

# Argument Mining for Ethical Chatbots

La costruzione di intelligenze artificiali in grado di prendere decisioni in accordo con principi etici è una sfida che deve essere affrontata allo scopo di migliorare l'affidabilità ed equità dell'AI.

Questo progetto mira a combinare *argument mining* e ragionamento argomentativo per assicurare un comportamento etico nel contesto di sistemi di dialogo quali *chatbot*. L'argomentazione è uno strumento potente per modellare conversazioni e dispute. *Argument mining*, ovvero la ricerca automatica di argomenti da input testuali, è uno strumento che si intende usare sia per l'analisi dell'input dell'utente, sia per il recupero di feedback adatti diretti all'utente. Lo scopo è estendere i classici framework di argomentazione computazionale con vincoli etici e/o morali, e con capacità di interazione in linguaggio naturale, in modo tale da guidare la conversazione tra chatbot e umani in accordo con i vincoli etici.

L'attività si inquadra nell'ambito del WP5 ("AI Ethics and Responsible AI") del progetto H2020 **HumanE-AI-Net** H2020-ICT-2018-20 / H2020-ICT-2019-3, Grant Agreement number: 952026 (<https://www.ai4eu.eu/humane-ai-net-0>). In particolare, si prevede la collaborazione con un altro partner del progetto Humane-AI-Net, vale a dire l'ICAR-CNR di Cosenza / gruppo di argomentazione computazionale, nel contesto del microprogetto collaborativo "*Ethical chatbots*", già parzialmente finanziato dal progetto europeo.

La presente attività formativa è incentrata sullo sviluppo delle tecniche di *argument mining* e della loro applicazione nel contesto dei *chatbot*.