

## ESSAI 2024 Course

**Title:** **Explainable AI via Argumentation: Theory & Practice**

**Intended Level:** Requires general background on Computing and AI

**Tutors:** [Antonis Kakas](#) and [Nikolaos Spanoudakis](#)

**Affiliations:** University of Cyprus and Technical University of Crete

Co-founders of the AI start-up company of [Argument Theory](#) in Paris.

### Motivation

Providing explanations for AI system results is now essential. AI systems must not only deliver accurate outputs but also ensure their results are interpretable and understandable. Clear explanations enhance usability, foster trust, and support both system development and acceptance. Explainable AI (XAI) systems can be naturally integrated with human experts and users in a collaborative and enhancing way.

Argumentation, a form of reasoning naturally linked to informative explanations, offers a robust framework for Explainable AI (XAI). Arguments supporting conclusions provide the basis of explanations, while counter arguments add contrastive elements, making the decision-making process transparent and verifiable. In addition, argumentation provides a logical form of reasoning that covers the wide spectrum of logical reasoning from formal strict reasoning to informal common-sense reasoning. As such, Argumentation forms a suitable Calculus for human-centric Explainable AI.

### What You'll Learn

- **The Role of Explanations in AI:** Understand why explanations are crucial for AI systems' transparency and user trust and the essential quality requirements for effective explanations in AI systems.
- **Argumentation Theory:** Explore the theoretical framework that underpins argumentation-based XAI.
- **Practical Application:** Discover how to apply argumentation theory in real-world XAI systems using practical tools and methodologies.
- **Hands-On Experience:** A strong component of the course is to gain hands-on experience with cutting-edge tools, enabling you to build XAI decision-making systems.

*You will build at least a nucleus system application of your choice. If you want, you can further expand on the features of your system and build a prototype application using APIs provided by the argumentation development tools presented in the course.*

### Connect to the Gorgias Cloud SaaS platform

The Gorgias Cloud platform is provided for non-commercial use by the Applied Mathematics and Computers Laboratory of the Technical University of Crete.

To register to the [Gorgias Cloud](#) platform, follow the steps and advice at the [link](#).

### **Connect to the RAISON no-code development AI platform**

[Argument Theory](#) offers the RAISON platform, a no-code symbolic artificial intelligence (AI) development platform specializing in the modeling and development of automated decision-making systems. The RAISON platform is currently in beta-testing.

Instructions for connecting to RAISON will be given during the lectures.

### **Optional Background reading**

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267, 1-38.

Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., ... & Seifert, C. (2023). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*, 55(13s), 1-42.

Vassiliades, A., Bassiliades, N., & Patkos, T. (2021). Argumentation and explainable artificial intelligence: a survey. *The Knowledge Engineering Review*, 36, e5.

Spanoudakis, N. I., Gligoris, G., Koumi, A., & Kakas, A. C. (2023). Explainable argumentation as a service. *Journal of Web Semantics*, 76, 100772.

Kakas, A. C., Moraitis, P., & Spanoudakis, N. I. (2019). GORGAS: Applying argumentation. *Argument & Computation*, 10(1), 55-81.

Dietz, E., Kakas, A., & Michael, L. (2022). Argumentation: a calculus for human-centric AI. *Frontiers in Artificial Intelligence*, 5.