

De Onderzoeksgroep  
Artificial Intelligence Lab

nodigt U graag uit op de openbare verdediging van het proefschrift van

## Inês Terrucha

ter behaling van de graad van Doctor in de wetenschappen

Gezamenlijk doctoraat met Universiteit Gent

Titel van het proefschrift:

**Delegatie van beslissingen over belangenconflicten  
aan autonome agenten**

Promotoren:

Prof. dr. ir. Pieter Simoens (UGent)

Prof. dr. Tom Lenaerts (VUB)

De verdediging heeft plaats op

**Maandag 26 augustus 2024 om 16u30 in  
auditorium P Jozef Plateau, Gent**

De verdediging kan via een livestream gevolgd  
worden: [https://teams.microsoft.com/l/meetup-](https://teams.microsoft.com/l/meetup-join/19%3ameeting_YmE3ODdjYzYwMjNTgzZj00YjUwLWE2NTUzZGI4MTIhNDQ2NzQz%40thread.v2/0?content=7b%22Tid%22%3a%22d7811cde-ecf-496c-8f91-a1786241b99c%22%2c%22Oid%22%3a%22c0e01f4e-fa5d-4f70-8f6e-8a9b2c75d2bc%22%2d)

[join/19%3ameeting\\_YmE3ODdjYzYwMjNTgzZj00YjUwLWE2NTUzZGI4MTIhNDQ2NzQz%40thread.v2/0?content=7b%22Tid%22%3a%22d7811cde-ecf-496c-8f91-a1786241b99c%22%2c%22Oid%22%3a%22c0e01f4e-fa5d-4f70-8f6e-8a9b2c75d2bc%22%2d](https://teams.microsoft.com/l/meetup-join/19%3ameeting_YmE3ODdjYzYwMjNTgzZj00YjUwLWE2NTUzZGI4MTIhNDQ2NzQz%40thread.v2/0?content=7b%22Tid%22%3a%22d7811cde-ecf-496c-8f91-a1786241b99c%22%2c%22Oid%22%3a%22c0e01f4e-fa5d-4f70-8f6e-8a9b2c75d2bc%22%2d)

### Samenstelling van de jury

Prof. dr. Hennie De Schepper (UGent, voorzitter)

Prof. dr. Ann Nowé (VUB, secretaris)

Prof. dr. Tony Belpaeme (UGent)

Prof. dr. Fernando Santos (Universiteit van  
Amsterdam, NL)

Prof. dr. Jeremy Pitt (Imperial College London, VK)

### Curriculum vitae

Inês Terrucha schreef zich in als gezamenlijk doctoraatsstudent aan de Universiteit Gent en de Vrije Universiteit Brussel om, onder begeleiding van Prof. Pieter Simoens (UGent) en Prof. Tom Lenaerts (VUB), onderzoek te verrichten. Tijdens haar doctoraats-onderzoek was Inês co-auteur van 4 publicaties in internationale tijdschriften. Wanneer Inês niet op kantoor is, bereidt ze zich voor op een Ironman-triathlon.

### Abstract van het doctoraatsonderzoek

Mensen vertrouwen steeds meer op autonome agenten om hen te helpen bij de meest uiteenlopende activiteiten. Het is echter nog steeds onduidelijk hoe de introductie van autonome agenten in onze strategische beslissingen de collectieve uitkomsten van belangenconflictsituaties zou kunnen beïnvloeden. Deze thesis onderzoekt deze vraag door een interdisciplinaire lens die gedragsproeven en evolutionaire speltheoretische modellen combineert. Uit de experimentele setting blijkt dat degenen die delegeren meer prosociaal gedrag vertonen.

Dit vertaalt zich echter niet noodzakelijkerwijs in hogere succespercentages bij het voorkomen van collectieve rampen. Ervan uitgaande dat dit te wijten zou kunnen zijn aan fouten bij het programmeren van hun agenten, wordt een model voorgesteld waarin delegatie wordt onderscheiden van niet-delegatie door de timing waarop fouten kunnen optreden: voordat het spel begint (delegatie) versus tijdens het spel (geen-delegatie). Dit model maakt het mogelijk de langetermijnimplicaties van delegatie aan onze samenleving te bestuderen, wat leidt tot de visie van hybride mens-AI-teams die samen sociale dilemma's aanpakken.

In het algemeen onthult dit proefschrift dat delegatie aan autonome agenten het potentieel heeft om onze samenleving ten goede te komen, desalniettemin pleiten we voor de gecontroleerde inzet van dergelijke agenten om dit positieve resultaat te garanderen.