

Chapter 4

Justin Beaver Stories: A conversational and empathic virtual animal in mixed reality technology

Historias del Castor Justin. Un animal virtual conversacional y empático en tecnología de realidad mixta

Alexandra Sierra Rativa, Aura Forero, Arbey Aragon, Andres Burbano, Nelson Arias, Marie Postma, and Menno van Zannen

SCAN THIS QR CODE TO WATCH A VIDEO ABOUT THIS CHAPTER
ESCANEE ESTE CODIGO QR PARA VER UN VÍDEO SOBRE ESTE CAPÍTULO

<https://www.youtube.com/watch?v=a2-F2o5lfio>



Justin Beaver Stories: A conversational and empathic virtual animal in mixed reality technology

Alexandra Sierra Rativa¹, Aura Forero², Arbey Aragon², Andres Burbano³,
Nelson Arias², Marie Postma⁴, and Menno van Zannen⁵

¹Tilburg University

profealexandراسierra@gmail.com

²Cetreal

arbey.aragon@gmail.com

aura.forero@cetreal.com

nelcetreal@gmail.com

³Canvar

andres.a.burbano@gmail.com

⁴Tilburg University

Marie.Postma@tilburguniversity.edu

⁵Sadilar

Menno.VanZaanen@nwu.ac.za

Abstract

In this appendix, this work describes a framework for the creation of a conversational character in a mixed reality empathetic experience. The framework allows for the synchronization of emotional animations of the virtual character in line with the character's dialogue text, with the aim to improve the users' empathetic experience. The dialogue is driven by a Natural Language Processing (NLP) pipeline, including automatic speech recognition, chat-bot, and text to speech generation micro-services. Within this framework, we present a holographic experience called "Justin Beaver Stories" using the Magic Leap one, HoloLens or Nreal mixed reality goggles to project the virtual character into the user's field of vision. This can be used to evaluate the impact of bringing a beaver to the user's environment instead of bringing the user to the beaver's natural environment. Interaction occurs by humanizing the beaver through human communications abilities, resulting in a conversational virtual beaver. The storyline describes the

beaver's lifestyle and problems, represented in a situation of distress. Positive experiences show the practical usability of the framework in the area of HCI.

Keywords: Conversational virtual character, mixed reality, emotional expressions, empathy, natural language processing, chat-bot, animal appearance.

INTRODUCTION

As animals do not communicate as humans do, people are incapable of speaking with animals, creating a language barrier between species. However, here we propose a framework that allows for the development of *conversational virtual characters* in a Mixed Reality (MR) environment, also allowing for animal shaped characters. The framework allows the generation and visualization of emotional responses in the virtual character to elicit empathetic reactions in humans towards the animal in an affective way (Zhao et al., 2019) using mixed reality devices. We describe the construction of a communication system for a virtual agent, which is capable of communicating, adapting, customizing, and offering a new “reality” to the users in their own environment.

The effectiveness of the framework relies heavily on whether empathetic behavior can be realized or not. Empathy is defined here as the ability to share the feelings and thoughts of other people, including their psychological states of pain or distress (Batson et al., 1987b; Clark et al., 2019b). It facilitates the process of social interactions and can repress antisocial behavior and aggression towards others (Stanger et al., 2012b). In recent years, a new construct, “dispositional empathy with nature” (Tam, 2013), has been defined in terms of the dispositional tendency to understand and share the emotional experience of the natural world. The development of this tendency plays a crucial role in environmentalism and in the assessment of environmental educational programs.

PREVIOUS WORK

Recent evidence suggests that one may experience empathy not just towards humans or animals in real life, but also towards virtual agents or virtual characters (Paiva, 2011; Kano and Morita, 2019). On the one hand, studies investigate the use of virtual agents for emotional and social support for mental health of users. For instance, De Gennaro, Krumhuber, and Lucas de Gennaro (2020) investigated effects of an empathetic chat-bot, which had more sensitive responses, such as “I’m sorry that this happened to you” compared to a control condition, and using more neutral responses as “Thank you for your feedback”. The results show that the empathetic chat-bot, compared to the control condition, elicited a more positive mood in the users. On the other hand, studies investigated the use of virtual agents in education. In a study with a virtual tutor called “Alice” (Oker et al., 2020), which uses empathetic feedback to stimulate motivation to learn, the conversational agent was manipulated on facial expressions and the textual feedback to the user on their learning performance. It showed that verbal feedback that was coherent with the agent’s face resulted in higher levels of empathy towards the agent.

The effect of expression of emotions in conversational character behavior with human forms has received much interest (Egges et al., 2004). It has been demonstrated that people respond more positively to agents that express emotions compared to those that do not (de Melo et al., 2015; de Gennaro et al., 2020). To improve the emotional appearance of the conversational agents in mixed and virtual reality setups, one should consider three key aspects (Mensio et al., 2018): *textual interaction*, *vocal interaction*, and *embodied agents*. For textual interaction, the agent should simulate emotions related to the text in a coordinated style, allowing the interlocutor to generate meaningful responses, for instance, by extracting emotions from text. Regarding vocal interaction, the emotion should be recognized (speech-to-text) and manipulated (text-to-speech) while emphasizing human imitation of the tone modulation, and aiming to recreate human voice expressions. Embodied agents control the movements of the body to mimic emotions that are expressed in the agent’s conversation. It is relevant to consider that, although these aspects have shown effective empathetic interactions in virtual human agents, it is unclear whether this has the same effect on virtual animal agents.

GENERAL FRAMEWORK

We propose a framework to enable investigation in the effectiveness of emotional appearance in interactive narratives with empathetic virtual conversational characters. This allows for a wide range of interaction channels, such as voice or text to interact with the conversational characters, but also other (sensory) inputs such as video, temperature measurements, and eye tracking (Haag et al., 2004; Mirsamadi et al., 2017; Marinoiu et al., 2018). Using multiple input sources might help detect user's emotions during the experience, which in turn may improve the character's emotional response and delivery of the narrative.

The framework is designed to support interactive narratives. Interactive narratives should provide textual interaction, expressive speech synthesis, and animated embodied agents. With respect to the textual interaction, emotions can be delivered together with the text of the narrative (as illustrated in Table 1). Furthermore, vocal interaction is ensured through expressive speech synthesis, which modulates the voice of the character by adjusting pitch, tone, cadence, or accent (Valle et al., 2020), adding a more natural expression. The virtual embodied character is also capable of expressing emotions using animations of the body synchronized with the narrative text.

Emotion	Sentence
Surprise	Ohhhh a beginner in this a job
Angry	The river is drying because of the humans
Neutral	Have you noticed the temperature change?
Fear	It is impossible! The forest is burning!
Sad	I want to cry! The trees are burning!
Happy	I can eat a willow tree and you?
Disgust	You are a very strange beaver

Table 1. Relation of emotions and text in the narrative.

The framework consists of a pipeline of interchangeable conceptual modules that detect and deliver emotions as represented in Figure 1. The pipeline starts with the user input acquisition, followed by the parallel execution of emotion extraction and Automatic Speech Recognition (ASR). At this point, the pipeline obtains a text tagged with emotions that serves as input to the Natural Language Processing (NLP) module. (Note that the framework supports multiple emotions per text fragment as illustrated in Figure 2.) The NLP module produces a synchronized sequence of texts and corresponding emotions to send to the user. This is converted using a Text-to-Speech (TTS) module maintaining the synchronization of emotions. Finally, it delivers the voice to the user, synchronized with emotion animations with the aim to reinforce the users' experience and provide natural human-to-machine interaction. This pipeline can be understood as an extension to conversational AI pipelines solutions like NeMo (Kuchaiev et al., 2019) from NVidia or Par-LAI (Miller et al., 2017). The modular setup of the framework allows for the continuous improvement by incorporating new software pieces once they are available.

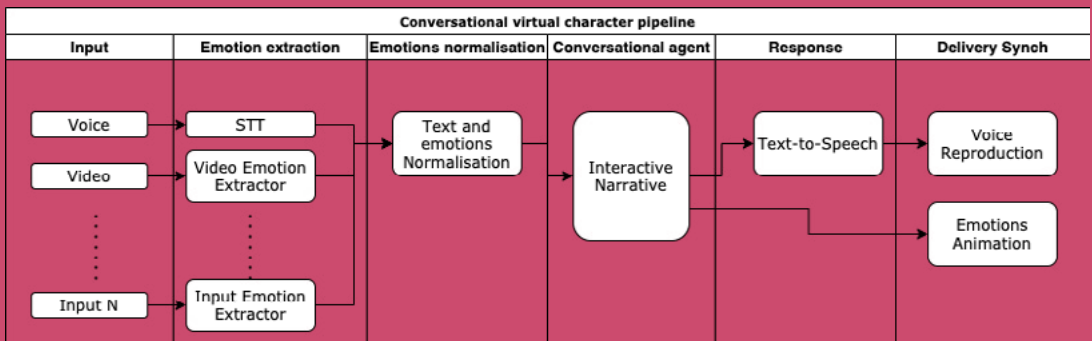


Figure 1. Conversational virtual character pipeline, which is divided in five stages: source inputs, emotions extraction, emotions normalization, chat-bot, response, and synchronized output. The first three handle input emotions and the last two handle output emotions.

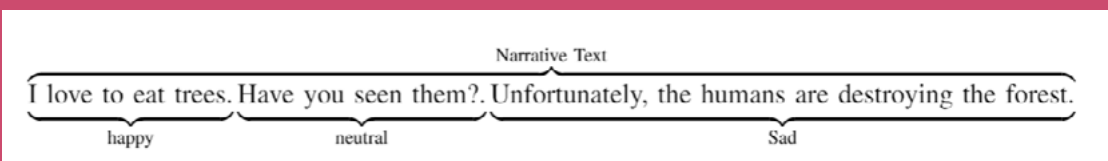


Figure 2. Tagging of multiple emotions per text fragment.

CASE STUDY

While the overall framework is important for the overall emotional experience, the delivery system plays a crucial role in its effectiveness. Here, we present a case study that implements the framework which includes the delivery of emotions using an augmented reality device.

The case study fits in the larger Justin Beaver project. The main objective of this project is to educate people in the beaver's environmental problems by developing an empathetic relationship with the users. Justin is a virtual beaver that inhabits several virtual universes with different capabilities. Two versions already existed: a "regular" computer version gamifying the experiences and a virtual reality version.

In the "regular" computer version, the appearance of the beaver is manipulated in terms of emotional facial expressions and body appearance. The results show that people reacted empathetically when the animal had a biological body and displayed emotional facial expressions, and with a robot representation without facial expressions. This indicates that body shape, and facial expressions have an effect on empathy towards virtual animal characters.

Experiments in the "Justin Beaver VR" virtual reality environment bring users to the beaver's natural environment. In first person view, users perform beavers' daily activities embodied as Justin. The experience then highlights two problems: deforestation and hunting (in the form of a virtual shooting, emphasized through haptic feedback). This study shows that the appearance of the animal is important when it comes to immersion and the perception of pain (during the virtual shooting). This effect was stronger when users embodied a beaver with natural appearance, compared to conditions of more artificial appearances (e.g., robot beavers, or amorphous figures like a marshmallow). The results, however, show that in this version no empathetic response is elicited compared to the "regular" computer version.

Finally, as is illustrated in the Figure 3, "Justin Beaver Stories" mixed reality environment, proposed here, we allow users to interact with Justin Beaver in its natural environment as a conversational virtual character

(Smid and Pandzic, 2002). Through interactive storytelling, Justin provides information on the nutritional behavior of beavers and also talks about wildfires due to global warming. This allows, for instance, experimentation on the effects of two aspects: emotional expression (of the character), and language interaction. Emotional expressions are implemented as animations with both facial and bodily movements. These animations are synchronized with the narrative of Justin. The language interaction relates to the communication pipeline that allows users to communicate with Justin via speech, using Natural Language Processing (NLP) micro-services. As a whole, these capabilities are used to induce empathetic reactions in the users, with the aim to expand the results of the Justin “regular” computer and VR experiences.

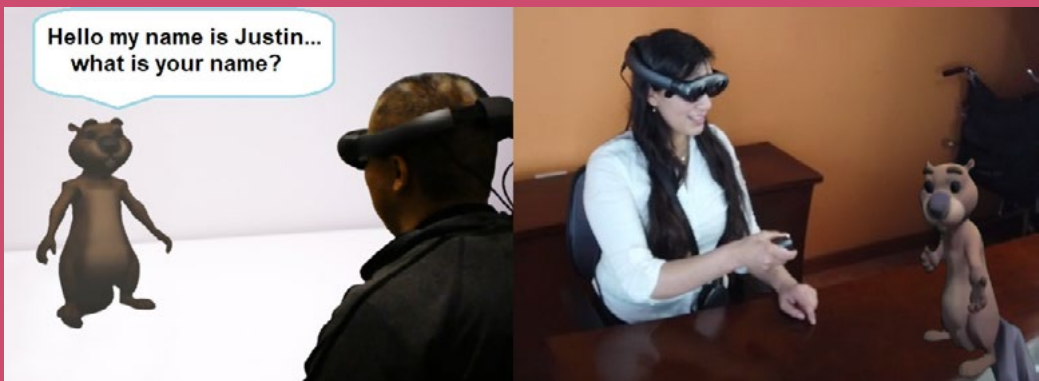


Figure 3. Virtual animal character simulation with the Magic Leap glasses.

Implementation

“Justin Beaver Stories” follows the pipeline using only voice as conversational input. Functionality is implemented using IBM Watson’s micro-services, while including automatic speech recognition (Anusuya and Katti, 2010), and generating input to its chat-bot (Setiaji and Wibowo, 2016), which simultaneously generates text for its text-to-speech system (Allen, 1976). The output is represented as a three-dimensional hologram, placing the character in the user’s real world. The hologram is projected using the Magic Leap mixed reality device (Bradski et al., 2019) utilizing its development framework. This device is capable of combining a virtual layer over the user’s reality and providing real-time interaction, taking into account the environment (registered in three dimensions through motion detection and 3D reconstruction) (Monyneaux et al., 2019). The character’s dialog delivery is enriched with the synchronization between emotional expressions and the generated speech (see Figure 4).



Figure 4. Interaction handling process.

Practically, the experience is implemented using three tiers (see Figure 4). This first tier is implemented using a unity application that runs on the Magic Leap one device, which presents a virtual layer and captures user's voice interactions. The second tier is an Electron web server, which receives the interactions and passes them through the third tier that relies on IBM Watson's natural language micro-services: Natural Language Understanding, Dialog Management, and Natural Language Generation modules (Park and Jeong, 2019). The results are sent back to the virtual agent. Finally, the agent outputs the audio combined with emotion animations.

Note that the animations of the face and the body of the animal (projecting the emotional content of the message) are synchronized in time with the dialogue. Table 1 illustrates how the emotions are annotated in the dialog text. Animations corresponding to the emotions synchronize the movement of the virtual character's eyebrows and different body parts. The transition between emotions is performed using an automatically interpolated transition between the different animations. A natural transition between emotions is pivotal for a more empathetic experience.

Visualization

In order to place the character in a virtual scene within the user's environment, we use the native 3D data framework from Magic Leap. The device recognizes the environment and builds a forest taking the physical constraints into account. The virtual world scales with respect to two constraints: the minimum size possible to facilitate control during the experience, and the maximum size of the hologram to fairly visualize the emotional body and facial expression animations.

The experience has been tested on five different scales (where scale indicates height, width, and depth in cm of the virtual world): 30, 50, 70, 100, and 150. After informal testing using laboratory members, an arbitrary selection of 70cm for the scale was made (as illustrated in Figure 3).

Communication capabilities

The virtual character is capable of communicating, by using the multitier architecture described earlier, and relying on several NLP microservices combined with an experience narrative that guides the user through an educational story on the beaver's life. Due to the interchangeability of the NLP modules, Justin can speak English and Spanish and can be easily ported to other languages by adding the relevant narrative.

As the aim of the narrative is to try to build a relationship between the beaver and the user to enable sharing information on the beaver's environment and dietary description. Each step in the narrative is composed of text (in the form of voice) and corresponding emotion-driven animations. Internally, the text is tagged with emotions (happy, sad, angry, disgusted, fear, surprised, and neutral) which are converted into corresponding animations.

Award recognition as the best research in VR/MR

As is illustrated in Figure 5, the demonstration of this project had the award recognition of the best revolution research in immersive technology (2021) in Laval Virtual, France.

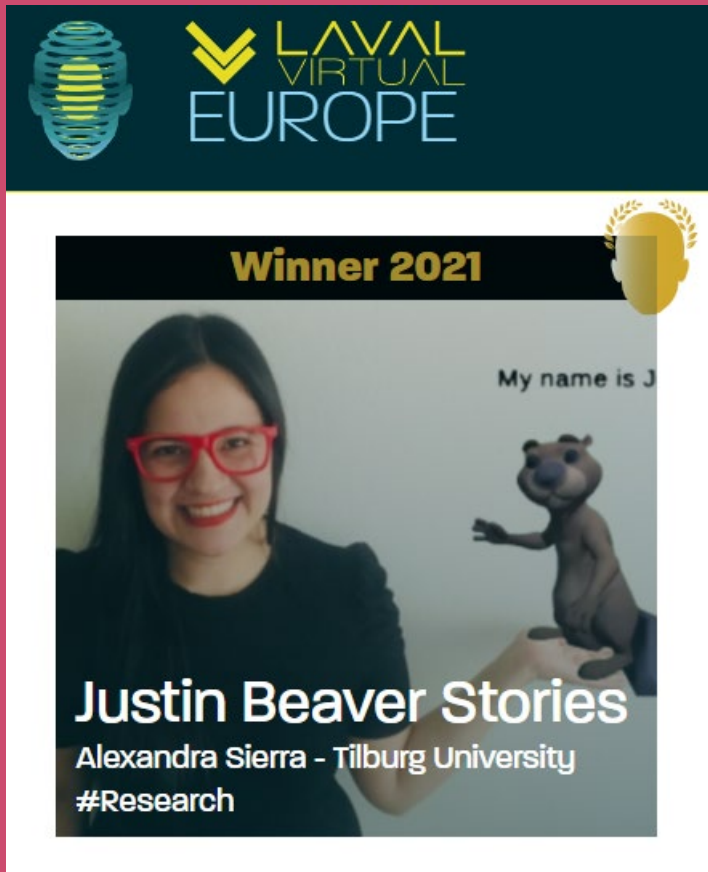


Figure 5. Award recognition at Laval Virtual.

(See more information:

<https://blog.laval-virtual.com/en/all-the-winners-of-the-laval-virtual-awards-2021/>)

CONCLUSIONS AND FUTURE WORK

We introduced a flexible framework that allows the development of interactive storytelling applications with a conversational character (in animal form) which incorporates emotional synchronization of visual and language channels. We have shown the effectiveness of the framework through a case study, “Justin Beaver Stories”, which extends earlier computer and virtual reality versions. This allows for research on the impact of both visual and emotional choices on the immersion and emotional impact. For future work, we will focus on the effects of manipulations of different communication channels, (e.g., to investigate the effect of aligned versus misaligned, or presence versus absence of emotional animation related to the text). As the framework allows experiments in a mixed reality, we can also compare these results to the computer and virtual reality experiments.

We can improve virtual animals’ performance using artificial intelligence. Artificial Intelligence could help this virtual animal be seen as a smart virtual agent capable of vocal interaction and creating interactive stories. Further research could be beneficial in exploring how such a virtual animal might help foster 21st-century learning skills in education through its active interactions with learners and empathic reactions toward nature.

REFERENCES

- Anusuya, M. A., & Katti, S. K. (2010). Speech recognition by machine, a review. *arXiv preprint arXiv:1001.2267*.
- Allen, J. (1976). Synthesis of speech from unrestricted text. *Proceedings of the IEEE*, 64(4), 433-442.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of personality*, 55(1), 19-39.
- Bradski, G. R., Miller, S. A., & Abovitz, R. (2019). *U.S. Patent No. 10,203,762*. Washington, DC: U.S. Patent and Trademark Office.
- Clark, M. A., Robertson, M. M., & Young, S. (2019). "I feel your pain": A critical review of organizational research on empathy. *Journal of Organizational Behavior*, 40(2), 166-192.
- De Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Frontiers in psychology*, 10, 3061.
- De Melo, C. M., Gratch, J., & Carnevale, P. J. (2014). Humans versus computers: Impact of emotion expressions on people's decision making. *IEEE Transactions on Affective Computing*, 6(2), 127-136.
- Kano, Y., & Morita, J. (2019, September). Factors Influencing Empathic Behaviors for Virtual Agents: Examining about the Effect of Embodiment. In *Proceedings of the 7th International Conference on Human-Agent Interaction* (pp. 236-238).
- Eqges, A., Kshirsagar, S., & Magnenat-Thalmann, N. (2004). Generic personality and emotion simulation for conversational agents. *Computer animation and virtual worlds*, 15(1), 1-13.
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004, June). Emotion recognition using biosensors: First steps towards an automatic system. In *Tutorial and research workshop on affective dialogue systems* (pp. 36-48). Springer, Berlin, Heidelberg.
- Kuchaiev, O., Li, J., Nguyen, H., Hrinchuk, O., Leary, R., Ginsburg, B., Krizan, S., Beliaev, S., Lavrukhin, V., Cook, J. & Castonguay, P., (2019). Nemo: a toolkit for building ai applications using neural modules. *arXiv preprint arXiv:1909.09577*.
- Marinoiu, E., Zanfir, M., Olaru, V., & Sminchisescu, C. (2018). 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2158-2167).

- Mensio, M., Rizzo, G., & Morisio, M. (2018, April). The rise of emotion-aware conversational agents: threats in digital emotions. In *Companion Proceedings of the The Web Conference 2018* (pp. 1541-1544).
- Mirsamadi, S., Barsoum, E., & Zhang, C. (2017, March). Automatic speech emotion recognition using recurrent neural networks with local attention. In *2017 IEEE International conference on acoustics, speech and signal processing (ICASSP)* (pp. 2227-2231). IEEE.
- Miller, A. H., Feng, W., Fisch, A., Lu, J., Batra, D., Bordes, A., Devi Parikh, & Jason Weston J. (2017). Parlai: A dialog research software platform. *arXiv preprint arXiv:1705.06476*.
- Monyneaux D. G., Steinbrucker F. T., Wu Z., Wei W., Min J., & Yifu Zhang (2019, June). Caching and updating of dense 3d reconstruction data. Library Catalog: Google Patents.
- Oker, A., Pecune, F., & Declercq, C. (2020). Virtual tutor and pupil interaction: A study of empathic feedback as extrinsic motivation for learning. *Education and Information Technologies*, 25(5), 3643-3658.
- Paiva, A. (2011). Empathy in social agents. *International Journal of Virtual Reality*, 10(1), 1-4.
- Park, K., & Jeong, Y. S. (2019, June). Indoor Dialog Agent in Mixed Reality (video). In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services* (pp. 708-709).
- Smid, K., & Pandzic, I. G. (2002, June). Conversational virtual character for the web. In *Proceedings of Computer Animation 2002 (CA 2002)* (pp. 240-247). IEEE.
- Setiaji, B., & Wibowo, F. W. (2016, January). Chatbot using a knowledge in database: human-to-machine conversation modeling. In *2016 7th international conference on intelligent systems, modelling and simulation (ISMS)* (pp. 72-77). IEEE.
- Stanger, N., Kavussanu, M., & Ring, C. (2012). Put yourself in their boots: Effects of empathy on emotion and aggression. *Journal of sport and exercise psychology*, 34(2), 208-222.
- Tam, K. P. (2013). Dispositional empathy with nature. *Journal of environmental psychology*, 35, 92-104.
- Valle, R., Shih, K., Prenger, R., & Catanzaro, B. (2020). Flowtron: an autoregressive flow-based generative network for text-to-speech synthesis. *arXiv preprint arXiv:2005.05957*.
- Zhao, Z., Han, F., & Ma, X. (2019, December). Live emoji: A live storytelling vr system with programmable cartoon-style emotion embodiment. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)* (pp. 251-2511). IEEE.

Historias del Castor Justin Beaver: Un animal virtual conversacional y empático en tecnología de realidad mixta

Alexandra Sierra Rativa¹, Aura Forero², Arbey Aragon², Andres Burbano³,
Nelson Arias², Marie Postma⁴, and Menno van Zannen⁵

¹Tilburg University
profealexandrasierra@gmail.com

²Cetreal
arbey.aragon@gmail.com
aura.forero@cetreal.com
nelcetreal@gmail.com

³Canvar
andres.a.burbano@gmail.com

⁴Tilburg University
Marie.Postma@tilburguniversity.edu

⁵Sadilar
Menno.VanZaanen@nwu.ac.za

Resumen

Este trabajo describe una estructura informática para la creación de un personaje en una experiencia empática de realidad mixta. La estructura permite la sincronización en línea de una animación emocional del personaje virtual con el texto de diálogo que éste expresa, con el objetivo de mejorar la empatía y la experiencia del usuario. El diálogo es controlado por un flujo de procesamiento de lenguaje natural que incluye reconocimiento automático de habla, chatbot, y microservicios de generación de diálogo verbal a escrito. Con esta estructura informática, presentamos una experiencia holográfica llamada "Historias del castor Justin Beaver", usando los dispositivos de realidad mixta Magic leap, Hololens o Nreal, y proyectamos un personaje virtual dentro del campo de vista del usuario. Esto se puede utilizar para evaluar el impacto de llevar un castor al entorno del usuario en lugar de llevarlo al entorno natural del castor. La interacción ocurre al humanizar al castor a través de las habilidades de comunicación humana, lo que resulta en un castor virtual conversacional. La historia

describe el estilo de vida y los problemas del castor, representados en una situación de angustia. Las experiencias positivas muestran la usabilidad práctica del marco en el área de HCI.

Palabras clave: Personaje conversacional virtual, realidad mixta, expresiones emocionales, empatía, procesamiento de lenguaje natural, chatbot, apariencia animal.

INTRODUCCIÓN

Como los animales no se comunican de la misma manera que los humanos, las personas son incapaces de hablar con los animales, lo que crea una barrera lingüística entre las especies. Sin embargo, aquí proponemos un marco que permite el desarrollo de personajes virtuales conversacionales en un entorno de Realidad Mixta (MR), permitiendo también personajes con forma de animales. El marco permite la generación y visualización de respuestas emocionales en el personaje virtual para provocar reacciones empáticas en humanos hacia el animal de forma afectiva (Zhao et al., 2019) utilizando dispositivos de realidad mixta. Describimos la construcción de un sistema de comunicación para un agente virtual; que es capaz de comunicarse, adaptarse, personalizar y ofrecer una nueva “realidad” a los usuarios en su propio entorno.

La eficacia de la estructura informática depende en gran medida de si el comportamiento empático se puede realizar o no. La empatía se define aquí como la capacidad de compartir los sentimientos y pensamientos de otras personas, incluidos sus estados psicológicos de dolor o angustia (Batson et al., 1987b; Clark et al., 2019b). Facilita el proceso de interacciones sociales y puede reprimir el comportamiento antisocial y la agresión hacia los demás (Stanger et al., 2012b). En los últimos años, un nuevo constructo, la “empatía disposicional con la naturaleza” (Tam, 2013), se ha definido en términos de la tendencia disposicional a comprender y compartir la experiencia emocional del mundo natural. El desarrollo de esta tendencia juega un papel crucial en el ambientalismo y en la evaluación de los programas de educación ambiental.

TRABAJO PREVIO

La evidencia reciente sugiere que uno puede experimentar empatía no solo hacia humanos o animales en la vida real, sino también hacia agentes virtuales o personajes virtuales (Paiva, 2011; Kano and Morita, 2019). Por un lado, los estudios investigan el uso de agentes virtuales para el apoyo emocional y social para la salud mental de los usuarios. Por ejemplo, De Gennaro, Krumhuber y Lucas (2020) investigaron los efectos de un chatbot empático, que tenía respuestas más sensibles, como “Lamento que te haya pasado esto” en comparación con una condición de control, y utiliza respuestas más neutras, como: “Gracias por sus comentarios”. Los resultados muestran que el chatbot empático, en comparación con la condición de control, provocó un estado de ánimo más positivo en los usuarios. Por otro lado, los estudios investigaron el uso de agentes virtuales en la educación. En un estudio con un tutor virtual llamado “Alice” (Oker et al., 2020), que utiliza retroalimentación empática para estimular la motivación para aprender, el agente conversacional fue manipulado en las expresiones faciales y la retroalimentación textual al usuario sobre su desempeño en el aprendizaje. Demostró que la retroalimentación verbal que era coherente con el rostro del agente daba como resultado niveles más altos de empatía hacia el agente.

El efecto de la expresión de emociones en el comportamiento del carácter conversacional con formas humanas ha recibido mucho interés (Egges et al., 2004). Se ha demostrado que las personas responden de manera más positiva a los agentes que expresan emociones en comparación con aquellos que no las expresan (de Melo et al., 2015; de Gennaro et al., 2020). Para mejorar la apariencia emocional de los agentes conversacionales en configuraciones de realidad virtual y mixta, se deben considerar tres aspectos clave (Mensio et al., 2018): interacción textual, interacción vocal y agentes personificados. Para la interacción textual, el agente debe simular emociones relacionadas con el texto en un estilo coordinado, permitiendo que el interlocutor genere respuestas significativas, por ejemplo, extrayendo emociones del texto. Con respecto a la interacción vocal, la emoción debe ser reconocida (de voz a texto) y manipulada (de texto a voz) mientras se enfatiza la imitación humana de la modulación del tono y se busca recrear las expresiones de la voz humana. Los agentes personificados controlan

los movimientos del cuerpo para imitar las emociones que se expresan en la conversación del agente. Es relevante considerar que, si bien estos aspectos han mostrado interacciones empáticas efectivas en agentes humanos virtuales, no está claro si esto tiene el mismo efecto en agentes animales virtuales.

MARCO GENERAL

Proponemos una estructura informática que permita investigar la efectividad de la apariencia emocional en narrativas interactivas con personajes conversacionales virtuales empáticos. Esto permite que una amplia gama de canales de interacción, como voz o texto, interactúen con los personajes conversacionales, pero también otras entradas (sensoriales) como video, mediciones de temperatura y seguimiento ocular tracking (Haag et al., 2004; Mirsamadi et al., 2017; Marinoiu et al., 2018). El uso de múltiples fuentes de entrada puede ayudar a detectar las emociones del usuario durante la experiencia, lo que a su vez puede mejorar la respuesta emocional del personaje y la entrega de la narrativa.

Emoción	Frase
Sorpresa	Oh un aprendiz en este trabajo.
Enojo	El río se está secando debido a los humanos
Neutral	¿Has notado el cambio de temperatura?
Miedo	¡Es imposible! El bosque se está quemando!
Tristeza	¡Quiero llorar, los árboles se están quemando!
Felicidad	Puedo comer un árbol completo, y tú?
Disgusto	¡Tú eres un castor extraño!

Tabla 1. Relación de emociones y texto en la narrativa.

La estructura está diseñada para respaldar narrativas interactivas. Las narrativas interactivas deben proporcionar interacción textual, síntesis de voz expresiva y agentes animados personificados. Con respecto a la interacción textual, las emociones se pueden transmitir junto con el texto de la narración (como se presenta en la Tabla 1). Además, la interacción vocal está asegurada a través de la síntesis del habla expresiva, que modula la voz del personaje ajustando el tono, la cadencia o el acento (Valle et al., 2020), agregando una expresión más natural. El personaje virtual personificado también es capaz de expresar emociones utilizando animaciones del cuerpo sincronizadas con el texto narrativo.

La estructura informática propuesta consiste en una distribución de módulos conceptuales intercambiables que detectan y entregan emociones como se representa en la Figura 1. La distribución de la estructura comienza con la adquisición de la entrada del usuario, seguida de la ejecución paralela de la extracción de emociones y el Reconocimiento Automático de Voz (ASR). En este punto de la distribución se obtiene un texto etiquetado con emociones que sirve como entrada para el módulo Procesamiento del lenguaje natural (NLP). (Tenga en cuenta que la estructura propuesta admite múltiples emociones por fragmento de texto, como se ilustra en la Figura 2). El módulo de NLP produce una secuencia sincronizada de textos y emociones correspondientes para enviar al usuario. Esto se convierte mediante un módulo de Texto-a-Voz (TTS) que mantiene la sincronización de las emociones. Finalmente, entrega la voz al usuario, sincronizada con animaciones correspondientes a las emociones con el objetivo de reforzar la experiencia de los usuarios y proporcionar una interacción natural entre humanos y máquinas. Esta distribución puede entenderse como una extensión de las soluciones de distribución de agentes conversacionales con IA como NeMo (Kuchaiev et al., 2019) de NVidia o Par-LAI (Miller et al., 2017). La configuración modular de la estructura propuesta permite la mejora continua mediante la incorporación de nuevas piezas de software una vez que están disponibles.

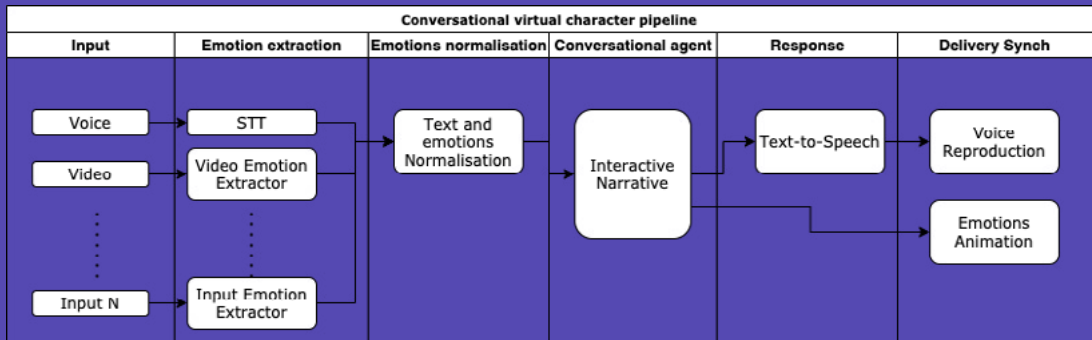


Figura 1. Distribución conversacional de personajes virtuales, se divide en cinco etapas: entradas de origen, extracción de emociones, normalización de emociones, chatbot, respuesta y salida sincronizada. Los primeros tres manejan las emociones de entrada y los dos últimos manejan las emociones de salida.

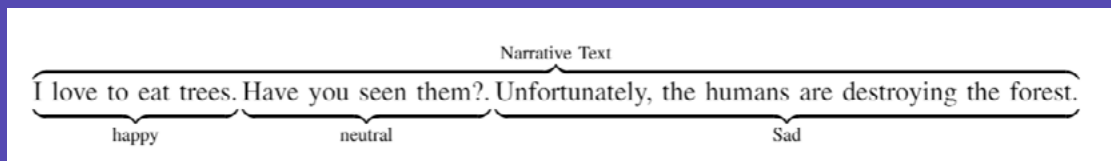


Figura 2. Etiquetado de múltiples emociones por fragmento de texto.

CASO DE ESTUDIO

Si bien el marco general es importante para la experiencia emocional general, el sistema de entrega juega un papel crucial en su efectividad. A continuación, presentamos un estudio de caso que implementa la estructura propuesta que incluye la entrega de emociones utilizando un dispositivo de realidad aumentada. El caso de estudio encaja en el proyecto más amplio del castor Justin Beaver. El principal objetivo de este proyecto es educar a las personas sobre los problemas ambientales del castor desarrollando una relación empática con los usuarios. Justin Beaver es un castor virtual que habita varios universos virtuales con diferentes capacidades.

Ya existían dos versiones: una versión informática “habitual ” la cual gamifica las experiencias y una versión de realidad virtual. En la versión de computadora “habitual”, la apariencia del castor se manipula en términos de expresiones faciales emocionales y apariencia corporal. Los resultados muestran que las personas reaccionaron con empatía cuando el animal tenía un cuerpo biológico y mostraba expresiones faciales emocionales, y con una representación de robot sin expresiones faciales. Esto indica que la forma del cuerpo y las expresiones faciales tienen un efecto en la empatía hacia los personajes animales virtuales.

Los experimentos en el entorno de realidad virtual “El Castor Justin Beaver VR” llevan a los usuarios al entorno natural del castor. En la vista en primera persona, los usuarios realizan las actividades diarias de los castores personificados como Justin Beaver. La experiencia luego resalta dos problemas: la deforestación y la caza (en forma de un disparo virtual, enfatizado a través de retroalimentación háptica). Este estudio muestra que la apariencia del animal es importante en lo que respecta a la inmersión y la percepción del dolor (durante el rodaje virtual). Este efecto fue más fuerte cuando los usuarios personifican un castor con apariencia natural, en comparación a la situación con condiciones de apariencia más artificial (por ejemplo, castores robot o figuras amorfas como un malvavisco). Los resultados, sin embargo, muestran que en esta versión ninguna respuesta empática es ilícita en comparación con la versión informática “habitual”.

Finalmente, en el entorno de realidad mixta “Historias del Castor Justin Beaver” (Justin Beaver Stories) como se ilustra en la Figura 3, propuesto aquí, permitimos a los usuarios interactuar con el castor Justin en su entorno natural como un personaje virtual conversacional (Smid and Pandzic, 2002). A través de la narración interactiva, Justin Beaver brinda información sobre el comportamiento nutricional de los castores y también habla sobre los incendios forestales debido al calentamiento global. Esto permite, por ejemplo, experimentar los efectos de dos aspectos: la expresión emocional (del personaje) y la interacción del lenguaje. Las expresiones emocionales se implementan como animaciones con movimientos faciales y corporales. Estas animaciones están sincronizadas con la narrativa de Justin. La interacción del lenguaje se relaciona con la distribución de comunicación que permite a los usuarios comunicarse con

Justin a través del habla, utilizando microservicios de procesamiento del lenguaje natural (NLP). En su conjunto, estas capacidades se utilizan para inducir reacciones de empatía en los usuarios, con el objetivo de ampliar los resultados de las experiencias de realidad virtual y de ordenador “habitual” de Justin.

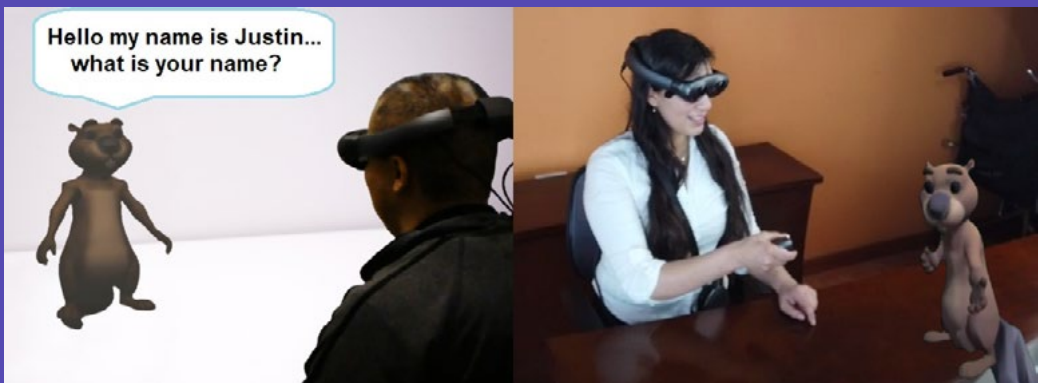


Figura 3. Simulación de personajes de animales virtuales llamada Historias del castor Justin Beaver.

Implementación

“Las historias del castor Justin Beaver” (Justin Beaver Stories) sigue el proceso utilizando solo la voz como entrada de conversación. La funcionalidad se implementa utilizando los microservicios de IBM Watson, al tiempo que incluye el reconocimiento automático de voz (Anusuya and Katti, 2010) y genera información para su chatbot (Setiaji and Wibowo, 2016), que simultáneamente genera texto para su sistema de conversión de texto a voz (Allen, 1976). La salida se representa como un holograma tridimensional, que coloca al personaje en el mundo real del usuario. El holograma se proyecta utilizando gafas de realidad mixta (por ejemplo, el

dispositivo de realidad mixta Magic Leap (Bradski et al., 2019), Hololens, Nreal) utilizando su estructura de desarrollo.

Este dispositivo es capaz de combinar una capa virtual sobre la realidad del usuario y brindar interacción en tiempo real, teniendo en cuenta el entorno (registrado en tres dimensiones mediante detección de movimiento y reconstrucción 3D) Monyneaux et al., 2019). La entrega del diálogo del personaje se enriquece con la sincronización entre las expresiones emocionales y el discurso generado (ver Figura 4).



Figura 4. Proceso de manejo de las interacciones.

Prácticamente, la experiencia se implementa utilizando tres niveles (ver Figura 4). Este primer nivel se implementa mediante una aplicación de unidad que se ejecuta en el dispositivo Magic Leap one, que presenta una capa virtual y captura las interacciones de voz del usuario. El segundo nivel es un servidor web Electron, que recibe las interacciones y las pasa a través del tercer nivel que se basa en los microservicios de lenguaje natural de IBM Watson: módulos de comprensión del lenguaje natural, gestión de diálogos y generación del lenguaje natural (Park and Jeong, 2019). Los resultados se envían de vuelta al agente virtual. Finalmente, el agente genera el audio combinado con animaciones emocionales.

Nótese que las animaciones de la cara y el cuerpo del animal (que proyectan el contenido emocional del mensaje) están sincronizadas en el tiempo con el diálogo. La Tabla 1 ilustra cómo se anotan las emociones en

el texto del diálogo. Las animaciones correspondientes a las emociones sincronizan el movimiento de las cejas del personaje virtual y diferentes partes del cuerpo. La transición entre emociones se realiza mediante una transición interpolada automáticamente entre las diferentes animaciones. Una transición natural entre las emociones es fundamental para una experiencia más empática.

Visualización

Para colocar al personaje en una escena virtual dentro del entorno del usuario, utilizamos el marco de datos 3D nativo de Magic Leap. El dispositivo reconoce el entorno y construye un bosque teniendo en cuenta las limitaciones físicas. El mundo virtual escala con respecto a dos limitaciones: el tamaño mínimo posible para facilitar el control durante la experiencia, y el tamaño máximo del holograma para visualizar de manera justa el cuerpo emocional y las animaciones de la expresión facial. La experiencia ha sido probada en cinco escalas diferentes (donde la escala indica altura, ancho y profundidad en cm del mundo virtual): 30, 50, 70, 100 y 150. Después de pruebas informales con miembros de laboratorio, una selección arbitraria de 70 cm para la escala (como se ilustra en la Figura 3).

Capacidades de comunicación

El personaje virtual es capaz de comunicarse mediante el uso de la arquitectura de varios niveles descrita anteriormente y dependiendo de varios microservicios de NLP combinados con una narrativa de experiencia que guía al usuario a través de una historia educativa sobre la vida del castor. Debido a la intercambiabilidad de los módulos de NLP, Justin Beaver puede hablar inglés y español y puede configurarse fácilmente a otros idiomas agregando la narrativa relevante.

Como el objetivo de la narrativa es tratar de construir una relación entre el castor y el usuario para permitir compartir información sobre el entorno del castor y la descripción de su alimentación. Cada paso de la narrativa se compone de texto (en forma de voz) y las correspondientes animaciones impulsadas por las emociones. Internamente, el texto está etiquetado con emociones (feliz, triste, enojado, disgustado, miedo, sorprendido y neutral) que se convierten en las animaciones correspondientes.

Premio a la mejor investigación in RV/RM

Como se ilustra en la Figura 5, la demostración de este proyecto tuvo el reconocimiento de premio a la mejor investigación revolucionaria en tecnología inmersiva (2021) en Laval Virtual, Francia.

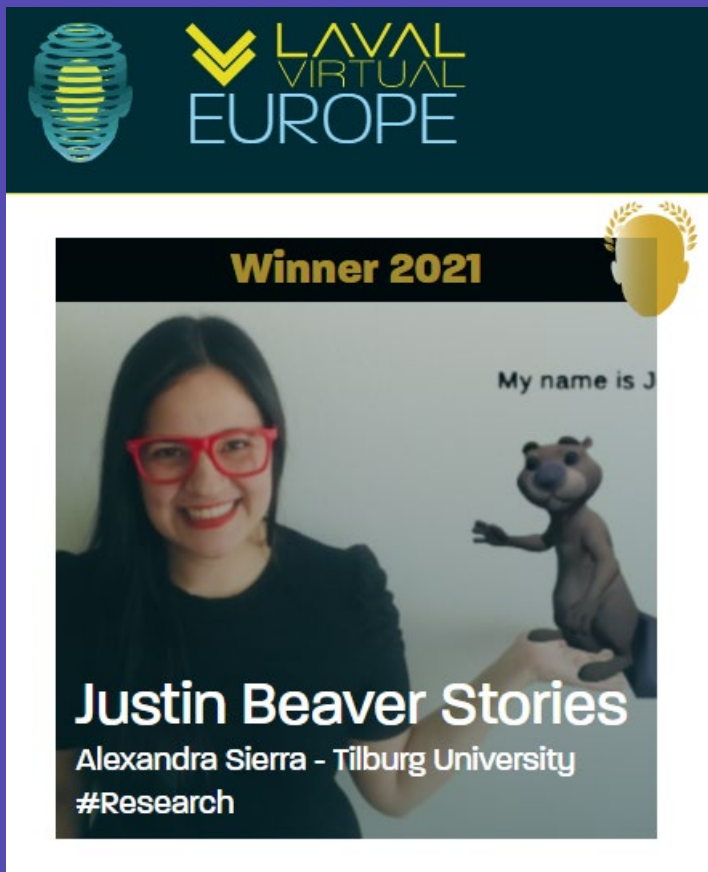


Figura 5. Premio en Laval Virtual.

(Más información:

<https://blog.laval-virtual.com/en/all-the-winners-of-the-laval-virtual-awards-2021/>)

CONCLUSIONES Y TRABAJO FUTURO

Introdujimos un marco flexible que permite el desarrollo de aplicaciones de narración interactiva con un carácter conversacional (en forma de animal) que incorpora sincronización emocional de canales visuales y de lenguaje. Hemos demostrado la efectividad del marco a través de un estudio de caso, “Historias de Justin Beaver” (Justin Beaver Stories), que amplía las versiones anteriores de computadora y realidad virtual. Esto permite investigar el impacto de las opciones tanto visuales como emocionales en la inmersión y el impacto emocional. Para trabajos futuros, nos centraremos en los efectos de la manipulación de diferentes canales de comunicación (p. ej., para investigar el efecto de alineación frente a desalineación, o presencia frente a ausencia de animación emocional relacionada con el texto). Como el marco permite experimentos en una realidad mixta, también podemos comparar estos resultados con los experimentos de computadora y realidad virtual.

Podemos mejorar el rendimiento de los animales virtuales utilizando inteligencia artificial. La inteligencia artificial podría ayudar a que este animal virtual sea visto como un agente virtual inteligente capaz de interactuar vocalmente y crear historias interactivas. La investigación adicional podría ser beneficiosa para explorar cómo un animal virtual de este tipo podría ayudar a fomentar las habilidades de aprendizaje del siglo XXI en la educación a través de sus interacciones activas con los alumnos y reacciones empáticas hacia la naturaleza.

REFERENCIAS

- Anusuya, M. A., & Katti, S. K. (2010). Speech recognition by machine, a review. *arXiv preprint arXiv:1001.2267*.
- Allen, J. (1976). Synthesis of speech from unrestricted text. *Proceedings of the IEEE*, 64(4), 433-442.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of personality*, 55(1), 19-39.
- Bradski, G. R., Miller, S. A., & Abovitz, R. (2019). *U.S. Patent No. 10,203,762*. Washington, DC: U.S. Patent and Trademark Office.
- Clark, M. A., Robertson, M. M., & Young, S. (2019). "I feel your pain": A critical review of organizational research on empathy. *Journal of Organizational Behavior*, 40(2), 166-192.
- De Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Frontiers in psychology*, 10, 3061.
- De Melo, C. M., Gratch, J., & Carnevale, P. J. (2014). Humans versus computers: Impact of emotion expressions on people's decision making. *IEEE Transactions on Affective Computing*, 6(2), 127-136.
- Kano, Y., & Morita, J. (2019, September). Factors Influencing Empathic Behaviors for Virtual Agents: Examining about the Effect of Embodiment. In *Proceedings of the 7th International Conference on Human-Agent Interaction* (pp. 236-238).
- Egges, A., Kshirsagar, S., & Magnenat-Thalmann, N. (2004). Generic personality and emotion simulation for conversational agents. *Computer animation and virtual worlds*, 15(1), 1-13.
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004, June). Emotion recognition using biosensors: First steps towards an automatic system. In *Tutorial and research workshop on affective dialogue systems* (pp. 36-48). Springer, Berlin, Heidelberg.
- Kuchaiev, O., Li, J., Nguyen, H., Hrinchuk, O., Leary, R., Ginsburg, B., Krizan, S., Beliaev, S., Lavrukhin, V., Cook, J. & Castonguay, P., (2019). Nemo: a toolkit for building ai applications using neural modules. *arXiv preprint arXiv:1909.09577*.
- Marinoiu, E., Zanfir, M., Olaru, V., & Sminchisescu, C. (2018). 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2158-2167).

- Mensio, M., Rizzo, G., & Morisio, M. (2018, April). The rise of emotion-aware conversational agents: threats in digital emotions. In *Companion Proceedings of the The Web Conference 2018* (pp. 1541-1544).
- Mirsamadi, S., Barsoum, E., & Zhang, C. (2017, March). Automatic speech emotion recognition using recurrent neural networks with local attention. In *2017 IEEE International conference on acoustics, speech and signal processing (ICASSP)* (pp. 2227-2231). IEEE.
- Miller, A. H., Feng, W., Fisch, A., Lu, J., Batra, D., Bordes, A., Devi Parikh, & Jason Weston J. (2017). Parlai: A dialog research software platform. *arXiv preprint arXiv:1705.06476*.
- Monyneaux D. G., Steinbrucker F. T., Wu Z., Wei W., Min J., & Yifu Zhang (2019, June). Caching and updating of dense 3d reconstruction data. Library Catalog: Google Patents.
- Oker, A., Pecune, F., & Declercq, C. (2020). Virtual tutor and pupil interaction: A study of empathic feedback as extrinsic motivation for learning. *Education and Information Technologies*, 25(5), 3643-3658.
- Paiva, A. (2011). Empathy in social agents. *International Journal of Virtual Reality*, 10(1), 1-4.
- Park, K., & Jeong, Y. S. (2019, June). Indoor Dialog Agent in Mixed Reality (video). In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services* (pp. 708-709).
- Smid, K., & Pandzic, I. G. (2002, June). Conversational virtual character for the web. In *Proceedings of Computer Animation 2002 (CA 2002)* (pp. 240-247). IEEE.
- Setiaji, B., & Wibowo, F. W. (2016, January). Chatbot using a knowledge in database: human-to-machine conversation modeling. In *2016 7th international conference on intelligent systems, modelling and simulation (ISMS)* (pp. 72-77). IEEE.
- Stanger, N., Kavussanu, M., & Ring, C. (2012). Put yourself in their boots: Effects of empathy on emotion and aggression. *Journal of sport and exercise psychology*, 34(2), 208-222.
- Tam, K. P. (2013). Dispositional empathy with nature. *Journal of environmental psychology*, 35, 92-104.
- Valle, R., Shih, K., Prenger, R., & Catanzaro, B. (2020). Flowtron: an autoregressive flow-based generative network for text-to-speech synthesis. *arXiv preprint arXiv:2005.05957*.
- Zhao, Z., Han, F., & Ma, X. (2019, December). Live emoji: A live storytelling vr system with programmable cartoon-style emotion embodiment. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)* (pp. 251-2511). IEEE.

ALEXANDRA SIERRA RATIVA



English

Alexandra was born in Bogota, Colombia. She is a researcher in immersive learning technologies (virtual reality, augmented reality and mixed reality) and conventional technologies (educational platforms, robotics and video games). She worked as a lecturer in Research and Media at Breda University of Applied Sciences and PhD candidate in Cognitive Sciences and Artificial Intelligence at Tilburg University, The Netherlands. She is magister in education at Universidad de los Andes with a Bachelor in electronics at Universidad Pedagógica Nacional. In 2021, she received an award for revolutionary research in immersive technology at Laval Virtual, France. She is a founding member of the Colombian Association of Immersive, Interactive, and Emerging Realities (XRCOL). Currently, she is organizing the international conference for teachers and researchers called XR ACADEMIA, and the book publication XR ACADEMIA.

Español

Alexandra Sierra Rativa nació en Bogotá, Colombia. Es investigadora en tecnologías de aprendizaje inmersivo (realidad virtual, realidad aumentada y realidad mixta) y tecnologías convencionales (plataformas educativas, robótica y videojuegos). Trabajo como profesora de Investigación y Medios en la Universidad de Ciencias Aplicadas de Breda y candidata a doctorado en Ciencias Cognitivas e Inteligencia Artificial en la Universidad de Tilburg, Países Bajos. Ella es Magíster en Educación de la Universidad de los Andes con Licenciatura en Electrónica de la Universidad Pedagógica Nacional. En 2021, recibió un premio por investigación revolucionaria en tecnología inmersiva en Laval Virtual, Francia. Ella es miembro fundador de la Asociación Colombiana de Realidades Inmersivas, Interactivas y Emergentes (XRCOL). Actualmente, coordina la conferencia internacional para profesores e investigadores llamada XR ACADEMIA, y la publicación del libro XR ACADEMIA.



ANDRES BURBANO



English

Andrés Burbano holds a Ph.D. in 3D Image Processing from the Paris-Saclay University (2018) in France, a Master's Degree in Systems and Networks from the Ecole centrale électronique de Paris (2013), and computer science engineer's degree at the Escuela Colombiana de Ingeniería "Julio Garavito" (2008). He divides his time as an activist and XR experience producer, IT expert and researcher. He is the director of CANVAR, co-founder of XRCOL, and director of Artizen's XR LATAM grant.

As an artist, he uses immersive technologies as a tool to produce cultural, social and research projects, in a transdisciplinary and collaborative practice. Since 2018, he has been introducing immersive technologies to artists and minorities through labs where everyone can experience, learn, and use them as a means of creation. He is currently a fellow of the Art of Practice Interdisciplinary Program at Sundance Institute.

As a technology expert, he has experience developing scalable applications that integrate artificial intelligence and image processing services. He has developed his experience in positions as CTO and Research director in French companies and is currently developing a platform to promote the CO2 emissions reduction in the transport of goods under the GreenTech French label.

As a researcher, he studies the impact of immersive technologies used on the creative process of performing artists. He has published scientific articles related to the evaluation of volumetric cameras and their use in the study of human behavior. He participates in Latin American Panels such as Stereopsia Latam (Chile), EXPYLAB (Paraguay), VR FEST MX (Mexico) and XR ACADEMIA (Colombia / The Netherlands).



Español

Andres Burbano tiene un Ph.D. en Procesamiento de Imágenes 3D de la Universidad París-Saclay (2018) en Francia, un Máster en Sistemas y Redes de la Ecole centrale électronique de París (2013), e Ingeniería de sistemas en Escuela de Ingeniería de Colombia “Julio Garavito” (2008). Divide su tiempo como activista y productor de experiencias XR, experto en TI e investigador. Co-fundador de XR COL, director de la beca XR LATAM de Artizen y director de CANVAR.

Como artista, utiliza tecnologías inmersivas como herramienta para producir proyectos culturales, sociales y de investigación, en una práctica transdisciplinar y colaborativa. Desde 2018, ha estado presentando las tecnologías inmersivas a artistas y minorías a través de laboratorios donde todos pueden experimentarlas, aprender y usarlas como medio de creación. Actualmente es miembro del Programa Interdisciplinario el Arte de la Práctica en Sundance Institute 2021-2022.

Como experto en tecnología, tiene experiencia en el desarrollo de aplicaciones escalables que integran inteligencia artificial y servicios de procesamiento de imágenes. Ha desarrollado su experiencia en puestos como CTO y director de Investigación en empresas francesas y actualmente está desarrollando una plataforma para fomentar la reducción de emisiones de CO2 en el transporte de mercancías bajo la etiqueta francesa GreenTech.

Como investigador, estudia el impacto de los usos de tecnologías inmersivas en el proceso creativo de artistas escénicos. Ha publicado artículos científicos relacionados con la evaluación de cámaras volumétricas y su uso en el estudio del comportamiento humano. Es editor de la conferencia internacional XR Academia. Dicta conferencias y participa en Paneles latinoamericanos tales como Stereopsia Latam (Chile), EXPYLAB (Paraguay), VR FEST MX (México) y XR ACADEMIA (Colombia/Holanda).

AURA FORERO



English

She was born in Bogotá, Colombia. She is Mechatronic Engineer and Master's in Biomedical Engineering from the National University of Colombia. She has worked on geomatics projects, such as precision agriculture and geolocation by using satellite images and teledetección. She worked in the area of medical imaging developing a model for movement correction to detect the default neural network in patients with Alzheimer's disease in functional magnetic resonance images in a resting state, integrating information from the nuclei of the basal ganglia. Currently, she is CEO of the company CETREAL SAS in Colombia, where mixed reality, artificial intelligence, analytics and blockchain projects are developed. She is also cofounder of XRCOL and currently is the vice president of the association.

Español

Nació en Bogotá, Colombia. Es egresada del programa de Ingeniería Mecatrónica de la Universidad Nacional de Colombia y Máster en Ingeniería Biomédica. Ha trabajado en proyectos de geomática, como agricultura de precisión y geolocalización. Trabajó en el área de imagenología médica desarrollando un modelo para la corrección de movimiento para detectar la red neuronal por defecto en pacientes con enfermedad de alzheimer en imágenes de resonancia magnética funcional en estado de reposo integrando información de los núcleos del ganglio basal. Actualmente es CEO de la empresa CETREAL SAS en Colombia, donde se desarrollan proyectos de realidad mixta, inteligencia artificial, analítica y blockchain.

ARBEBY ARAGON



English

Arbey Aragon was born in Colombia. He studied at the Manuel Cepeda Vargas High School. He has a bachelor's degree in mechatronics engineering from the National University of Colombia. He has expertise in C++ and designing low-cost flight simulators. He is an entrepreneur with his company CETREAL. In 2020, he participated in two events at MIT Latin America AI and Real Mixed reality. Also, he won an Artificial Intelligence hackathon with a Latin American team at MIT University with a project about crime prediction in São Paulo, Brazil. In 2021, he won an award for revolutionary research in immersive technology in Laval Virtual with Justin Beaver Stories. Finally, in 2022 he participated with CETREAL and VisMarCo in the Space Hackathon in Colombia. His project won third place in this competition.

Español

Arbey Aragón nació en Colombia. Estudió en el Liceo Manuel Cepeda Vargas. Tiene una licenciatura en ingeniería mecatrónica de la Universidad Nacional de Colombia. Tiene experiencia en C++ y en el diseño de simuladores de vuelo de bajo costo. Actualmente es empresario y su empresa se llama CETREAL. En 2020, él participó en dos eventos en MIT Latinoamérica AI y Real Mixed reality. Además, ganó un hackathon de Inteligencia Artificial con un equipo latinoamericano de la Universidad MIT con un proyecto sobre la predicción del crimen en São Paulo, Brasil. En 2021, ganó un premio por la investigación revolucionaria en tecnología inmersiva en Laval Virtual con Justin Beaver Stories. Finalmente, en 2022 participó con CETREAL y VisMarCo en el Space Hackathon en Colombia. Su proyecto ganó el tercer lugar en esta competencia.



NELSON ALBERTO ARIAS ARIAS



English

Nelson Alberto Arias Arias, was born in Jenesano-Boyacá, Colombia. He studied elementary school and high school in the field, standing out as one of the best students, he represented his school in physics and mathematics Olympiad, achieving important qualifications for his school. After finishing his secondary studies in the countryside, he moved to Bogotá, the capital city of Colombia, where he applied to the National University of Colombia and began to study Mechatronics Engineering, during his bachelor studies at the University he focused on automation, robotics and artificial intelligence applications. After completing his studies in Mechatronics Engineering, he worked as an Automation Engineer at the oil company Mansarovar Energy Colombia Ltda, later he worked as a Software Developer Engineer in the bunker of the Attorney General's Office in Bogotá. Then, he became an associated member of CETREAL S.A.S. as an entrepreneur and developer of artificial intelligence, virtual reality, augmented reality and mixed reality and currently continues working with CETREAL S.A.S.

Español

Nelson Alberto Arias Arias, Nació en Jenesano-Boyacá, Colombia. Estudió la primaria y bachillerato en el campo destacándose como uno de los mejores estudiantes, representó a su colegio en olimpiadas de física y matemáticas logrando importantes calificaciones para su colegio. Luego de terminar sus estudios de bachillerato en el campo se trasladó a Bogotá, se presentó a la Universidad Nacional de Colombia y entró a estudiar Ingeniería Mecatrónica, durante sus estudios de Ingeniería en la Universidad se enfocó en automatización, robótica y aplicaciones de inteligencia artificial. Después de culminar sus estudios en Ingeniería Mecatrónica trabajó como Ingeniero en automatización en la empresa petrolera Mansarovar Energy Colombia Ltda, posteriormente trabajó



como Ingeniero desarrollador de Software en el búnker de la fiscalía general de la nación en Bogotá, luego se asocia con CETREAL S.A.S. como emprendedor y desarrollador de inteligencia artificial, realidad virtual, realidad aumentada y realidad mixta y actualmente continúa con CETREAL S.A.S.

MENNO VAN ZAAZEN



English

Menno van Zaanen was born in 's-Gravenhage, the Netherlands. He received an Msc in computer science from the Vrije University in Amsterdam, the Netherlands in 1997 and an MA in computational linguistics in 1998 from the University of Amsterdam, the Netherlands. Menno received his PhD in computer science from Leeds University, UK in 2002. He worked as research fellow at the University of Amsterdam, Tilburg University (both the Netherlands) and Macquarie University (Australia) as well as a guest researcher at the University of Groningen (the Netherlands). After that, he worked as a researcher, lecturer and assistant professor at Tilburg University, the Netherlands. Currently, he works as a professor in digital humanities at the South African Centre for Digital Language Resources (SADiLaR) which is hosted at North-West University, South Africa.

His current research interests deal with unsupervised grammatical inference and applied machine learning to sequential data (in particular natural language). Among others, he worked on multi-modal structuring of data (music) and multi-modal information retrieval, as well on proofing tools (spelling checkers for African languages). More recently, he has been concentrating on applying digital techniques in the field of humanities. Menno is a founding member of the International Community in Grammatical Inference (ICGI), which he chaired from 2007-2010. Currently, he is chair of the Digital Humanities Association of Southern Africa (DHASA) and member of the Constituent Organization Board of the international Alliance of Digital Humanities Organizations (ADHO).

Español

Menno van Zaanen nació en la ciudad de 's-Gravenhage, Países Bajos. En 1997, Menno es magíster en ciencias de la computación de la Universidad de Vrije localizada en Amsterdam, Holanda. En 1998, obtuvo su segunda maestría como lingüística computacional en la Universidad de Amsterdam. En 2002, Menno obtuvo su doctorado en ciencias de la computación de la Universidad de Leeds, Reino Unido. Trabajó como investigador asociado en la Universidad de Amsterdam, la Universidad de Tilburgo (ambos Países Bajos) y la Universidad Macquarie (Australia). Fue investigador invitado en la Universidad de Groningen (Países Bajos). Luego trabajó como investigador, conferencista y profesor asistente en la Universidad de Tilburgo. Actualmente, trabaja como profesor de humanidades digitales en el Centro Sur Africano de Recursos Digitales en Lenguaje (SADiLaR) que se encuentra ubicado en la Universidad Norte Oeste de Suráfrica. Sus intereses de investigación actuales se ocupan de la inferencia gramatical no supervisada y el aprendizaje automático aplicado a datos secuenciales (en particular, el lenguaje natural). Así mismo, trabajó en la estructuración multimodal de datos (música) y la recuperación de información multimodal, así como en herramientas de corrección (correctores ortográficos para idiomas africanos). Más recientemente, se ha concentrado en la aplicación de técnicas digitales en el campo de las humanidades. Es miembro fundador de la Comunidad Internacional de Inferencia Gramatical (ICGI), que presidió de 2007 a 2010. Actualmente, es presidente de la Asociación de Humanidades Digitales de África Austral (DHASA) y miembro del Consejo de Organización Constituyente de la Alianza Internacional de Organizaciones de Humanidades Digitales (ADHO).

MARIE POSTMA



English

Marie Postma is associate professor in Cognitive Science and Artificial Intelligence at Tilburg University in the Netherlands where she established several educational programs in the area of data science and artificial intelligence. She obtained her PhD at the Institute of Logic, Language and Computation at the University of Amsterdam. Prior to taking up a position in Tilburg, she worked at the IRIT institute in Toulouse, France. The research topics she works on center around modelling of attentional and perceptual phenomena with applications in education. In 2018, she was awarded a grant to examine the use of virtual reality for the purposes of STEM education.

Español

Marie Postma es profesora asociada de Ciencia Cognitiva e Inteligencia Artificial en la Universidad de Tilburgo en los Países Bajos, donde estableció varios programas educativos en el área de ciencia de datos e inteligencia artificial. Obtuvo su doctorado en el Instituto de Lógica, Lenguaje y Computación de la Universidad de Amsterdam. Antes de ocupar un puesto en Tilburgo, trabajó en el Instituto IRIT en Toulouse, Francia. Los temas de investigación en los que trabaja se centran en el modelado de fenómenos de atención y percepción con aplicaciones en la educación. En 2018, recibió una subvención para examinar el uso de la realidad virtual con fines de educación STEM.

