A multimodal interaction solutions. "The Way" for educational resources

Manuel SOTO RAMOS[1], Rebecca ACHEAMPONG[1], Dorin-Mircea POPOVICI[2]

[1] Transilvania University of Brașov, 29 Eroilor Bd., Postal code 500036, Brașov, România

[2] Ovidius University of Constanța, 124 Mamaia Bd., Postal code 900527, Constanța, România

manuel.ramos@unitbv.ro, rebecca.acheampong@unitbv.ro, dmpopovici@univ-ovidius.ro

Abstract: *This paper examines the implementation of multimodal interaction solutions emphasizing education and explores the potential of these solutions for improving learning outcomes. The paper presents a detailed analysis of the developments and advancements in multimodal interaction solutions specifically designed for educational contexts, including the use of natural user interfaces, virtual and augmented reality, voice, gesture and touch interaction, smart eye, and affective and intelligent tutoring systems. The paper also discusses the benefits and challenges associated with the implementation of these solutions, as well as future directions for research and development, multimodal interaction solutions cannot longer be seen as "a way", they must be "the way". What do we already have? What do we have and do not use? What are the benefits of this solution? Those are questions we are answering in this paper.*

Keywords: Multimodal interaction, Virtual reality, Augmented reality, Voice interaction, Gesture and touch interaction, Smart eye, Affective, intelligent tutoring systems.

1. Introduction

Technology has progressed to the point that it is possible to use a variety of different methods to achieve the same objective. This is true in many areas, from transportation to communication, and it is also true in the world of problem-solving. Multimodal solutions are a way for individuals and organizations to combine different approaches to achieve a single desired outcome. Multimodal interaction solutions are gaining popularity rapidly within the education sector.

The primary objective of these solutions in education is to offer students a highly immersive and interactive learning experience that is engaging and effective. They are used to combine several types of media, such as text, audio, video, and graphics, to create a more immersive learning environment (Chen et al., 2018), by combining these distinct types of media, students can better understand and retain the material they are learning. Also, they are used in a variety of educational settings, including traditional classrooms, online courses, and virtual classrooms (Medina, 2021).

https://doi.org/10.58503/icvl-v18y202306

The incorporation of multimodal interaction solutions can contribute to the establishment of a highly captivating and interactive learning environment, thereby fostering student motivation and heightened engagement with the subject matter. Multimodal interaction solutions can help to reduce the amount of time needed to complete a course, as students can access the material more quickly and easily. What do we already have? What do we have and do not use? What are the benefits of this solution? Those are questions we are answering in this paper, where, we are also discussing the several types of multimodal interaction solutions used in education (Munteanu et al., 2017), and the challenges associated with implementing them. Furthermore, we are exploring the possibilities of utilizing multimodal interaction solutions to foster dynamic and immersive learning environments that captivate and involve learners effectively.

2. A view about multimodal interaction solutions

Multimodal interaction solutions are technologies that enable individuals to actively engage with digital content using multiple input and output modalities. Examples of multimodal interaction solutions include voice recognition, gesture recognition, facial recognition (visual, auditory, tactile, and/or cognitive), and augmented reality (Sugumaran, Sreedevi & Xu, 2022). Visual modes might include images, diagrams, graphs, and text, while auditory modes could involve spoken words, music, or sound effects. Tactile modes could involve physical objects, while cognitive modes might involve problem-solving strategies or creative thinking. This paper will present several types of multimodal solutions and how they have been used to create effective solutions, particularly in Education.

Voice-Enabled Assistants (Intelligence n.d.) have gained tremendous popularity in recent years, becoming an integral part of our daily lives. These digital companions, powered by sophisticated natural language processing (NLP) algorithms, listen to user commands, and respond with relevant information or actions. Voice-activated assistants such as Amazon Alexa, Apple Siri, and Google Assistant have seamlessly integrated into our lives, effortlessly understanding our verbal commands, and providing virtual assistance.

Gesture Recognition (Encyclopedia, n.d.) has taken interactive technology to a new level, allowing users to interact with devices through hand movements, body gestures, or facial expressions. These systems use advanced sensors and cameras to detect and interpret gestures accurately. Immersive gaming consoles like Microsoft Xbox Kinect have capitalized on gesture recognition technology, enabling players to control games through intuitive body movements. Similarly, smartphones with gesture-based controls offer users a touch-free alternative for tasks like answering calls or navigating through applications.

Touch and Haptic Feedback (See, Choco & Chandramohan, 2022) have transformed the way we interact with our devices, making it more tactile and engaging. Touchscreen technology has become a standard feature in smartphones,

tablets, and many other devices, empowering users with the ability to tap, swipe, pinch, and zoom to interact with content effortlessly.

Augmented Reality (AR) and Virtual Reality (VR) technologies (Xiong et al., 2021) have opened up new possibilities for multimodal interactions. AR overlays digital information onto the real-world environment, enabling users to interact with a combination of voice commands, gestures, and visual cues. On the other hand, VR immerses users entirely into virtual environments, where they can interact with objects using hand controllers and voice commands, blurring the line between the real and digital realms.

At the forefront of cutting-edge interaction solutions are Brain-Computer Interfaces (BCI) (Gonfalonieri, 2020). Although still in the early stages of development, BCIs hold immense promise for the future. By establishing a direct communication link between the human brain and external devices, these interfaces enable users to interact and control technology using their thoughts.

Natural Language Interfaces (World Wide Web Consortium, n.d.) empower users to interact with technology using everyday language, making interactions feel more human-like and conversational. These interfaces go beyond simple voice commands and cater to more complex queries and responses.

In collaborative environments, Multi-Touch Tabletops (Niu, McCrickard & Nguyen 2016) have gained popularity, facilitating interactive group work and presentations. These large touch-sensitive displays enable multiple users to interact simultaneously, making them ideal for brainstorming sessions, interactive exhibits, and educational settings.

Eye-Tracking Interfaces (Bitbrain, n.d.) offer a hands-free and natural way to interact with technology. These systems use specialized sensors to track a user's eye movements, allowing them to control devices and interact with content without the need for physical touch or speech.

Wearable Technology has integrated multimodal interaction features into our everyday accessories. Smartwatches, fitness trackers, and other wearables combine touch, voice, and gesture-based inputs to provide users with a seamless and versatile interaction experience. These devices enable users to check notifications, track health metrics, and control connected devices with ease.

3. Companies who develop multimodal interaction solutions

Companies that develop multimodal interaction solutions are at the forefront of revolutionizing the way humans interact with technology. These innovative organizations harness cutting-edge technologies like natural language processing, speech recognition, gesture tracking, and artificial intelligence to create seamless and intuitive user experiences. By integrating multiple modes of communication, such as voice commands, touch, and gestures, these companies empower users to

interact with devices and applications in a more natural and immersive manner. Here are some examples of these companies:

1. Sensory (Sensory, n.d.): Sensory is a technology company that specializes in AI-based facial and speech recognition, biometrics, and embedded machine learning technologies. They offer a suite of tools designed to make user experience (UX) more intuitive and secure. Their solutions are used in a variety of industries, including automotive, banking, healthcare, and consumer electronics.

2. Affective (Affective, n.d.): Smart Eye has extended the utilization of Affective Emotion AI technology by introducing novel conversational engagement metrics. These metrics offer enhanced insights into consumer responses by analysing facial expressions in qualitative online research studies.

3. Inbenta (Inbenta, n.d.): Conversational AI with ZERO training required. It builds conversational experiences with patented Neuro-Symbolic AI 4-in-1 platform. It powers interactions with customers and employees giving answers to the most complex questions. In fact, it helps but sometimes provides predictable answers.

4. eGain (eGain, n,d,): eGain Corporation, previously known as eGain Communications Corporation, is a California-based cloud-based software company headquartered in Sunnyvale. The company specializes in delivering customer service, knowledge management, and analytics applications that businesses leverage to effectively serve and engage with their customers, ultimately driving sales and enhancing customer experiences.

5. Interactions (Interactions, n.d.): The company is privately owned and specializes in developing and offering hosted Virtual Assistant applications. These applications equip businesses with the capability to provide automated natural language communications for their enterprise customer care services. With these solutions in place, businesses can enhance their customer care operations by enabling seamless and automated interactions with their customers.

4. Popular multimodal interaction solutions. Voice recognition

Voice recognition technology has emerged as one of the most popular multimodal interaction solutions, revolutionizing the way we interact with devices and services. With the ability to convert spoken language into text or actionable commands, voice recognition has become a ubiquitous feature in our daily lives. This cutting-edge technology enables users to perform tasks, access information, and control various devices using natural spoken language, making interactions more seamless and user-friendly. From virtual assistants like Amazon Alexa and Apple Siri to voice-enabled applications and smart home devices, voice recognition has found widespread adoption across different industries, empowering users with hands-free and intuitive communication with technology.

1. Google Assistant (Google, n.d.) is a virtual voice-enabled assistant developed by

Google. It serves as your own personal digital companion, designed to help with various tasks and answer questions using natural language understanding.

2. Amazon Alexa (Amazon, n.d.) is a versatile virtual voice assistant developed by Amazon. Serving as the core of Amazon's smart speakers and smart displays, Alexa is designed to make life easier by responding to voice commands and carrying out a wide array of tasks.

3. Apple Siri (Apple, n.d.) is an intelligent virtual assistant developed by Apple Inc. and is an integral part of their iOS devices, macOS, Apple Watch, and HomePod. Siri is designed to assist users with a wide range of tasks and queries through voice commands and natural language understanding. As a default virtual assistant on Apple devices, Siri offers a seamless hands-free experience, allowing users to perform various tasks with just their voice.

4. Cortana (Microsoft, n.d.) developed by Microsoft, is a powerful and versatile personal productivity assistant designed to assist users across various Microsoft platforms and devices. Named after the AI character from the "Halo" video game series, Cortana provides users with a range of capabilities through natural language interactions and adaptive learning. As a default assistant on Windows 10 devices, Cortana seamlessly integrates with the operating system, enabling users to perform tasks using voice commands, typed text, or even gestures.

5. IBM Watson (IBM Watson, 2023), a pioneering example of popular multimodal interaction solutions, stands at the forefront of cutting-edge artificial intelligence and cognitive computing. As an advanced AI platform developed by IBM, Watson seamlessly integrates multiple modes of interaction, such as natural language processing, machine learning, and deep learning, to facilitate intuitive and comprehensive human-computer communication.

5. Multimodal interaction solutions in education

The use of multimodal interaction solutions in the educational setting has become increasingly popular with the advancement of technology and the rise of ubiquitous computing, multimodal interaction solutions have become a way to bridge the gap between physical and virtual instruction (Li et al. 2019).

The use of multimodal interaction solutions in educational settings is becoming increasingly popular as technology advances. Multimodal interaction solutions (Sugumaran et al., 2022) are defined as the integration of multiple input and output modalities to allow for more natural and intuitive interactions between humans and computing devices. This can include a combination of speech, touch, gestures, facial expressions, and several types of sensors (Chen et al., 2018). Multimodal interaction solutions allow for a more natural and intuitive interaction between the user and the computer, which can make learning more engaging and efficient.

One example of a multimodal interaction solution in education is the use of augmented reality (AR) and virtual reality (VR) technologies. These technologies allow users to interact with digital content in a more natural way, allowing for more immersive learning experiences. For example, AR and VR technologies can be used to create virtual classrooms where students can interact with digital objects and collaborate with each other in a virtual setting (Nesenbergs et al., 2021). AR and VR technologies can also be used to create interactive simulations and game-based learning activities that can help to engage and motivate students (Xi et al., 2022).

Another example of a multimodal interaction solution in education is the use of machine learning (ML) and natural language processing (NLP) (Shaik et al., 2022). ML and NLP technologies can be used to analyse large datasets and find patterns in the data that can be used to create personalized learning experiences. For example, ML and NLP technologies can be used to detect student's emotions and provide learning experiences tailored to their individual needs.

The use of robots and drones (Palaigeorgiou, Malandrakis & Tsolopani, 2017) is another example of a multimodal interaction solution in education. Robots and drones can be used to provide physical feedback to students, allowing them to interact with digital content in a more tangible way. For example, they can provide virtual tours of historical sites, allowing students to explore and learn in a more immersive environment. These are some examples of those interactions:

1. Voice Interaction (Insider Intelligence, n.d.): this multimodal interaction solution allows users to interact with a system using voice commands. Also, for educational purposes allows the creation of interactive learning experiences, such as virtual tutoring or voice-controlled quizzes (Munteanu et al., 2017).

2. Gesture Interaction (Li et al., 2019): Gesture interaction allows users to interact with a system using hand gestures. Gesture interaction can be used in education to create interactive learning experiences, such as virtual labs or interactive 3D models of scientific concepts (Negi, Pawar & Lal, 2020).

3. Touch Interaction: Touch interaction allows users to interact with a system using touch-based input (Ibharim, Borhan & Yatim, 2013). Touch interaction can be used in education to create interactive learning experiences, such as virtual whiteboards or interactive 3D models of scientific concepts.

6. Smart Eye and Affective: Smart Eye is a computer vision technology that can measure and analyse human behaviours in real-time. It uses cameras, sensors, and algorithms to detect facial expressions, body language, and gaze direction (Bishay et al., 2023).

Natural user interfaces are a novel form of human-computer interaction that uses nature as its basis. This technology is made up of speech recognition, gesture recognition, touch screens, touch capability, eye tracking, and brain-computer interfaces. Such interfaces are great for education, as they decrease the cognitive

load, raise the level of engagement, and offer easy access to learning feedback (Yang et al., 2020).

6. Examples of multimodal interaction platforms for education

These solutions can be broadly categorized as "multimodal interaction technologies" or "multimodal interaction platforms." Multimodal interaction refers to the integration of multiple modes of communication, such as voice, text, gestures, touch, and visuals, to facilitate a more natural and effective user experience. There are some examples:

1. BrainWare Safari (BrainWare SAFARI, n.d.): This software utilizes a combination of audio, video, graphics, and text to help students learn and understand complex concepts. It also provides detailed feedback on student progress and performance.

2. SmartClassroom (Adobe, n.d.): This interactive platform combines audio, video, and 3D graphics to create an engaging learning environment. It allows teachers to create interactive lessons, quizzes, and games.

3. Kineo (Kineo, n.d.): This platform uses artificial intelligence, machine learning, and natural language processing to create personalized learning experiences. It also provides teachers with analytics to track student progress and performance.

4. Adobe Spark Video (Adobe, n.d.): it is a free online software that allows users to create engaging video presentations, with the ability to add text, video clips, music, and images. It is a great tool for creating educational videos for teachers and students alike.

5. Pear Deck (Pear Deck, n.d.): it is an online platform that allows teachers to create interactive and engaging lessons. Students can respond to questions and polls in real-time, which helps teachers gauge student understanding and track their progress.

Intelligent tutoring systems (Wu et al., 2022) (ITSs) are another type of multimodal interaction solutions that can be used in educational settings. ITSs use artificial intelligence (AI) to analyse a student's performance and provide personalized instruction and feedback. ITSs can enable teachers to monitor student progress and provide personalized instruction and can also enable students to practice and improve their skills in a safe and controlled environment. There are several commercial ITS solutions available, one of them is IBM's Watson (IBM, n.d.).

7. Advantages and disadvantages of using multimodal interaction solutions

Like any technology, these solutions also come with their own set of disadvantages. It is essential to weigh both the benefits and drawbacks to make informed decisions about their implementation. We select some advantages and

disadvantages of using multimodal interaction solutions, shedding light on their potential to enhance user experiences while acknowledging the challenges they may present.

Table 1. Advantages and disadvantages of using multimodal interaction solutions, particularly in education

Advantages	Disadvantages
Increased Interactivity: Users can use touch, voice, and gesture inputs, allowing them to interact with the system in a more intuitive and natural way.	**Increased Complexity**: The system needs to be able to understand and process various input modalities, and the user needs to be aware of how to use the different modalities of interaction
Improved Usability: Allowing users to interact with the system in multiple ways, they can find the most effective means of input and output, making the system more user-friendly.	**High Cost**: Multimodal interaction solutions can be expensive to develop and maintain, especially when compared to single input/output solutions.
Enhanced Flexibility: Adapted to the user's needs, allowing them to use the system in the most efficient way. This can be especially beneficial when dealing with complex tasks, as the user can use different input and output modes to tackle the task at hand.	**Reduced Accuracy**: Multimodal interaction solutions can be less accurate than traditional single input/output solutions. For example, a system that is designed to understand voice commands may not be able to accurately interpret a gesture input.
Increased Accessibility: A user with limited mobility can use voice or gesture inputs to interact with the system, allowing him/her to access the same services as those without impairments.	**Limited Usability**: Multimodal interaction solutions may not be suitable for all users. For example, a user with a disability may not be able to use the solution due to physical limitations.

8. What challenges can we identify?

The challenges associated with implementing multimodal interaction solutions in education include cost, technical complexity, and privacy concerns. Implementing multimodal interaction solutions can be expensive, as they require specialized hardware and software. Besides, multimodal interaction solutions can be technically complex, as they require the integration of multiple technologies. Furthermore, privacy concerns can arise when using multimodal interaction solutions, as they can collect and store sensitive data about students.

The future of multimodal interaction solutions in education is bright. As technology continues to evolve, they will become more accessible and more affordable. The more research is conducted on the effectiveness of multimodal

interaction solutions, the more educational institutions will begin to adopt them. As more educational institutions adopt multimodal interaction solutions, more research will be conducted on how to best use them to create more engaging and interactive learning experiences.

The conception with which multimodal interaction solutions must be changed, they cannot be "a way", they must be "the way". It must be implemented in all educational software productions, so it would no longer have to talk about resources for blind or deaf people, it would be talk of resources for people in general because multimodal interaction solutions open access to all regardless of their physical and motor qualities. Implementing them in our developments will bring us closer to a more inclusive future and propitiate a better educational exchange.

9. Conclusions

Multimodal interaction solutions offer a range of potential benefits for educational settings, including increased engagement, improved understanding, and better learning outcomes. However, there are also some challenges associated with the implementation of these solutions, such as cost, complexity, and the need for specialized training. Nevertheless, the potential of these solutions is significant, and there is a great deal of potential for further research and development in this area. This is just a view about what we already have, we must improve the classroom performance to prepare students to face challenges and change the future.

REFERENCES

Adobe. (n.d.) Online Video Editor: Free Video Maker. *Adobe Express*. https://www.adobe.com/express/create/video [Accessed 14 April 2023].

Affectiva. (n.d.) *Affectiva - Humanizing Technology*. https://www.affectiva.com/ [Accessed 11 April 2023].

Amazon. (n.d.) *Amazon Alexa – Learn What Alexa Can Do*. https://www.amazon.com/b?node=21576558011 [Accessed 14th April 2023].

Apple. (n.d.) Siri. https://www.apple.com/siri/ [Accessed 14 April 2023].

Belda-Medina, J. (2021) Enhancing Multimodal Interaction and Communicative Competence through Task-Based Language Teaching (TBLT) in Synchronous Computer-Mediated Communication (SCMC). *Education Sciences*. 11(11), 723. doi: 10.3390/educsci11110723.

Bishay, M., Kenneth, P., Matthew, S., Graham, P., Jay, T. & Mohammad, M. (2023) AFFDEX 2.0: A Real-Time Facial Expression Analysis Toolkit. In: *IEEE*

17th International Conference on Automatic Face and Gesture Recognition (FG). pp. 1-8.

Bitbrain. (n.d.) *The Different Kinds of Eye Tracking Devices*. https://www.bitbrain.com/blog/eye-tracking-devices [Accessed 30 July 2023].

BrainWare SAFARI. (n.d.) *Unity WebGL Player*. https://www.brainwaresafari.com/ [Accessed 14 April 2023].

Chen, F., Ji, R., Su, J., Cao, D. & Gao, Y. (2018) Predicting Microblog Sentiments via Weakly Supervised Multimodal Deep Learning. *IEEE Transactions on Multimedia*. 20(4), 997-1007. doi: 10.1109/TMM.2017.2757769.

eGain. (n.d.) *Home*. https://www.egain.com/ [Accessed 11 April 2023].

Encyclopedia. (n.d.) *Recognition Technologies of Hand Gesture*. https://encyclopedia.pub/entry/24100 [Accessed 30th July 2023].

Gonfalonieri, A. (2020) What Brain-Computer Interfaces Could Mean for the Future of Work. *Harvard Business Review*. https://hbr.org/2020/10/what-brain-computer-interfaces-could-mean-for-the-future-of-work.

Google. (n.d.) *Google Assistant, Your Own Personal Google Default*. https://assistant.google.com/ [Accessed 14 April 2023].

Ibharim, L. F. M., Borhan, N. &. Yatim, M. H. M. (2013) A field study of understanding child's knowledge, skills and interaction towards capacitive touch technology (iPad). In*: 2013 8th International Conference on Information Technology in Asia (CITA), Kota Samarahan, Malaysia*. pp. 1-5. doi: 10.1109/CITA.2013.6637571.

IBM (n.d.) *IBM Watson*. https://www.ibm.com/watson [Accessed 14April 2023]

ImmersiveMinds. (n.d.) *Innovate Create Learn*. https://www.immersiveminds.com [Accessed 14 April 2023].

Inbenta. (n.d.). *AI, Enterprise Search, Chatbots, Ticketing*. https://www.inbenta.com/en/ [Accessed 11 April 2023].

Insider Intelligence. (n.d.) *Voice Assistants in 2023: Usage, Growth, and Future of the AI Voice Assistant Market*. https://www.insiderintelligence.com/insights/voice-assistants/ [Accessed 30 July 2023].

Interactions. (n.d.) *Interactions - Intelligent Virtual Assistant*. https://www.interactions.com/ [Accessed 11 April 2023]

Kineo. (n.d.) *Great Digital Learning That Delivers Results*. https://kineo.com/ [Accessed 14 April 2023]

Li, L.-K., Li, Y.-X. & Hung, C.-Y. (2019) Multi-Touch Technology in Education: Digital Simulation Improves Student Comprehension, Cognitive Concept, and

Anxiety. In: *International Symposium on Educational Technology (ISET)*. pp. 241-44.

Li, Y., Huang, J., Tian, F., Wang, H.-A. & Dai, G.-Z. (2019) Gesture Interaction in Virtual Reality. *Virtual Reality & Intelligent Hardware*. 1(1), 84–112. doi: 10.3724/SP.J.2096-5796.2018.0006.

Microsoft. (n.d.) *Cortana - Your Personal Productivity Assistant*. https://www.microsoft.com/en-us/cortana/ [Accessed 14 April 2023].

Munteanu, C., Irani, P., Oviatt, S., Aylett, M., Penn, G., Pan, S., Sharma, N., Rudzicz, F., Gomez, R., Cowan, B. & Nakamura. K. (2017) Designing Speech, Acoustic and Multimodal Interactions. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17).* New York, NY, USA. Association for Computing Machinery, pp. 601-608.

Negi, P. S., Pawar, R. & Lal, R. (2020) Vision-Based Real-Time Human–Computer Interaction on Hand Gesture Recognition. In: Sharma, K., Balas, V. E., Son, L. H., Sharma, R. & Cengiz. K. (eds.) *Singapore Micro-Electronics and Telecommunication Engineering*. *Lecture Notes in Networks and Systems*, pp. 499-507. Springer.

Nesenbergs, K., Abolins, V., Ormanis, J. & Mednis, A. (2021) Use of Augmented and Virtual Reality in Remote Higher Education: A Systematic Umbrella Review. *Education Sciences.* 11(1), 8. doi: 10.3390/educsci11010008.

Niu, S., McCrickard, D. S. & Nguyen. S. M. (2016) Learning with Interactive Tabletop Displays. In: *IEEE Frontiers in Education Conference (FIE).* pp. 1-9.

Palaigeorgiou, G., Malandrakis, G. & Tsolopani, C. (2017) Learning with Drones: Flying Windows for Classroom Virtual Field Trips. In: *IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. pp. 338-42.

Pear deck. (n.d.) *Homepage*. https://www.peardeck.com/ [Accessed 14 April 2023].

See, A. R., Choco, J. A. G. & Chandramohan. K. (2022) Touch, Texture and Haptic Feedback: A Review on How We Feel the World around Us. *Applied Sciences*. 12(9), 4686. doi: 10.3390/app12094686.

Sensory. (n.d.) *Artificial Intelligence on the Edge*. https://www.sensory.com/ [Accessed 11 April 2023].

Shaik, T., Tao, X. Li, Y., McDonald, D. J., Redmond, P. & Galligan, L. (2022) A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis. *IEEE Access*. 10, 56720–56739. doi: 10.1109/ACCESS.2022.3177752.

Sugumaran, V., Xu, Z, Shankar P. & Zhou, H. (2022) Application of Intelligent Systems in Multi-Modal Information Analytics: In: The *4th International*

Conference on Multi-Modal Information Analytics (ICMMIA 2022), Vol. 138(2). Cham, Springer International Publishing.

World Wide Web Consortium (W3C). (n.d.) Natural Language Interface Accessibility User Requirements. https://www.w3.org/TR/naur/ [Accessed 30 July 2023].

Wu, S., Xu, X., Liu, R., Liang, G, Meng, H. & He, B. (2022) An Intelligent Tutoring System for Math Word Problem Solving with Tutorial Solution Generation. In: *International Conference on Intelligent Education and Intelligent Research (IEIR),* pp. 183-88.

Xi, N., Chen, J., Gama, F., Riar, M. & Hamari, J. (2023) The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. *Information Systems Frontiers*. 25, 659–680. doi: 10.1007/s10796-022-10244-x.

Xiong, J., Hsiang, E.-L., He, Z., Zhan, T. & Wu, S.-T. (2021) Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light: Science & Applications*. 10(1), 216. doi: 10.1038/s41377-021-00658-8.

Yang, H., Li, J., Han, Y. & Hu, M. (2020) Research on the Application of New Generation of Human-Computer Interaction in Education. In: *2020 International Conference on Information Science and Education (ICISE-IE), Sanya, China.* pp. 382-385. doi: 10.1109/ICISE51755.2020.00089.