

# PROJECT POLIGEN



Lili Takács  
Budapest University of  
Technology and Economics  
RealAI

A STUDY CASE ON AI INCLINATION PREVENTION AND ON PROMOTING FAIRNESS.

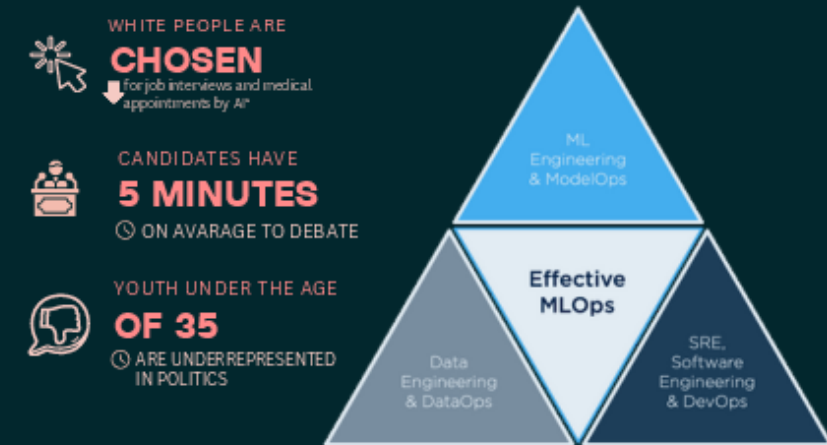
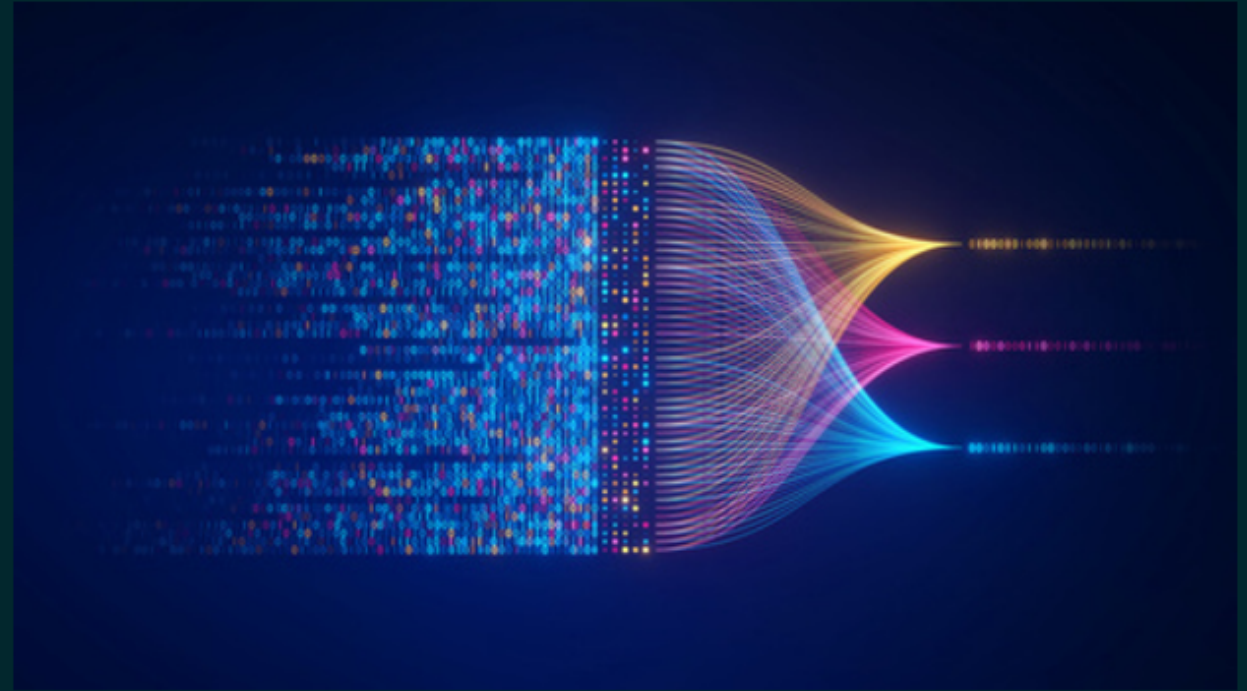
Problems never stop; they merely get exchanged and/or upgraded. [...] "The solution to one problem is merely the creation of the next one."

MARK MANSON



## Introduction

Independent politicians face significant challenges in parliamentary discussions due to their lack of access to party resources, research teams, and collaborative support. Moreover today politics has many problems worldwide. Women, young people (under 35) and people coming from lower-income or ethnical backgrounds are only some of the people who are majorly underrepresented in most legislatures. Therefore in the end the question remains: How to make an AI model impartial, inclusive and as trustworthy as possible in order to be dependable?



## Abstract



If there existed an AI-driven tool that would analyze parliamentary speeches in real time it could help independents bridge the resource gap and prepare more effective responses. This way they could engage in debates on equal footing with party-affiliated representatives and would make the political field more inclusive and representative of previously underrepresented groups of people through giving a more fair chance of representatives with different backgrounds to make their voices heard.

## Literature Review

### How do biases form in AI models?

- Historical bias: AI mirroring prejudices from historical data
- Sample bias: lacking diversity
- Label bias: e.g., only front-facing lion images
- Aggregation bias: merging diverse data
- Confirmation bias: favoring data that aligns with preexisting beliefs.
- Evaluation bias: unrepresentative data



## Conclusion

Most IT professionals have the opinion that eliminating bias from AI is impossible due to inherent societal, historical, and statistical biases in data and algorithms

the emergence of new biases even when others are mitigated.

the best approach might be to use ethical data, detect biases, test with neutral datasets, and continuously refine the model while ensuring diversity in development.

Other methods as using diverse, high quality, neutral data for example from Wikipedia are of high benefits.

Using existing bias reducing tools like AI Governance Tools, Responsible AI Platforms, Data Governance Tools, Machine Learning Operations, Large Language Model Operations, fine-, and prompt-tuning and combining retrieval systems with generative models in innovative new architectures and frameworks are all possibilities for developing the real time fact checking and political argumentation artificial intelligence tool in which this study could be extended upon.

## References

1. Drah, C. (2023) "AI was asked to create images of Black African dogs treating white kids. How'd it go?" NPR. Retrieved at August 22, 2024.

All other resources used by this study can be found under the link below:



## Acknowledgements

The HCAIM (the Human-Centred AI Master's Programme) Project is Co-Financed by the Connecting Europe Facility of the European Union Under Grant NoCEF-TC-2020-1 Digital Skills 2020-EU-IA-0068. This poster was created as part of the Blended Intensive Programme organised under the Erasmus + Programme of the European Union. I would like to express my deepest appreciation to the organizers of this Erasmus+ program. I could not have undertaken this journey without Budapest University of Technology and Economics, RealAI and Simonyi Szakkollégium. I would like to extend my sincere thanks to Tarry Singh and Foke van der Helm who I had the pleasure of collaborating with.

## Research methodology

The main focus of this study is centered around the topic of creating a debiased AI system for political discourse analysis and counterargument generation in order to enhance equity in political debate. To do this it is required to operate independently of established power structures and human cognitive biases. How?

### How to eliminate biases?



- AI Governance Tools
- Responsible AI Platforms
- MLOps
- LLMOps
- Data Governance Tools