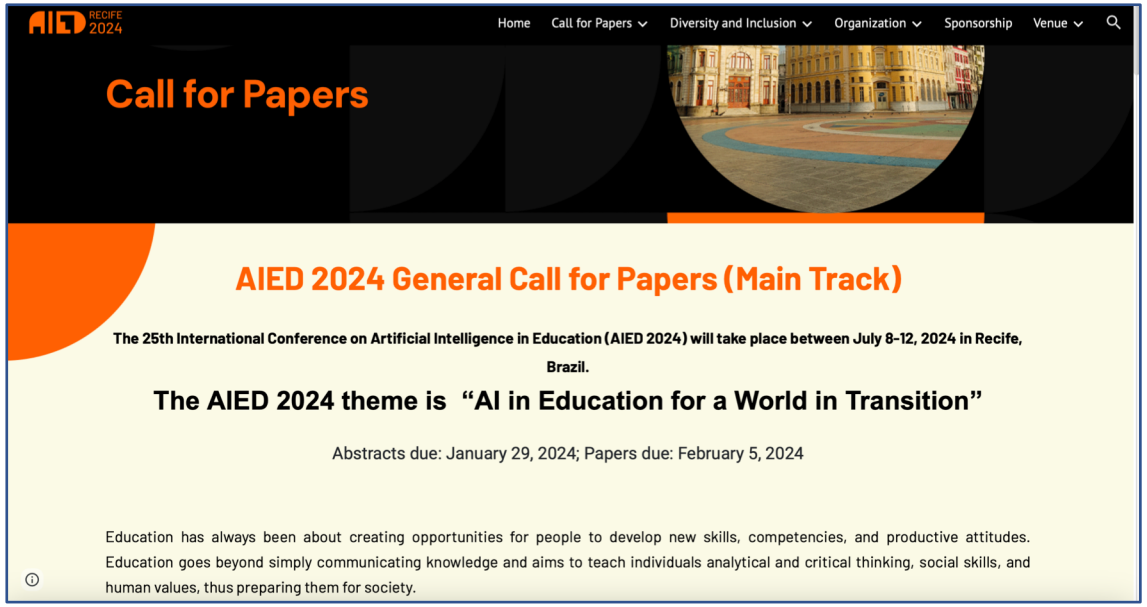
- Talk of the town

EDM 2024: the 17th International Conference on Educational Data Mining

Atlanta, Georgia, USA, July 14—17, 2024

New tools, new prospects, new risks – educational data mining in the age of generative AI

Educational Data Mining is a leading international forum for high-quality research that mines datasets to answer educational research questions, including exploring how people learn and how they teach. These data may originate from a variety of learning contexts, including learning and information management systems, interactive learning environments, intelligent tutoring systems, educational games, and data-rich learning activities. Educational data mining considers a wide variety of types of data, including but not limited to log files, student-produced artifacts, discourse, learning content and context, sensor data, and multi-resource and multimodal streams. The overarching goal of the Educational Data Mining research community is to support



General Call

The 2024 edition of ***The International Conference on Learning Analytics & Knowledge (LAK24)*** will take place in Kyoto, Japan. LAK24 is organized by the Society for Learning Analytics Research (SoLAR) with Kyoto University. LAK24 is a collaborative effort by learning analytics researchers and practitioners to share the most rigorous cutting edge work in learning analytics.

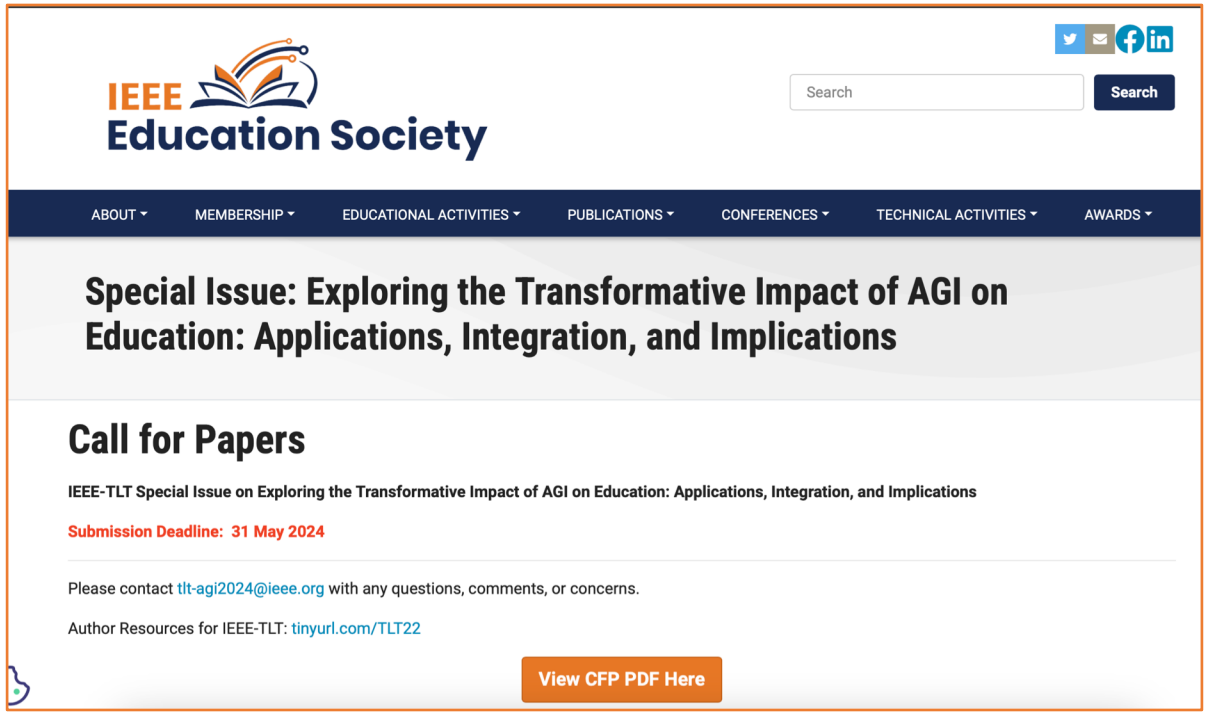
Learning Analytics in the Age of Artificial Intelligence

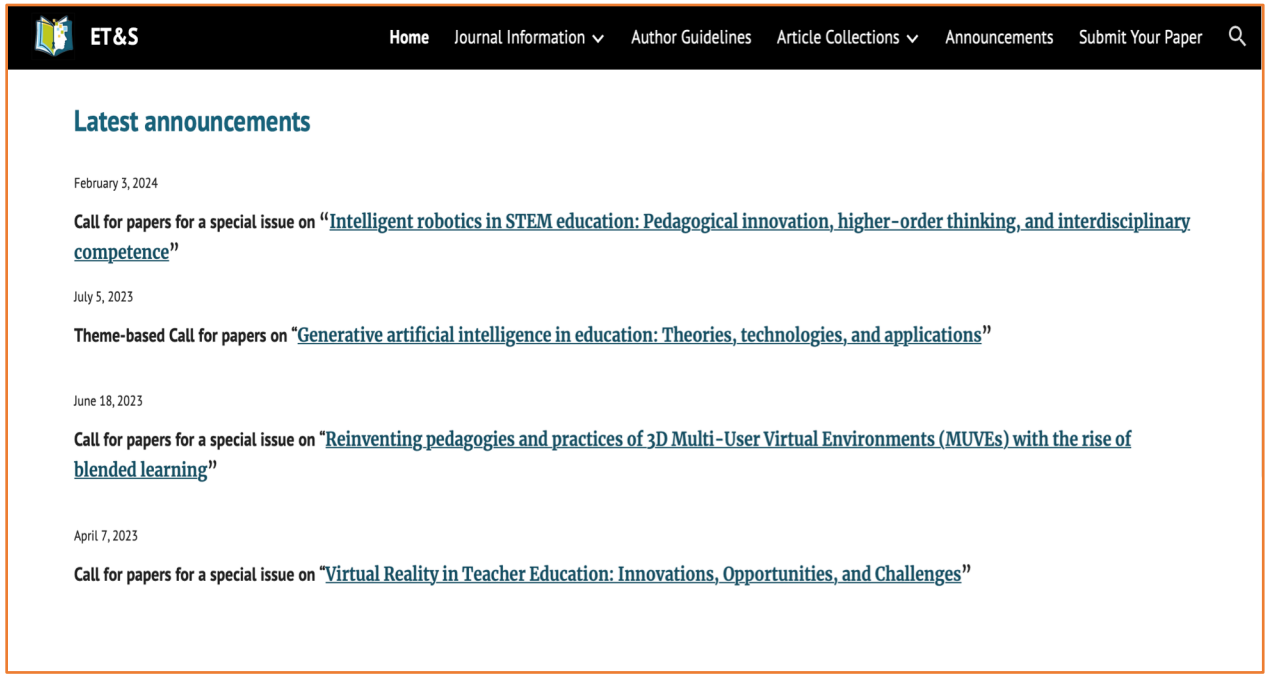
The theme for the 14th annual LAK conference is ***Learning Analytics in the Age of Artificial Intelligence***. Artificial intelligence has been relevant for learning analytics since the early days of the field. This has mostly been manifested by building upon the algorithms of machine learning to analyze data about learners and learning environments. The conversations about artificial intelligence in education used to be mostly contained within specialized communities of practitioners and researchers. Since late 2022, this has rapidly changed. Discourse in mainstream media and among the general public has been dominated by the coverage of the developments in generative artificial intelligence. The notable examples are such technologies as ChatGPT and DALL-E that harness the power of deep learning algorithms to generate impressively human-like text and images based on relatively simple human prompts. These technologies have given some glimpses about the emerging age of artificial intelligence. The prominence of artificial intelligence has also opened profound debates about implications on education from the need to develop relevant literacies to work with artificial intelligence to challenging the established notions of assessment in education. Through the theme of the 14th annual LAK conference, we encourage the authors to consider implications for learning analytics and the role the field can play in the age of artificial intelligence.

The LAK conference is intended for both researchers and practitioners. We invite both researchers and practitioners of learning analytics to come and join a proactive dialogue around the future of learning analytics and its practical adoption. We further extend our invite to educators, leaders, administrators, government and industry professionals interested in the field of learning analytics and its related disciplines.

教育データにAIを使用

- Talk of the town





なぜTalk of the Townでしょうか？

- Personalized Learning
  - 学習者の強み(strength)、短所(weakness)、学習の好み(learning styles)を理解できる
- Intelligent Tutoring Systems
  - AIを活用したインテリジェントな個別指導システムは仮想メンターとして機能し、学習者にリアルタイムのフィードバック、ガイダンス、サポートを提供することができる
  - AI-enabled ITSシステムでは、生徒が苦労している領域を特定し、課題を克服するために的を絞った介入を提供できる
- Enhanced Content Creation
  - 自然言語処理 (NLP)を使用すると、AI が人間のような応答と説明(human-like responses and explanations)を生成できるため、オンライン学習がより魅力的でインタラクティブになる
- Smart Data Analysis
  - AIは、教育者や管理者がデータに基づいた意思決定を行えるようにする傾向、パターン、洞察を特定できる
- Adaptive Assessments
  - 学習者の回答に基づいて質問の難易度を動的に調整する
  - 適応型評価はテストの不安を軽減し、学習者の能力をより公平に評価する

AI-powered LAツール開発・評価

Author(s) and Year	Country	Goals and Objectives	Participants	Data Sources	Techniques	Tools	Ethical Procedures	Results
Michos and Hernández-Leo (2018) [46]	Spain	To support community awareness to facilitate teachers' learning design process using a dashboard with data visualizations.	23 PSTs, 209 ISTs	• ILDE dashboard: profile views, comments, created designs, re-used designs, and edits.	LA dashboard	The Integrated Learning Design Environment (ILDE) dashboard IBM SPSS 22 Heidi SQL and Tableau	n.d.	The ILDE dashboard can provide an understanding of the social presence in the community of teachers. Visualization was the most commonly used feature. There were time constraints.
Montgomery et al. (2019) [47]	Canada	To examine the relationships between self-regulated learning behaviors and academic achievements.	157 PSTs	• Self-regulated behaviors: • Activating: online access location, day of the week, time of day • Sustaining: access frequency • Structuring: average logins per week, exam review patterns, number of reviewed quizzes/day	LA	Moodle platform	n.d.	84.5% of PTSs' access to the platform took place off-campus. The strongest predictors for student success were the access day of the week and access frequency.
Newmann et al. (2021) [48]	Germany	To support PSTs' self-study using chatbots as a tool to scale mentoring processes.	19 PSTs	• Social bot: user intentions, bot messages • System Usability Scale: frequency, ease of use, confidence, consistency	AI NLP	Chatbots: Feedbot for self-study, Litbot for mentoring students' reading	n.d.	Promising results that bear the potential for digital mentoring to support students.
Post (2019) [49]	USA	To challenge PSTs to analyze and interpret data on students' online behavior and learning.	n.d. PSTs	• Learning action logs about search terms, visited websites, time spent on each website, and the order in which sites were visited	LA	Thinking app (Chrome extension) that tracks online behaviors	n.d.	PSTs lacked media literacy skills. Online assignments promoted student-centered learning and critical thinking. The prevalence of multitasking was highlighted.
Pu et al. (2021) [50]	Malaysia	To design a service-learning-based module training AI subjects (SLBM-TAIS).	60 PSTs	• Psychological variables: • Practical knowledge: educational beliefs, interpersonal relationships, teaching strategies, self-reflection • Motivation: intrinsic motivation, extrinsic motivation, amotivation • Other: gender, teaching experience, average academic performance	AI	The SLBM-TAIS educational module	n.d.	The SLBM-TAIS was effective in training PSTs to teach AI subjects to primary school students. The SLBM-TAIS module influences situational knowledge, teaching strategies, and both intrinsic and extrinsic motivation.
Sasmoko et al. (2019) [51]	Indonesia	To determine teacher engagement using artificial neural networks.	10,642 ISTs	• Based on the Indonesian Teacher Engagement Index (ITEI): positive psychology, positive education, teacher performance, nationalistic character, and leadership engagement	AI machine learning (ANN)	Django: a website framework for Python Chart.js for data visualization. MongoDB as the database	Not applicable	The ANN classification accuracy was 97.65%, proving the reliability of the instruments and websites; however, this still requires further testing in terms of both ease of use and trials with diverse data.
Sun et al. (2019) [52]	China	To investigate changes in PSTs' concept of engagement, analyzing data recorded during PSTs' discussions via an MOOC platform.	53 PSTs	• Discussion data • Dimensions based on Bloom's taxonomy: remember, understand, apply, analyze, evaluate, create	LA	MOOC platform	n.d.	The most frequent discussion behaviors were evaluated (31.52%) and analyzed (27.77%). PSTs with an analytical style implemented multiple strategies for learning.
Vazhayil et al. (2019) [53]	India	To introduce AI literacy and AI thinking to in-service secondary school teachers.	34 ISTs	• Types of AI tasks: text recognition, sentiment analysis, image classification, categorical/numerical data	AI	IBM Watson AI model Mitsuku chatbot Google AI experiment named Emoji Scavenger Hunt Scratch	15 ISTs consented to recorded video testimonials	77% appreciated peer teaching, 41% preferred the game-based approach, and 24% were concerned about internet access. The best strategy was embracing creative freedom and peer teaching to boost learners' confidence.

Salas-Pilco, Sdenka Zobeida, Kejiang Xiao, and Xinyun Hu. "Artificial intelligence and learning analytics in teacher education: A systematic review." *Education Sciences* 12.8 (2022): 569.

AI-powered LAツール開発・評価

Author(s) and Year	Country	Goals and Objectives	Participants	Data Sources	Techniques	Tools	Ethical Procedures	Results
Wulff et al. (2020) [54]	Germany	To employ AI algorithms for classifying written reflections according to a reflection-supporting model.	17 PSTs	• Reflection elements: circumstances, description, evaluation, alternatives, consequences	AI natural language processing	• Doc2Vec features • Four classifiers: decision trees, multinomial logistic regression, multinomial naïve Bayes, stochastic gradient descent	PSTs provided informed consent	The multinomial logistic regression was the most suitable classifier (0.63). Imprecise writing was a barrier to accurate computer-based classification.
Yang et al. (2020) [55]	China	To enhance self-directed reflective assessment (SDRA) using LA.	47 PSTs	• Epistemic agency, democratic knowledge, improvable ideas, reflective and transformative assessment, and community knowledge	LA	• Knowledge Forum (online notes)	Ethical approval was obtained from the hosting institution	SDRA fostered PSTs' collective empowerment, as reflected by their collective decision making, synthesis of ideas, and "rising above" ideas.
Yilmaz and Yilmaz (2020) [56]	Turkey	To examine PSTs' perceptions of personalized recommendations and feedback based on LA.	40 PSTs	• LMS log data: date, login frequency, views per week, participation in discussions	LA	• Moodle LMS platform	(Voluntary participation)	LA helped to identify learning deficiencies, provided self-assessment and personalized learning, improved academic performance, and instilled a positive attitude toward the course.
Yoo and Rho (2020) [57]	Korea	To determine ISTs' training and professional development using ML.	2933 ISTs, 177 principals	• Based on the Teaching and Learning International Survey (TALIS) 2013: types of activities, participation rates, intensity of participation, mentoring and induction programs	AI machine learning	• Group Mnet technique (glmnet package). • R software	Not applicable	Identified 18 predictors of ISTs' professional development. Found 11 new predictors related to ISTs' pedagogical preparedness, feedback, and participation.
Zhang J. et al. (2021) [58]	China	To build an intelligent assessment system of PSTs teaching competency.	240 PSTs	• PSTs' teaching competency framework (six dimensions): professional foundation, instructional design, teaching implementation, technology application, teaching evaluation, reflective development	AI machine learning	• AI tools: Back Propagation (BP) neural network • Delphi and Analytic Hierarchy Process (AHP) methods • Matlab software	n.d.	The trained model can be used to evaluate PSTs' competency on a large scale, its relative error was small between 0–0.2.
Zhang S. et al. (2021) [59]	China	To automatically detect the discourse characteristics of in-service teachers from online textual data.	1834 ISTs	• Discourse characteristics: number of posts per teacher, length of post per teacher, much or little new information, high or low topic relevance	AI natural language processing	• Word2vec toolkit to generate lexical vectors based on AI-NLP	Ethical approval from the institution	New and relevant information was posted at the beginning of the online discourse. Cluster analysis showed three different posts: relevant topic with new information, another with little new information, and a less relevant topic with little new information.
Zhao et al. (2021) [60]	China	To improve the outdoor learning experience and build a learning resource based on ontology information retrieval.	38 PSTs	• Vision-based mobile augmented reality from the university campus (e.g., plants, flowers, trees) through scene detection, retrieval, superposition, visualization, and interaction	AI vision-based mobile augmented reality (VMAR)	• MobileNetV2 network: a lightweight convolutional neural network by Google for mobile devices	n.d.	PSTs perceived the usability as good; it was preferred by younger users, and had a positive impact on learning. The average precision of retrieval based on keywords (97.46%) and ontology (90.85%) signified good performance.

Note. CDA = classroom discourse analyzer, DT = decision tree, FFTrees = fast-and-frugal trees, GBTD = gradient-boosted decision trees, KBSD = Knowledge–Behavior–Social Dashboard, ITEI = Indonesian Teacher Engagement Index, RF = random forests, NLP = natural language processing, SLBM-TAIS = service-learning-based module training AI subjects, SVM = support-vector machine, WISE = web-based inquiry science environment.

Salas-Pilco, Sdenka Zobeida, Kejiang Xiao, and Xinyun Hu. "Artificial intelligence and learning analytics in teacher education: A systematic review." *Education Sciences* 12.8 (2022): 569.

ミシガン大学のUniversity-wide AI Platform

- アメリカのミシガン大学がAIを活用したcustom AI platformを構築した
- [Getting Started With U-M Maizey]

Custom GenAI Services for the U-M Community

ITS offers a generative AI platform available to all active U-M faculty, staff, and students on the Ann Arbor, Flint, and Dearborn campuses and Michigan Medicine. These service offerings are equitable, accessible, and support everything from basic consumer usage to advanced research and experimentation.

U-M GPT

Our most accessible ITS AI service is U-M GPT, a tool that provides access to popular hosted AI models such as Azure OpenAI and U-M hosted open-source large language models.

- User-friendly interface that allows faculty, staff, and students to engage in chat-based queries and benefit from the expertise of GenAI technology.
- U-M GPT enhances teaching, learning, research, and collaboration, providing a valuable and equitable resource for the entire university community.
- U-M GPT has been designed to be accessible, including for use with screen readers.

U-M Maizey

Empowers users to extract valuable insights, discover patterns, and gain deeper knowledge from the available datasets.

U-M GPT Toolkit

designed for those who require full control over their AI environments and models.

U-M GenAI Videos

U-M GPT Wants You to Have a Great 2023 Winter Break

December 4, 2023

U-M Maizey Lets You Build a Teaching Assistant AI in Minutes

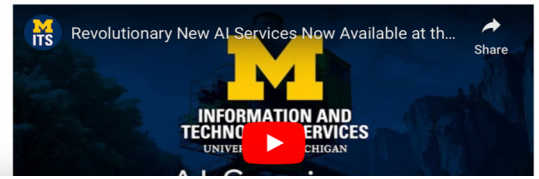
November 21, 2023

How Faculty Can Use the U-M Maizey AI Tool to Improve the Classroom Experience

October 20, 2023

Revolutionary New AI Services Now Available at the University of Michigan

August 25, 2023



LA研究でAIモデルの連携

- 生成AI (Generative Artificial Intelligent)
- 画像生成AI
- 大規模言語モデル (Large Language Model, LLM)
- Explainable AI
- Artificial General Intelligence
- Self-aware AI

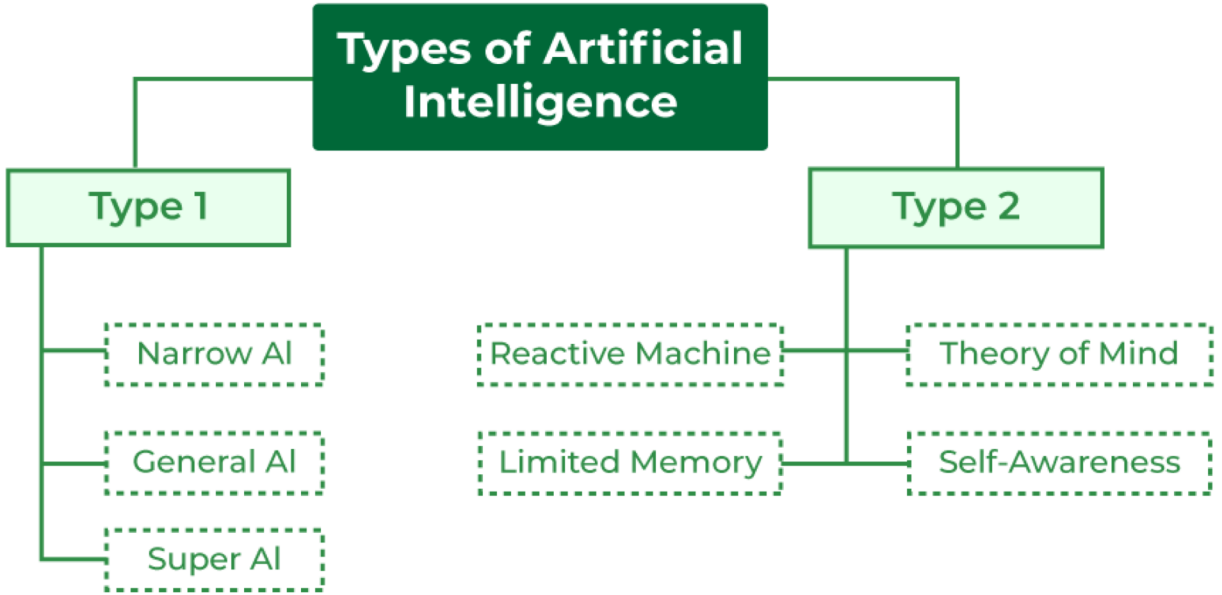


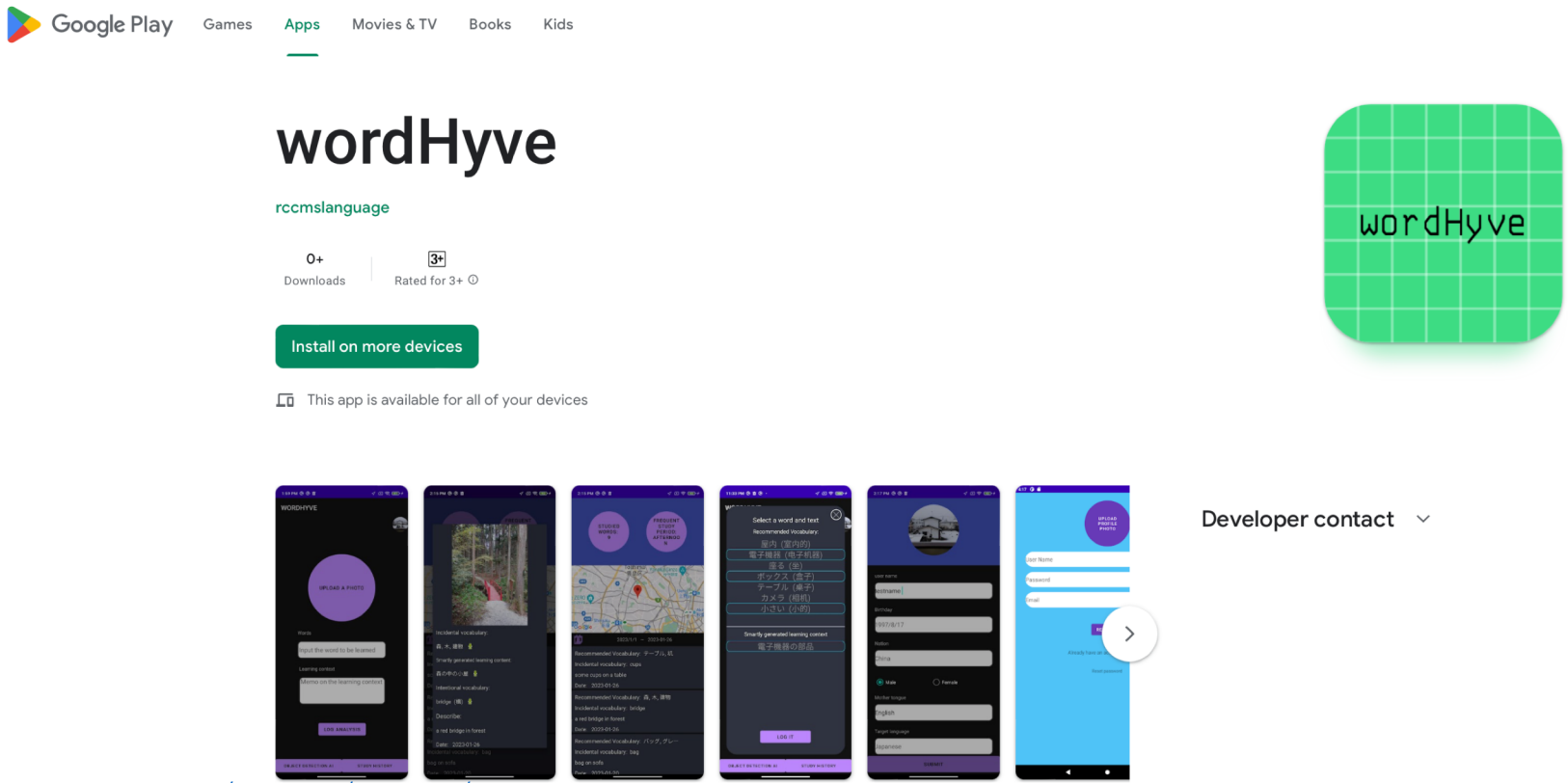
Photo source: https://www.geeksforgeeks.org/

自分のLA研究にAIの可能性

- プロジェクトのご紹介
  1. Ubiquitous Learning Analyticsに関する研究
     - 語彙学習支援システムWordhyveアプリケーション
  2. Multimodal Learning Analyticsに関する研究
     - MOEMOシステム
     - Teacher-facing LAD(ラーニングアナリティクスダッシュボード)

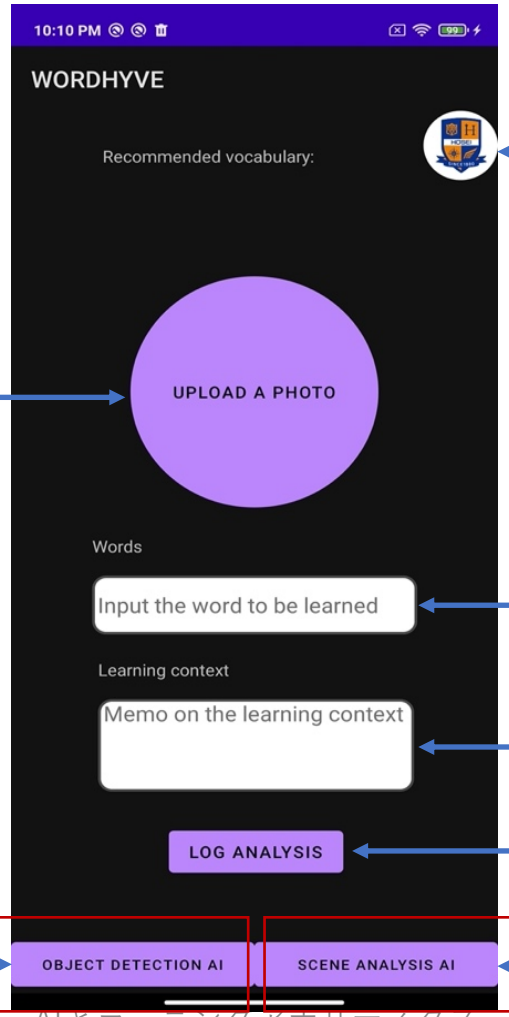
Ubiquitous Learning Analyticsに関する研究プロジェクトの報告

- Wordhyveアプリケーションを開発した
- WordhyveをGoogle Playにリリースした



https://play.google.com/store/apps/details?id=com.hosei.myapplication&pli=1

Wordhyveアプリ



画像のアップロード
(カメラからでもOK)

プロファイル設定

学びたい単語
(native・L2でもOK)

メモ

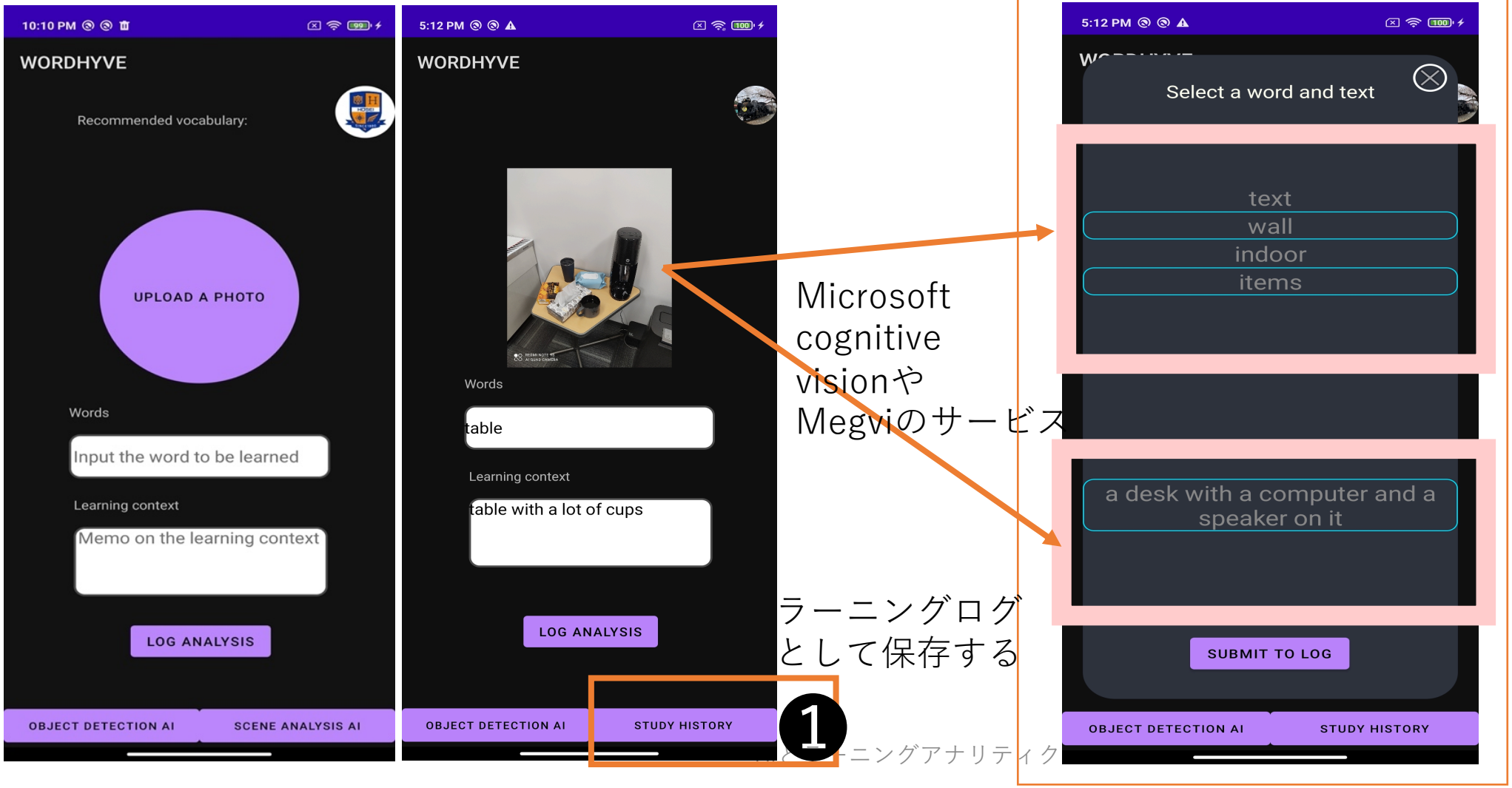
ログ分析

リアルタイムで
object detectionの機能

リアルタイムで
scene analysisの機能

AIとラーニングアナリティクス

Wordhyveアプリ



Wordhyveアプリ

- Wordhyveアプリケーションを用いることでどのような学習者(年齢，性別，言語)がどのような語彙を学習したのかの学習ログを収集する
- その収集した学習ログと様々な生成AI及び画像生成AIを用いて
    1. 個々の学習者のプロファイル情報や学習内容に応じた適切な画像を推薦するContext-specific Appropriate Images方式と
    2. 適切な学習ノートを推薦するGenerative Learning Contexts方式を実装したシステムを研究開発する

アプローチ

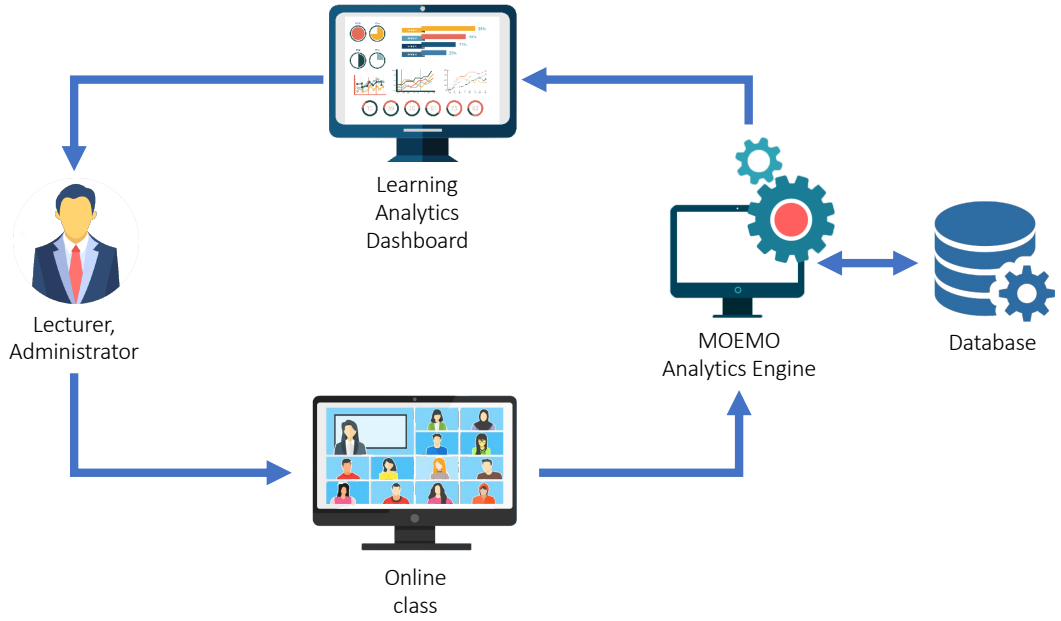
- 学習者の興味・国籍ごとに学習した単語に対する画像の正解データと不正解データを作成し，生成AIの技術 (GPT-4など) を用いて学習者のCultural Backgroundごとの単語に対する画像の解釈をWordhyveアプリ上に提示
- 画像生成AIの技術(Stable Diffusionなど)を用いて学習者の学習コンテキスト(学習者の国籍, 学習場所, 時間など)から適切な画像を作成・推薦
- 大規模言語モデル(Meta社のLlama-2など)を使用することで学習ログのテキストデータから学習者の学習状況(知識レベル，エモーション，インタラクション情報など)を分析して単語学習ノートを生成

→ **Small to Mid-scale教育データにAIを使用**

Accepted at
ARLE IFTE2024
June 18-22, 2024

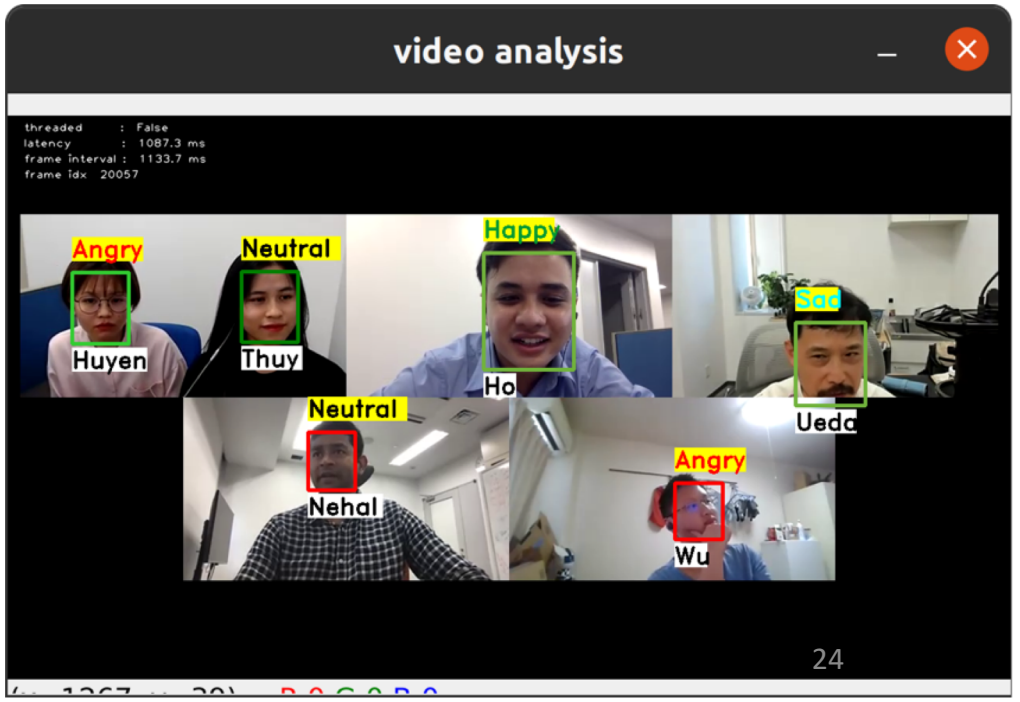
Multimodal Learning Analyticsに関する研究プロジェクトの報告

- MOEMOシステムの目的
  - ZoomやWebexビデオを分析して学生のaffective states(感情)やエンゲージメントやコンセントレーションのレベルを予測する
  - その結果を教員・担当の先生にリアルタイムで報告する



MOEMOでは

- 7種類のaffective states (happy, sad, angryなど)が分かる
- 5種類のエンゲージメント(highly engaged, low engaged, disengagedなど)が分かる
- 2種類のコンセントレーション(focusedかdistracted)が分かる
- 学生のクラスターが分かる
- After-classレポートの作成すること



https://doi.org/10.3390/s23094243

MOEMOのサマリ

- AI-based behavior recognitionテクニックを使用しClassroom monitoringができる
- オンライン授業やZoom breakout roomで使うことになっている
- Teacher-facingダッシュボードで学生のaffective statesをモニタリングすることができる

My thoughts

- AIは学習プロセスの理解に影響がある
- Small to Mid-scaleデータセットには役に立つ
- AIは人間の目には感知できないパターンや手がかりを識別できる
- Learning is contextual and personal- two students are not same
- AIを活用する際には、新ラーニングセオリー かハイブリッドラーニングセオリーを活用することが重要
  - 例えば、Self-directed learning, Metacognition, Social emotional learning, Constructivism learning theoryなど

Watchouts

- 学習ログにAIを統合する際の注意点
  - Transparency issues
  - Over-reliance
  - Bias
  - Security concerns