

VIDEO/Podcasts/vlogs This video and all related information and materials ("materials") are owned by Carnegie Mellon University. These materials are provided on an "as-is" "as available" basis without any warranties and solely for your personal viewing and use. You agree that Carnegie Mellon is not liable with respect to any materials received by you as a result of viewing the video, or using referenced web sites, and/or for any consequence or the use by you of such materials. By viewing, downloading and/or using this video and related materials, you agree that you have read and agree to our terms of use (<http://www.sei.cmu.edu/legal/index.cfm>).  
DM22-0102

**Script:** The Current State of Explainable AI

**SME:** Violet Turri

**Moderator:** Rachel Dzombak

**Interview Conducted:** 01/27/2022 at 1 p.m. ET (Remote)

**[Recorded Intro.]**

**Rachel Dzombak:** Hi and welcome to the SEI Podcast Series. I am Dr. Rachel Dzombak, and I am a digital transformation lead in the SEI's AI Division. Today I am joined by my colleague, Violet Turri, to discuss the current state of Explainable Artificial Intelligence, including its challenges and limitations.

Violet recently wrote a blog post outlining the strengths and practices of Explainable AI, also known as XAI, which she described as a "powerful tool to help develop user trust in AI systems and address rising ethical and legal concerns."

Welcome, Violet.

## Violet: Responds.

- 1. Rachel:** Let's start off by having you tell us a little bit about yourself and the work that you do here at the SEI.
- 2. Rachel:** Can you define Explainable Artificial Intelligence for us, perhaps with some real-world examples of the legal and ethical concerns that XAI can help address?
- 3. Rachel:** In your blog post you write that XAI comprises the processes and methods that allow human users to comprehend and trust the results and output created by machine learning algorithms. Tell us more about how these processes and methods work. Exactly how do these processes and methods give users a greater sense of trust in machine learning outcomes?
- 4. Rachel:** The [U.S. government](#) has identified explainability as a key tool for developing trust and transparency in AI systems, and we have also witnessed a rising demand for XAI within other government sectors. The U.S. Department of Health and Human Services lists an effort to “promote ethical, trustworthy AI use and development,” including explainable AI, as one of the focus areas of their [AI strategy](#).

How is the SEI's AI Division working to help the DoD address this rising demand for XAI? Does this research contribute to the human-centered pillar of AI engineering?

- 5. Rachel:** Despite a growing demand for explainable artificial intelligence in fields such as medicine and finance, there are still a number of limitations. Tell us about some of these limitations that you have identified and ways that they can be mitigated.
  
- 6. Rachel:** One thing that we like to address in our podcasts is transition. If am I leading an effort to adopt XAI within my organization, where should I start? What resources are available to me?

**Rachel:** Violet, thank you for being here to talk with us about explainable AI. And to our listeners, thanks for joining us today. We will include links in our transcript to all resources mentioned in this podcast.

As always, if you have any questions, please don't hesitate to email us at [info@sei.cmu.edu](mailto:info@sei.cmu.edu).

Thank you.

[Recorded Outro]