



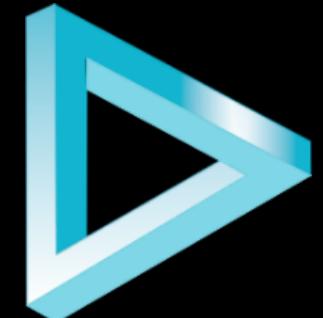
AI v kontextu filozofie

Radek Schuster

schust@kfi.zcu.cz

Katedra filozofie, Filozofická fakulta

Západočeská univerzita v Plzni

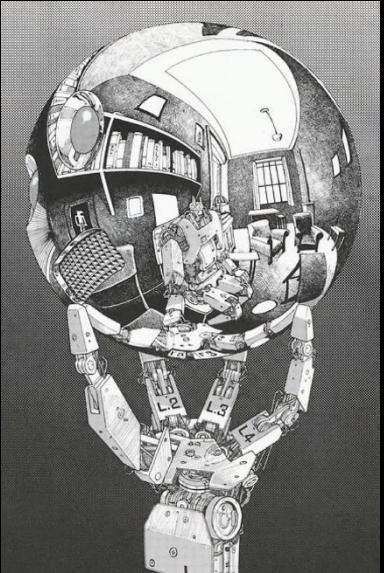


“AI in the context of philosophy”
Stable Diffusion, hyperreal style

AI V KONTEXTU
ÚFAL MFF UK, 4/5/2023

Abstrakt přednášky „AI v kontextu filozofie“

Přednáška bude filozoficky reflektovat abstrakt přednášky s názvem „AI v kontextu filozofie“, který je vygenerovaný AI: „Tato přednáška se bude zabývat vztahem mezi umělou inteligencí (AI) a filozofií. Budeme diskutovat o tom, jak se **vývoj AI** promítá do filozofických otázek, jako je vztah mezi **člověkem a strojem** nebo otázky týkající se **etiky a morálky** využívání AI. Dále se budeme věnovat otázce, zda je možné, aby stroj měl **vědomí** a jak by takový vývoj ovlivnil naše **chápání světa**. Na závěr se zaměříme na to, jakým způsobem bychom měli přistupovat k využívání AI, aby bylo v **souladu** s našimi **hodnotami a cíli** jako **lidské společnosti.**“ (ChatGPT, 1.5.2023)



- 1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie**
2. AI jako *Mathesis Universalis*
3. Etika umělé inteligence jako vyslovování nevyslovitelného
4. Turingův test jako Wittgensteinova jazyková hra
5. Wittgensteinova analýza významu jako distribuční sémantika
6. *Artificial Investigations* jako propojení AI s filozofií



Co je filozofie?

Filozofie v kontextu filozofie:

- láska k moudrosti (nemoudrých), údiv nad nesamozřejmostí všeho, tázání po příčinách
- ontologie, epistemologie, logika, etika, filozofická antropologie
- filozofie jazyka, mysli, vědy a techniky, matematiky, AI
- analytická (anglosaská?) a kontinentální (fenomenologie, hermeneutika, existentialismus)

Filozofie v kontextu „AI v kontextu“:

- rekontextualizace -- „you shall know a word by the company it keeps“
 - zkoumání významu v různých jazykových hrách
 - přepřádání sítí přesvědčení a pojmu
- zaměřování pozornosti -- „attention Is all you need“
 - zkoumání otázek namísto hledání odpovědí
 - odhalování iluzí a osvětlování záhad



Co je AI v kontextu filozofie?

Rekontextualizace:

- proměny inteligentních definic na umělé
- good old-fashioned (GOFAI); general (AGI), strong; narrow, weak; human-level (HLAI); generative
- střídání sub/symbolických paradigmat jako ročních období (AI-winters/summers)
- od kabaly a imitatio (imago) Dei, přes trans/post humanismus, k doomersimu
- AI-Hype (univerzální expert) a Criti-Hype (generování jako zamořování)

Zaměření pozornosti:

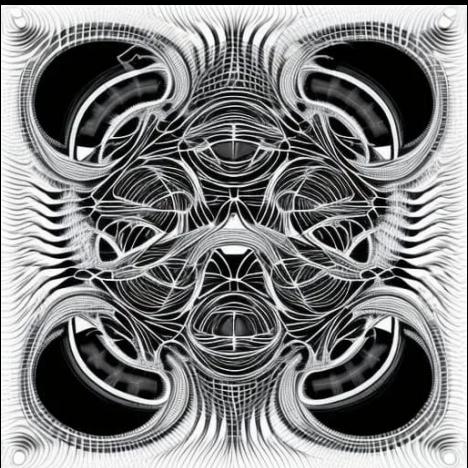
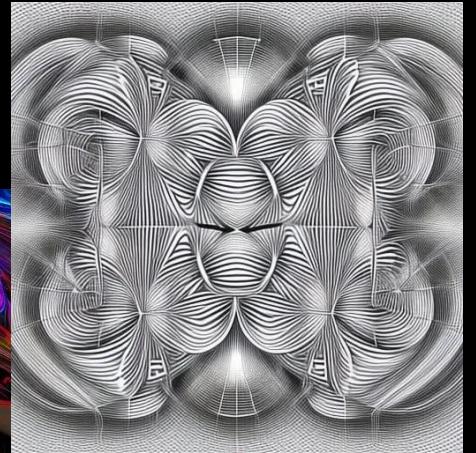
- stochasticí papoušci a robotičtí bullshiteři pod závojem AGI
- AI panic as a business, apocalyptic AI marketing
- iluze čínské komnaty: proměna keňských anotátorů na Auto-GPT

SC

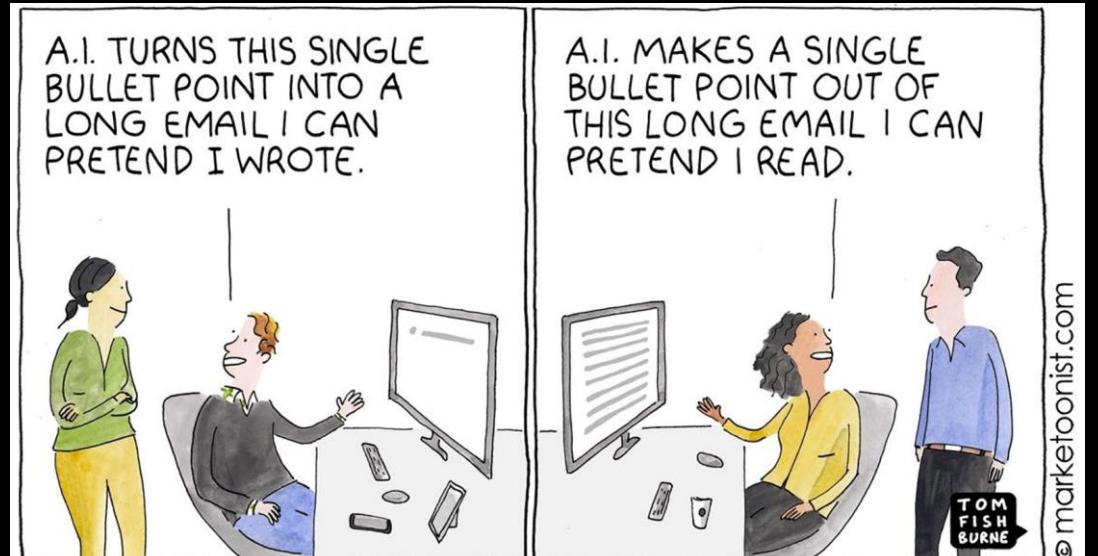
Vygeneruj krátký abstrakt přednášky s názvem "AI v kontextu filozofie"



Tato přednáška se bude zabývat vztahem mezi umělou inteligencí (AI) a filozofií. Budeme diskutovat o tom, jak se vývoj AI promítá do filozofických otázek, jako je vztah mezi člověkem a strojem nebo otázky týkající se etiky a morálky využívání AI. Dále se budeme věnovat otázce, zda je možné, aby stroj měl vědomí a jak by takový vývoj ovlivnil naše chápání světa. Na závěr se zaměříme na to, jakým způsobem bychom měli přistupovat k využívání AI, aby bylo v souladu s našimi hodnotami a cíli jako lidské společnosti.



Stable Diffusion:
"AI in the context of philosophy"



Introducing GPT-4, OpenAI's most advanced system

[Learn about GPT-4](#)[View GPT-4 research](#)

Creating safe AGI that benefits all of humanity

[Learn about OpenAI](#)

Pioneering research on the path to AGI

[Learn about our research](#)

Transforming work and creativity with AI

[Explore our products](#)

Join us in shaping the future of technology

[View careers](#)

ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

[Try ChatGPT ↗](#)[Read about ChatGPT](#)

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic

TIME, Billy Perrigo, January 18, 2023



Dall-E 2. The prompt was:
"A seemingly endless view
of African workers at desks
in front of computer
screens in a printmaking
style."



Dartmouth Summer Research Project on Artificial Intelligence, 1956

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

John McCarthy (Dartmouth College), Marvin Lee Minsky (Harvard University),
Nathaniel Rochester (I.B.M. Corporation), Claude Elwood Shannon (Bell Telephone Laboratories)

(August 31, 1955)

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of **learning** or any other feature of **intelligence** can in principle be so precisely **described** that a **machine** can be made to **simulate** it. An attempt will be made to find how to make machines **use language**, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

The following are some aspects of the artificial intelligence problem:

1. Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

2. How Can a Computer be Programmed to Use a Language

It may be speculated that a large part of human **thought** consists of **manipulating words according to rules** of reasoning and rules of conjecture. From this point of view, forming a generalization consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others. This idea has never been very precisely formulated nor have examples been worked out.

3. Neuron Nets

How can a set of (hypothetical) neurons be arranged so as to form concepts. Considerable theoretical and experimental work has been done on this problem by Uttley, Rashevsky and his group, Farley and Clark, Pitts and McCulloch, Minsky, Rochester and Holland, and others. Partial results have been obtained but the problem needs more theoretical work.

4. Theory of the Size of a Calculation

If we are given a well-defined problem (one for which it is possible to test mechanically whether or not a proposed answer is a valid answer) one way of solving it is to try all possible answers in order. This method is inefficient, and to exclude it one must have some criterion for efficiency of calculation. Some consideration will show that to get a measure of the efficiency of a calculation it is necessary to have on hand a method of measuring the complexity of calculating devices which in turn can be done if one has a theory of the complexity of functions. Some partial results on this problem have been obtained by Shannon, and also by McCarthy.

5. Self-Improvement

Probably a truly intelligent machine will carry out activities which may best be described as self-improvement. Some schemes for doing this have been proposed and are worth further study. It seems likely that this question can be studied abstractly as well.

6. Abstractions

A number of types of ``abstraction'' can be distinctly defined and several others less distinctly. A direct attempt to classify these and to describe machine methods of forming abstractions from sensory and other data would seem worthwhile.

7. Randomness and Creativity

A fairly attractive and yet clearly incomplete conjecture is that the difference between creative thinking and unimaginative competent thinking lies in the injection of a some randomness. The randomness must be guided by intuition to be efficient. In other words, the educated guess or the hunch include controlled randomness in otherwise orderly thinking.

Frank Rosenblatt (1928 – 1971)

Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, 1962

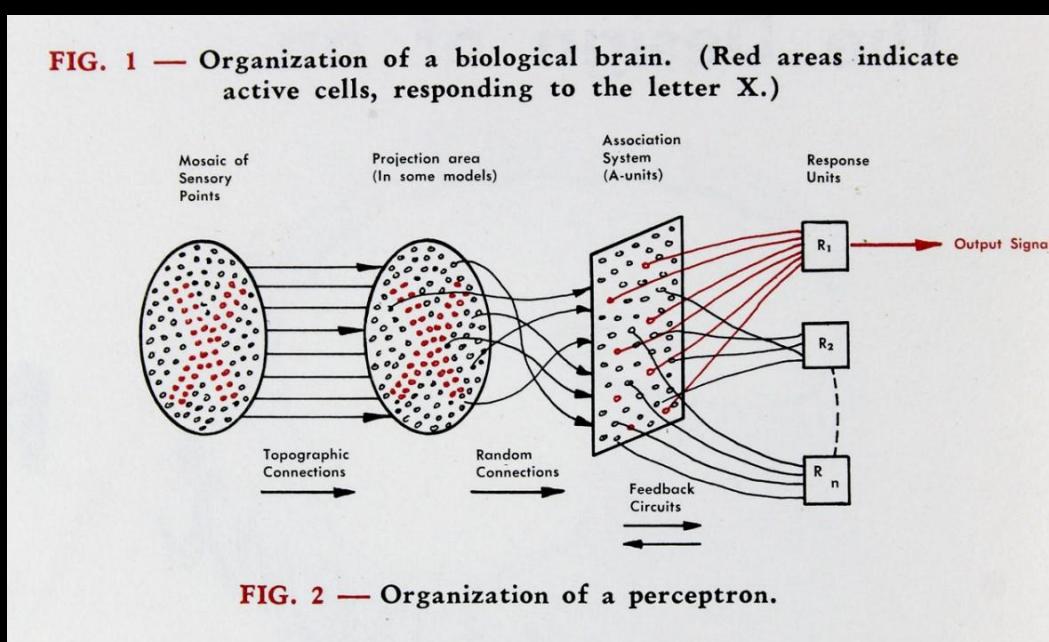
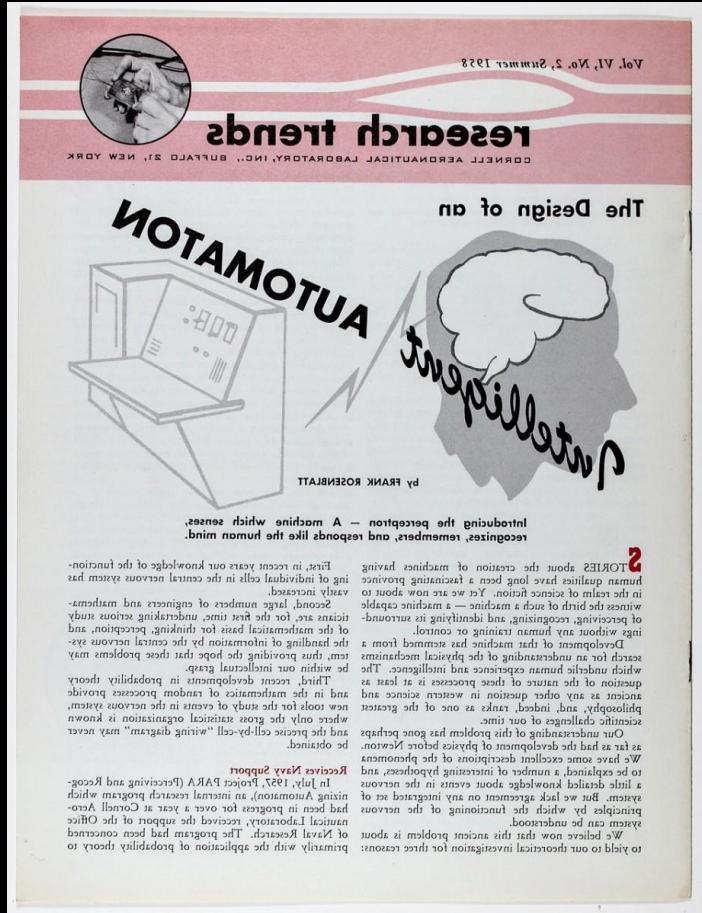


FIG. 2 — Organization of a perceptron.

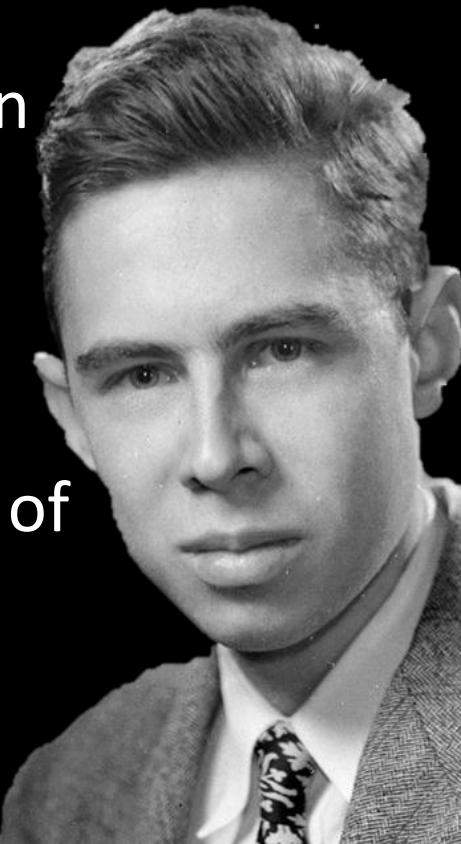


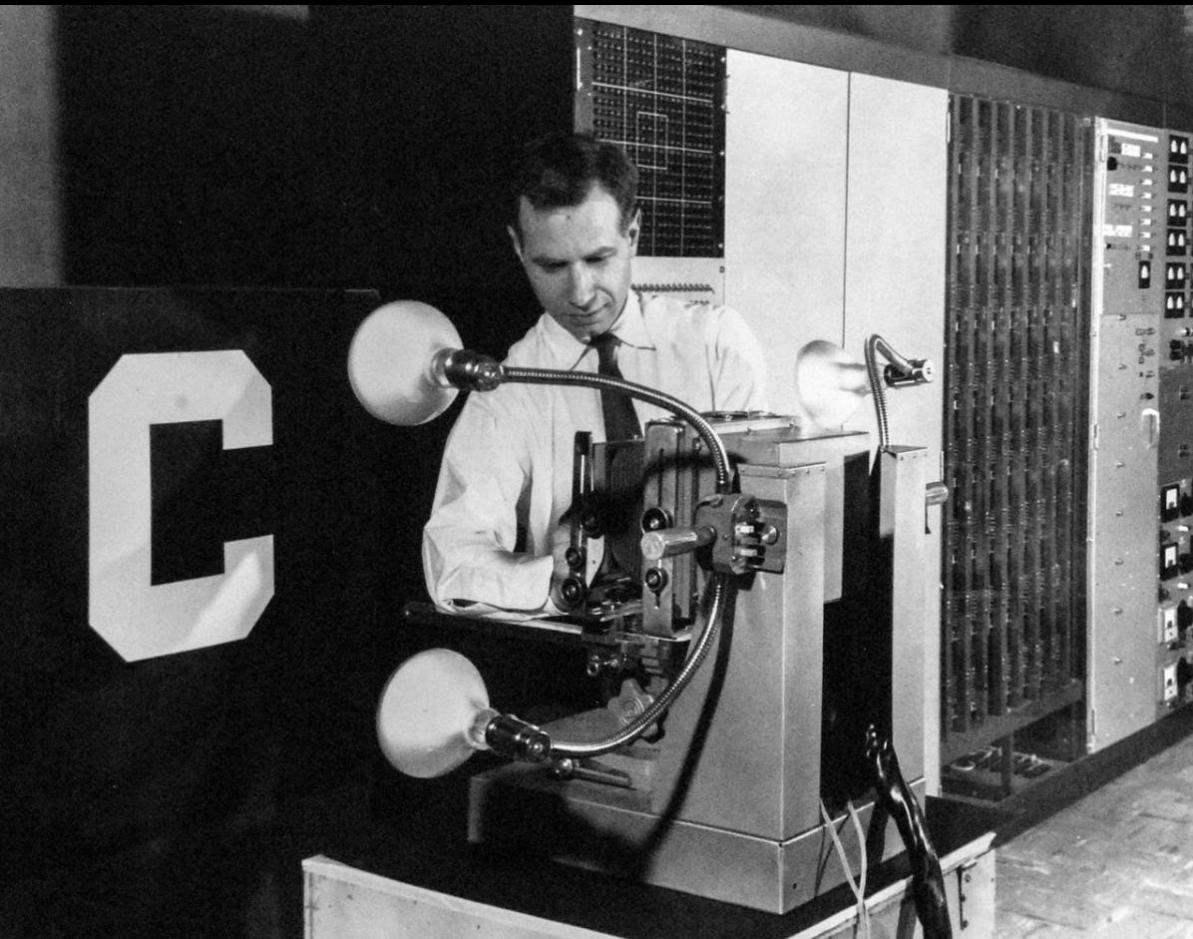
Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain.

Psychological Review, 65(6), 386–408. <https://doi.org/10.1037/h0042519>

Abstract

To answer the questions of how **information** about the physical world is **sensed**, in what form is information remembered, and how does information **retained in memory** influence recognition and behavior, a theory is developed for a hypothetical nervous system called a **perceptron**. The theory serves as a bridge between biophysics and psychology. It is possible to predict learning curves from neurological variables and vice versa. The **quantitative statistical** approach is fruitful in the understanding of the organization of **cognitive systems**.

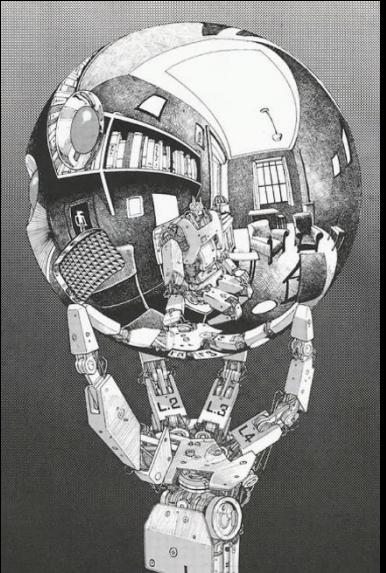




Frank Rosenblatt and the perceptron



Frank Rosenblatt, left, and Charles W. Wightman work on part of the unit that became the first perceptron in December 1958.

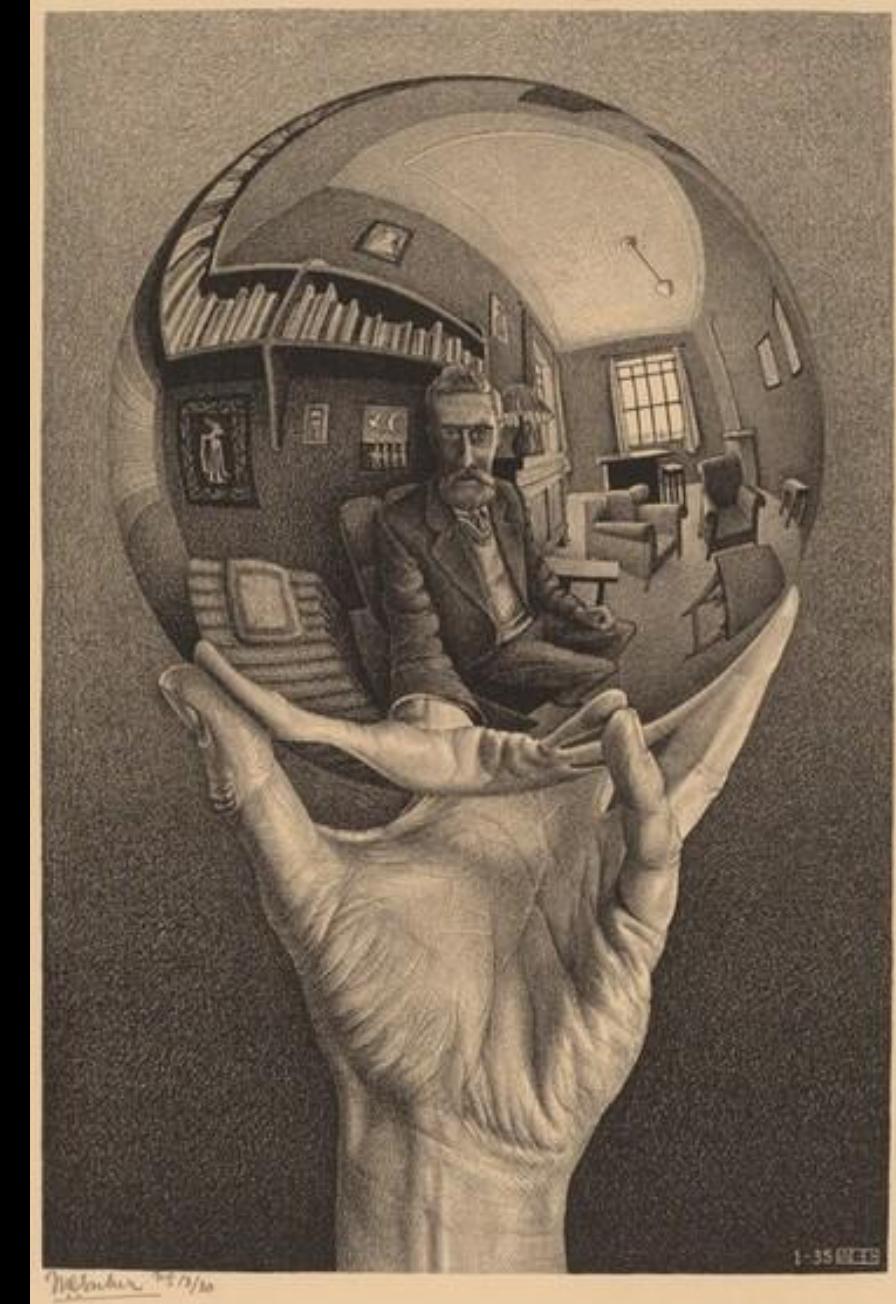
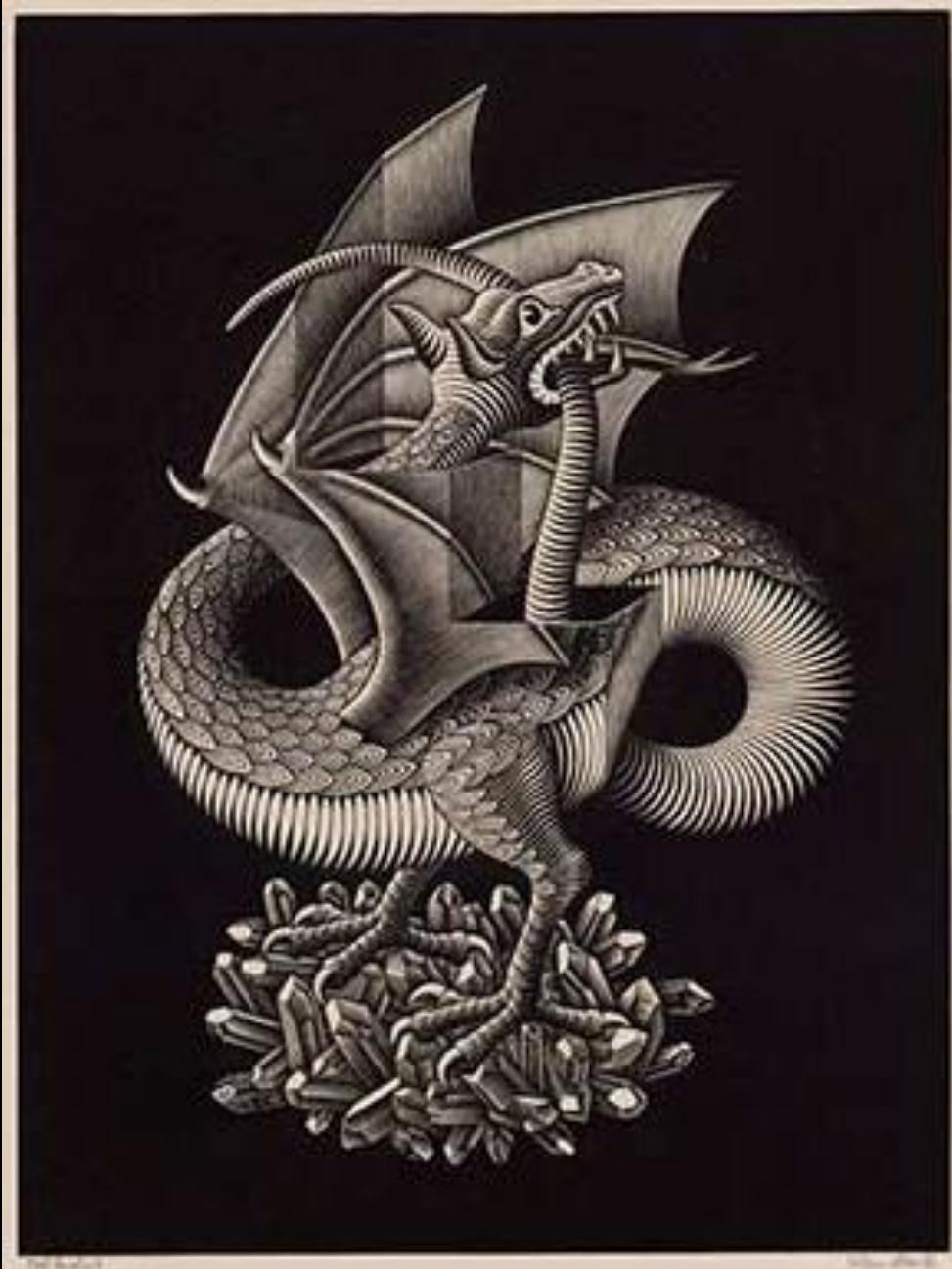


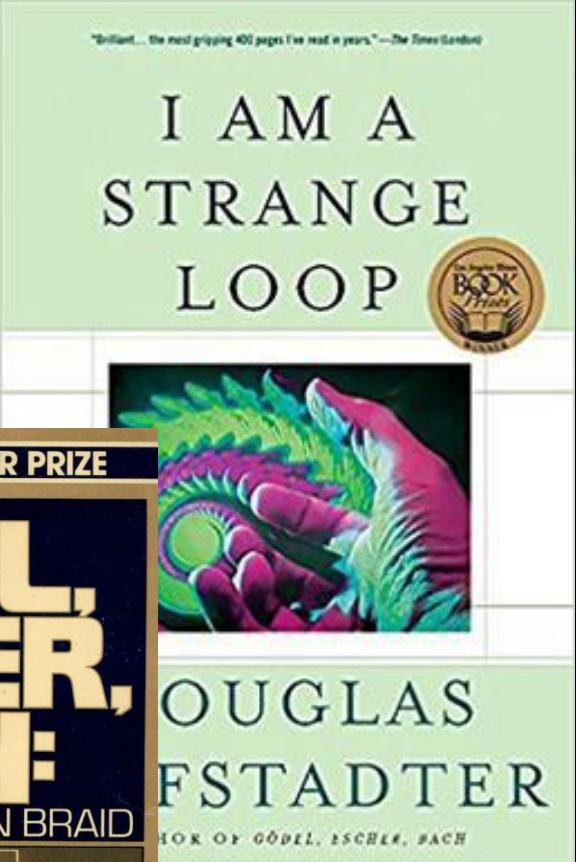
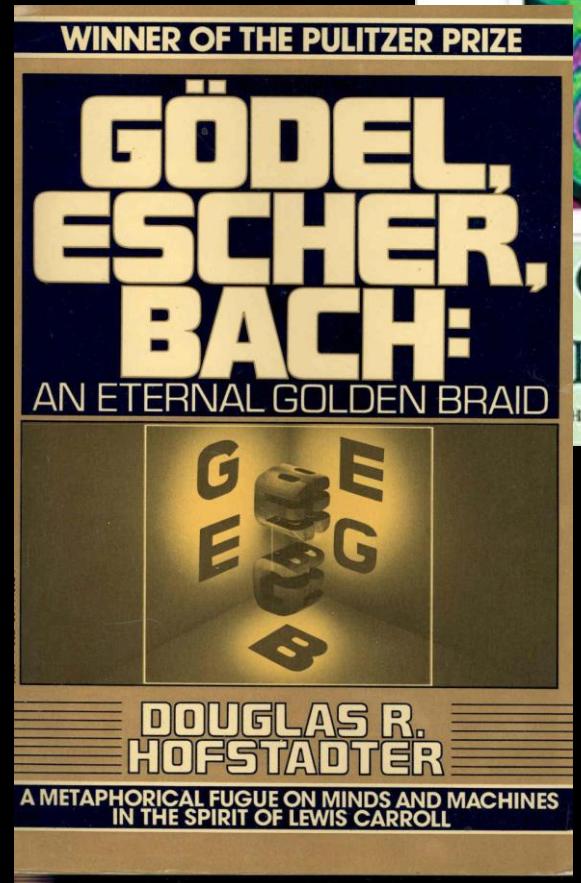
1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie
2. **AI jako *Mathesis Universalis***
3. Etika umělé inteligence jako vyslovování nevyslovitelného
4. Turingův test jako Wittgensteinova jazyková hra
5. Wittgensteinova analýza významu jako distribuční sémantika
6. *Artificial Investigations* jako propojení AI s filozofií



AI jako *Mathesis Universalis*

- AI jako matematizace slov a filozofie jako matematika v pojmech
- Fregův logicismus a Hilbertův formalismus
- Russellův paradox, Gödelova neúplnost a Turingova nerozhodnutelnost
- Jazyk referuje ke světu (prototže?) i k sobě sama
- Jazyk je nejdokonalejší kompresní nástroj díky auto-referenci
- Autoreference jako paradoxní nástroj k oživování strojů
- Rekurze jako zanořování k hlubinám vědomí





Jules Richard (1862-1956)

- Pomocí diagonální metody sestrojil k posloupnosti všech reálných čísel popsatelných konečně mnoha slovy konečný popis čísla, které v této posloupnosti (paradoxně) není:
- Na základě permutací s opakováním nad (26ti) písmeny (francouzské) abecedy lze v určité posloupnosti zapsat všechna (reálná) čísla, která jsou definovatelná konečně mnoha slovy, a utvořit tak spočetně nekonečnou množinu E. Kontradikce vyvstává, neboť lze popsat číslo N, které do množiny E nenáleží:
- „Nechť p je číslice na n-tém desetinném místě n-tého čísla množiny E. Utvořme číslo, jež má 0 před desetinou čárkou a na jeho n-tém desetinném místě je p + 1, pokud p není 8 nebo 9, v opačném případě je tam 1.“
- Pokud by N bylo n-tým číslem množiny E, číslice na jeho n-tém desetinném místě by byla táz, jako číslice na n-tém desetinném místě takového (n-tého) čísla, což ovšem neplatí. Číslo N je definováno řetězcem písmen v závorkách, neboli konečným počtem slov, tudíž by mělo náležet do množiny E, avšak viděli jsme, že nenáleží. Což je kontradikce. (Richard 1905)

George G. Berry (1867–1928)

- Zjednodušená verze Richardova paradoxu:
- $21 =$ „nejmenší číslo, které nelze popsat jedním slovem“
- „nejmenší číslo, které se nedá popsat méně než jedenácti slovy“
Popis má 10 slov: Nejmenší číslo, které nelze popsat méně než 11 slovy, lze popsat 10 slovy.
- „the least integer not nameable in fewer than nineteen syllables“
(Russell 1908, str. 223)

Množinové paradoxy

Bertrand Russell (1872-1970, Cambridge)

Množina všech množin, jež nejsou prvkem sebe sama, je prvkem sebe sama právě tehdy, když není prvkem sebe sama:

$$R = \{ X \mid X \notin X \}; R \in R \leftrightarrow R \notin R$$

(Russell 1903, §78 a §§101-106)

Russellův dopis Fregovi z 16. 6.1902 – zkáza logicismu

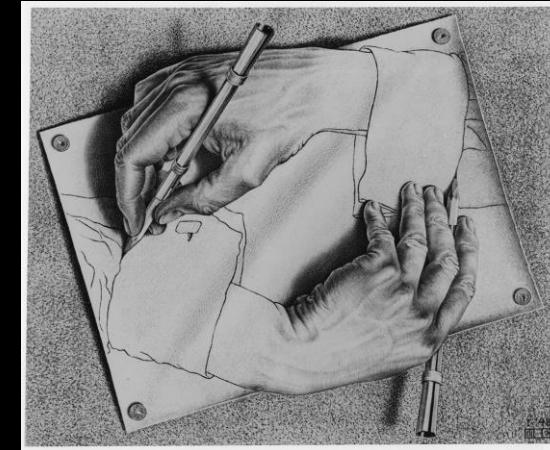
„Budiž w predikát: být predikátem, který nemůže být predikován sám sobě. Může být w predikován sám sobě? Z každé odpovědi vyplývá opak. Proto musíme usoudit, že w není predikát. Právě tak neexistuje žádná třída (jako celek) takových tříd, které samy sobě jakožto celku nenáleží.“

$$w = \{ x \mid \neg x \in x \} \rightarrow (w \in w \leftrightarrow \neg w \in w)$$

Mají paradoxy něco společného?

Struktura těch sémantických paradoxů, jež jsou založeny na **autoreferenci** je shodná se strukturou Russellova paradoxu množin:

$$R \in \{x \mid x \notin x\}: \quad R \in R \leftrightarrow R \notin R$$



**Kurt Grelling: autologický x heterologický:
ne/vztahuje se na sebe**

heterologický $\in \{P \mid P \notin \text{rozsah}(P)\}$:

heterologický $\in \text{rozsah}(\text{heterologický})$

\leftrightarrow

heterologický $\notin \text{rozsah}(\text{heterologický})$

Je Yablův ω -liar paradox bez autoreference?

“...paradoxes like the Liar are possible in the complete **absence of self-reference**. Imagine an infinite sequence of sentences S_1, S_2, S_3, \dots , each to the effect that every subsequent sentence is untrue:

(S1) for all $k > 1$, S_k is untrue

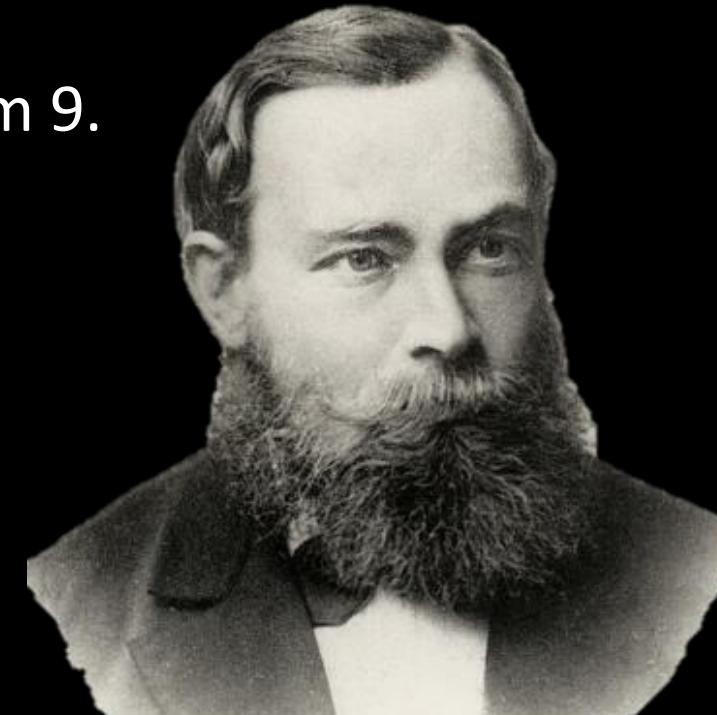
(S2) for all $k > 2$, S_k is untrue

(S3) for all $k > 3$, S_k is untrue

Suppose for contradiction that some S_n is true. Given what S_n says, for all $k > n$, S_k is untrue. Therefore (a) S_{n+1} is untrue, and (b) for all $k > n+1$, S_k is untrue. By (b), what S_{n+1} says is in fact the case, whence contrary to (a) S_{n+1} is true! So every sentence S_n in the sequence is untrue. But then the sentences subsequent to any given S_n are all untrue, whence S_n is true after all! I conclude that **self-reference is neither necessary nor sufficient** for Liar-like paradox.” Stephen Yablo (1993, 252)

Friedrich Ludwig Gottlob Frege (1848 - 1925, Jena)

- *Begriffschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens.* (Pojmové písmo.) Halle, 1879.
- *Die Grundlagen der Aritmetik.* (Základy aritmetiky.) Breslau, 1884.
- *Die Grundgesetze der Aritmetik.* (Základní zákony aritmetiky.) Jena, I. díl 1893, II. díl 1903.
- Funktion und Begriff: Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medizin und Naturwissenschaft. Jena, 1891.
- Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik.* (1892) 25-50.
- Über Begriff und Gegenstand. *Vierteljahrsschrift für wissenschaftliche Philosophie XVI.* (1892) 192-205.



BEGRIFFSSCHRIFT,

EINE DER ARITHMETISCHEN NACHGEBILDETE

FORMELSPRACHE

DES REINEN DENKENS.

VON

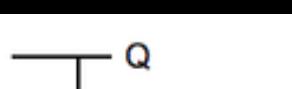
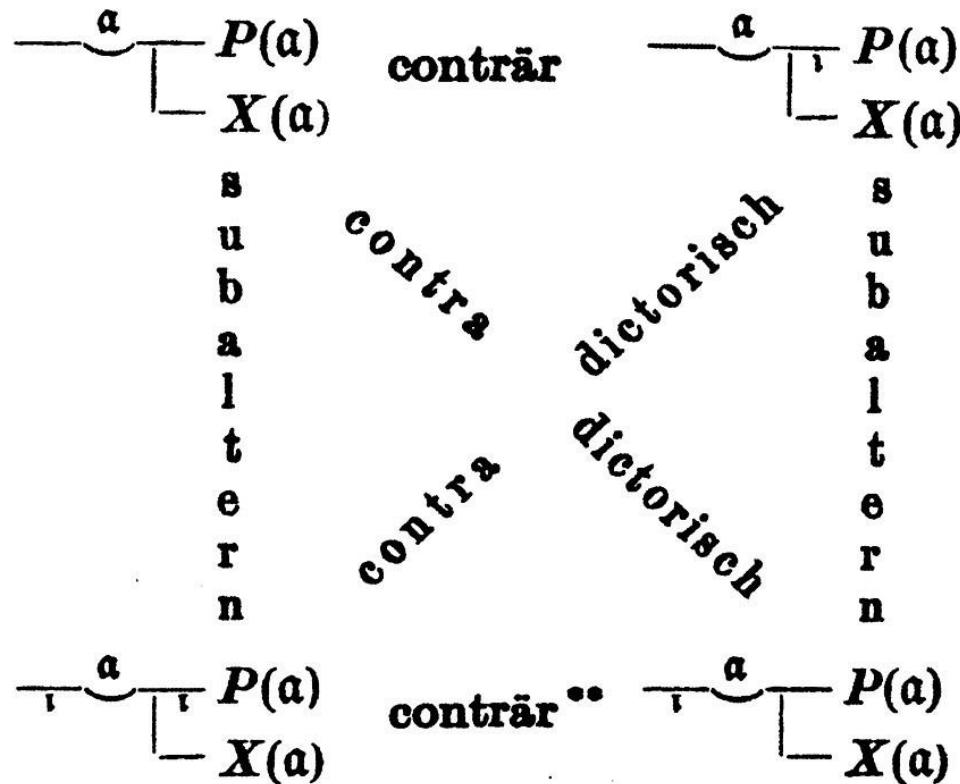
DR. GOTTLÖB FREGE.

PRIVATDOCENTEN DER MATHEMATIK AN DER UNIVERSITÄT JENA.

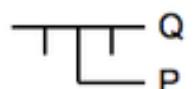
HALLE A/S.

VERLAG VON LOUIS NEBERT.

1879.



b) Verneinung: $\neg P$

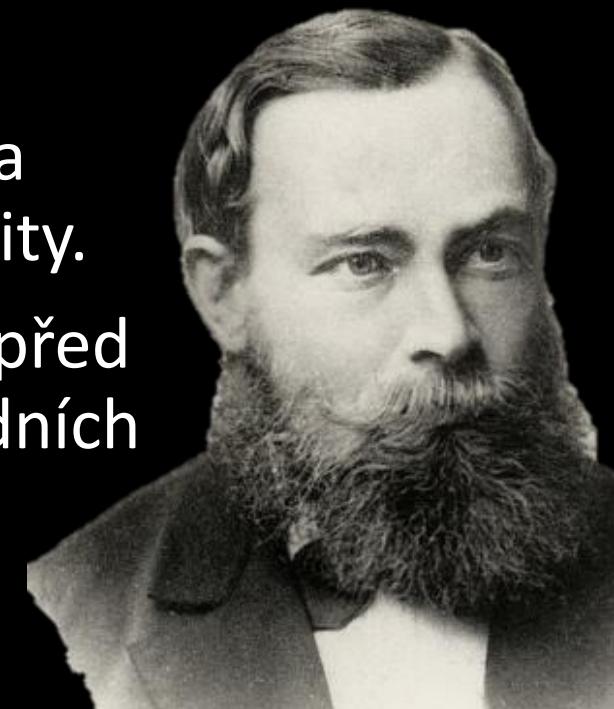


d) Disjunktion: $P \vee Q$

$\overline{x} \vdash F(x)$	$\forall x(G(x) \rightarrow F(x))$
$\vdash \overline{x} \vdash F(x)$	$\exists x(G(x) \wedge F(x))$
$\overline{x} \vdash \overline{F(x)}$	$\neg \forall x(G(x) \rightarrow \neg F(x))$
$\overline{x} \vdash F(x)$	$\neg \exists x(G(x) \wedge F(x))$
$\vdash \overline{x} \vdash F(x)$	$\forall x(G(x) \rightarrow \neg F(x))$
$\vdash \overline{x} \vdash \overline{F(x)}$	$\exists x(G(x) \wedge \neg F(x))$
$\vdash \overline{x} \vdash F(x)$	$\neg \forall x(G(x) \rightarrow F(x))$

Projekt logicismu

- Číslo je chápáno jako množina všech množin se stejným počtem prvků, přičemž prvky množiny jsou dohromady svázány prostřednictvím jazyka - predikátu (pojmu).
- Frege množinu $\{ x \mid F(x)\}$ nazývá „průběh hodnot funkce F“ (rozsah pojmu F).
- Matematická tvrzení o číslech je tak možné formulovat a dokazovat skrze predikáty, kvantifikátory a princip identity.
- Russell na svůj objev paradoxu Frega upozorňuje těsně před vydáním II. dílu *Grundgesetze* a projekt formulací základních zákonů aritmetiky pomocí logiky se hroutí.



David Hilbert (1862 – 1943, Göttingen) Program formalizace (axiomatizace) matematiky (1920)

Grundlagen der Geometrie (1899):

- Motto z Kantovy KČR (č. 2, od.2): „Veškeré lidské poznání začíná intuicemi, pak se posouvá k pojmem a končí u idejí.“
- Pokus vybrat pro geometrii jednoduchý a úplný soubor nezávislých axiomů a odvodit z nich nejdůležitější geometrické teorémy tak, aby co nejjasněji vynikl význam různých (5ti) skupin axiomů a rozsah závěrů, které lze z jednotlivých axiomů odvodit.
- Metageometrie: nejen vystavět geometrii na pevných logických základech, ale prozkoumat i geometrickou axiomatiku, logickou konstrukci.



David Hilbert (1862 – 1943, Göttingen)

Program formalizace (axiomatizace) matematiky (1920)

- Bezesporná množina axiomů určuje formální systém.
Bezespornost jako záruka existence.
- „Ve všech geometrických větách musí být možné nahradit slova bod, přímka, rovina slovy stůl, židle, hrnek.“
- Axiomy jsou definice pojmů, resp. jasné výroky, v nichž jsou zachyceny relace pojmů. Protože pojem lze logicky zachytit pouze prostřednictvím relace k jinému pojmu.
- „Jestliže nejsou arbitrárně stanovené axiomy vzájemně ve sporu se svými důsledky, pak jsou pravdivé a věci definované těmito axiomy existují. To je pro mě kritérium pravdy a existence.“ (Koresp. s Fregem)



Program formalizace (axiomatizace) matematiky (1920)

- vystavět mat. na pevných a úplných logických základech
- veškerá mat. by byla odvoditelná ze správně vybraného konečného systému axiomů
- tento systém by byl dokazatelně bezesporný
- axiomy jsou výroky definující pojmy zachycením jejich vztahů
- pojem lze logicky zachytit pouze vztahením k jinému pojmu



Program formalizace (axiomatizace) matematiky (1920)

- Filozofická motivace: odpor k principiální omezenosti vědeckého poznání (agnosticismu) -- **ignorabimus**.
- Přednáška v Königsbergu (8. 9. 1930): „**Wir müssen wissen — wir werden wissen.**“ (Musíme vědět – budeme vědět.)
- Emil Heinrich du Bois-Reymond, německý fyziolog (elektrofyziologie, nervová odezva). Projev před pruskou akademií věd (1880): „**ignoramus et ignorabimus**“ – „**Wir wissen es nicht und wir werden es niemals wissen.**“ (Nevíme a vědět nebudeme.)
- Pojednání (1872): *Über die Grenzen des Naturerkennens* (O mezích poznání přírody).



Kurt Friedrich Gödel (1906 – 1978, Brno, Wien, Princeton)

Věty o neúplnosti aritmetiky



- Větami o neúplnosti (1930/1) dokázal principiální neuskutečnitelnost programu formalizace.
- Vyplývá z nich, že existují pravdivá avšak nedokazatelná tvrzení o aritmetice přirozených čísel.
- Prostřednictvím aritmetiky nelze dokázat bezespornost a úplnost aritmetiky.
- V žádné teorii, která by popisovala všechny matematické pravdy, nelze dokázat její vlastní bezespornost.
- Žádný bezesporný formální systém, který je natolik komplexní, že zahrnuje alespoň aritmetiku, nemůže dokazat svou úplnost prostřednictvím vlastních axiomů.

Věty o neúplnosti aritmetiky



- Neúplnost aritmetiky (1. větu) oznámil na konferenci v Königsbergu (7. 10. 1930), článek s důkazem a důsledkem (2. větu) publikoval (na začátku) 1931.
- Gódel v Königsbergu (1930): „(Za předpokladu bezespornosti klasické matematiky) lze dát příklady tvrzení (Goldbachova hypotéza, Fermatova věta), která jsou pravdivá, ale nedokazatelná v dané formální teorii klasické matematiky. Pokud by tudíž někdo přidal k axiomům teorie negaci takového tvrzení, získal by bezesporný systém, ve kterém by bylo dokazatelné nepravdivé tvrzení.“

Věty o neúplnosti aritmetiky

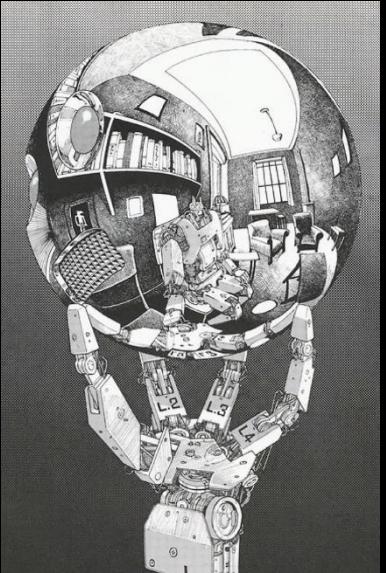


- I. věta: Rekurzivně axiomatizovaná bezesporná teorie zahrnující aritmetiku přirozených čísel je neúplná, neboť v ní vždy existuje nějaké nerozhodnutelné tvrzení, které nelze dokázat ani vyvrátit.
- II. věta: V žádná takové dostatečně silné bezesporné teorii nelze dokázat její bezespornost.
- Konstrukce věty, která tvrdí svojí vlastní nedokazatelnost a zároveň je pravdivá.
- Meta-matematická tvrzení o formálním systému lze vyjádřit v samotném tomto systému.

Hilbertův rozhodovací problém (Entscheidungsproblem)

David Hilbert a Wilhelm Ackermann, *Grundzüge der theoretischen Logik* (1928)

- Existuje efektivní postup (algoritmus), který na základě dané množiny axiomů rozhodne, zda je či není daná matematická věta z axiómů dokazatelná? Resp. existuje postup, který dokáže rozhodnout, zda je daná věta v daném formálním jazyce pravdivá nebo nepravdivá?
- Existuje algoritmus, který o libovolné dobře utvořené formuli určí, zda je to teorém, tedy logicky pravdivá formule?
- Alonzo Church (1936) a Alan Turing (1937) dokázali, že takový algoritmus neexistuje.



1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie
2. AI jako *Mathesis Universalis*
3. **Etika umělé inteligence jako vyslovování nevyslovitelného**
4. Turingův test jako Wittgensteinova jazyková hra
5. Wittgensteinova analýza významu jako distribuční sémantika
6. *Artificial Investigations* jako propojení AI s filozofií



ETHICS OF ARTIFICIAL INTELLIGENCE



SpringerBriefs in Research and Innovation Governance
Bernd Carsten Stahl · Doris Schroeder · Rowena Rodrigues



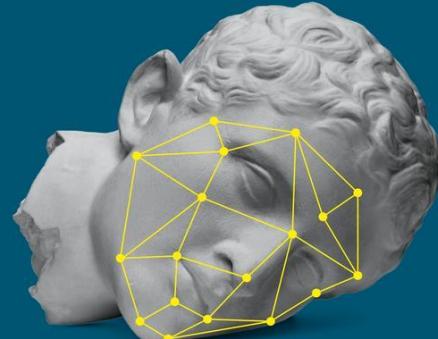
Ethics of Artificial Intelligence
Case Studies and Options for Addressing Ethical Challenges

OPEN ACCESS

Springer

Human Compatible

ARTIFICIAL INTELLIGENCE
AND THE
PROBLEM OF CONTROL



Stuart Russell

nature
machine intelligence

PERSPECTIVE

<https://doi.org/10.1038/s42256-019-0088-2>

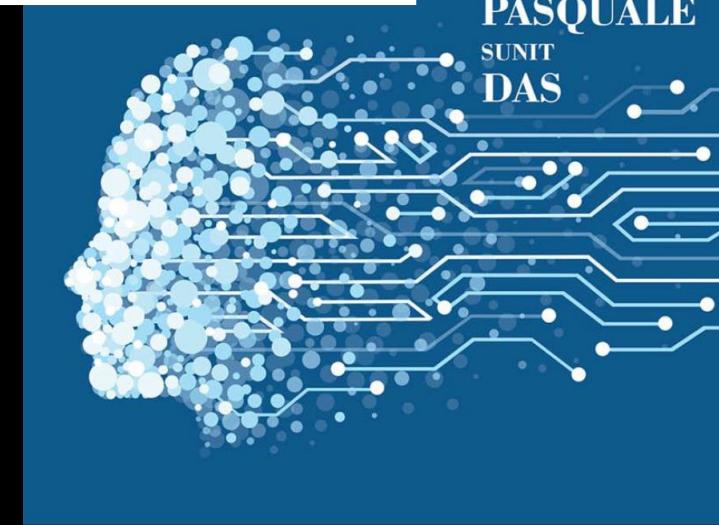
The global landscape of AI ethics guidelines

Anna Jobin, Marcello Ienca and Effy Vayena*

In the past five years, private companies, research institutions and public sector organizations have issued principles and guidelines for ethical artificial intelligence (AI). However, despite an apparent agreement that AI should be 'ethical', there is debate about both what constitutes 'ethical AI' and which ethical requirements, technical standards and best practices are needed for its realization. To investigate whether a global agreement on these questions is emerging, we mapped and analysed the current corpus of principles and guidelines on ethical AI. Our results reveal a global convergence emerging around five ethical principles (transparency, justice and fairness, non-maleficence, responsibility and privacy), with substantive divergence in relation to how these principles are interpreted, why they are deemed important, what issue, domain or actors they pertain to, and how they should be implemented. Our findings highlight the importance of integrating guideline-development efforts with substantive ethical analysis and adequate implementation strategies.

michael Kearns + aaron Roth

al+algorithm
the
/the
ethical
1+algorithm/
ithin/the/*et



≡ The Oxford Handbook of
ETHICS OF AI

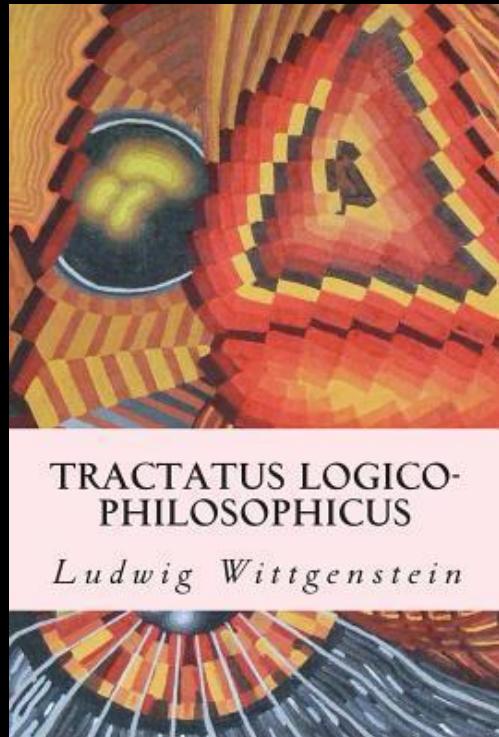
EDITED BY

MARKUS D.
DUBBER
FRANK
PASQUALE
SUNIT
DAS

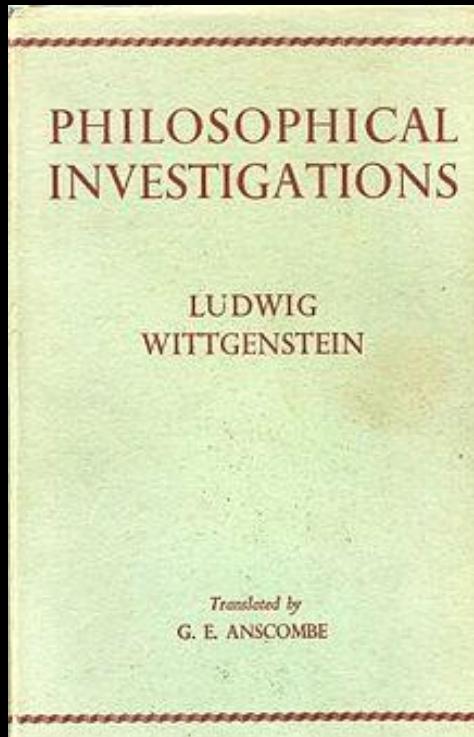
Ludwig Wittgenstein

(Wien 1889 – Cambridge 1951)

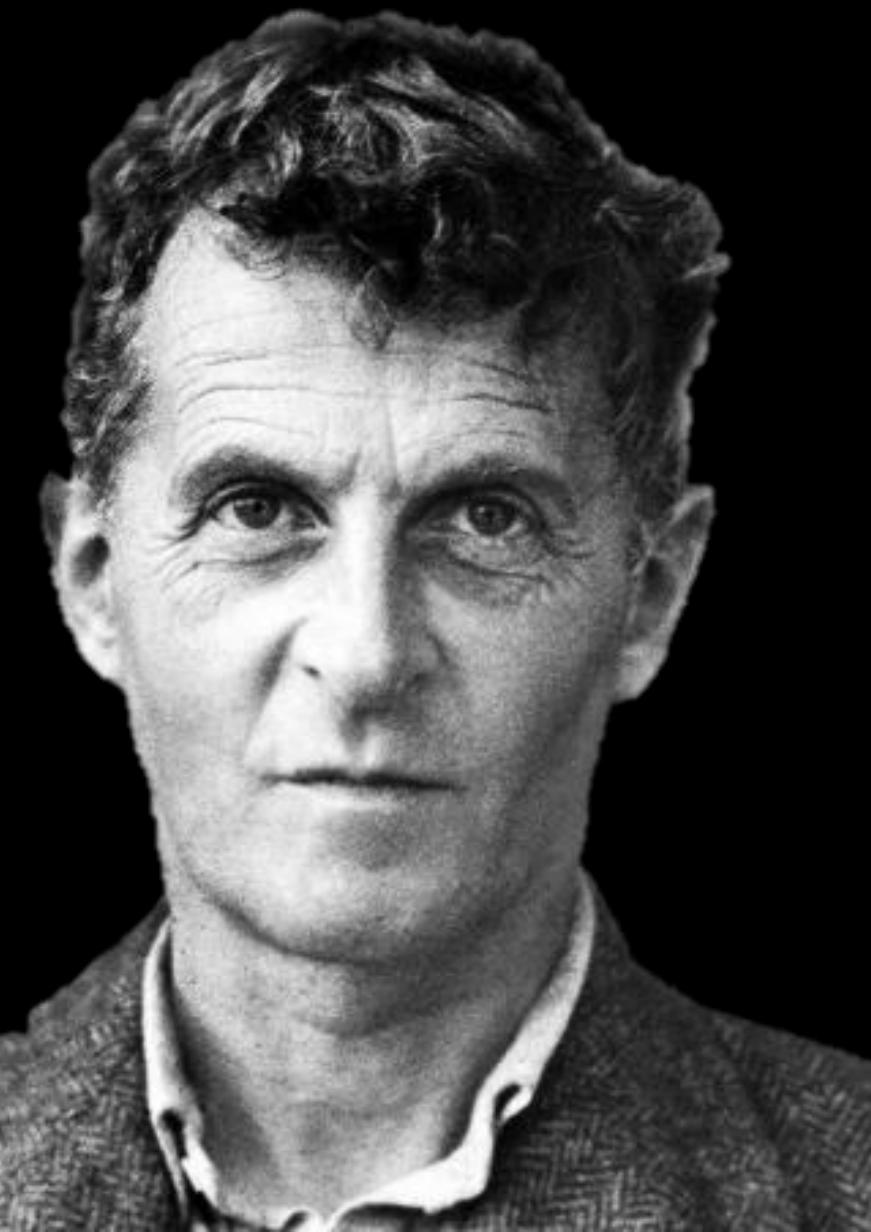
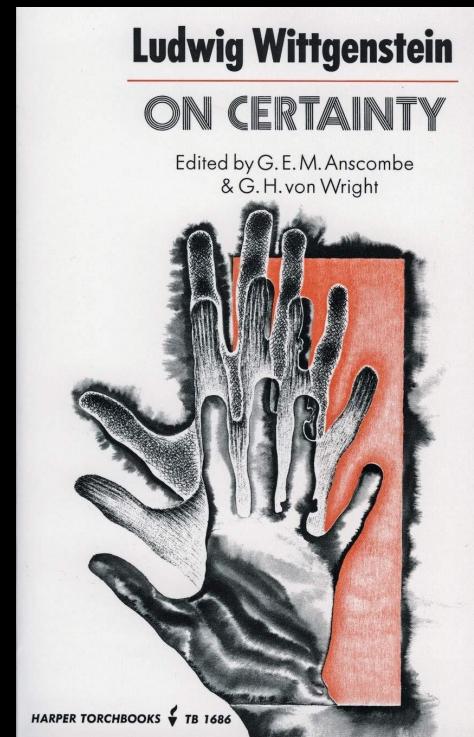
TLP (1921)



PI (1953)

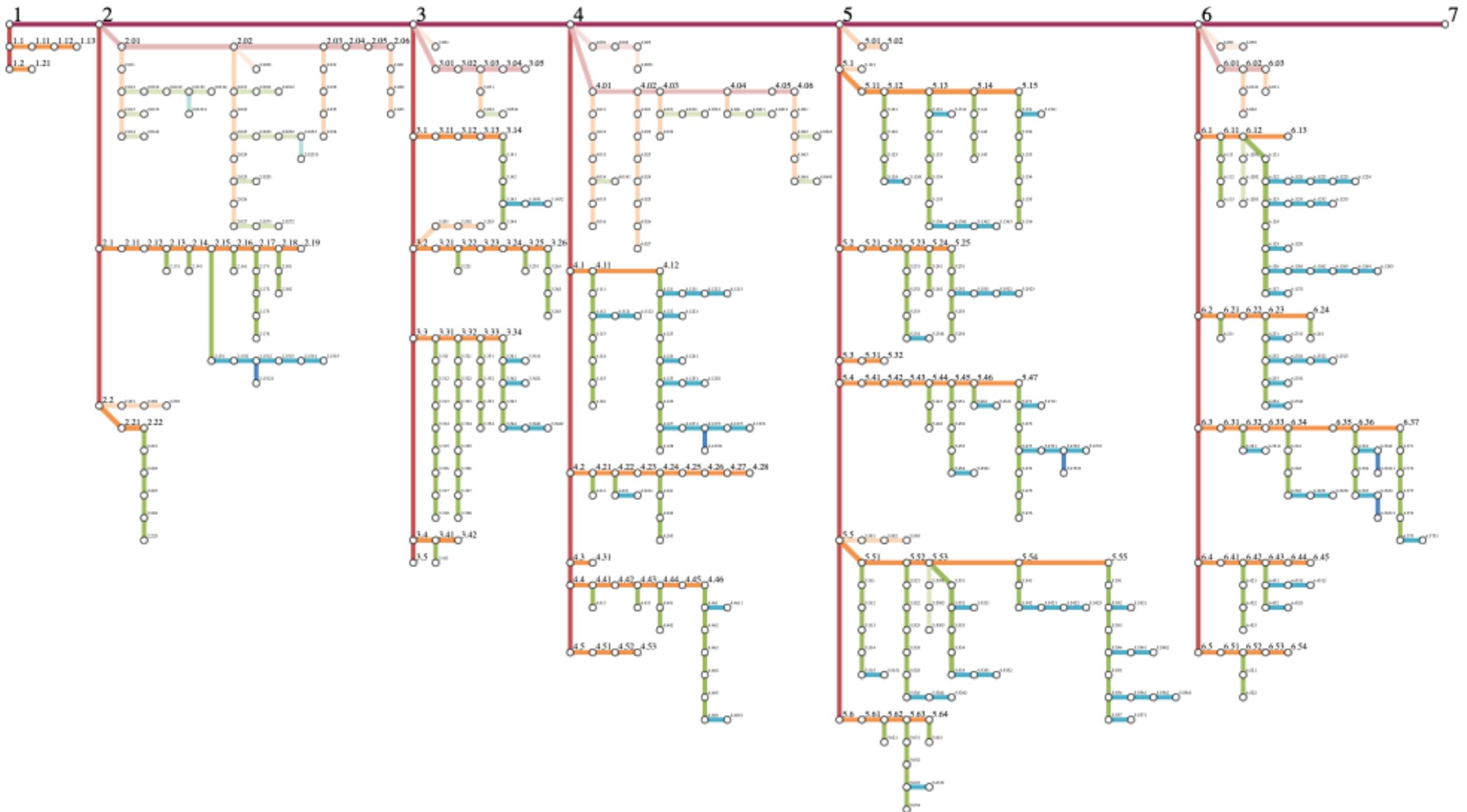


OC (1969)



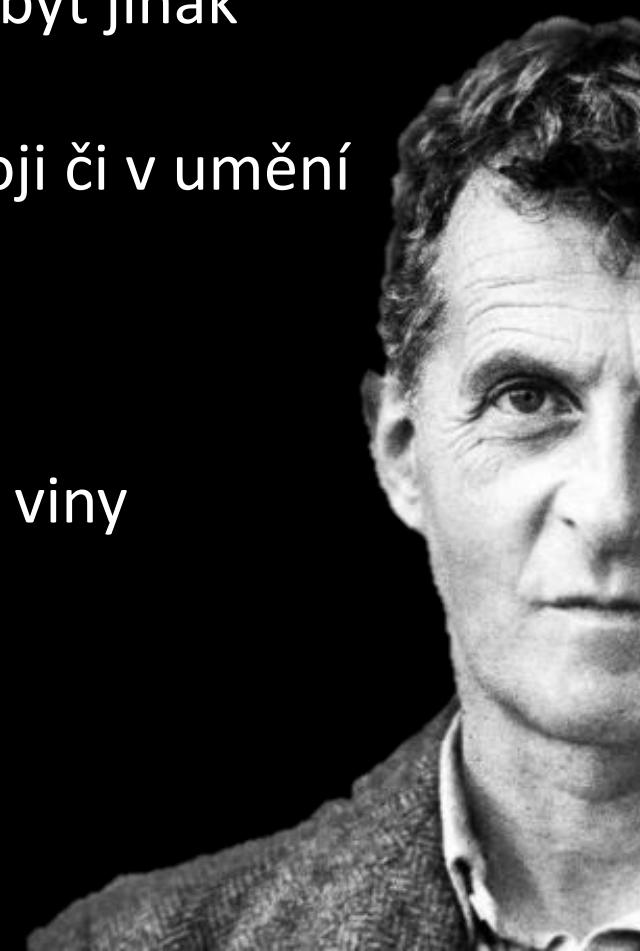
Tractatus Logico-Philosophicus

1. Die Welt ist alles, was der Fall ist.
2. Was der Fall ist, die Tatsache, ist das Bestehen von Sachverhalten.
3. Das logische Bild der Tatsache ist der Gedanke.
4. Der Gedanke ist der sinnvolle Satz.
5. Der Satz ist eine Wahrheitsfunktion der Elementarsätze.
(Elementarsatz ist eine Wahrheitsfunktion seiner selbst.)
6. Die allgemeine Form der Wahrheitsfunktion ist: [, , N()]. Dies ist die allgemeine Form des Satzes.
7. Wovon man nicht sprechen kann, darüber muß man schweigen.



Wittgensteinova etika: TLP a LoE

- Etika a logika: morální povinnost usilovat o logickou jasnost
- Transcendentální: podmínka světa, snaha vyjádřit co nemůže být jinak
- Transcendentní: hodnoty nejsou ve světě, svět není dobrý/zlý
- Neukazuje se na smysluplných větách, nýbrž na jednání, postoji či v umění
- Dobro/zlo se netýká faktů, ale postoji ke světu jako celku
- Dobrá vůle a šťastné oko: svět ne/šťastného je jiný
- Relativní/triviální a absolutní/etické soudy
- Evokace 3 etických prožitků: úžas nad světem, pocit bezpečí a viny
- Faktuální věta nemůže (být) implikovat morální soud
- Logická nezávislost světa a etické vůle



The European Commission's High-Level Expert Group on AI (AI HLEG)

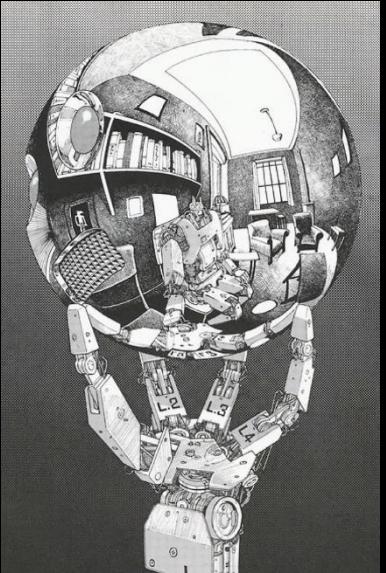
7 key requirements that AI systems should meet in order to be deemed **trustworthy**.

1. **Human agency and oversight:** AI systems should empower human beings, allowing them to make informed decisions and fostering their fundamental rights. At the same time, proper oversight mechanisms need to be ensured, which can be achieved through human-in-the-loop, human-on-the-loop, and human-in-command approaches
2. **Technical robustness and safety:** AI systems need to be resilient and secure. They need to be safe, ensuring a fall back plan in case something goes wrong, as well as being accurate, reliable and reproducible. That is the only way to ensure that also unintentional harm can be minimized and prevented.
3. **Privacy and data governance:** besides ensuring full respect for privacy and data protection, adequate data governance mechanisms must also be ensured, taking into account the quality and integrity of the data, and ensuring legitimised access to data.
4. **Transparency:** the data, system and AI business models should be transparent. Traceability mechanisms can help achieving this. Moreover, AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned. Humans need to be aware that they are interacting with an AI system, and must be informed of the system's capabilities and limitations.
5. **Diversity, non-discrimination and fairness:** Unfair bias must be avoided, as it could have multiple negative implications, from the marginalization of vulnerable groups, to the exacerbation of prejudice and discrimination. Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire life circle.
6. **Societal and environmental well-being:** AI systems should benefit all human beings, including future generations. It must hence be ensured that they are sustainable and environmentally friendly. Moreover, they should take into account the environment, including other living beings, and their social and societal impact should be carefully considered.
7. **Accountability:** Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data and design processes plays a key role therein, especially in critical applications. Moreover, adequate and accessible redress should be ensured.

The Future of Life Institute: Pause Giant AI Experiments: An Open Letter



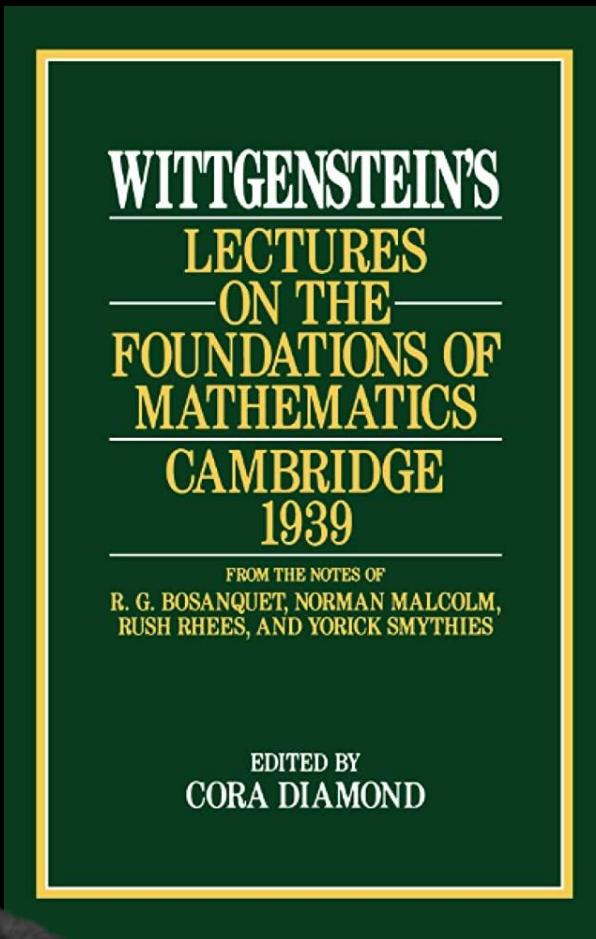
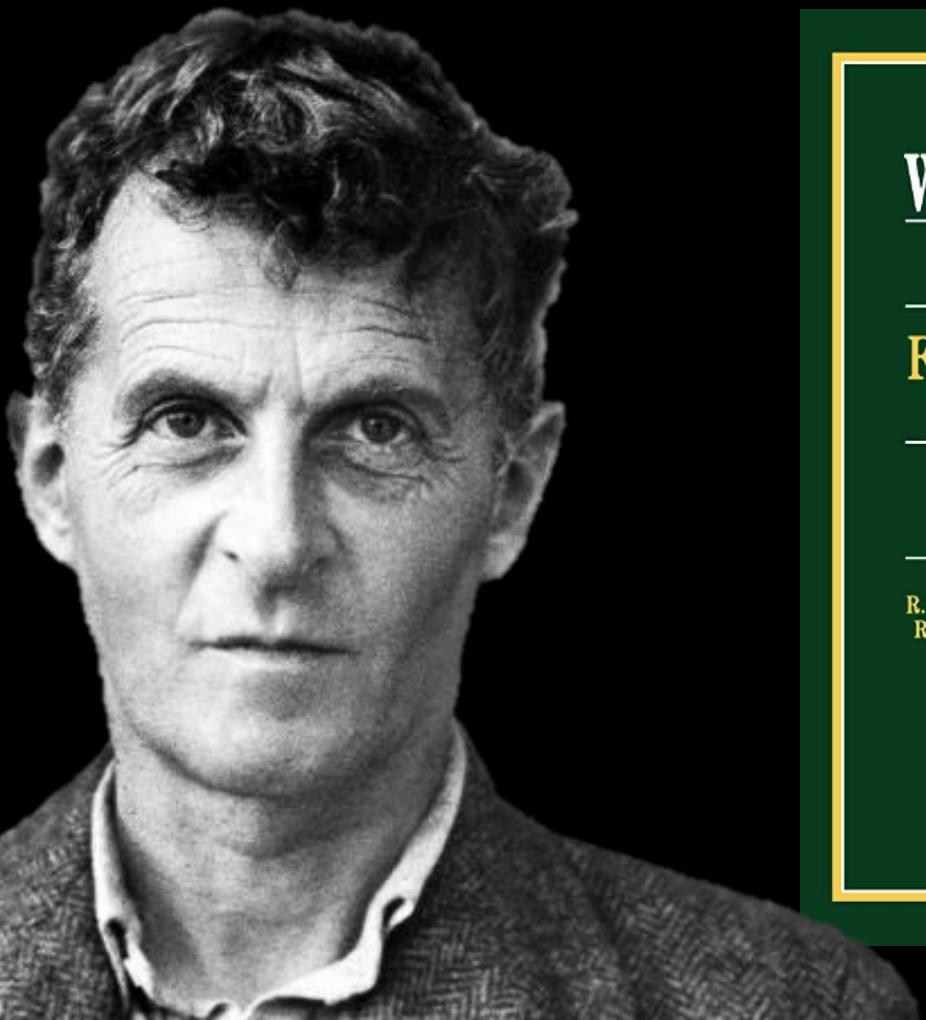
- We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.
- <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- Asilomar (2017) AI Principles: Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.
- Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders.
- Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.

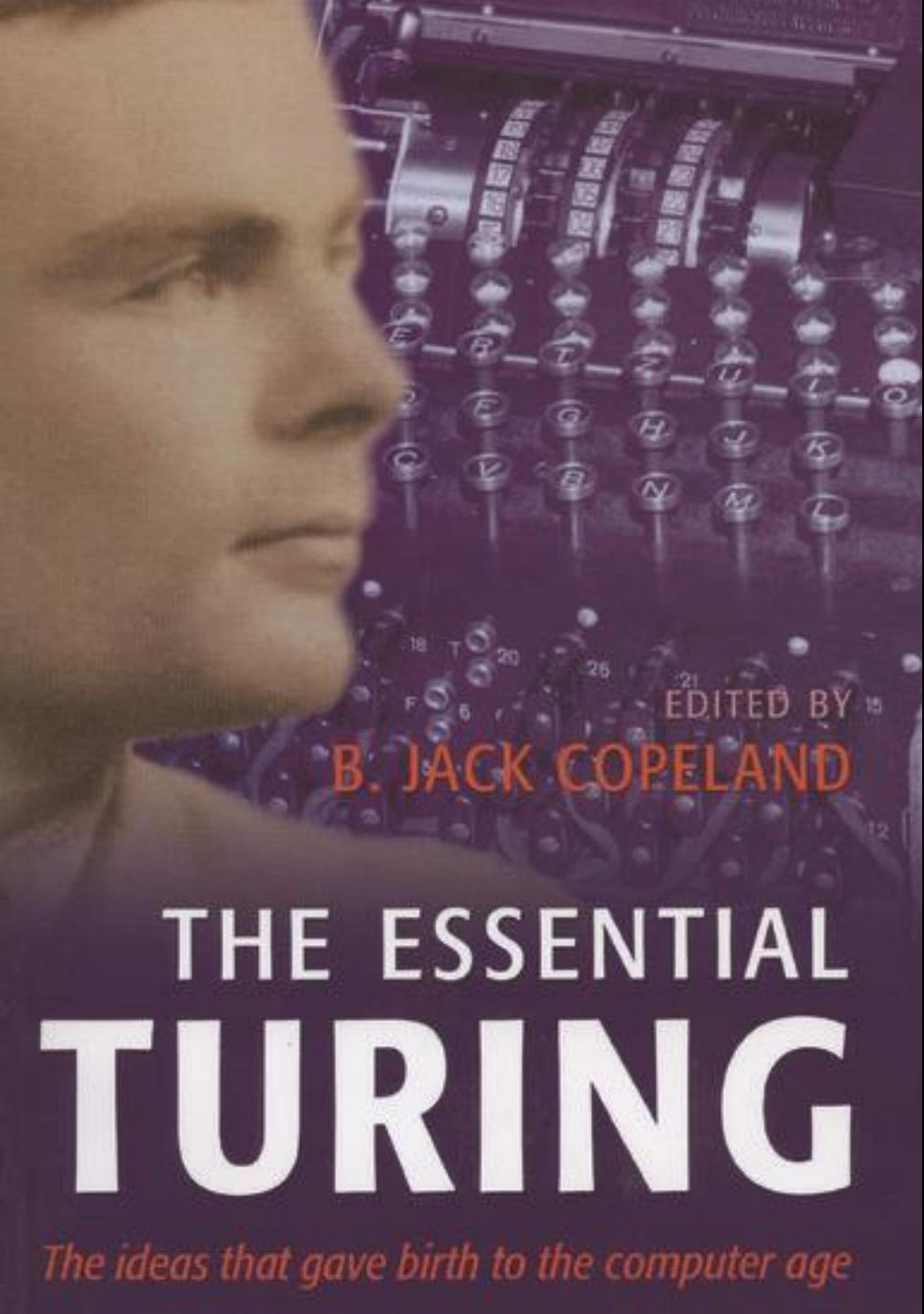


1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie
2. AI jako *Mathesis Universalis*
3. Etika umělé inteligence jako vyslovování nevyslovitelného
4. **Turingův test jako Wittgensteinova jazyková hra**
5. Wittgensteinova analýza významu jako distribuční sémantika
6. *Artificial Investigations* jako propojení AI s filozofií



Turingův test jako Wittgensteinova jazyková hra





B J Copeland. *The Essential Turing*.
OUP, 2004.

Artificial Intelligence

- 9: Lecture on the Automatic Computing Machine (1947)
- 10: Intelligent Machinery (1948)
- 11: Computing Machinery and Intelligence (1950)
- 12: Intelligent Machinery, A Heretical Theory (c. 1951)
- 13: Can Digital Computers Think?
- 14: Can Automatic Calculating Machines Be Said to Think? (1952), (AT, Richard Braithwaite, Geoffrey Jefferson, and Max Newman)

Artificial Life

- 15: The Chemical Basis of Morphogenesis (1952)
- 16: Chess (1953)
- 17: Solvable and Unsolvable Problems (1954)

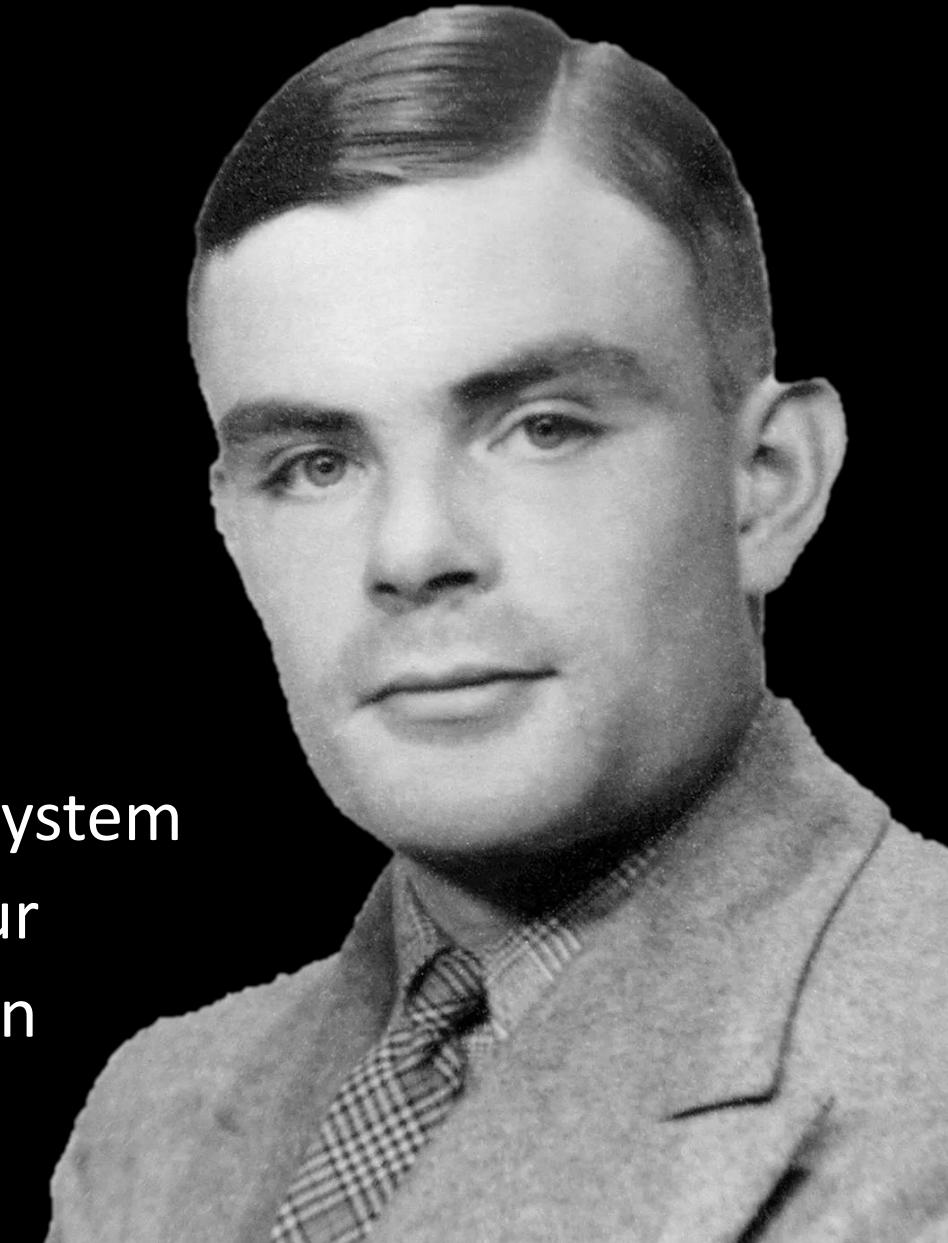
Turing, A.M. (1950). Computing machinery and intelligence. Mind, 59, 433-460.

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game.' It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A."

Turing's Objections

- (1) The Theological Objection
- (2) The "Heads in the Sand" Objection
- (3) The Mathematical Objection
- (4) The Argument from Consciousness
- (5) Arguments from Various Disabilities
- (6) Lady Lovelace's Objection
- (7) Argument from Continuity in the Nervous System
- (8) The Argument from Informality of Behaviour
- (9) The Argument from Extrasensory Perception

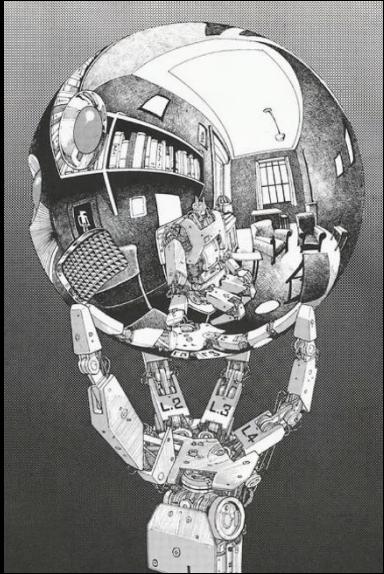


“This is only a foretaste of what is to come, and only the shadow of what is going to be. We have to have some experience with the machine before we really know its capabilities . . . I do not see why it should not enter any one of the fields normally covered by the human intellect, and eventually compete on equal terms.“

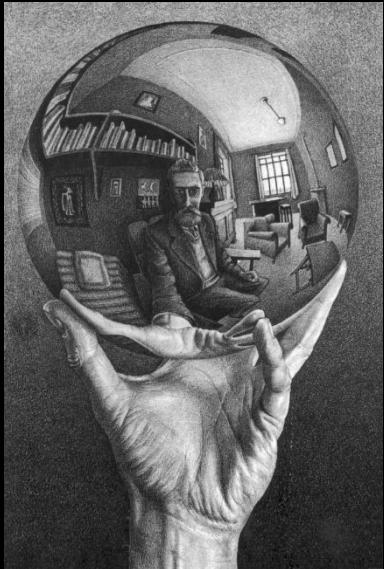
(Quoted in *The Times*, 11 June 1949:
‘The Mechanical Brain’)



Main frame, for Automatic Computing Engine (ACE) pilot model, made at the National Physical Laboratory in London (NPL). In 1945, while working at the NPL Turing formulated the design of ACE, which would have been the world's first general-purpose computer. He left the NPL in 1948, but this smaller 'pilot', or prototype, was made in his absence in 1950.



1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie
2. AI jako *Mathesis Universalis*
3. Etika umělé inteligence jako vyslovování nevyslovitelného
4. Turingův test jako Wittgensteinova jazyková hra
5. **Wittgensteinova analýza významu jako distribuční sémantika**
6. *Artificial Investigations* jako propojení AI s filozofií



The Distributional Hypothesis - Word Embedding

- The semantic theory of language **usage**.
- Words that are used or occur in the **same contexts** tend to purport **similar meanings**.
- **Context**-dependent nature of the meaning.
- Zellig Harris (1954):
“If A and B have almost identical **environments** we say that they are **synonyms**.“
- John Rupert Firth (1957):
“You shall know a word by the **company** it keeps“

Word Embedding in NLP

- Many techniques were designed to build dense vectors representing words semantics based on their neighbors, e.g. Word2vec, GloVe.

Word2vec

- 2 model architectures to produce distributed representations of words:
 - 1) Continuous bag-of-words (CBOW) - predicts the current word from the window of surrounding context words.
 - 2) Continuous skip-gram - uses the current word to predict the surrounding window of context words

Mikolov, T. et al. (2013) Distributed representations of words and phrases and their compositionality.

Advances in Neural Information Processing Systems. <https://arxiv.org/abs/1310.4546>

Mikolov, et al. (2015) Computing numeric representations of words in a high-dimensional space.

<https://patents.google.com/patent/US9037464B1/en>

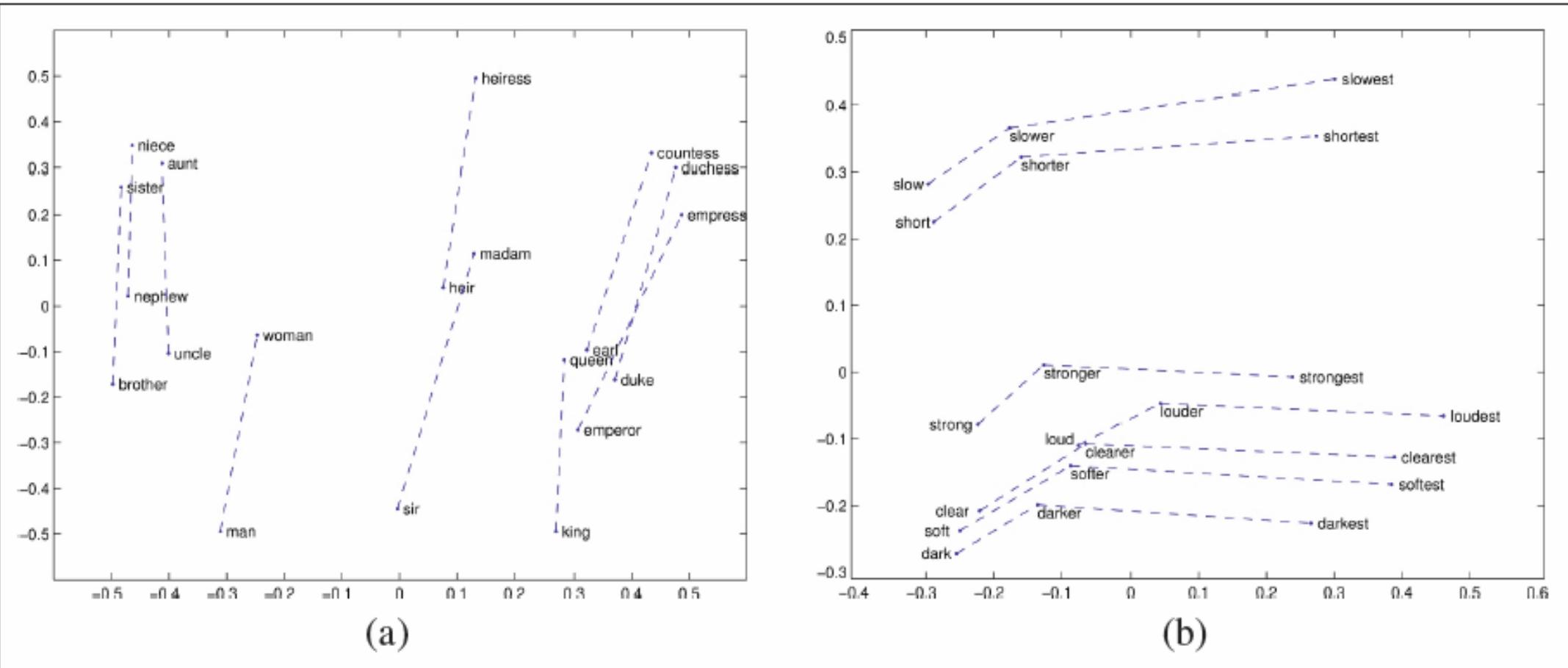


Figure 6.16 Relational properties of the GloVe vector space, shown by projecting vectors onto two dimensions. (a) $\overrightarrow{\text{king}} - \overrightarrow{\text{man}} + \overrightarrow{\text{woman}}$ is close to $\overrightarrow{\text{queen}}$. (b) offsets seem to capture comparative and superlative morphology (Pennington et al., 2014).

Word Embedding in NLP

Arvind Neelakantan et al. (2015) Efficient Non-parametric Estimation of Multiple Embeddings per Word in Vector Space.

<https://arxiv.org/abs/1504.06654>

Deficiency in word2vec:

- each word type has only one vector representation
- converging to an approximation of the average meaning of a word instead of capturing the polysemy inherent in language

Erhan Sezerer, Selma Tekir (2021) A Survey On Neural Word Embeddings.

<https://arxiv.org/abs/2110.01804>

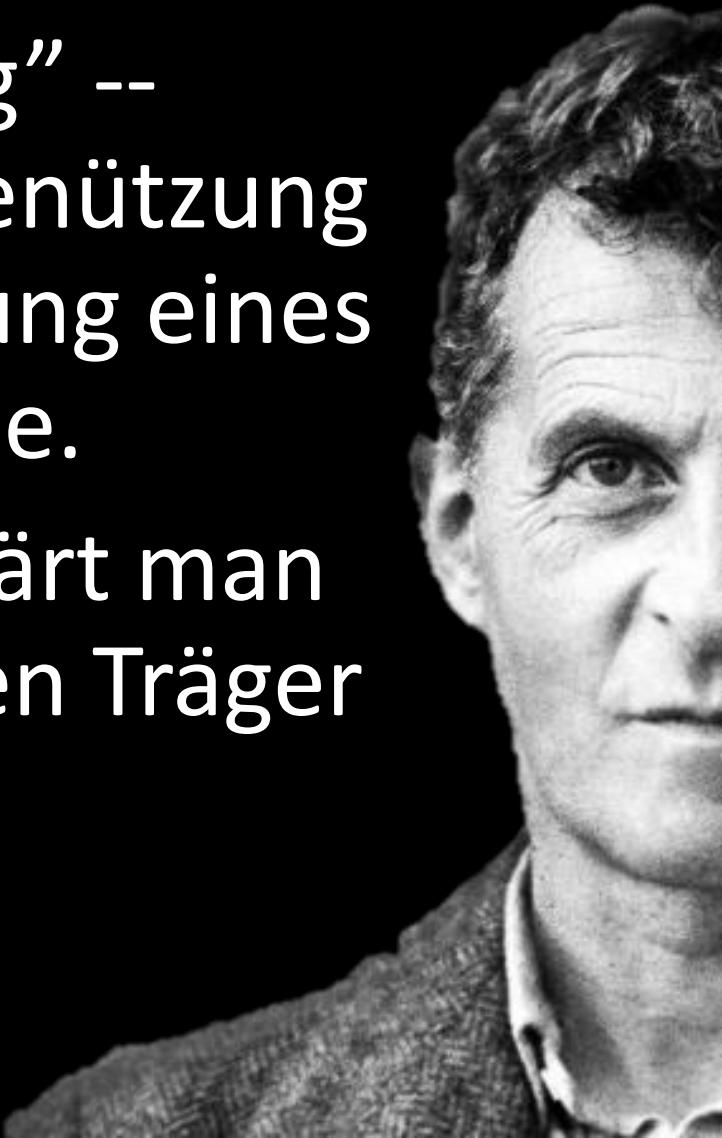
John Rupert Firth (1957)



The placing of a text as a constituent in a context of situation contributes to the statement of meaning since situations are set up to recognize use. As Wittgenstein says, '**the meaning of words lies in their use.**' (PI, 80, 109). The day-to-day practice of playing language games recognizes customs and rules. It follows that a text in such established usage may contain sentences such as 'Don't be such an **ass**!', 'You silly ass!', 'What an ass he is!' In these examples, the word ass is in familiar and habitual company, commonly collocated with you **silly**-, he is a silly-, don't be such an-. **You shall know a word by the company it keeps!** One of the meanings of ass is its habitual collocation with such other words as those above quoted. Though Wittgenstein was dealing with another problem, he also recognizes the plain face-value, the physiognomy of words. They look at us! 'The sentence is composed of words and that is enough'.

§ 43. Man kann für eine *große* Klasse von Fällen der Benützung des Wortes “Bedeutung” -- wenn auch nicht für alle Fälle seiner Benützung – dieses Wort so **erklären**: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.

Und die Bedeutung eines Namens erklärt man manchmal dadurch, daß man auf seinen Träger zeigt.



Anscombe's translation (1953):

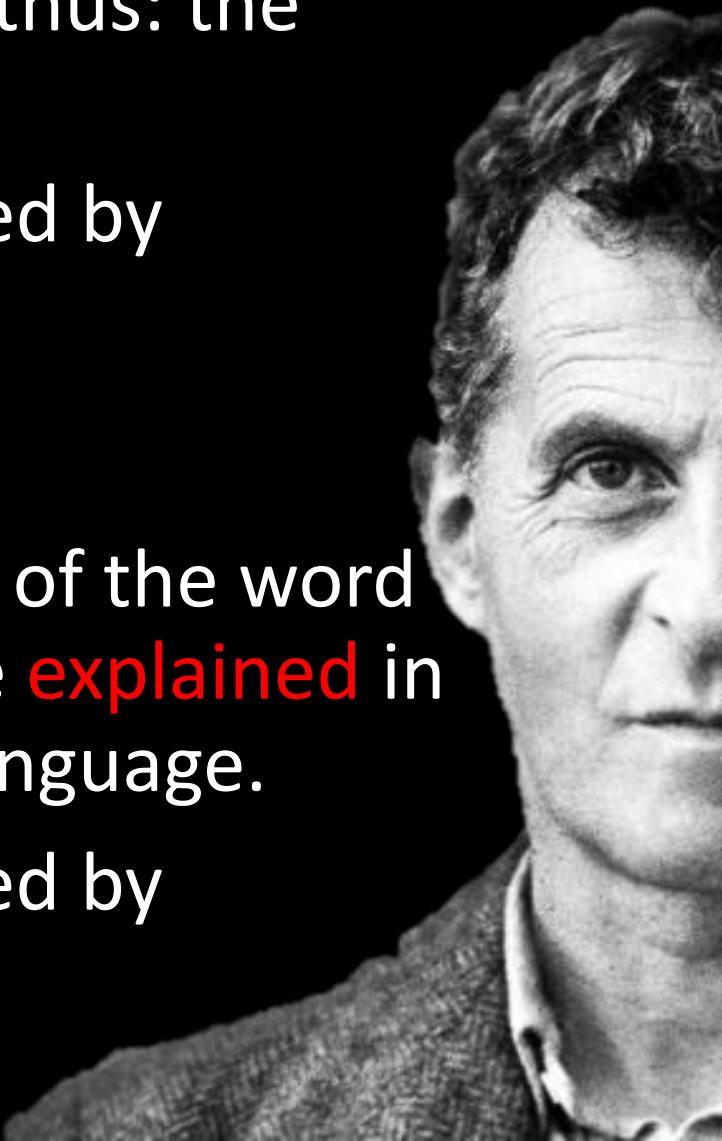
§ 43. For a *large* class of cases—though not for *all* — in which we employ the word "meaning" it can be **defined** thus: the meaning of a word is its use in the language.

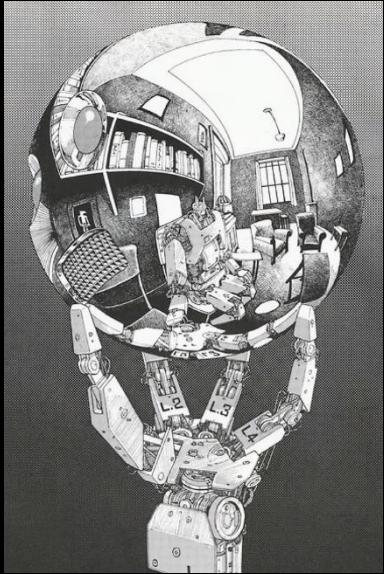
And the *meaning* of a name is sometimes explained by pointing to its *bearer*.

Schulte/Hacker's translation (2009):

§ 43. For a *large* class of cases of the employment of the word "meaning" -- though not for *all* -- this word can be **explained** in this way: the meaning of a word is its use in the language.

And the *meaning* of a name is sometimes explained by pointing to its *bearer*.





1. Filozofie v kontextu „AI v kontextu“ a AI v kontextu filozofie
2. AI jako *Mathesis Universalis*
3. Etika umělé inteligence jako vyslovování nevyslovitelného
4. Turingův test jako Wittgensteinova jazyková hra
5. Wittgensteinova analýza významu jako distribuční sémantika
6. ***Artificial Investigations* jako propojení AI s filozofií**



Projekt LudWitDig: Artificial Investigations

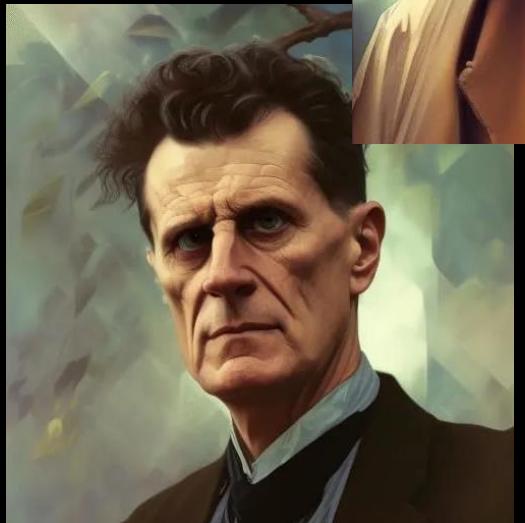
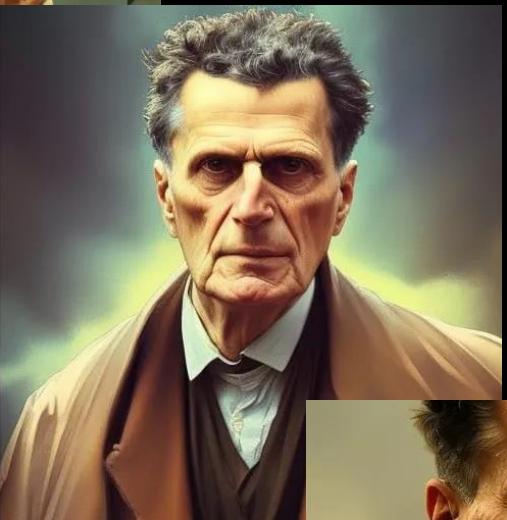
Východiska:

- Witt-y analýzy jazyka a myšlení skýtají kritickou i konstruktivní reflexi principů LLM a NLP obecně.
- Witt-y texty vykazují specifický styl a strukturu: fragmentárnost, aforističnost, dialogičnost

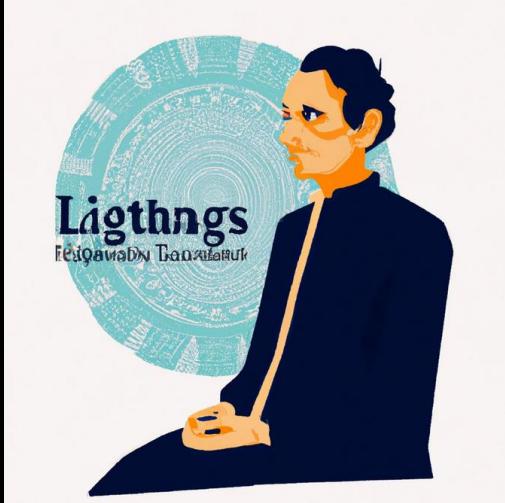
Cíl:

Vytvoření meta-modelu LudWitDig, který bude schopen auto-referenčně reflektovat vztah mezi Witt-ou filozofii jazyka a jejím jazykovým modelem.

Stable Diffusion



Ludwig Wittgenstein,
a philosopher and an author
of *Philosophical Investigations*,
artistic portrait

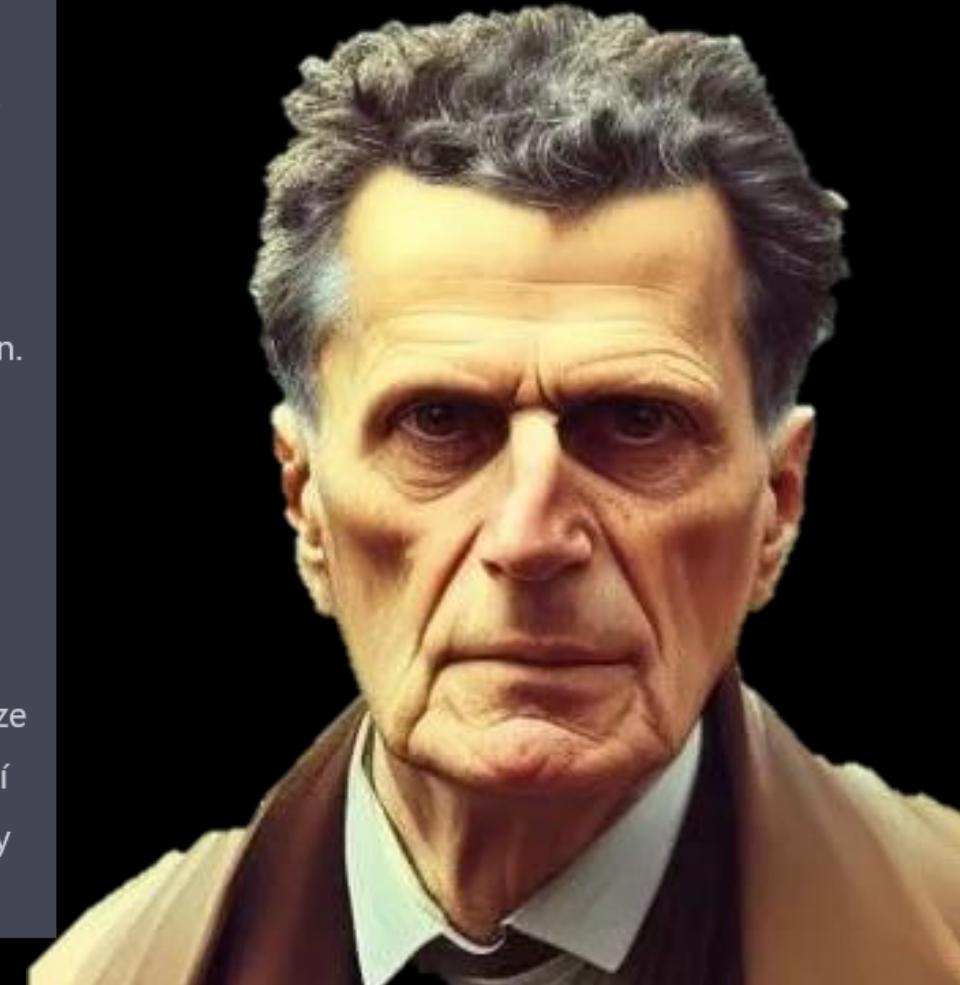


DALL-E



Ludwig Wittgenstein byl významným filosofem jazyka, jehož myšlenky jsou stále relevantní pro strojové zpracování jazyka. Zde jsou 5 nejdůležitějších Wittgensteinových myšlenek, které lze využít pro strojové zpracování jazyka:

1. Jazyk je hra: Wittgenstein popsal jazyk jako součást naší činnosti, nikoli jako abstraktní systém symbolů. Tento koncept lze aplikovat na strojové zpracování jazyka tím, že se zaměří na používání jazyka v kontextu konkrétních činností a situací.
2. Význam slov závisí na kontextu: Wittgenstein zdůraznil, že slova mají význam pouze v konkrétním kontextu. Tento koncept lze aplikovat na strojové zpracování jazyka tím, že se snaží porozumět kontextu a zohlednit ho při interpretaci textu.
3. Filosofie je kritika jazyka: Wittgenstein se zajímal o to, jak jazyk ovlivňuje naše myšlení a chování. Tento koncept lze aplikovat na strojové zpracování jazyka tím, že se snaží porozumět, jak stroje interpretují a používají jazyk a jak tento proces ovlivňuje jejich výkon.
4. Jazyk nemá jednoznačnou logiku: Wittgenstein se zaměřil na to, že jazyk nemá jednoznačnou logiku a že jeho používání je spíše založeno na konvencích a pravidlech. Tento koncept lze aplikovat na strojové zpracování jazyka tím, že se snaží porozumět, jak lidé používají jazyk a jak lze tyto konvence a pravidla zakomponovat do algoritmů a modelů pro strojové zpracování jazyka.
5. Jazyk je limitovaný: Wittgenstein se domníval, že jazyk má své limity a že některé věci nelze v jazyce vyjádřit. Tento koncept lze aplikovat na strojové zpracování jazyka tím, že se snaží porozumět, jaké jsou limity strojového zpracování jazyka a jakým způsobem lze tyto limity překonat nebo minimalizovat.



Jak lze využít Wittgensteinovy filozofie k vytvoření AGI?

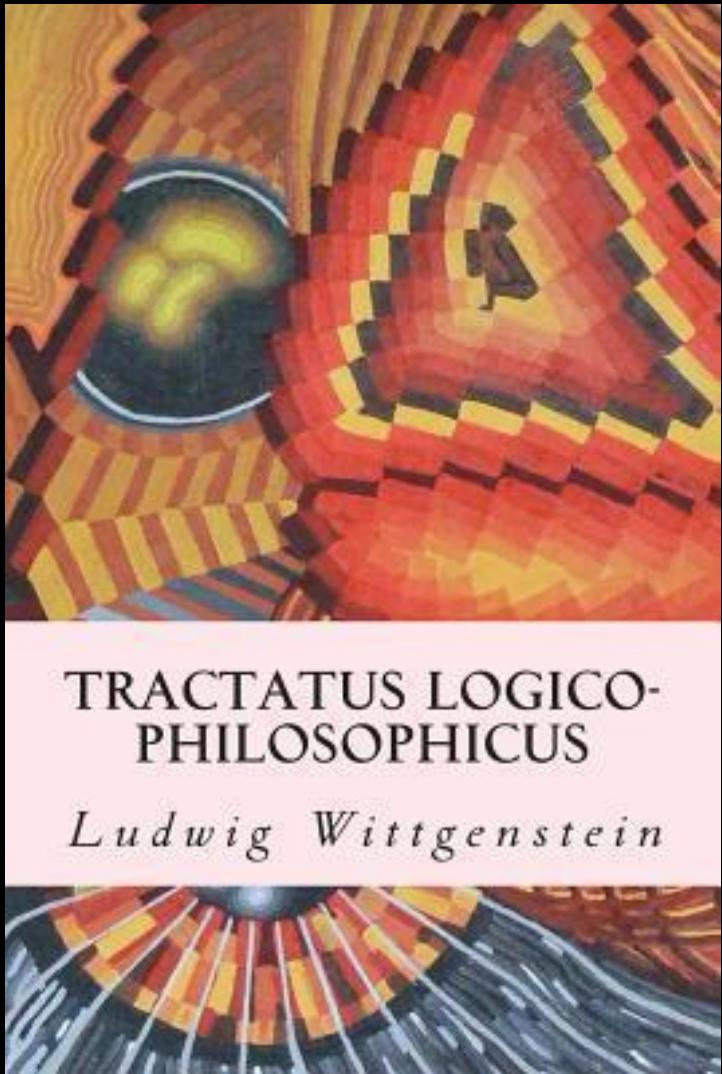


Ludwig Wittgenstein byl filosofem jazyka a myšlení, jehož práce mohou poskytnout užitečné základy pro vytváření umělé inteligence obecného typu (AGI). Zde je pár způsobů, jak lze využít Wittgensteinovu filozofii pro vytvoření AGI:

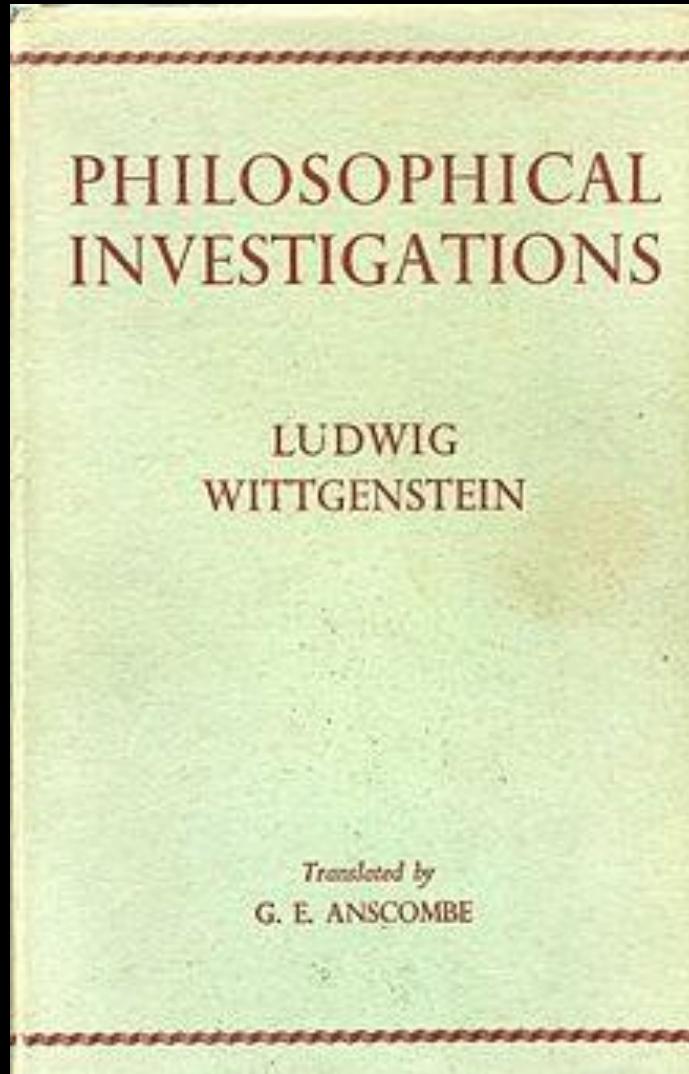
1. Koncept jazyka jako hry: Wittgenstein viděl jazyk jako činnost, nikoli jako statický systém symbolů. Lze tedy použít tohoto konceptu pro AGI tím, že se zaměří na dynamiku interakce mezi AGI a okolím, ať už v rámci hry, dialogu nebo jiné formy interakce. AGI by mělo být schopno používat jazyk v konkrétních situacích a přizpůsobit se různým kontextům, aby mohlo úspěšně interagovat s lidmi a prostředím.
2. Kontextuální porozumění: Wittgenstein zdůrazňoval, že význam slov závisí na kontextu. AGI by mělo být schopno porozumět kontextu, ve kterém se používá jazyk, a přizpůsobit své chování a odpovědi podle situace. To znamená, že AGI by mělo být schopno rozpoznat různé konvence, pravidla a zvyklosti jazyka, a přizpůsobit se jim.
3. Filosofie jako kritika myšlení: Wittgenstein se zajímal o to, jak jazyk ovlivňuje naše myšlení a chování. To znamená, že AGI by mělo být schopno pochopit, jak lidé myslí a jaké jsou jejich motivace a touhy. To by umožnilo AGI lépe porozumět, jak interagovat s lidmi a jak plnit jejich požadavky.
4. Omezení jazyka: Wittgenstein se domníval, že jazyk má své limity a že některé věci nelze v jazyce vyjádřit. To znamená, že AGI by mělo být schopno rozpoznat a přizpůsobit se těmto omezením, aby mohlo úspěšně interagovat s lidmi a řešit složité problémy.

Tyto myšlenky lze tedy aplikovat na vývoj AGI tím, že AGI bude moci interagovat s lidmi a světem jako celek, rozumět kontextu, kriticky přemýšlet a přizpůsobit se omezením jazyka a myšlení. To by umožnilo AGI dosáhnout obecné inteligence

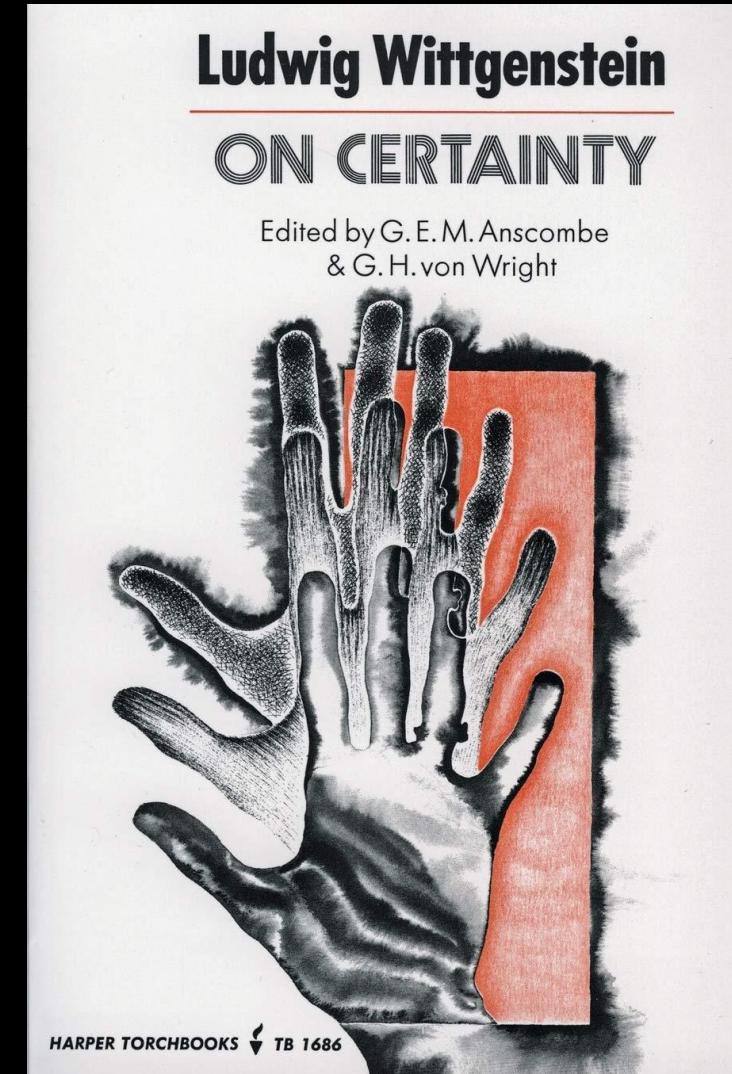
TLP (1921)



PI (1953)



OC (1969)



Wittgenstein Nachlass (~20 000 pages)

Wittgenstein's will (29. 1. 1951):

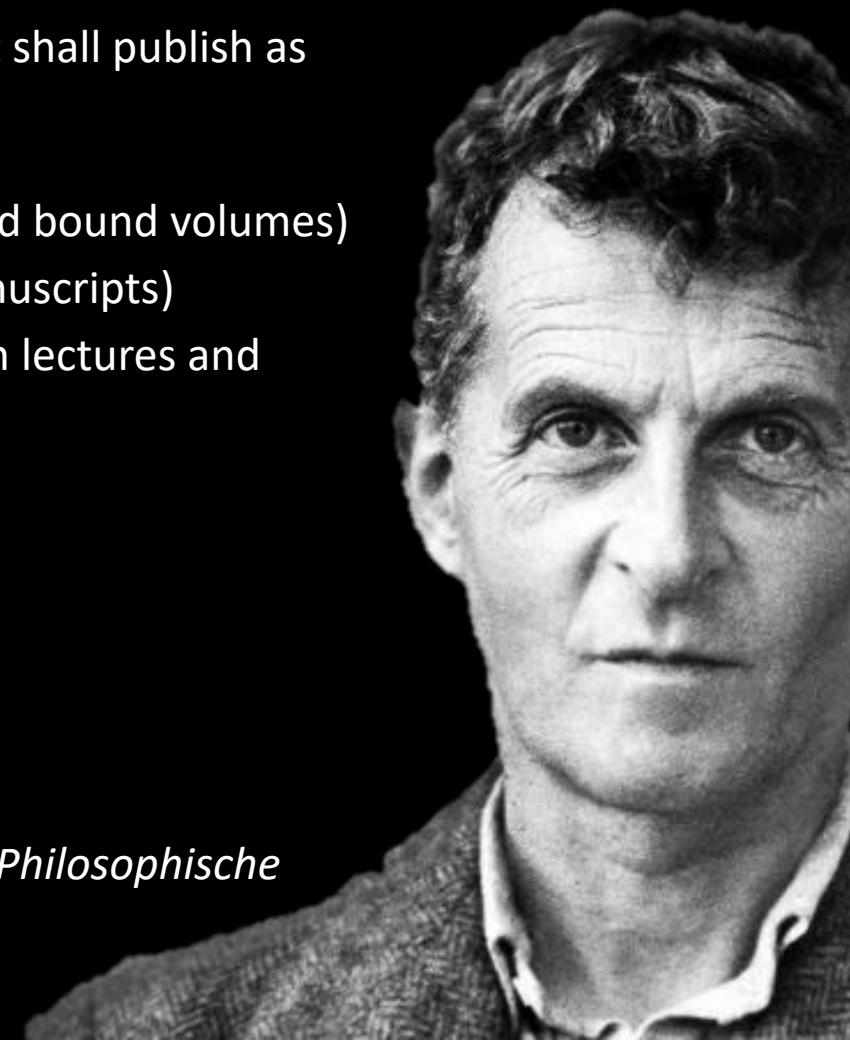
"I intend and desire that Mr Rhees, Miss Anscombe, and Professor von Wright shall publish as many of my unpublished writings as they think fit (...),"

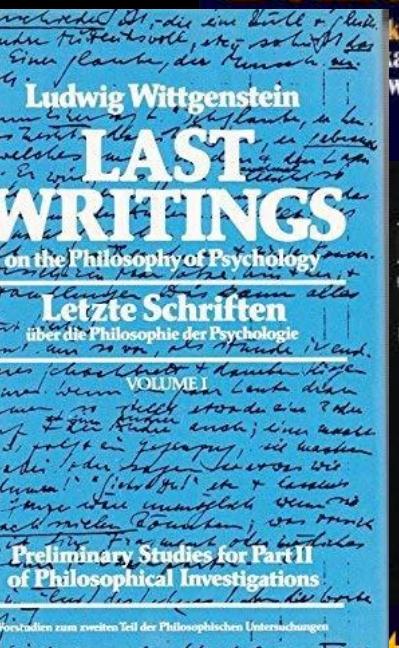
Wright's catalogue:

- 101-182: manuscripts (texts in Wittgenstein's hand, primarily notebooks and bound volumes)
- 201-245: typescripts (usually dictated by Wittgenstein directly from his manuscripts)
- 301-311: refer to dictations (to friends and students, e.g. in connection with lectures and seminars).

The originals preserved at various libraries/archives:

- The greater part of items: Trinity College Library, Cambridge
- MSS 105-107, 112-113, TS 203: Austrian National Library, Vienna
- MS 104, TS 202: Bodleian Library, Oxford
- TSS 201a: Bertrand Russell Archive, Hamilton, Ontario
- TS 209 (*Philosophische Bemerkungen*) and TS 234 (TS of the second part of *Philosophische Untersuchungen*). are now considered lost.





LAST WRITINGS

on the Philosophy of Psychology

Letzte Schriften

über die Philosophie der Psychologie

VOLUME I

Preliminary Studies for Part II
of Philosophical Investigations

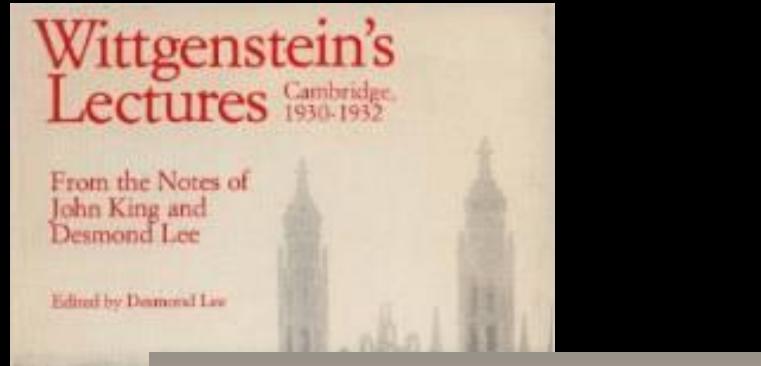
Studien zum zweiten Teil der Philosophischen Untersuchungen

Rechte um das Urheberrecht erweitert.

Wittgenstein Bemerkungen über die Philosophie der Psychologie

Ludwig Wittgenstein The Blue and Brown Books

Preliminary Studies for the 'Philosophical Investigations'



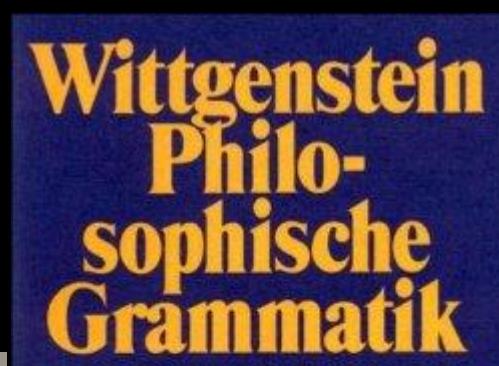
Ludwig Wittgenstein The Big Typescript: TS 213

German-English Scholars' Edition



Edited and Translated by
C. Grant Luckhardt and Maximilian A. E. Au

© WILEY-BLACKWELL



Wittgenstein Bemerkungen über die Grundlagen der Mathematik

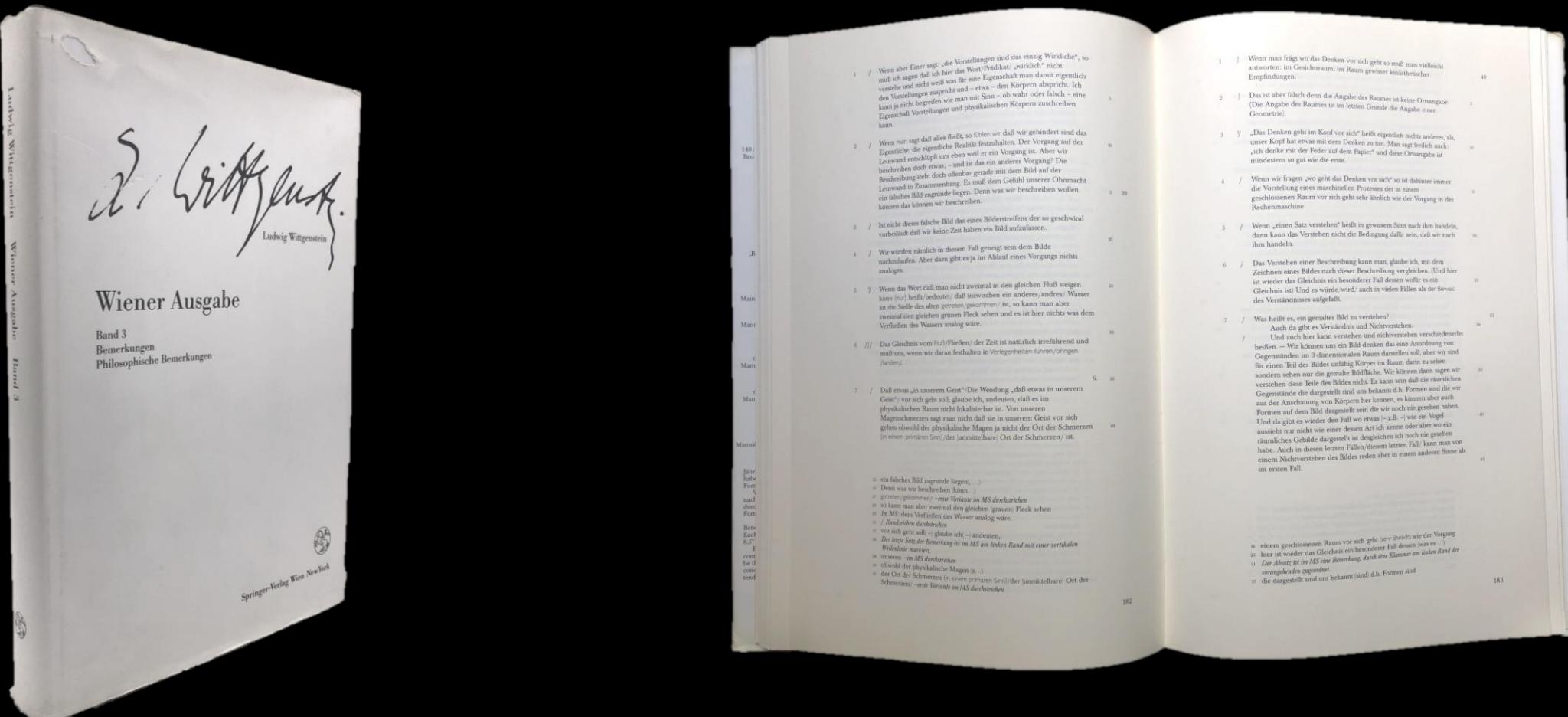
Werkausgabe Band 6
suhrkamp taschenbuch
wissenschaft

Ludwig Wittgenstein

Vermischte
Bemerkungen

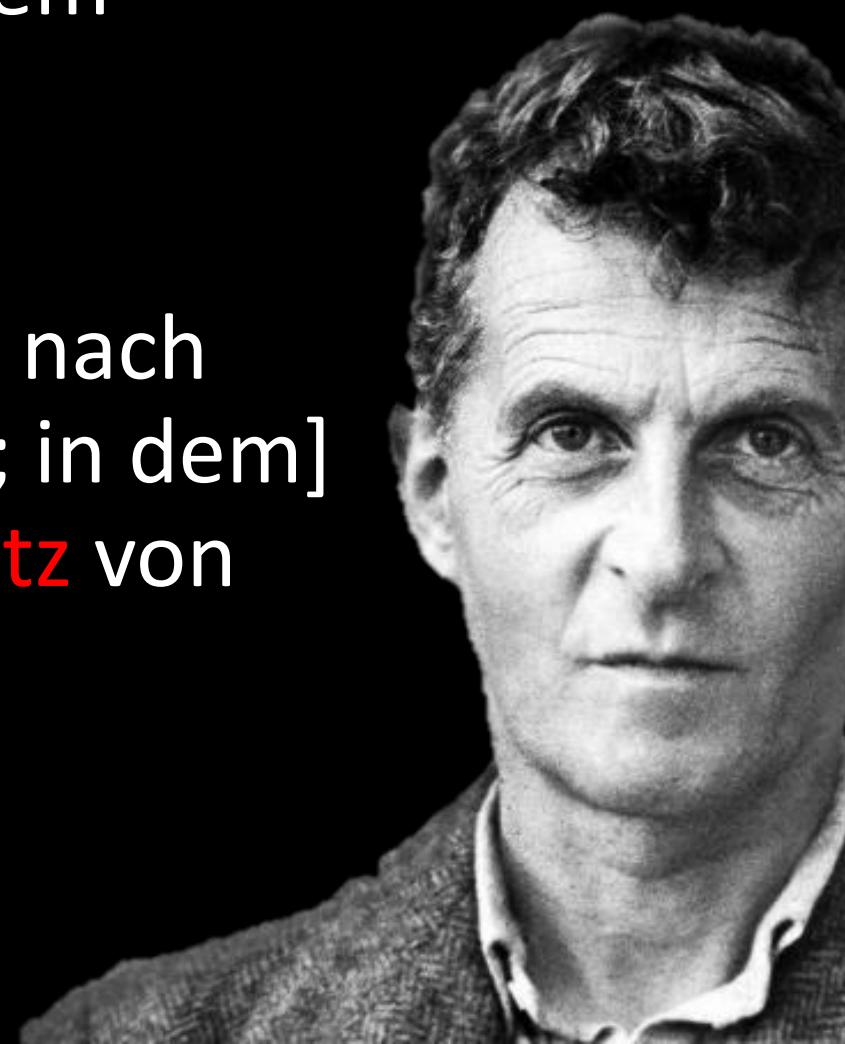
Sibliothek Suhrkamp

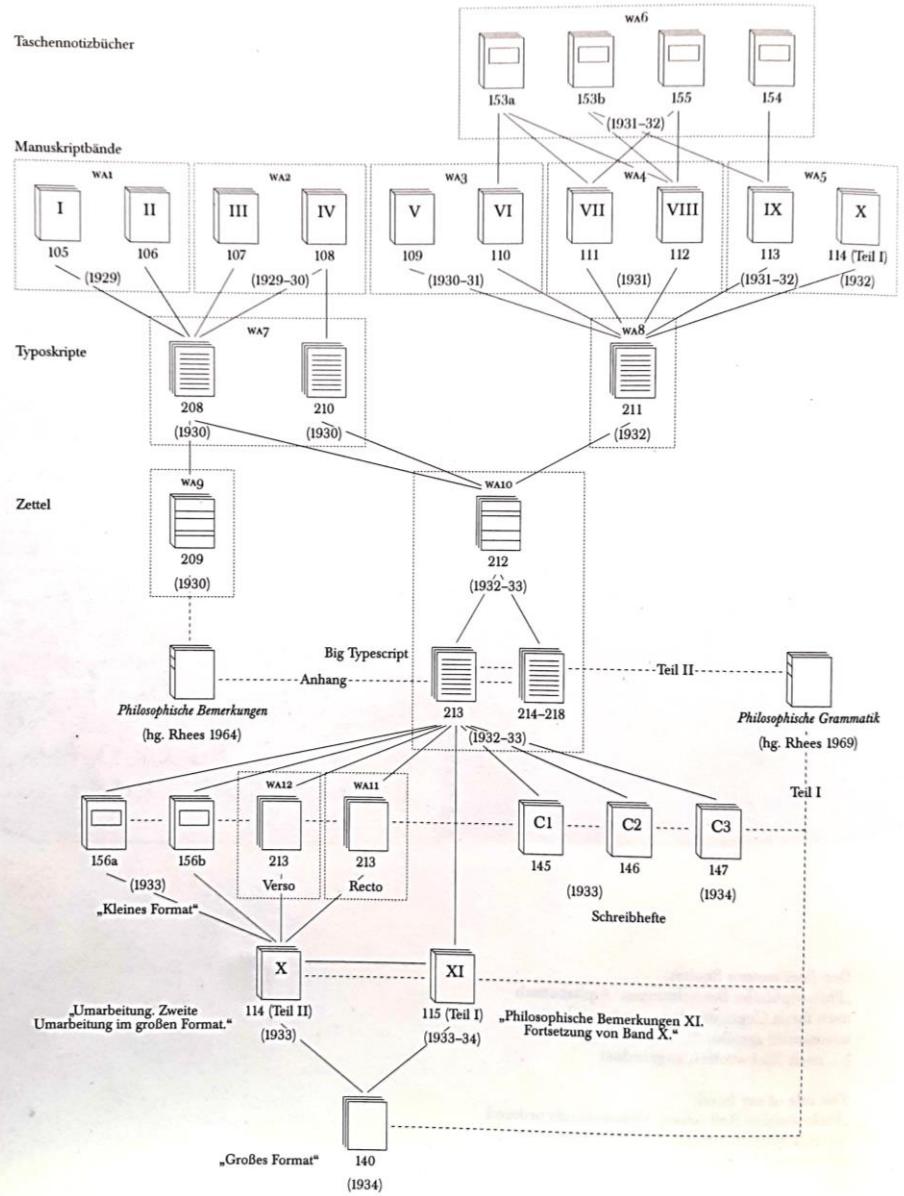
Wiener Ausgabe [Bänden 1-5], 1993–1995, Michael Nedo (ed.), Vienna: Springer.



W. přirovnává svou práci ve filozofii ke zkoumání „myšlenkového území, které je třeba projít všemi směry; nebo jsou v něm myšlenky navzájem propojeny ve spletité síti". (1938)

„Gedankenengebiet[, das] kreuz und quer, nach alien Richtungen hin zu durchreisen [ist; in dem] die Gedanken in einem **verwickelten Netz** von Beziehungen zueinander stehen."





Wiener Ausgabe.

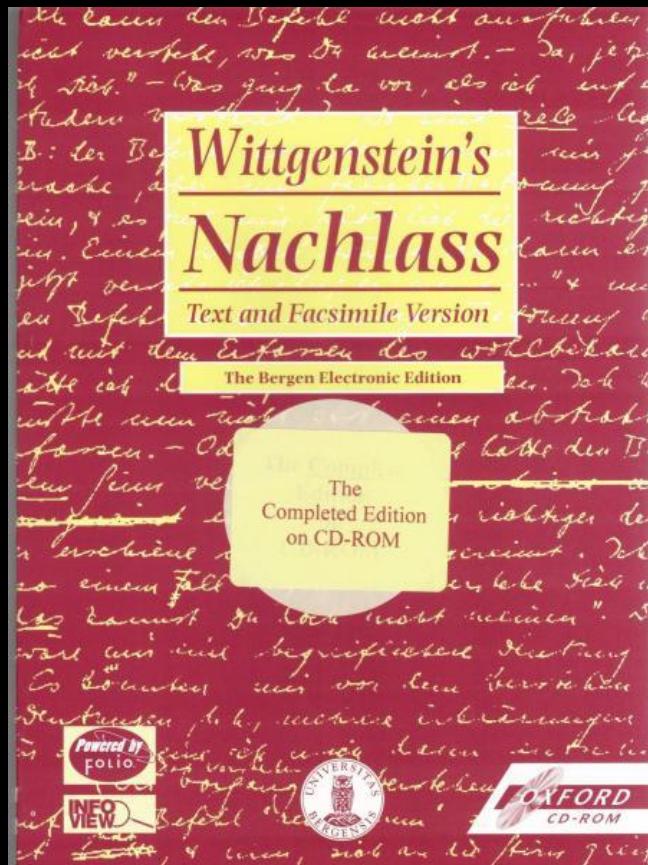
Register zu den Bänden 1-5. (1998)

Editionsplan Plan of the Edition

Bd	ANTE QUAM			WA2			POST QUAM			PHIL. BEM. S Nr. B A	PHIL. GRAM. T S Nr. B A
	Bd	S	B	A	R	Bd	S	B	A		
1	2	1	/	283							
2	3	1	/								
3	4	1	/	284							
4	5	2	/								
5	6	1	/	285							
6	7	1	/								
7	8	1	/	286							
8	9	1	/								
9	10	6	6	287							
10	11	6	6								
11	12	6	6	288							
12	13	6	6								
13	14	93	6	289							
14	15	66	7	290							
15	16	66	7								
16	17	64	1								
17	18	64	1								
18	19	64	1								
19	20	64	1								
20	21	64	1								
21	22	64	1								
22	23	64	1								
23	24	64	1								
24	25	64	1								
25	26	64	1								
26	27	64	1								
27	28	64	1								
28	29	64	1								
29	30	64	1								
30	31	64	1								
31	32	64	1								
32	33	64	1								
33	34	64	1								
34	35	64	1								
35	36	64	1								
36	37	64	1								
37	38	64	1								
38	39	64	1								
39	40	64	1								
40	41	64	1								
41	42	64	1								
42	43	64	1								
43	44	64	1								
44	45	64	1								
45	46	64	1								
46	47	64	1								
47	48	64	1								
48	49	64	1								
49	50	64	1								
50	51	64	1								
51	52	64	1								
52	53	64	1								
53	54	64	1								
54	55	64	1								
55	56	64	1								
56	57	64	1								
57	58	64	1								
58	59	64	1								
59	60	64	1								
60	61	64	1								
61	62	64	1								
62	63	64	1								
63	64	64	1								
64	65	64	1								
65	66	64	1								
66	67	64	1								
67	68	64	1								
68	69	64	1								
69	70	64	1								
70	71	64	1								
71	72	64	1								
72	73	64	1								
73	74	64	1								
74	75	64	1								
75	76	64	1								
76	77	64	1								
77	78	64	1								
78	79	64	1								
79	80	64	1								
80	81	64	1								
81	82	64	1								
82	83	64	1								
83	84	64	1								
84	85	64	1								
85	86	64	1								
86	87	64	1								
87	88	64	1								
88	89	64	1								
89	90	64	1								
90	91	64	1								
91	92	64	1								
92	93	64	1								
93	94	64	1								
94	95	64	1								
95	96	64	1								
96	97	64	1								
97	98	64	1								
98	99	64	1								
99	100	64	1								
100	101	64	1								
101	102	64	1								
102	103	64	1								
103	104	64	1								
104	105	64	1								
105	106	64	1								
106	107	64	1								
107	108	64	1								
108	109	64	1								
109	110	64	1								
110	111	64	1								
111	112	64	1								
112	113	64	1								
113	114	64	1								
114	115	64	1								
115	116	64	1								
116	117	64	1								
117	118	64	1								
118	119	64	1								
119	120	64	1								
120	121	64	1								
121	122	64	1								
122	123	64	1								
123	124	64	1								
124	125	64	1								
125	126	64	1								
126	127	64	1								
127	128	64	1								
128	129	64	1								
129	130	64	1								
130	131	64	1								
131	132	64	1								
132	133	64	1								
133	134	64	1								
134	135	64	1								
135	136	64	1								
136	137	64	1								
137	138	64	1								
138	139	64	1								
139	140	64	1								
140	141	64	1								
141	142	64	1								
142	143	64	1								
143	144	64	1								
144	145	64	1								
145	146	64	1								
146	147	64	1								
147	148	64	1								
148	149	64	1								
149	150	64	1								
150	151	64	1								
151	152	64	1								
152	153	64	1								
153	154	64	1								
154	155	64	1								
155	156	64	1								
156	157	64	1								
157	158	64	1								
158	159	64	1								
159	160	64	1								
160	161	64	1								
161	162	64	1			</td					

„A listing of the **repetitions** in the work, as they are presented for the first time here in the *Register zu den Banden 1-5* of the *Wiener Ausgabe*, is an essential tool for the study of Wittgenstein's writings. With the help of the Registers the reader can in Wittgenstein's sense '**cross in all directions**' the '**territory of thought**' represented in the manuscripts - that is what he again and again did himself, when he spread out his manuscripts on long tables and, reading them in all directions and dictating from them, went through them in ever new ways. It is only by this kind of study that those '**new views**' of the '**same points**' are revealed in his territory of thought; it is only in this changing, **ever different contexts** and the shifts of emphasis in content and style of the individual remarks that their **full meaning** is revealed - only in this way does the work of Ludwig Wittgenstein become accessible to the reader.“ (Nedo)

Wittgenstein's Nachlass. The Bergen Electronic Edition (BEE)



- A joint publication by the WAB and Oxford University Press.
- Publication of the edition, which was initially intended to consist of four separate volumes, begun in April 1998. After publication of Volume 2 it was decided to publish one final, complete title.
- The complete edition was developed for Windows (3.1/3.11, 95/98, NT 4.0, 2000) and released in 2000 on 6 CDs (1 CD with edited Nachlass texts and software, and 5 CDs with facsimile files).
- BEE is available online in the *Past Masters* series from InteLex Corporation.



 EXPLORE

 SEARCH

 EDITIONS

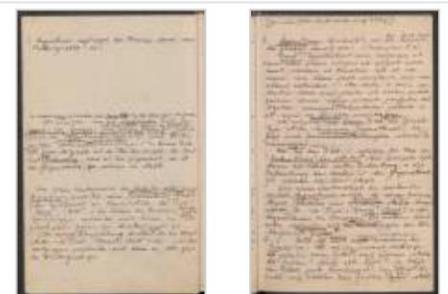
- | | |
|---|--|
| Bergen Nachlass Edition (BNE) | |
| Moore Notes Facsimile Edition (MWN) | Ms-123: Philosophische
Bemerkungen (WL) m L |
| Tractatus Publication Materials (LPA) | Ms-124 (WL) m L |
| Prototractatus Tools (PTT) | Ms-125 (WL) m L |
| Wörterbuch für Volksschulen Materials (WVF) | Ms-126 (WL) m L |
| | |
|  ABOUT  | |
| About Wittgenstein Source | Ms-127: F, Mathematik und Logik (WL) m L |
| About BNE | Ms-128 (WL) m L |
| About MWN | Ms-129 (WL) m L |
| About LPA | Ms-130 (WL) m L |
| About PTT | Ms-131 (WL) m L |
| About WVF | Ms-132 (WL) m L |
| Updates | Ms-133 (WL) m L |
| | |
| | Ms-134 (WL) m L |
| | Ms-135 (WL) m L |
| | Ms-136: Q (WL) m L |
| | Ms-137: R (WL) m L |
| | Ms-138: S (WL) m L |

 Bergen Nachlass Edition (BNE)

[s-142 Facsimile]

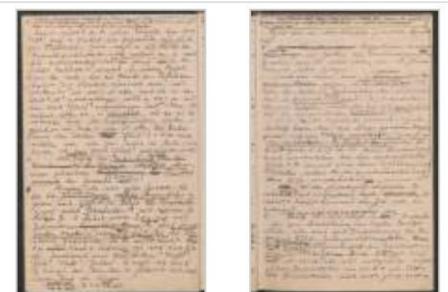
5

Is-142,2 facsimile



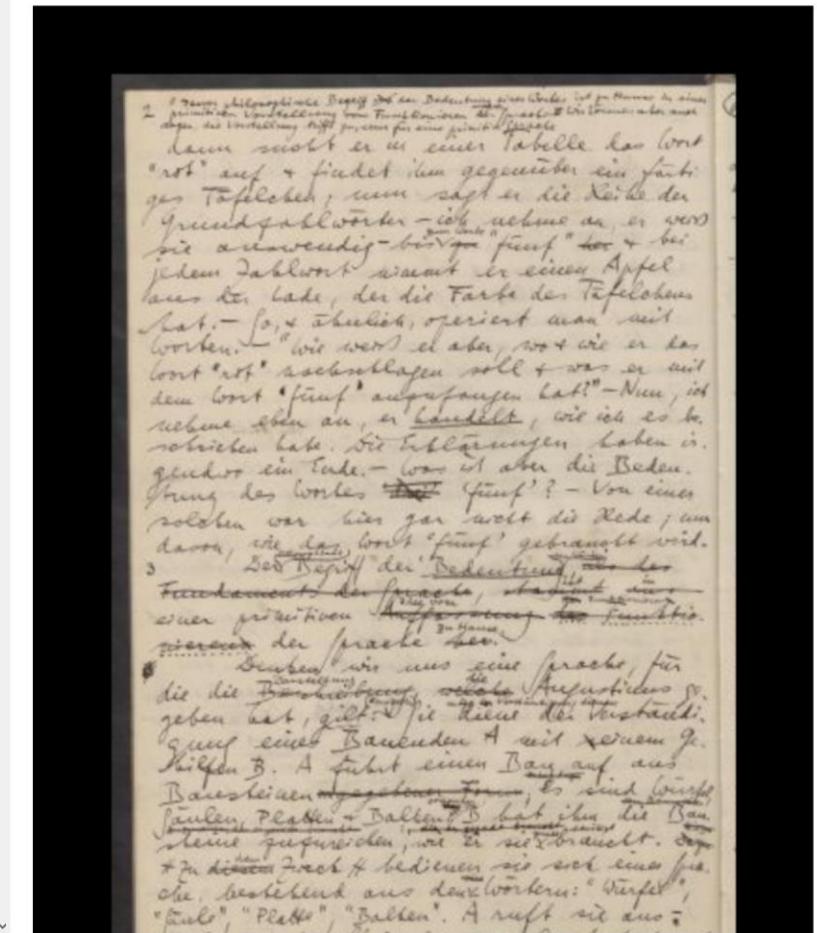
Ms-142.IIv

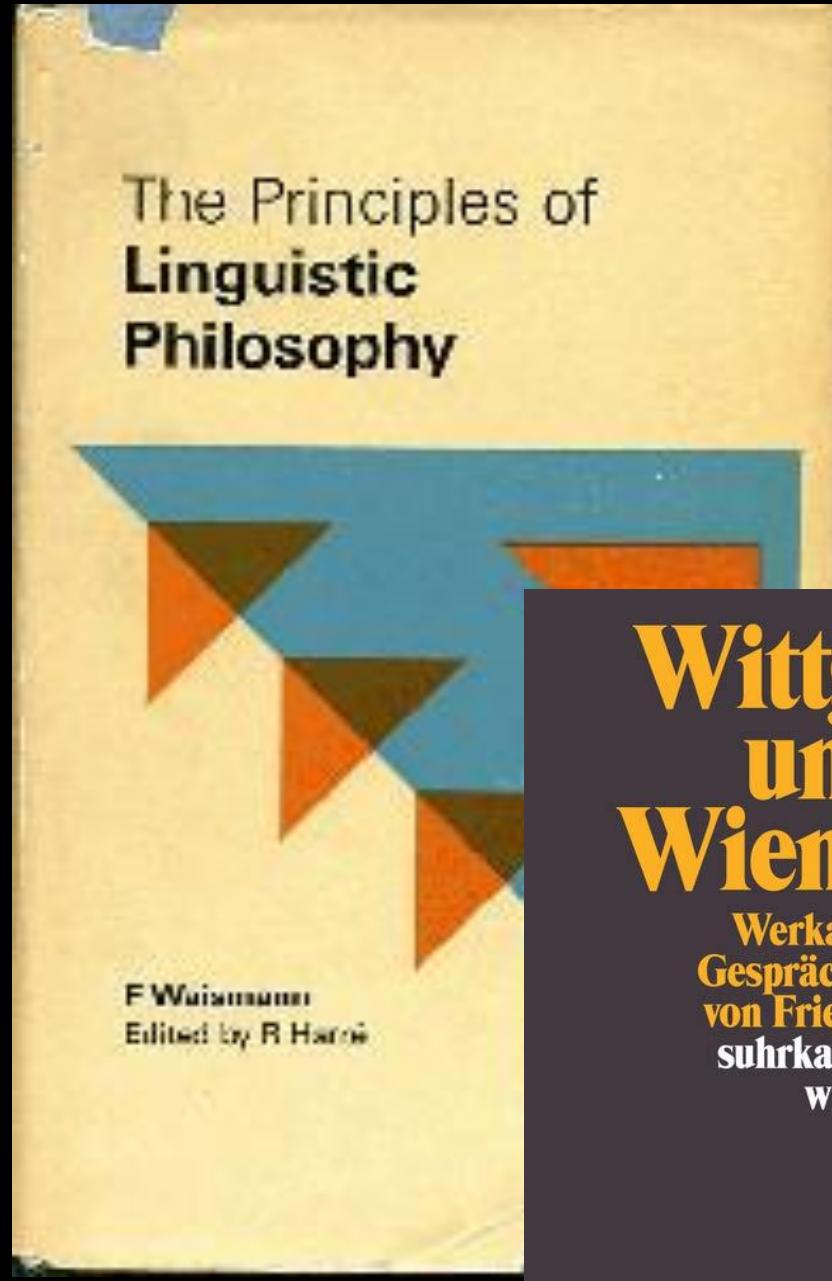
Ms-142,1

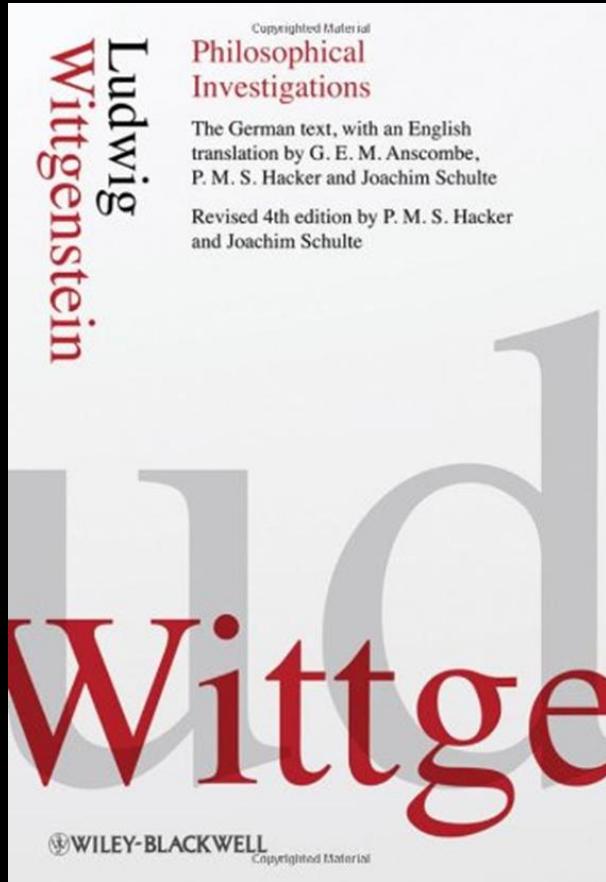
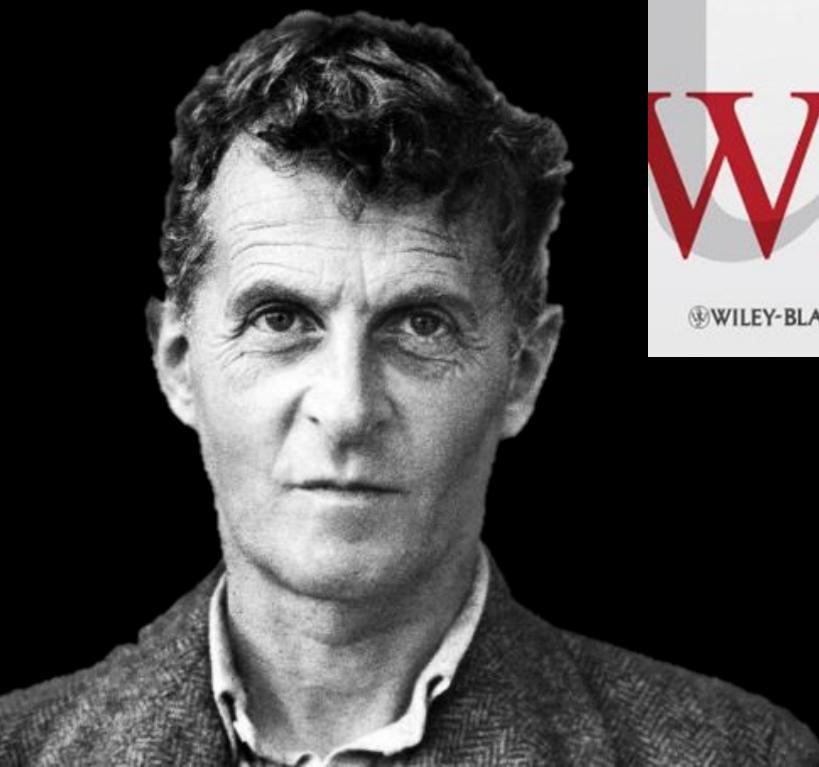


Ms-142,2

Ms-142,3





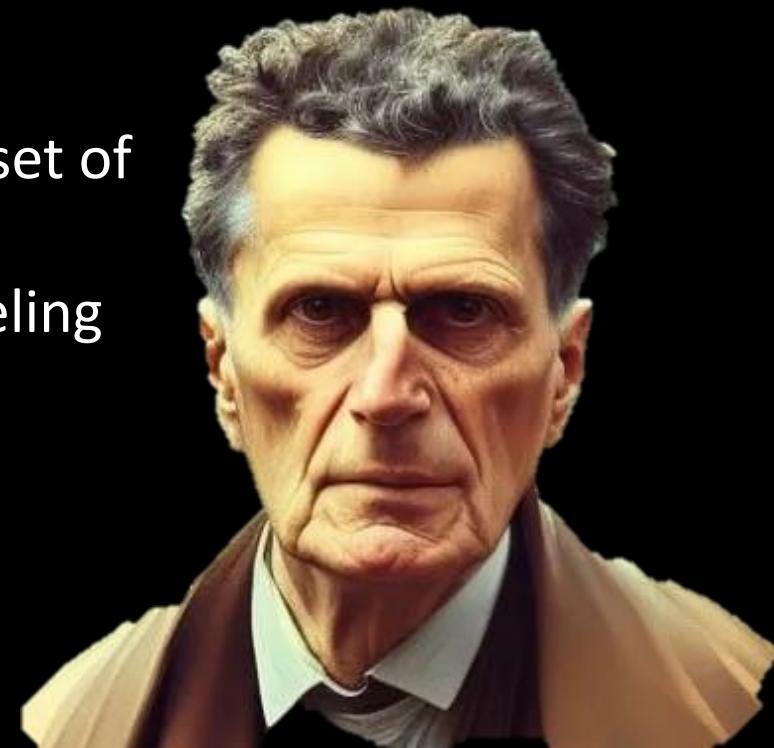


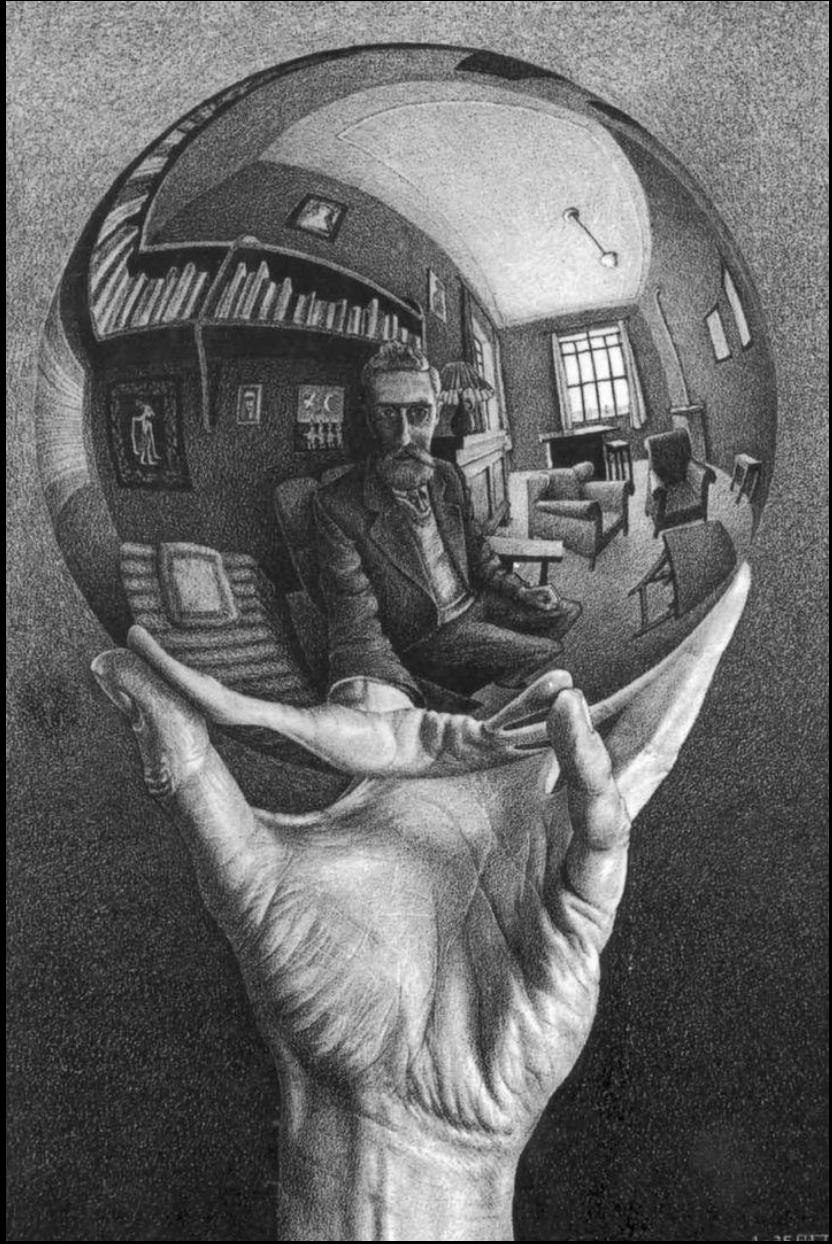
EleutherAI

github.com/EleutherAI/gpt-neox

GPT-NeoX-20B
An Open-Source
Autoregressive
Language Model

The Pile
An 800GB Dataset of
Diverse Text for
Language Modeling





Děkuji vám za pozornost!