

Enchanted Determinism: Power without Responsibility

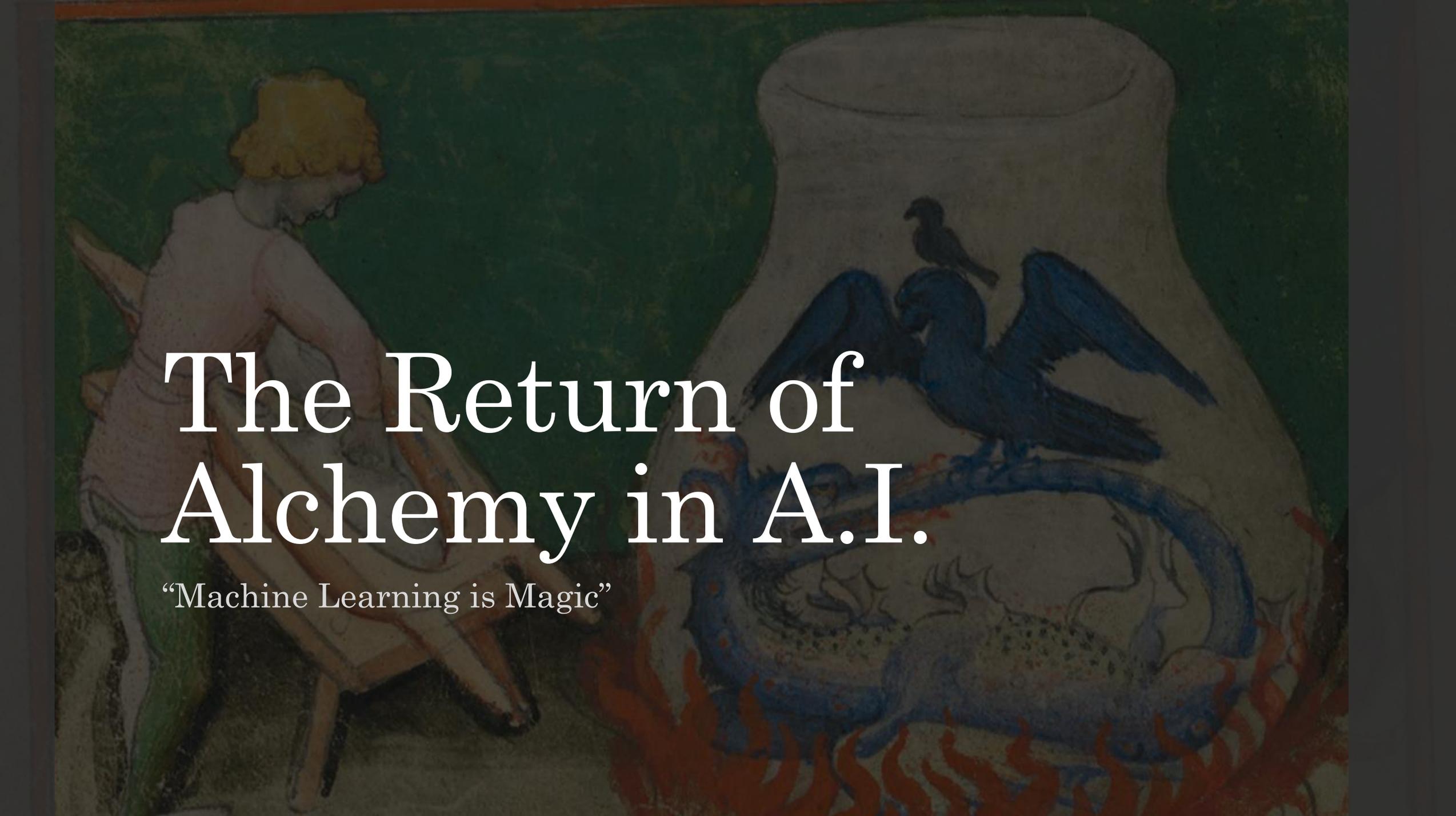
Written by Alexander Campolo and Kate Crawford

Presentation by Alex Johnson, Dominic Reilly, and Alexander Good

Introduction

- Deep Learning is a powerful tool for making predictions and classifications
 - Pattern recognition on a mass-scale
- "Enchanted Determinism" disconnects users from the inner workings of the AI
 - Determinism is mystified by creators and observers alike
- Opaque inner-workings leads to even more enchantment
 - AI cannot explain itself
- Impact on accountability for corporations and developers
 - Assuming accuracy without responsibility





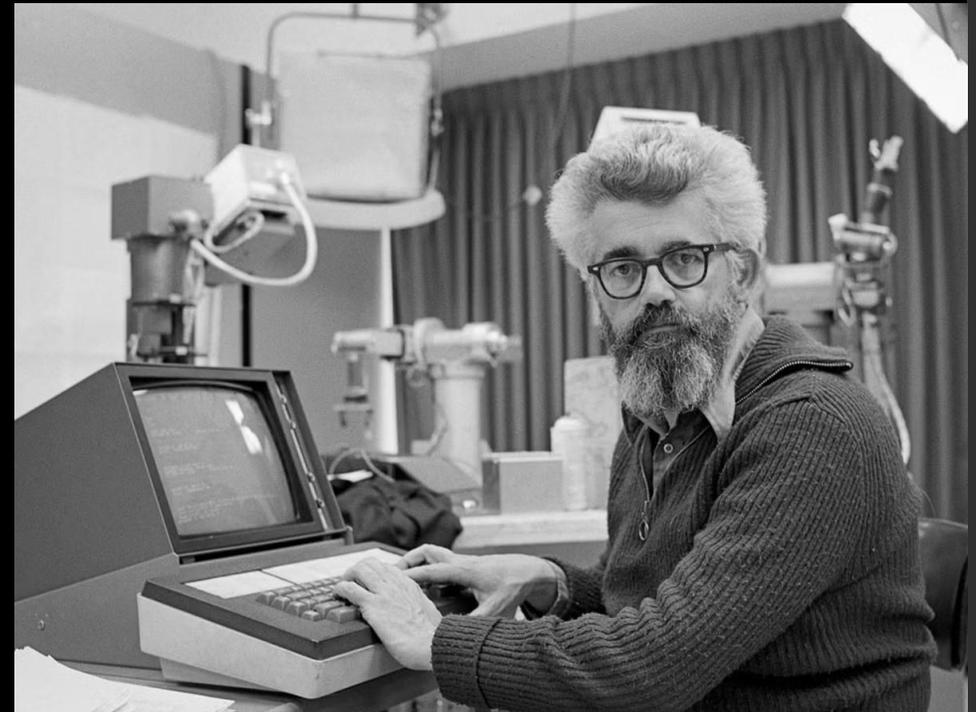
The Return of Alchemy in A.I.

“Machine Learning is Magic”

Are Mind and Machine equal?

- Before 1961: A consensus of separate human tasks and machine tasks
- 1961 Rebuttal by McCarthy: Separation of human and machine tasks is an illusion
- Some tasks were more complex, but will eventually be figured out by machines

John McCarthy, father of AI



No! AI is like Alchemy

- Hubert Dreyfus criticized McCarthy's assertion
- Excessive optimism driven by too simple and problematic metaphors about intelligence
- Compared to Alchemy: early success justified and continued by those favoring adventure over patience
- AI is the “modern alchemy”: prescientific and magical

-57-

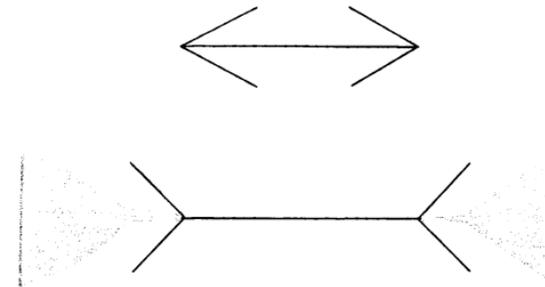


Fig. 1--Muller-Lyer Illusion

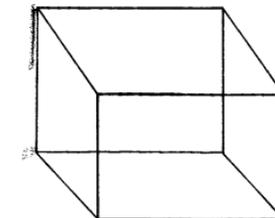


Fig. 2--Necker Cube

Present Day: Still “magic”

- Recent focus on machine learning techniques: AI systems beginning to be deployed internationally
- Deep neural networks can effectively predict images, but the theoretical explanation is missing
- Many researchers continue to ask: Is machine learning science or alchemy?
- Discourse of machine learning “magic” seen all around us



Enchanted Determinism

- A discourse presenting deep learning techniques as magical and outside our current scientific knowledge
- Deterministic by detecting patterns in data and giving access to findings that wouldn't be found otherwise
- E.g., machine learning does a great job! But *why*?
- Veil of “magic” distances AI designers and corporations from ethical and legal liability

Enchantment and Disenchantment

To be or not to be explainable

Magic in Novelty

- New technologies tend to be presented as magical
- Benefit for those looking for a profit
- Magic narrative builds hype, presents machine learning systems as “superhuman”
- Enchanted Determinism operates when the systems “succeed”



TIME FOR
SOMETHING
NEW!

Max Weber's Theory of Disenchantment

- Diagnosis of Western Modernity
- Describes a decline in mystical/religious forces, rise of “rationalization” and “intellectualization”
- Disenchantment: No “magical” forces in play,
- One can master all aspects of something by calculation



Deep learning is enchanted AND disenchanted??

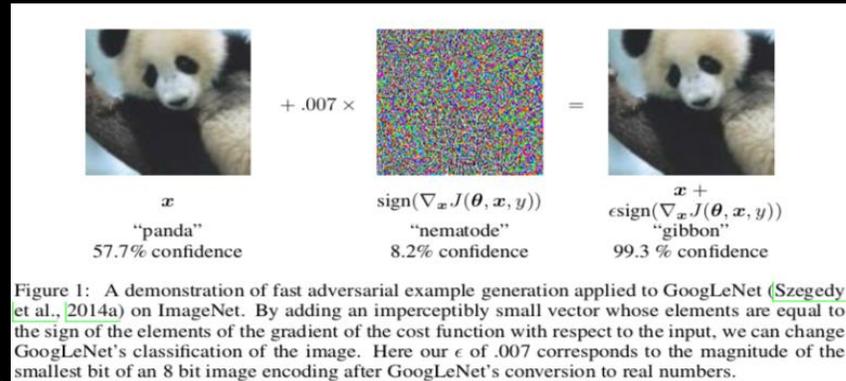
- Disenchanted that DL systems work to master or control new domains through calculation
- Promise of higher accuracy and elimination of human biases
- Enchanted that these systems are discussed in a optimistic discourse describing them as magical and superhuman
- Covers up harmful processes that machine learning systems may create or enforce

Deep Learning's Enchanted Epistemology

- Deep learning models do not utilize human epistemologies, but instead build upon relations found in datasets, and tie these together to make more complex relations
- Deep learning models do not use human a priori assumptions, and will not necessarily understand ideas like eyes or other facial characteristics despite learning what a face is
- There exists a mismatch between mathematical optimization and the demands of human reasoning
- Deep learning's effectiveness has resulted in less cause for improvement in causal understanding available

Producing Optimism: The "Black Art"

- Understanding and technological progress is uncoupled
- Deep learning has far less agency, but shapes our world and can strengthen relationships and hierarchies it finds
- Deep learning can create things seen as "sublime" but does so using simple optimizations, rather than grand intuition or strategy
- It is incredibly easy to fool these relational models in ways humans could not be
- Accuracy does not imply a true understanding of the tasks



Enchantment and Disenchantment: The Worst of Both Worlds?

- AI viewed as genius
- AI trained with statistics does not wield the social or historical context to understand these statistics
 - Over-policing
- Some work is being done to "explain" AI results
 - New, inherently interpretable models
 - Tracing steps of AI to "explain"
- Enchanted determinism leads people to not question AI
- Pattern recognizing AI can assume race, gender, youth, etc are clear-cut defining categories
- These issues are often justified by the quest for "superhuman" accuracy

Conclusion

- AI is seen as both an inerrant, deterministic machine and a mystical, incomprehensible force
- This leads consumers, lawmakers, developers, and corporations to blindly rely on AI tools
- AI is susceptible to human bias as it is trained with human data and recognizes patterns
- If left unchecked, we will continue to blindly rely on AI and justify its mistakes
- We should continue to pursue methods of "explanatory" AI

Discussion Time!