

Sveučilište u Zagrebu

Fakultet hrvatskih studija Sveučilišta u Zagreb

Luka Poslon

**ETIČKO-EPISTEMOLOŠKI ASPEKTI  
PRIMJENE UMJETNE INTELIGENCIJE  
U ZDRAVSTVU**

DOKTORSKI RAD

Zagreb, 2025.

Sveučilište u Zagrebu

Fakultet hrvatskih studija Sveučilišta u Zagreb

Luka Poslon

**ETIČKO-EPISTEMOLOŠKI ASPEKTI  
PRIMJENE UMJETNE INTELIGENCIJE  
U ZDRAVSTVU**

DOKTORSKI RAD

Mentori:

izv. prof. dr. sc. Sandro Skansi

izv. prof. dr. sc. Anto Čartolovni

Zagreb, 2025.

University of Zagreb

The Faculty of Croatian Studies at the University of Zagreb

Luka Poslon

**ETHICAL-EPISTEMOLOGICAL ASPECTS  
OF THE APPLICATION OF ARTIFICIAL  
INTELLIGENCE IN HEALTHCARE**

DOCTORAL THESIS

Supervisors:

Associate Professor Sandro Skansi

Associate Professor Anto Čartolovni

Zagreb, 2025.

## Sažetak:

Ovaj doktorski rad istražuje etičke i epistemološke izazove primjene sustava umjetne inteligencije (dalje: UI) u zdravstvu s posebnim naglaskom na objašnjivu umjetnu inteligenciju (dalje: OUI). Zbog sve veće primjene složenih sustava UI-e koji često rade na principu „crne kutije“, što otežava razumijavanje rada sustava UI-e i otežava pravovaljani informirani pristanak, javlja se potreba za razvojem metoda koje omogućuju veću razinu transparentnosti i objašnjivosti. Rad naglašava nužnost ljudskog nadzora, validacije i verifikacije sustava kako bi se osigurala preciznost predviđanja sustava UI-e i zaštitila prava pacijenta pri donošenju odluka. Posebna pažnja rada posvećena je izazovima objašnjivosti, povjerenja, pristranosti, informiranog pristanka i odgovornosti koji nastaju na presjecištu zdravstva, računarstva i filozofije.

Rad analizira temeljne pojmove UI-e, prikazuje prednosti koje UI donosi u zdravstvu, kao što su brža dijagnostika i personalizirana terapija, ali i rizike kao što su pogreške, pristranosti i neobjašnjiva predviđanja. Znanstveni doprinos rada očituje se u metodološki utemeljenom razgraničenju ključnih pojmova objašnjivosti, interpretabilnosti i transparentnosti, ali i u originalnom pojmovnom okviru za sustavno razumijevanje OUI-e te epistemološkoj analizi pouzdanosti. Empirijski dio rada donosi prikaz primjene UI-e u transplantaciji bubrega kao studiju slučaja koja pokazuje kako tehnološke odluke imaju duboke etičke posljedice. Time ovaj rad doprinosi jasnijem razumijevanju UI-e i postavlja temelje za etički utemeljenu i epistemološki pouzdanu primjenu UI-e u budućnosti u zdravstvenoj skrbi.

Ključne riječi: umjetna inteligencija, objašnjiva umjetna inteligencija, crna kutija, etika umjetne inteligencije, zdravstvo, medicina, transplantacija bubrega

## Abstract:

This doctoral dissertation explores the ethical and epistemological challenges involved in application of the artificial intelligence (AI) in healthcare, with special emphasis on explainable artificial intelligence (xAI). Many AI systems operate as “black boxes,” obscuring their processes and complicating valid informed consent, highlighting the need for methods that enhance transparency and explainability. The dissertation emphasizes human oversight,

validation, and verification to ensure accurate AI predictions and protect patient rights in decision-making. Special attention is given to issues of explainability, trust, bias, informed consent, and responsibility at the intersection of healthcare, computing, and philosophy.

The dissertation examines core AI concepts and highlights the benefits AI can provide in healthcare, from accelerated diagnostics to individualized therapies, while simultaneously acknowledging the risks posed by errors, biases, and non-transparent predictions. Its scientific contribution lies in a methodologically grounded distinction between explainability, interpretability, and transparency, alongside an original framework for systematically understanding xAI and an epistemological analysis of its reliability. The empirical section presents the application of AI in kidney transplantation as a case study, illustrating how technological decisions can entail profound ethical consequences. By doing so, the dissertation fosters a deeper comprehension of AI and provides a framework for its ethically grounded and epistemologically reliable integration into future healthcare.

Keyword: artificial intelligence, explainable artificial intelligence, black box, AI ethics, healthcare, medicine, kidney transplantation