# Demonstrating the Rapid Integration & Development Environment (RIDE): Embodied Conversational Agent (ECA) and Multiagent Capabilities

Demonstration Track

Arno Hartholt, Ed Fast, Andrew Leeds, Kevin Kim, Andrew Gordon, Kyle McCullough, Volkan Ustun, Sharon Mozgai
USC Institute for Creative Technologies, Play Vista, CA, USA
{hartholt,fast,leeds,kkim,gordon,mccullough,ustun,mozgai}@ict.usc.edu

## ABSTRACT

We demonstrate the Rapid Integration & Development Environment (RIDE), a research and development platform that enables rapid prototyping in support of multiagents and embodied conversational agents. RIDE is based on commodity game engines and includes a flexible architecture, system interoperability, and native support for artificial intelligence and machine learning frameworks.

## KEYWORDS

Multiagents; Embodied Conversational Agents; Virtual Humans; Toolkits, Integration; Artificial Intelligence; Machine Learning

**ACM Reference Format:**
Arno Hartholt, Ed Fast, Andrew Leeds, Kevin Kim, Andrew Gordon, Kyle McCullough, Volkan Ustun, Sharon Mozgai. 2022. Demonstrating the Rapid Integration & Development Environment (RIDE): Embodied Conversational Agent (ECA) and Multiagent Capabilities: Demonstration Track. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

## 1 INTRODUCTION

The Rapid Integration & Development Environment (RIDE) is a research & development (R&D) platform that initially grew out of the US Army's desire to prototype the next generation training and simulation system. As such, RIDE combines many simulation capabilities into a single framework, including synthetic real-world terrain, support for artificial intelligence (AI) and machine learning (ML) frameworks, networked multiplayer, Experience Application Programming Interface (xAPI) logging, Distributed Interactive Simulation (DIS) messaging, a unified web service interface, and multi-platform support. Many of these capabilities support rapid prototyping needs beyond the military, in particular for creating scenarios with multiple scripted agents embedded in synthesized terrain as a starting point to train more advanced behavior models.

This demonstration of RIDE focuses on the overall architecture of RIDE as well as three core pillars: synthesized terrain, AI and ML support, and embodied conversational agents (ECAs).

*Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.* 

## 2 RIDE ARCHITECTURE

RIDE has been designed and developed from the ground up to facilitate rapid prototyping specifically for simulation researchers and developers, following these guidelines:

(1) Leverage real-time game engine technologies that provide core capabilities, including rendering, physics, and audio.
(2) Abstract away from specific game engines in order to provide simulation researchers and developers with the concepts they are most familiar with.
(3) Provide a drag-and-drop development environment that offers reusable blueprints of commonly used functionalities.
(4) Integrate core functionalities into a common framework in order to add combinatorial value.
(5) Offer all of the above through a principled Application Programming Interface (API) that unifies all RIDE functionality in a consistent and easy to use manner.

In order to support a large ecosystem that contains many different developers, researchers, technologies, applications, and organizations, RIDE follows a layered architecture, see Figure 1. The *Engine Layer* allows RIDE to leverage robust gaming technologies that provide common capabilities, including rendering, physics, animation, pathfinding, User Interface (UI), audio, and network protocols. RIDE is agnostic to any specific game or simulation engine, which avoids vendor lock-in and enables researchers and developers to create simulation scenarios without the need to have experience with a specific engine. Currently, the main target is Unity, through the C# RIDE API, with key capabilities ported to Unreal Engine using C++. The *Middleware Layer* abstracts and augments the Engine Layer with simulation-specific capabilities, including synthesized terrain, AI agent behaviors, combat system, scenario system, ML interfaces, and networked multi-user capabilities. The *Project Layer* allows researchers and developers to leverage RIDE as a foundation for their own projects. This incubation layer validates the middleware and allows for robust capabilities to flow back into RIDE and benefit the entire community.

## 3 RIDE CAPABILITIES

### 3.1 Synthesized Terrain

Terrain data is a critical component of many simulation systems. Due to the nature of real-world terrain data, especially for military areas, sourcing existing 3D Geospatial data is a resource intensive process both in time, money, and human effort. While there are

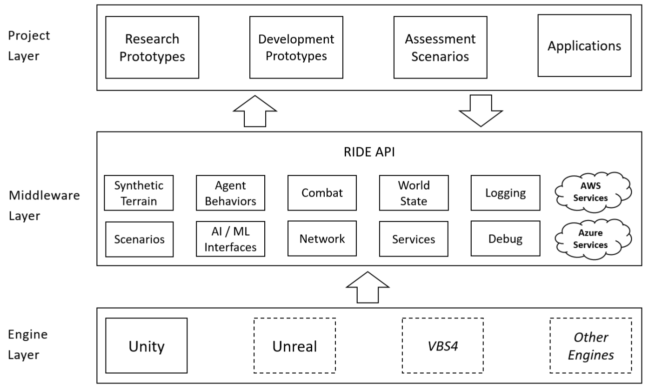


**Figure 1: The multi-layer RIDE architecture.**

existing databases of high-quality data, that data is not often simulation ready or of a level of fidelity that could support ground level first person systems as would be necessary for specific use cases such as human training systems or autonomous vehicles simulators. Researchers at USC ICT have developed a fully automated pipeline that takes in high-resolution data from sources such as Small Unmanned Aerial System (sUAS) [2] or Microsoft Bing [3], and outputs high-resolution 3D models ready for simulation [1]. The resulting terrain is used in RIDE with agent navigation, tree replacement, detailed measurement, deformation, etc.

### 3.2 AI and ML

A serious bottleneck is the ability to create, collect, analyze, and validate quality data that is relevant and readily accessible, in sufficient quantities. RIDE addresses this challenge by combining relevant capabilities into a principled architecture and API, shaped directly by AI R&D needs. As such, it provides a natural environment for AI and ML experiments, augmenting general data science needs.

RIDE enables users to author, test, validate, train, and execute agent behaviors. To this end, initial agent behaviors can be scripted with either state machines or behavior trees in order to bootstrap experiments. Provided behaviors include movement, attacking, and formations, tied to dedicated combat and health systems. Alternatively, networked multiplayer allows multiple teams of users to control avatars in a common synthetic environment, each from their own location and type of hardware (e.g., desktop, mobile).

Both human avatars and bots can be mixed, gathering data to train ML models. RIDE has been used to create and run models in TensorFlow, PyTorch and custom solutions. Results can either directly be run in RIDE through the inference model, by providing individual commands to agents, or by generating behavioral code that leverages atomic actions. See [7] for more details.

### 3.3 Embodied Conversational Agents

ECAs are interactive, digital characters that perceive real humans and respond appropriately, both verbally and nonverbally. They act as social interface agents that add a social component to the environments in which they are embedded. They provide a standardized experience across users and can be omnipresent and indefatigable in their roles. ECAs have been shown to improve user's perception of their environment [11], increase interaction time [12], and improve learning outcomes [16].

ICT has widely been recognized as one of the leaders in ECA R&D, including basic research in cognitive architectures [14], audio-visual sensing [15], and character animation simulation [17], as well as applied prototypes for leadership development [10], information dissemination [13], job interview training [8], and life-long learning [18]. Our approach is highly interdisciplinary with a strong focus on integrating both theory and technology into common frameworks [6] [4]. This work has resulted in the Virtual Human Toolkit, a collection of modules, tools, and libraries designed to aid and support researchers and developers with the creation of virtual human conversational characters [9]. The Toolkit is freely available for academic and government purpose use. Recent efforts have resulted in the ability to develop virtual humans for a range of hardware platforms, including web, mobile, AR and VR [5].

Ongoing work integrates these efforts with RIDE, revolving around architecture, capabilities, and content. The Toolkit architecture is modular, using a combination of message passing and direct data streams to facilitate inter-module communication, augmented with a microservices architecture. RIDE has integrated its messaging protocol (VHMsg) built on top of ActiveMQ. This allows interfacing with any of the main Toolkit modules and the capabilities they represent, including natural language processing (NLP) and nonverbal behavior generation. In addition, by leveraging RIDE's dedicated web services system, commodity AI-related services are provided, including audio-visual sensing, speech recognition, NLP, and text-to-speech generation. Nonverbal behavior realization (e.g., lip-sync, facial expressions, conversational gestures) is directly integrated within RIDE. These capabilities allow agents to observe real human users, create its communicative intent, and realize that intent verbally and nonverbally, synchronized in real-time.

## 4 DISCUSSION

We demonstrate RIDE, the Rapid Integration & Development Environment, as a foundational platform for R&D in the field of multiagents and embodied conversational agents. RIDE's focus on rapid prototyping, system interopability, and architectural flexibility, combined with a range of integrated features (e.g., synthesized terrain, native AI and ML support, multiplatform support), results in a platform that enables researchers and developers to quickly create virtual environments. RIDE is freely available for US government purpose or through a custom license[1]. RIDE is early work and therefore one of its limitations is varied levels of maturity for its features. Future work will extend the number of agents that can run simultaneously, integrate the ability to rapidly create self-avatars, and expand support for Unreal Engine.

## ACKNOWLEDGMENTS

We thank all of our partners at USC ICT, DEVCOM STTC, and ONR. Part of the efforts depicted were sponsored by the US Army under contract number W911NF-14-D-0005. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

[1] https://ride.ict.usc.edu/getting-started/obtain-ride

## REFERENCES


[1] Meida Chen, Andrew Feng, Ryan McAlinden, and Lucio Soibelman. 2020. Photogrammetric point cloud segmentation and object information extraction for creating virtual environments and simulations. *Journal of Management in Engineering* 36, 2 (2020), 04019046.

[2] Meida Chen, Andrew Feng, Kyle McCullough, Pratusha Bhuvana Prasad, Ryan McAlinden, and Lucio Soibelman. 2020. 3D photogrammetry point cloud segmentation using a model ensembling framework. *Journal of Computing in Civil Engineering* 34, 6 (2020), 04020048.

[3] Meida Chen, Andrew Feng, Kyle McCullough, Pratusha Bhuvana Prasad, Ryan McAlinden, and Lucio Soibelman. 2020. Semantic Segmentation and Data Fusion of Microsoft Bing 3D Cities and Small UAV-based Photogrammetric Data. *arXiv preprint arXiv:2008.09648* (2020).

[4] David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, et al. 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 1061–1068.

[5] Arno Hartholt, Ed Fast, Adam Reilly, Wendy Whitcup, Matt Liewer, and Sharon Mozgai. 2020. Multi-Platform Expansion of the Virtual Human Toolkit: Ubiquitous Conversational Agents. *International Journal of Semantic Computing* 14, 3 (2020).

[6] Arno Hartholt, Jonathan Gratch, Lori Weiss, et al. 2009. At the virtual frontier: Introducing Gunslinger, a multi-character, mixed-reality, story-driven experience. In *International Workshop on Intelligent Virtual Agents*. Springer, 500–501.

[7] Arno Hartholt, Kyle McCullough, Ed Fast, Andrew Leeds, Sharon Mozgai, Tim Aris, Volkan Ustun, Andrew S. Gorden, and Christopher McGroarty. 2021. Rapid Prototyping for Simulation and Training with the Rapid Integration Development Environment (RIDE). In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*.

[8] Arno Hartholt, Sharon Mozgai, and Albert" Skip" Rizzo. 2019. Virtual job interviewing practice for high-anxiety populations. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. 238–240.

[9] Arno Hartholt, David Traum, Stacy C Marsella, Ari Shapiro, Giota Stratou, Anton Leuski, Louis-Philippe Morency, and Jonathan Gratch. 2013. All together now, Introducing the Virtual Human Toolkit. In *International Workshop on Intelligent Virtual Agents*. Springer, 368–381.

[10] Matthew J Hays, Julia C Campbell, Matthew A Trimmer, Joshua C Poore, Andrea K Webb, and Teresa K King. 2012. *Can role-play with virtual humans teach interpersonal skills?* Technical Report. UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES INST FOR CREATIVE TECHNOLOGIES.

[11] W Lewis Johnson, Jeff W Rickel, James C Lester, et al. 2000. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial intelligence in education* 11, 1 (2000), 47–78.

[12] H Chad Lane, Dan Noren, Daniel Auerbach, Mike Birch, and William Swartout. 2011. Intelligent tutoring goes to the museum in the big city: A pedagogical agent for informal science education. In *International Conference on Artificial Intelligence in Education*. Springer, 155–162.

[13] Albert Rizzo, Belinda Lange, John G Buckwalter, Eric Forbell, Julia Kim, Kenji Sagae, Josh Williams, JoAnn Difede, Barbara O Rothbaum, Greg Reger, et al. 2011. SimCoach: an intelligent virtual human system for providing healthcare information and support. (2011).

[14] Paul S Rosenbloom, Abram Demski, and Volkan Ustun. 2016. The Sigma cognitive architecture and system: Towards functionally elegant grand unification. *Journal of Artificial General Intelligence* 7, 1 (2016), 1–103.

[15] Stefan Scherer, Stacy Marsella, Giota Stratou, Yuyu Xu, Fabrizio Morbini, Alesia Egan, Albert Rizzo, and Louis-Philippe Morency. 2012. Perception Markup Language: Towards a Standardized Representation of Perceived Nonverbal Behaviors. Springer, Berlin, Heidelberg, 455–463. https://doi.org/10.1007/978-3-642-33197-8_47

[16] Noah L Schroeder, Olusola O Adesope, and Rachel Barouch Gilbert. 2013. How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research* 49, 1 (2013), 1–39.

[17] Ari Shapiro. 2011. Building a character animation system. In *International conference on motion in games*. Springer, 98–109.

[18] William R Swartout, Benjamin D Nye, Arno Hartholt, Adam Reilly, Arthur C Graesser, Kurt VanLehn, Jon Wetzel, Matt Liewer, Fabrizio Morbini, Brent Morgan, et al. 2016. Designing a personal assistant for life-long learning (PAL3). In *The Twenty-Ninth International Flairs Conference*.