

Mental health: A bet on responsible AI

July 14, 2026

FBK develops intelligent agents that analyze therapeutic conversations to identify signs of psychological vulnerability and support the work of professionals. The model, validated in an international competition and presented at AIME 2026, opens new perspectives for digital mental health.

Fondazione Bruno Kessler has confirmed its position among the leading contributors to European research on artificial intelligence applied to mental health. A team from the **Intelligent Digital Agents Unit (IDA)** of the **FBK [Center for Digital Health & Wellbeing](#)** won first place in the **MentalRiskES 2026** international competition, organized as part of IberLEF **2026**, one of the leading scientific events dedicated to natural language processing and its applications. In previous editions, the event has featured tasks in the clinical and biomedical fields, addressing challenges such as the anonymization of medical documents, the extraction and standardization of information from health-related texts, the recognition of biomedical entities, and the early identification of mental health risk signals.

The research team, composed of Patrizio Bellan, Piergiorgio Maruotti, Leonardo Sanna, and Mauro Dragoni, achieved the best result in the task dedicated to the **early identification of symptoms of psychological distress through the analysis of real therapeutic conversations**.

The challenge involved 17 international teams and focused on a topic of growing social importance: developing systems capable of supporting the **identification of signals associated with mental health by analyzing dialogues between patients and healthcare professionals**. This is a particularly complex area, where interpreting emotional, behavioral, and relational elements normally requires specialized expertise and significant clinical experience.

The system developed by the team is based on a multi-agent reasoning approach designed to **analyze complex problems by combining multiple expert perspectives**. Rather than entrusting the entire evaluation to a single language model, the approach creates a **group of artificial agents**, each with a specific role and a distinct perspective on the problem. This makes it possible to break down the analysis according to different expert viewpoints: the same

therapeutic conversation is examined from complementary perspectives, for example by focusing on emotional signals, behavioral aspects, dialogue coherence, or contextual clues that emerge progressively throughout the interaction.

Each agent produces its own reasoned evaluation, contributing a partial but specialized analysis. These evaluations are not simply combined but are compared and integrated through a decision aggregation phase, in which the different observations are brought together into a more robust and balanced final assessment. In this way, **multi-agent reasoning reduces dependence on a single model interpretation and makes the decision-making process more transparent, because the** final result can be traced back to the evaluations produced by the individual agents.

A particularly relevant aspect is that the system operated **in zero-shot mode: it** was not trained on the competition's specific data or on labeled examples of the task to be solved. To make it work, the system was provided only with the task instructions and the operational context in which it had to operate, without any additional training or adaptation on the competition data. Decisions were then generated from the instructions assigned to the agents, their expert perspectives, and the system's ability to reason about the conversation, demonstrating how a multi-agent organization can support analysis even in the absence of task-specific training data.

In addition, the data used in the competition came from **real therapeutic conversations and were validated by experts in the field**, making the results particularly significant from both scientific and practical application perspectives.

Potential applications include improving psychotherapy and counseling services, enabling the early identification of vulnerable individuals, and supporting patient-professional matching processes. In all of these areas, artificial intelligence can help make services more accessible, personalized, and efficient while keeping the healthcare professional at the center of the care pathway.

The research underlying the system was presented in early July at AIME **2026, the 24th International Conference on Artificial Intelligence in Medicine**, held at the University of Ottawa, Canada. The conference is one of the leading international forums for advancing research and applications in artificial intelligence for biomedicine, clinical practice, and healthcare, bringing together researchers, clinicians, and innovators working on methods and solutions for medicine and digital health.

PERMALINK

<https://magazine.fbk.eu/en/news/mental-health-a-bet-on-responsible-ai/>

TAGS

- #artificialintelligence
- #digitalhealthwellbeing
- #disagio psicologico
- #ida

- #mental health
- #MentalRiskES 2026
- #multi-agents
- #psychological vulnerability

AUTHORS

- Giovanna Rauzi