Team Noyce: Alzheimer's Insights AI Prize Application

Main

What is your project's core hypothesis, and what data supports it?

Outline the innovative AI or biomedicine methods you will use, explaining how these—supported by relevant evidence—advance ADRD research. Address compatibility with AD Workbench. Provide strong scientific evidence and peer-reviewed references to support your claims. Include your plans for validating robustness and reliability and discuss any technical risks or potential biases. Include any existing product demos. (4,000 characters)

Currently, no comprehensive multimodal dataset exists for training AI systems to manage dementia-related agitation. While resources like DementiaBank (Lanzi et al., 2023) provide structured clinical interviews, they lack naturalistic caregiver-patient interactions during actual agitation episodes. Without such data, developing agentic AI for dementia care remains impossible. Our goal is to create the first privacy-preserving, clinically validated multimodal dataset capturing the communicative dynamics of agitation, enabling the global research community to advance agentic AI for empathetic dementia care.

We hypothesize that the absence of shareable, ecologically valid dementia interaction data is the main bottleneck preventing agentic AI development. Synthetic, harmonized datasets approximating real agitation dynamics can support systems capable of reasoning and acting autonomously. By aligning synthetic data with real-world distributions while protecting privacy, we accelerate research that would otherwise require years of ethical and logistical effort.

Since July 2025, our multidisciplinary team, spanning clinical nursing, AI, linguistics, and robotics, has collected ADRD clinic and simulation data streams (audio, video), processed into valence and arousal values. These remain private under HIPAA and GDPR but serve as calibration ground truth. Unlike existing corpora focused on diagnosis, we capture real caregiving moments when agitation arises. In parallel, we curated a library of clinically validated personas and scenarios, such as repetitive questioning or asking to go home, from which multi-turn dialogues are generated and validated.

Our framework integrates two linguistic lenses. Drawing from Heidi Hamilton (e.g., Hamilton, 2008, 2019), we analyze turn-level interaction to model agitation as a co-constructed process. Concurrently, using Louise Cummings' work in clinical pragmatics (Cummings, 2009; 2014; 2015; 2017), we identify markers like topic deviation and repair failure. These are linked to agitation scores derived from video.

Computer vision models detect behaviors like pacing and restlessness, mapped to the Cohen-Mansfield Agitation Inventory (CMAI), which categorizes 29 agitated behaviors (Cohen-Mansfield, 1986), and the Neuropsychiatric Inventory (NPI), which assesses 12 behavioral domains in dementia (Cummings et al., 1994). These tools allow us to link subjective symptoms to observable multimodal signals.

Synthetic dialogues are rendered using prosody-tuned Text-to-Speech (TTS) that approximates elderly vocal patterns. Speech Synthesis Markup Language (SSML) allows precise control over pitch, tempo, and pause to express emotion. Acoustic-prosodic features, including F0 range, jitter, shimmer, energy, and speech rate, are informed by studies on how prosodic variation reflects emotional expression in speech (Frick, 1985; Kehrein, 2002; Pereira & Watson, 1998). Audio and affective curves are aligned with movement data via Dynamic Time Warping (DTW). Fidelity is measured using Maximum Mean Discrepancy (MMD), Fréchet Audio Distance (FAD), and expert review.

Validation is led by a multidisciplinary panel of clinicians and linguists. They co-develop a rubric assessing sociolinguistic and clinical fidelity, which guides a fine-tuned LLM for first-pass annotation. A 10–15% sample is then audited by human experts. Inter-rater reliability is ensured using Fleiss' κ and Intraclass Correlation (ICC), validating both technical accuracy and contextual authenticity.

If synthetic audio lacks aging timbre, we will refine SSML parameters or train on elderly corpora. If multimodal fusion introduces bias, we will use late-fusion or ablation. All artifacts, including dataset specs, model cards, and annotations, will be shared via AD Workbench. While this phase focuses on dataset construction, the system is explicitly built to support future agentic AI deployment in real-time care settings.

(3,876/4,000)

—

Reference List (will submit this as a separate file)

Cohen-Mansfield, J. (1986). Agitated behaviors in the elderly. I. A conceptual review. *Journal of the American Geriatrics Society*, 34(10), 711–721.

Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*, 44(12), 2308–2308.

Cummings, L. (2009). *Clinical pragmatics*. Cambridge University Press.

Cummings, L. (2014). Pragmatic disorders across the life span. In *Pragmatic Disorders* (pp. 31–65). Dordrecht: Springer Netherlands.

Cummings, L. (2015). Theory of mind in utterance interpretation: The case from clinical pragmatics. *Frontiers in Psychology*, 6, 1286.

Cummings, L. (Ed.). (2017). *Research in clinical pragmatics* (Vol. 11). Springer.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin*, 97(3), 412–429.

Hamilton, H. E. (2008). Language and dementia: Sociolinguistic aspects. *Annual Review of Applied Linguistics*, 28, 91–110.

Hamilton, H. E. (2019). *Language, dementia and meaning making: Navigating challenges of cognition and face in everyday life*. Palgrave Macmillan.

Kehrein, R. (2002, April). The prosody of authentic emotions. In *Speech Prosody 2002, International Conference*.

Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. L. (2023). DementiaBank: Theoretical rationale, protocol, and illustrative analyses. *American Journal of Speech-Language Pathology*, 32(2), 426–438.

Pereira, C., & Watson, C. I. (1998, November). Some acoustic characteristics of emotion. In *ICSLP*.

Schiffrin, D., Tannen, D., & Hamilton, H. E. (Eds.). (2008). *The handbook of discourse analysis*. John Wiley & Sons.

Project milestone

Provide an overview of your project's key milestones, emphasizing how your approach is expected to enable researchers to accelerate their research and desired outcomes compared to traditional research workflows.

Additionally, explain how you will utilize the AIzInsights Prize’s core support (funding, data, etc.) to advance your project, ensuring efficient use of resources. Describe any further support you may require beyond the prize. (max 2,000 characters)

Our project delivers essential training data for agentic AI in dementia care, capturing the dynamics of real agitation episodes. We will deliver this in two phases aligned with major ADRD conferences.

By CTAD 2025 (December 1), we will release Dataset v1.0: 1,000+ validated, text-based synthetic multi-turn dialogues grounded in clinical personas and scenarios. This dataset is tailored for agitation management, with annotations for emotional valence, triggers, and de-escalation strategies to enable supervised training of conversational AI agents.

By AD/PD 2026 (March 17–21), we will release Dataset v2.0, the complete multimodal version, including prosody-tuned synthetic audio with elderly vocal characteristics and synchronized behavioral annotations. This version will be the first to capture agitation's full communicative spectrum. Metadata will support training via behavioral cloning, inverse reinforcement learning, or hybrid approaches, accelerating development versus collecting real-world patient data.

We will use the AlzInsights AI Prize's $10,000 for scientific rigor and open community access. Expert validation ($6,500) will support honoraria for our 13 clinical and linguistic reviewers, who will develop rubrics and evaluate dialogue realism and fidelity, ensuring the dataset's authenticity. Dissemination and computation costs will support presentations at CTAD and AD/PD ($2,500) and cloud-based generation and LLM validation ($1,000).

All foundational infrastructure, including GPU resources, workspace, and clinical partnerships, is provided in-kind by UC Irvine. The prize funds will directly support validation and dissemination, maximizing research community benefit. Post-release, we plan to seek NIH funding to expand the dataset and build reference implementations of agentic systems.

(1,973/2,000)

Team expertise

Summarize your team's multidisciplinary expertise, highlighting strengths in AI engineering, ADRD biology, computational biology, biomedicine, and any experience with productization or open-source communities. Please address any gaps in expertise and how you plan to overcome them. (max 3,000 characters)

Dr. Cheonkam Jeong, a Ph.D. in linguistics and Postdoctoral Scholar at UCI Nursing, leads agentic AI design. Her research focuses on the discourse-phonetics interface, combining conversation analysis with LLM-based generation for dementia communication. She develops multimodal models integrating prosody, pragmatics, and patient response trajectories for real-time interaction. She specializes in LLM fine-tuning for empathetic communication and reinforcement learning frameworks for empathetic dialogue generation.

Dr. Adeline Nyamathi, founding Dean of UCI School of Nursing and globally recognized behavioral scientist, provides clinical leadership. Her expertise in vulnerable populations, including ADRD, is grounded in decades of NIH-funded research. Her involvement in empathetic patient-robot interaction studies, including as corresponding author on our JMIR publication, ensures the project remains clinically meaningful and ethically grounded.

Mahyar Abbasian, a Ph.D. candidate in Computer Science, leads the generative AI and dialogue modeling pipeline. He created openCHA, an open-source conversational agent framework with multi-step reasoning capabilities, and has published in npj Digital Medicine on healthcare-focused LLM architectures.

Fatemeh (Saba) Azizabadi Farahani, a Ph.D. student in Electrical Engineering and Computer Science, directs multimodal signal fusion. Her research focuses on aligning heterogeneous clinical data streams, including text, biosignals, and motion data, and developing algorithms for synchronized audio-video processing in healthcare contexts.

Ali Ahmad, a senior robotics engineer at Xavor, manages integration of agentic AI into embodied systems. As lead developer of the ROS-based Care Robot platform, he ensures dialogue systems can be deployed in real-time robotic environments with seamless hardware-software coordination.

Farhan Azhar, a senior AI consultant at Xavor, leads the prosody-tuned synthetic speech pipeline. He brings hands-on experience developing elderly voice synthesis and emotion-driven speech modulation from previous UCI collaborations on human-robot interaction.

Muhammad Rafay Shafqat, a machine learning engineer at Xavor, leads the computer vision pipeline. He builds video-based behavioral classifiers for agitation detection, developing edge-deployable models calibrated to CMAI and NPI scoring systems to ensure clinical relevance and objectivity.

Potential gaps in our team's expertise are addressed through coordination between UCI and Xavor, combining academic innovation with industry-grade technical execution. Our cross-functional collaboration, spanning nursing, AI, linguistics, and robotics, ensures that both the data and its downstream use are grounded in rigorous, domain-specific insight. This foundation uniquely positions our team to deliver a dataset that enables meaningful breakthroughs in dementia AI research, notably in developing truly agentic, multimodal systems for empathetic care.

(2,988/3,000)

Supplementary documents

Create a Dropbox folder that includes a 1-page CV for each team member (required).
Attach a viewable link to your team’s Dropbox folder below.
You may also include the following in the same folder (optional but strongly encouraged):

- Institutional or industry support letters
- Concept video demos