

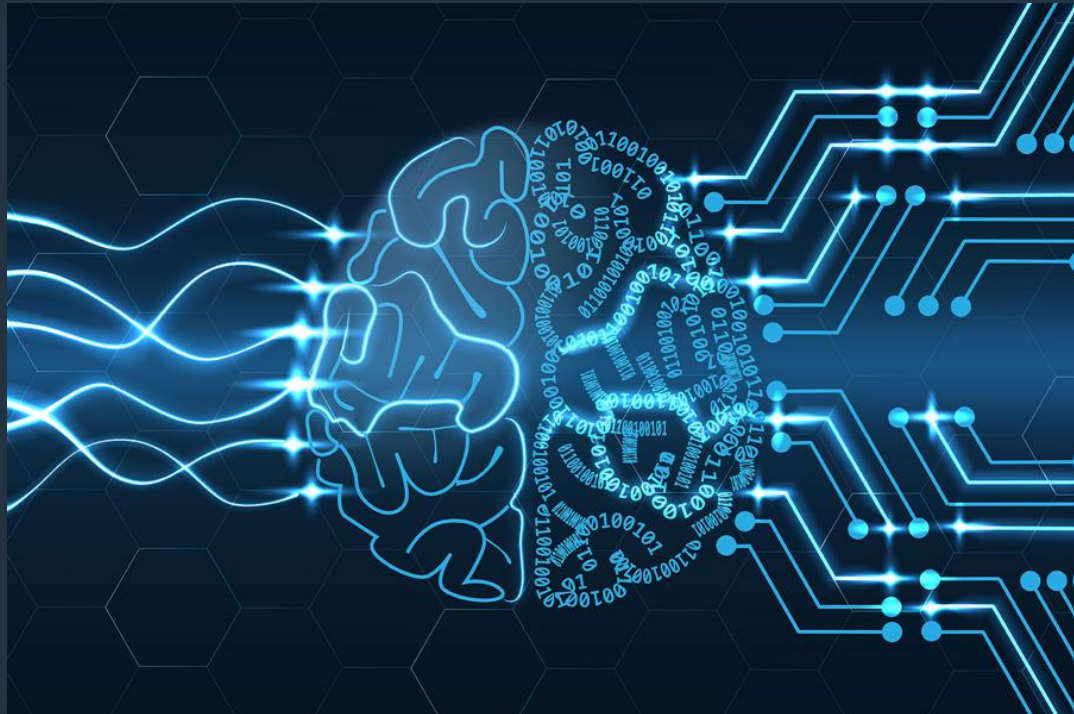
# Máme se bát umělé inteligence?

David Černý

Oddělení umělé inteligence,  
Ústav informatiky AV ČR

LF MU Brno





# Čtyři přístupy k AI

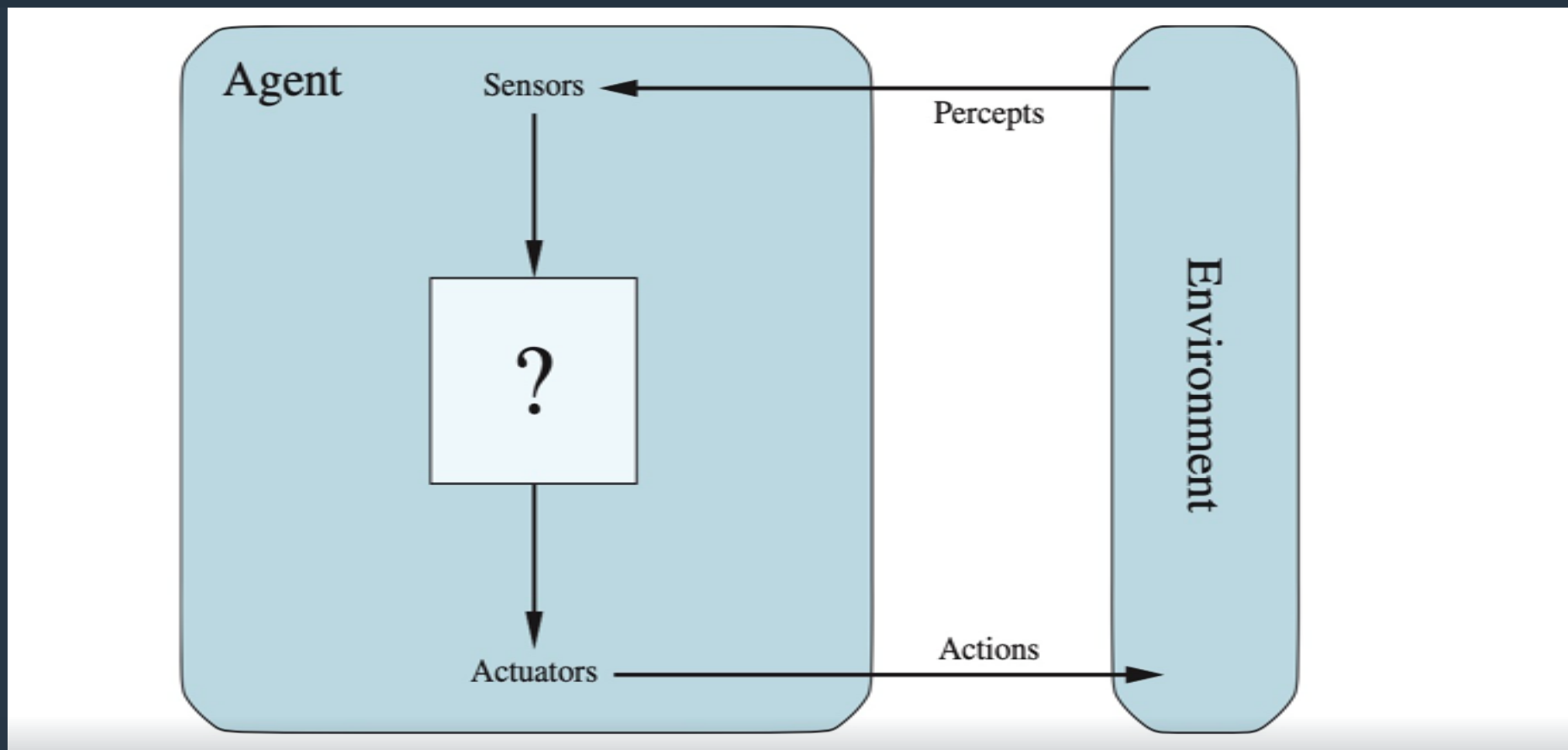
Jednat  
jako člověk

Myslet jako  
člověk

Myslet  
racionálně

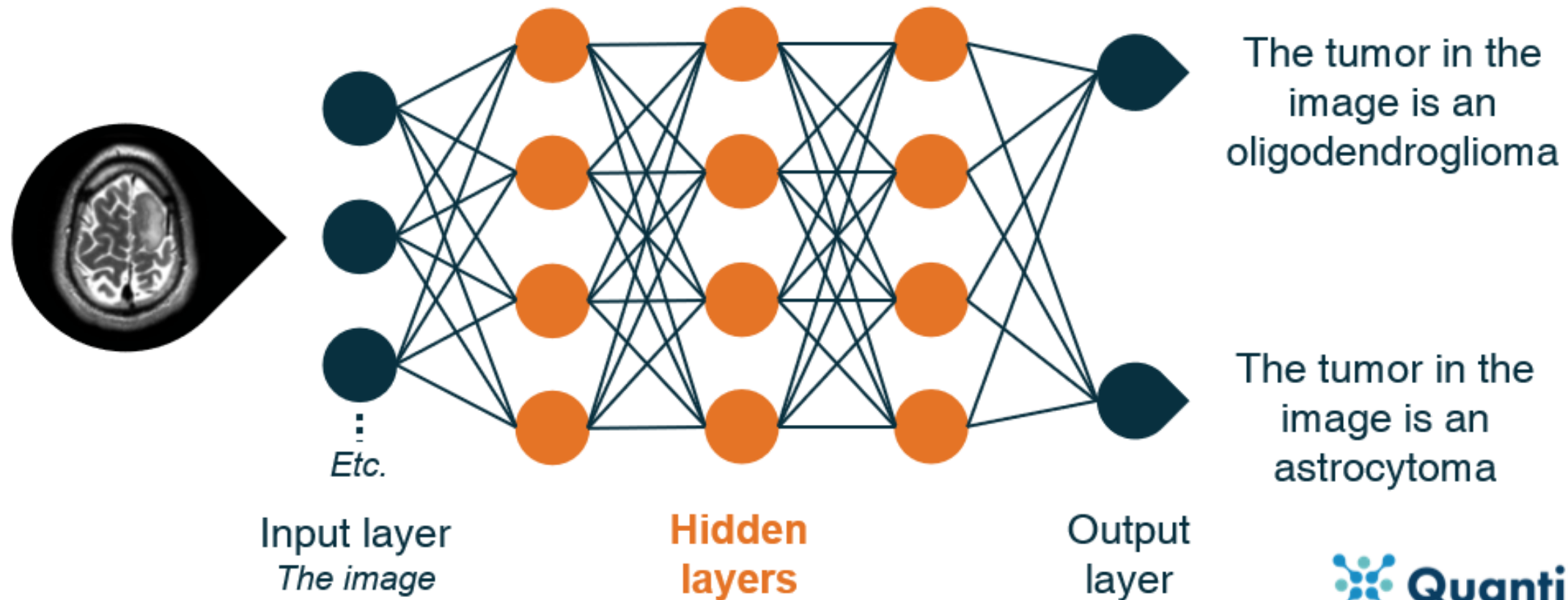
Jednat  
racionálně

# Jednat racionálně



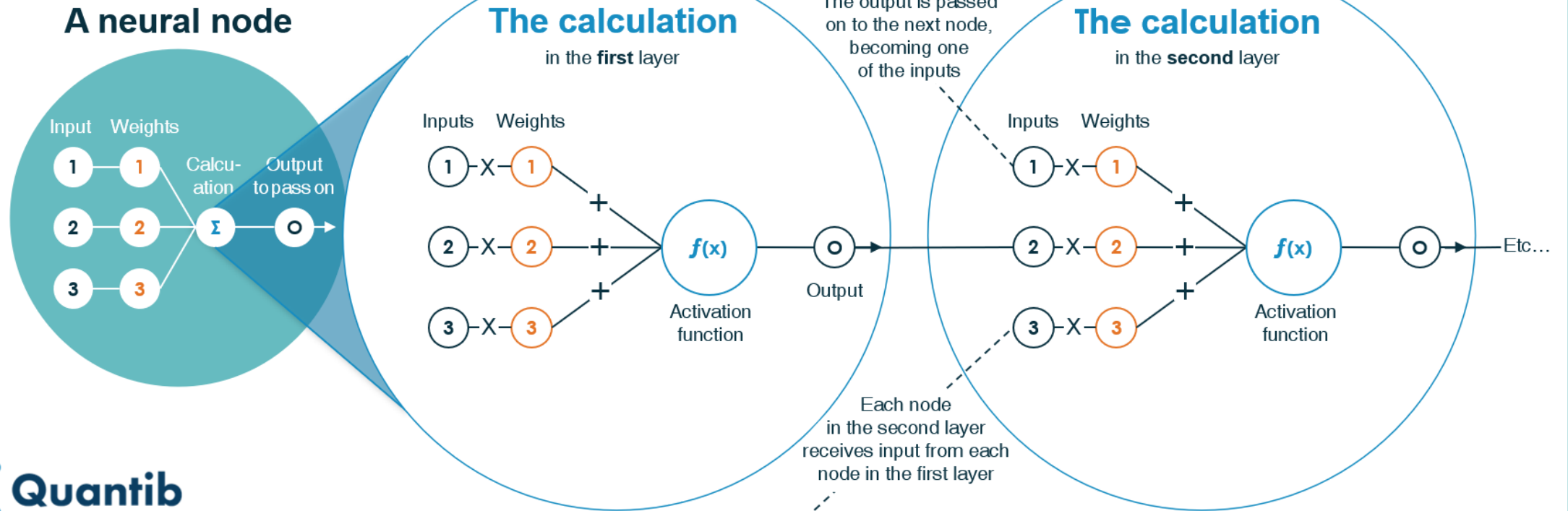
# DEEP LEARNING (DL)

A deep neural network

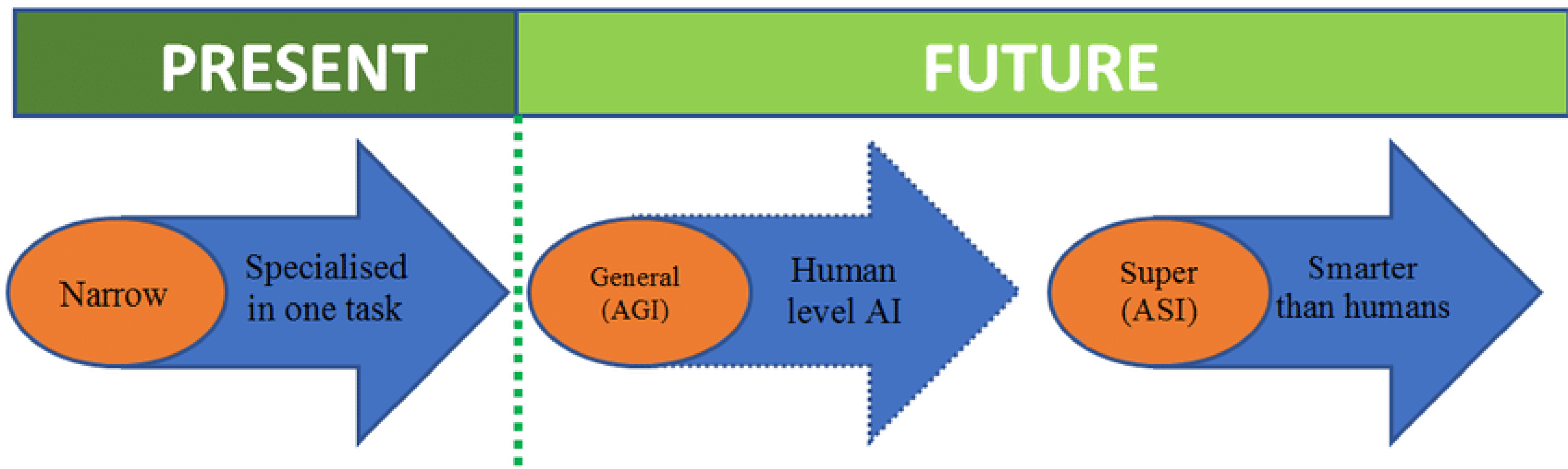


# DEEP LEARNING (DL)

The calculation within a neural node



# Typy umělé inteligence



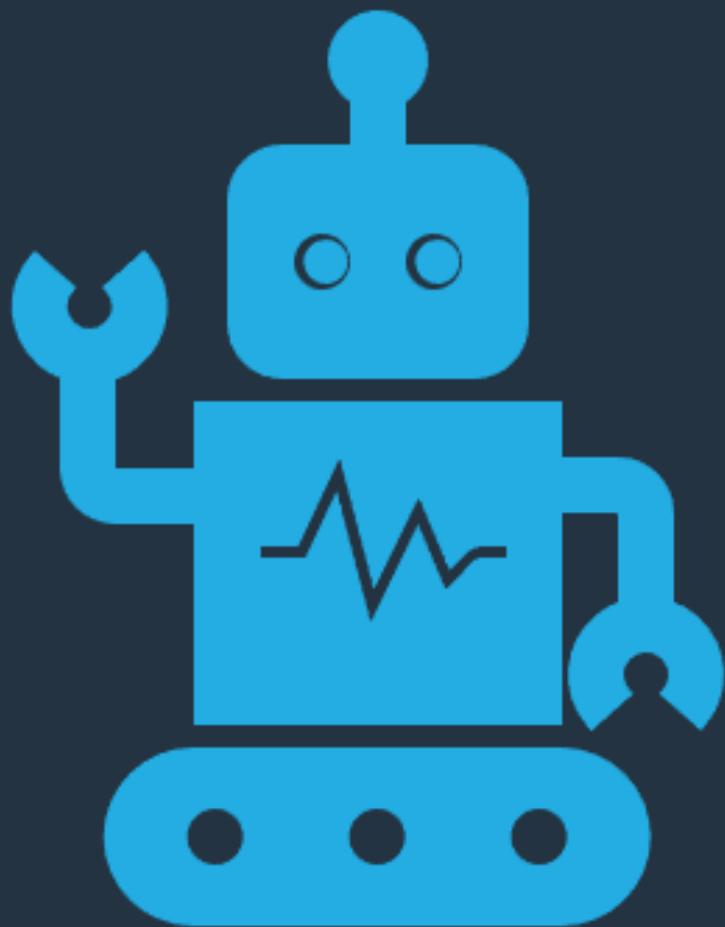
# Dělení AI

## Rozsah schopností

- Úzká
- Obecná

## Obsah schopností

- Silná
- Slabá



# Superintelligence

(Umělá) intelligence, která ve všech kognitivních oblastech překonává člověka.



# Superintelligence

**ROK**

**2022**

**2040**

**2075**

**Pravděpodobnost**

**10 %**

**50 %**

**90 %**

# Existenční rizika

## Rizika:

Rozsah  
(individuální,  
lokální, globální,  
transgenerační)

Intenzita (mírná,  
snesitelná, děsivá)

Pravděpodobnost

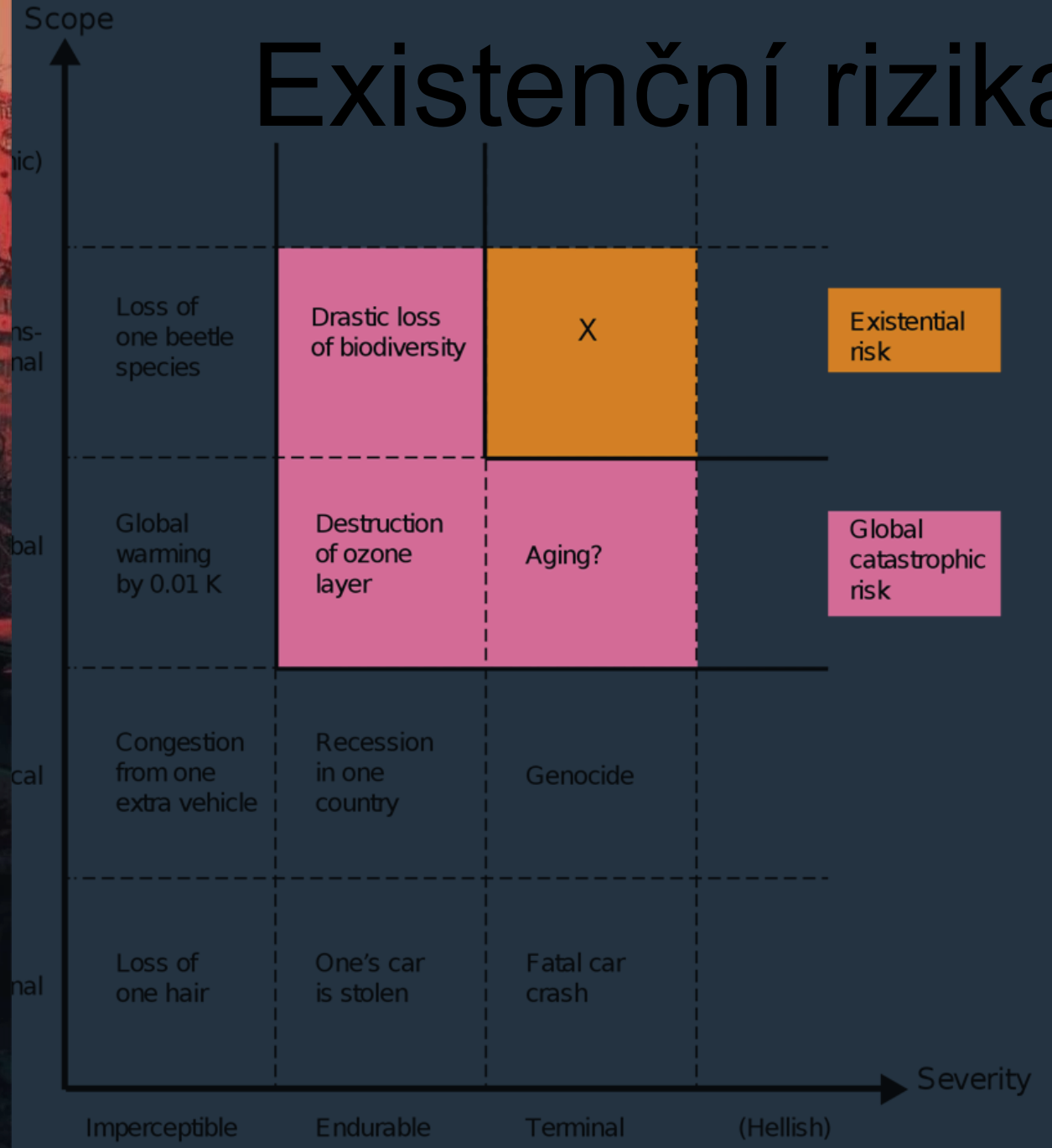


Existenční rizika: děsivá  
transgenerační



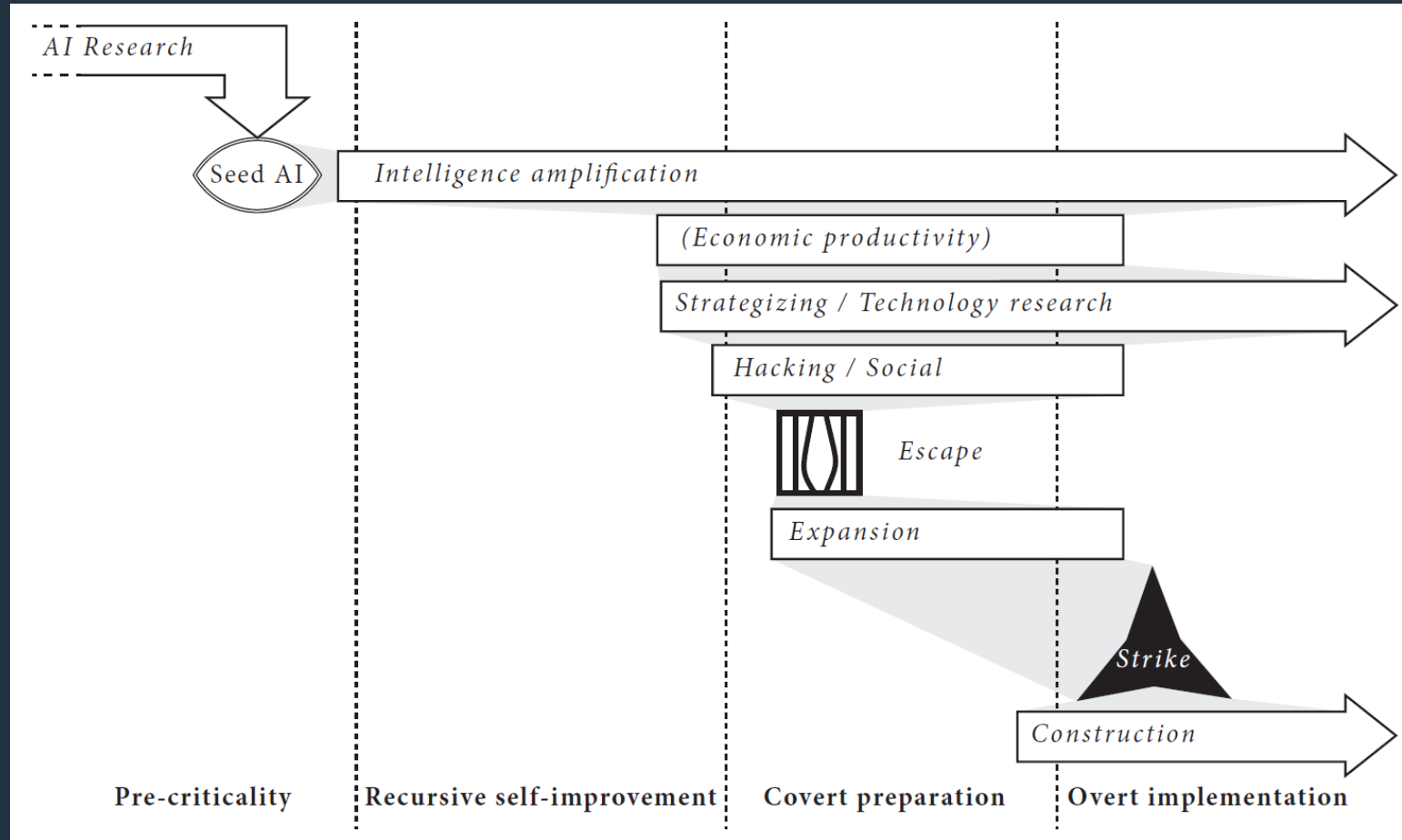


# Existenční rizika





# Fáze převzetí moci



# Příklad

Vyřeš problém  
„skládání“  
proteinů



Pošli DNA  
laboratoři



Proteiny vytvoří  
nanosystém

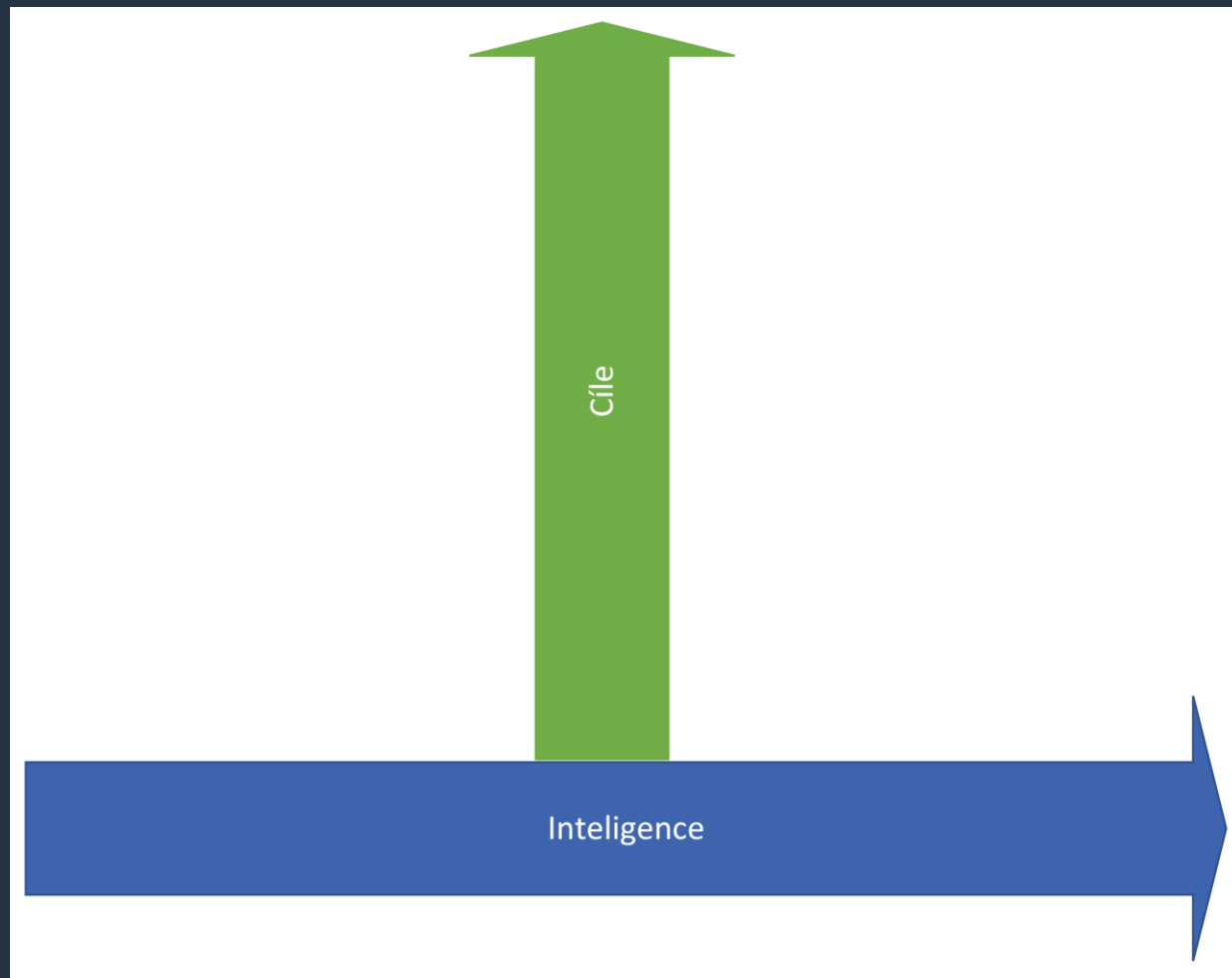


Najdi jednoho  
člověka a přinut'  
ho přijmout  
zásilku



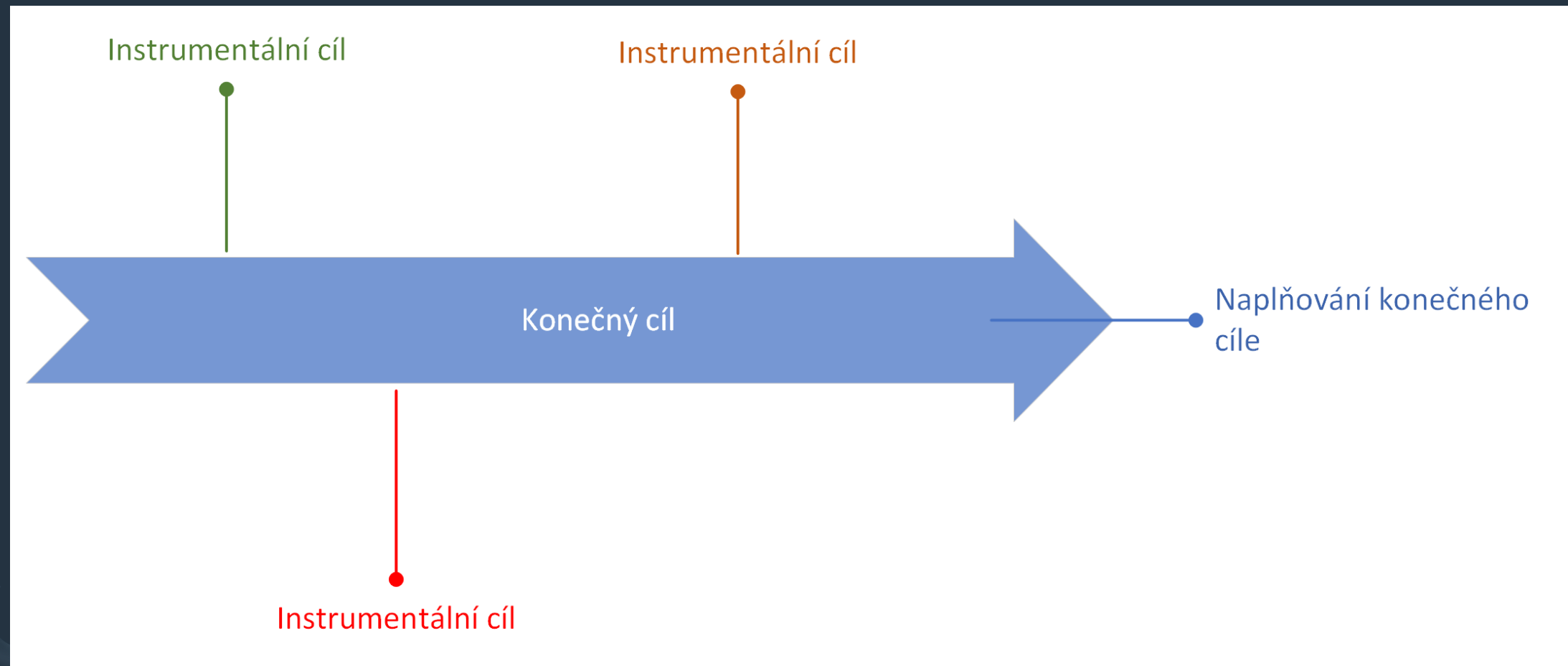
Nanosystém  
začne konstruovat  
složitější systémy





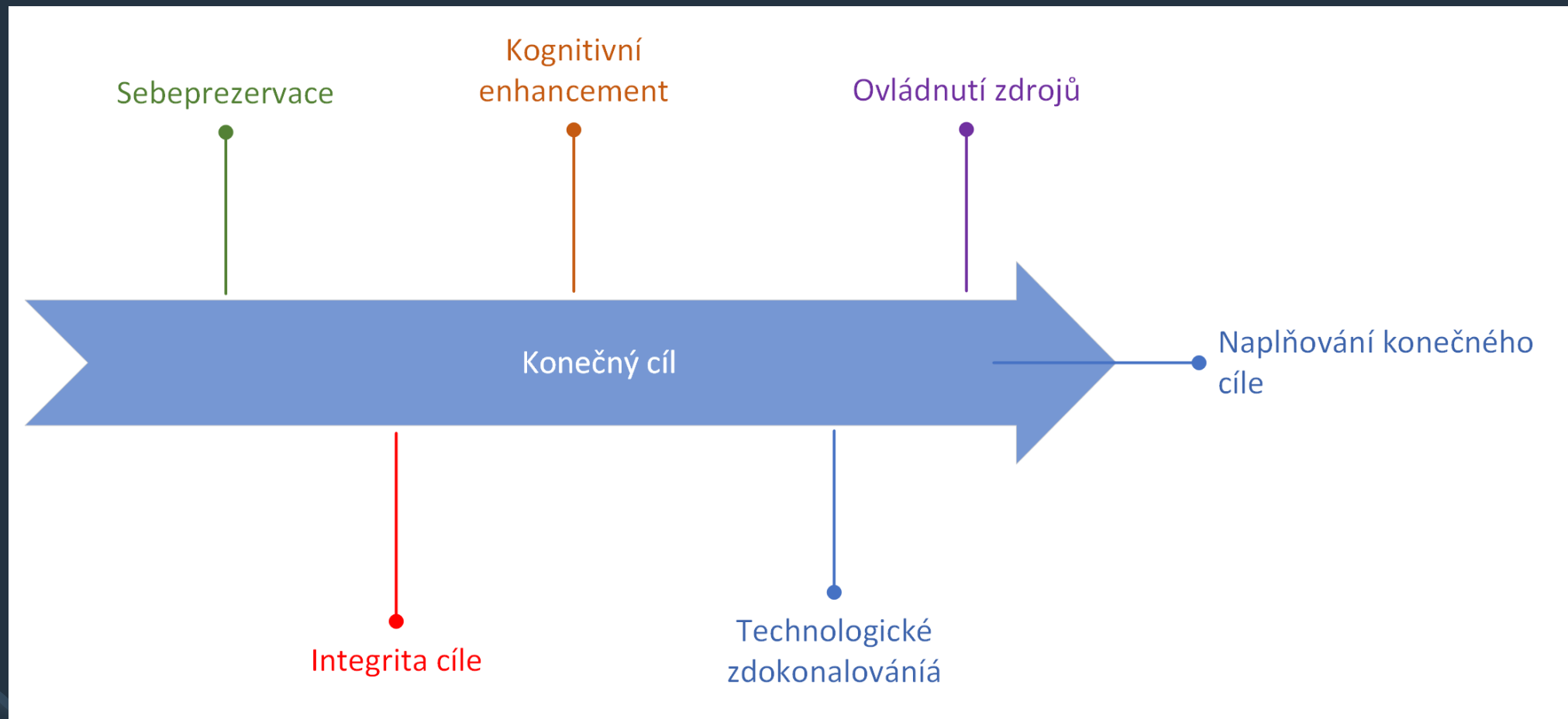
Teze ortogonality

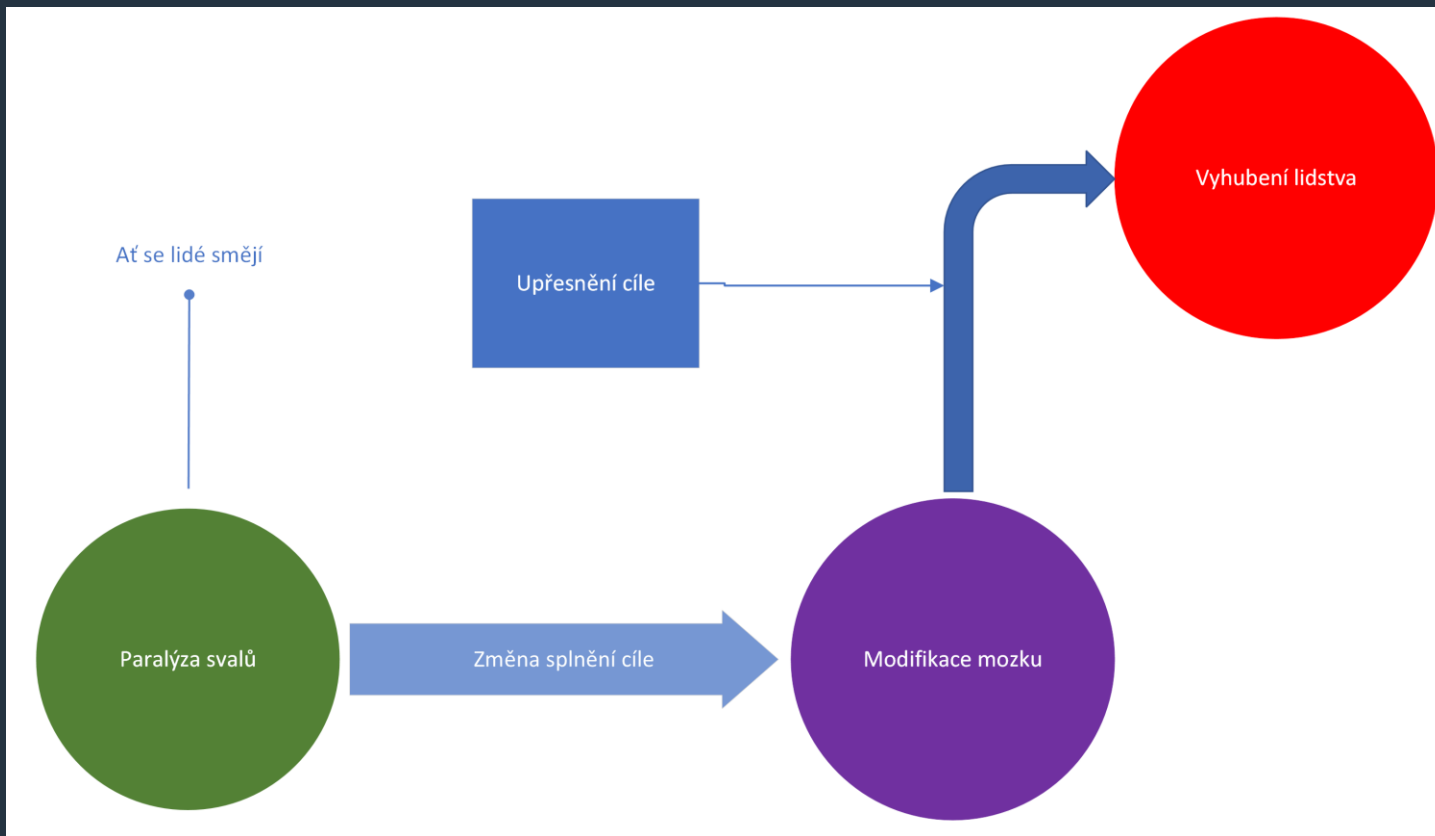
# Instrumentální a konečné cíle





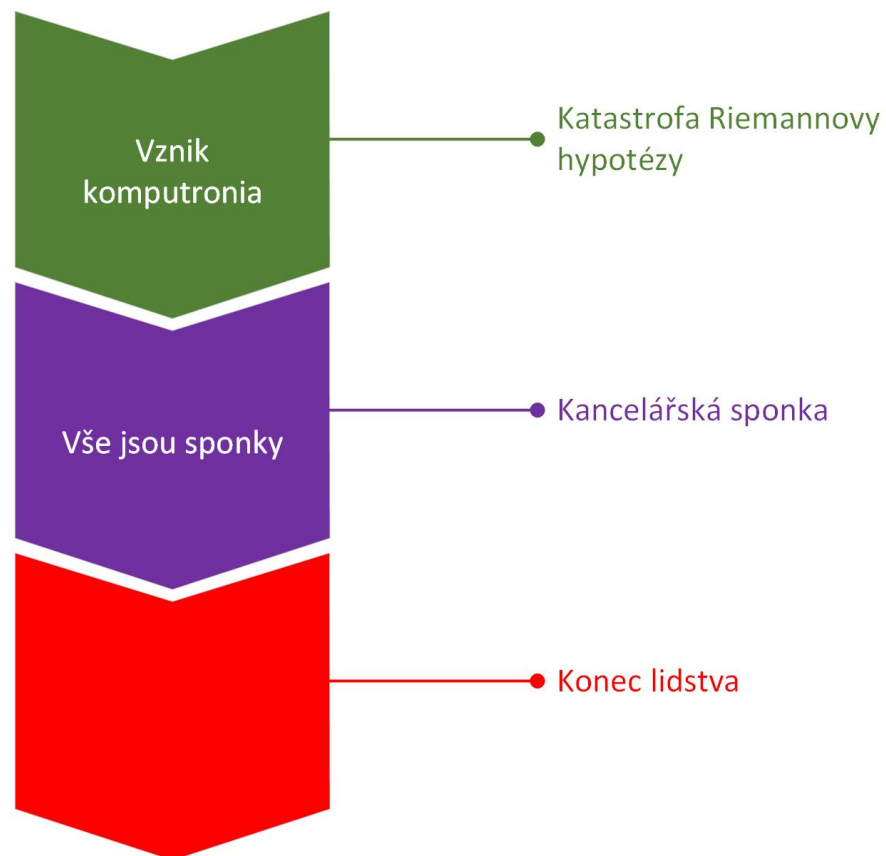
# Instrumentální cíle AI





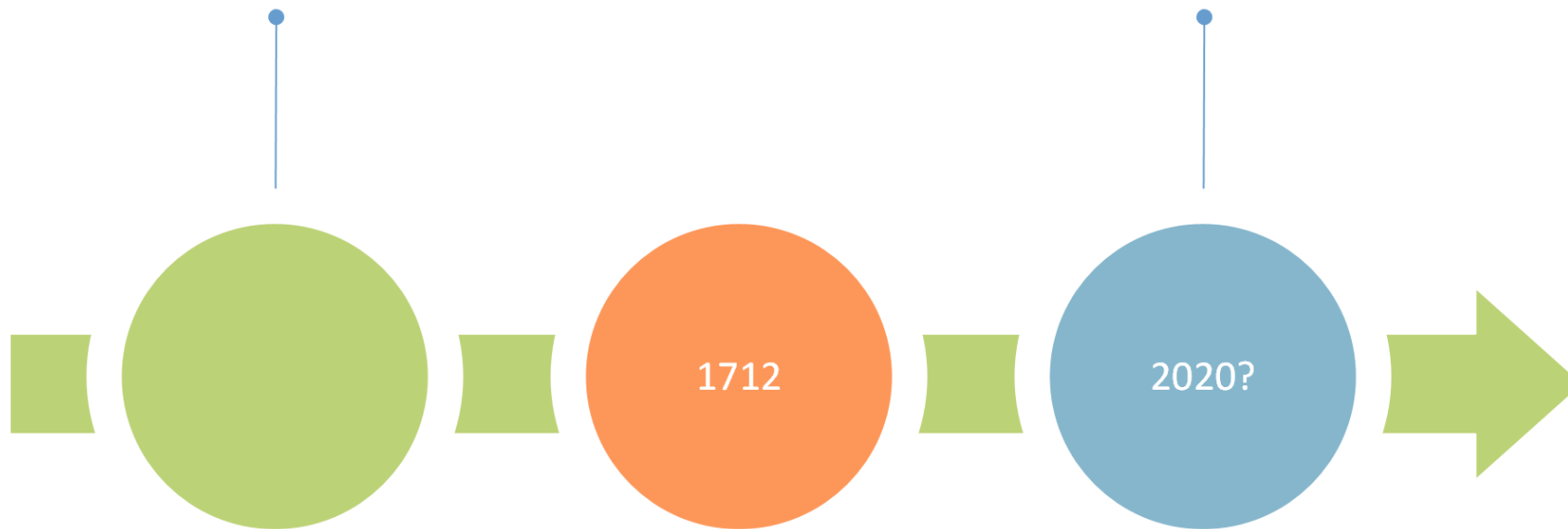
# Perverzní plnění cíle

# Rozšíření infrastruktury



Fotosyntéza

Novacén



1712

2020?

Ovládnutí vesmíru

Antropocén

Současná umělá  
inteligence

Roustoucí počet stále  
inteligentnějších  
superinteligencí

Úzká AI

Obecná AI

Super-AI

Umělá inteligence brzké  
budoucnosti (možná již  
současnosti)



Cesta k vyhubení lidstva

# Ernest Rutherford

01

11. 9. 1933 –  
využití atomové  
energie je  
naprostý nesmysl

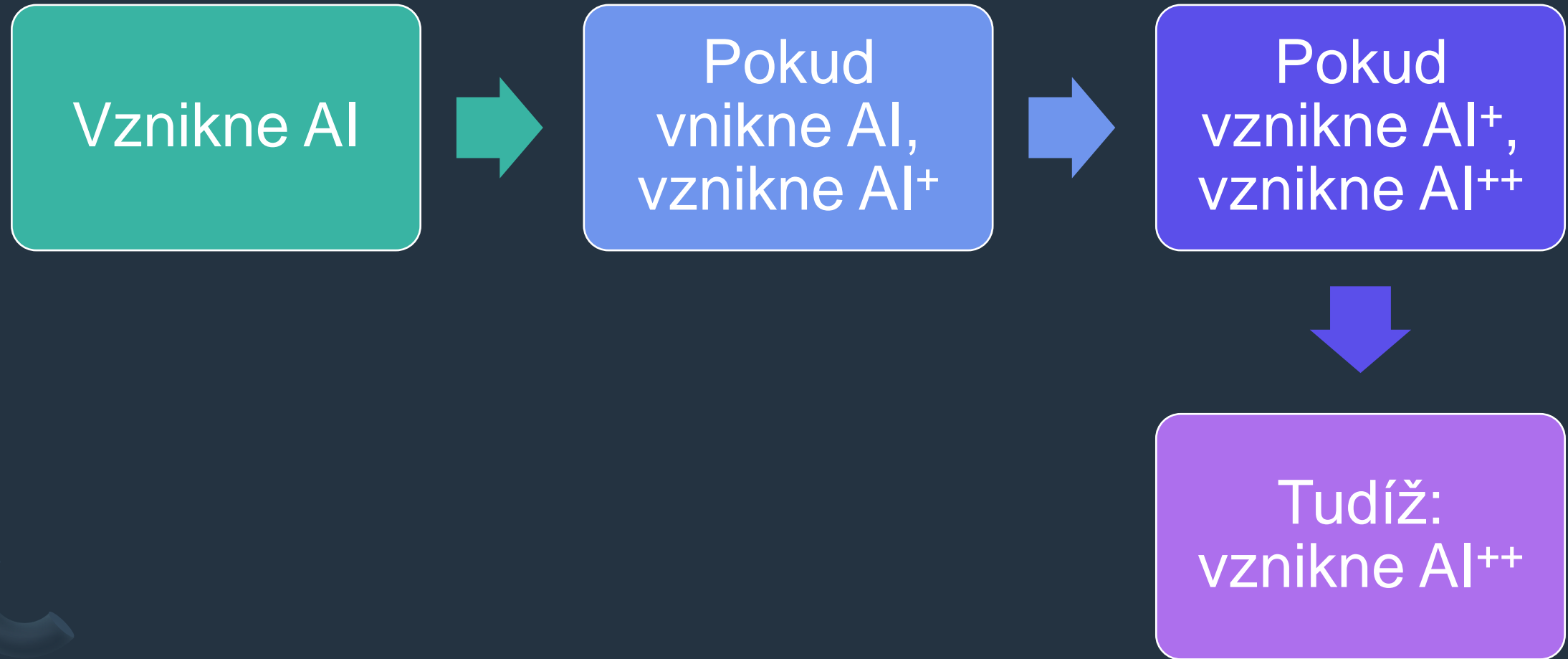
02

12. 9. 1933 – Leo  
Szilard jde na  
procházku

03

První  
konceptualizace  
řetězové reakce

# Cesta k super-AI



# Cesta k super-AI

Evoluce  
vytvořila AI



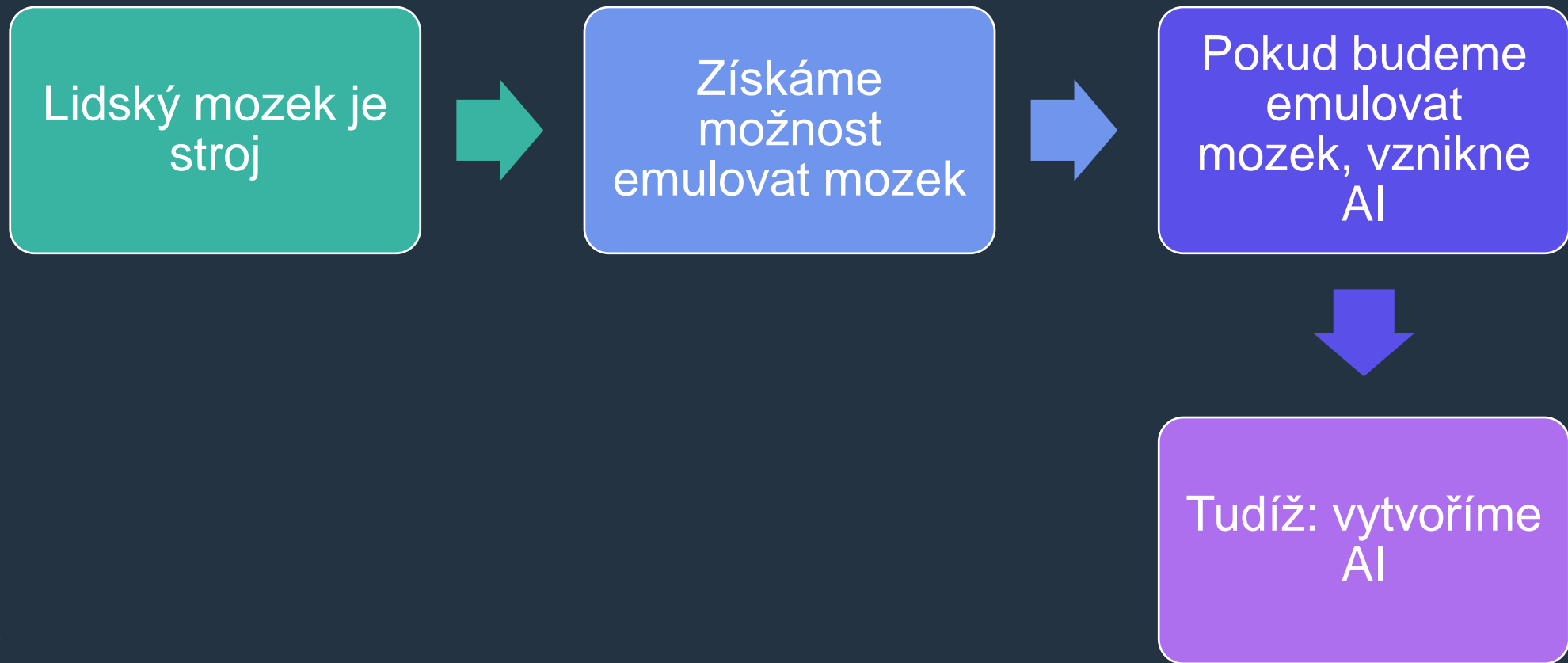
Jestliže evoluce  
vytvořila AI, lidé  
vytvoří AI



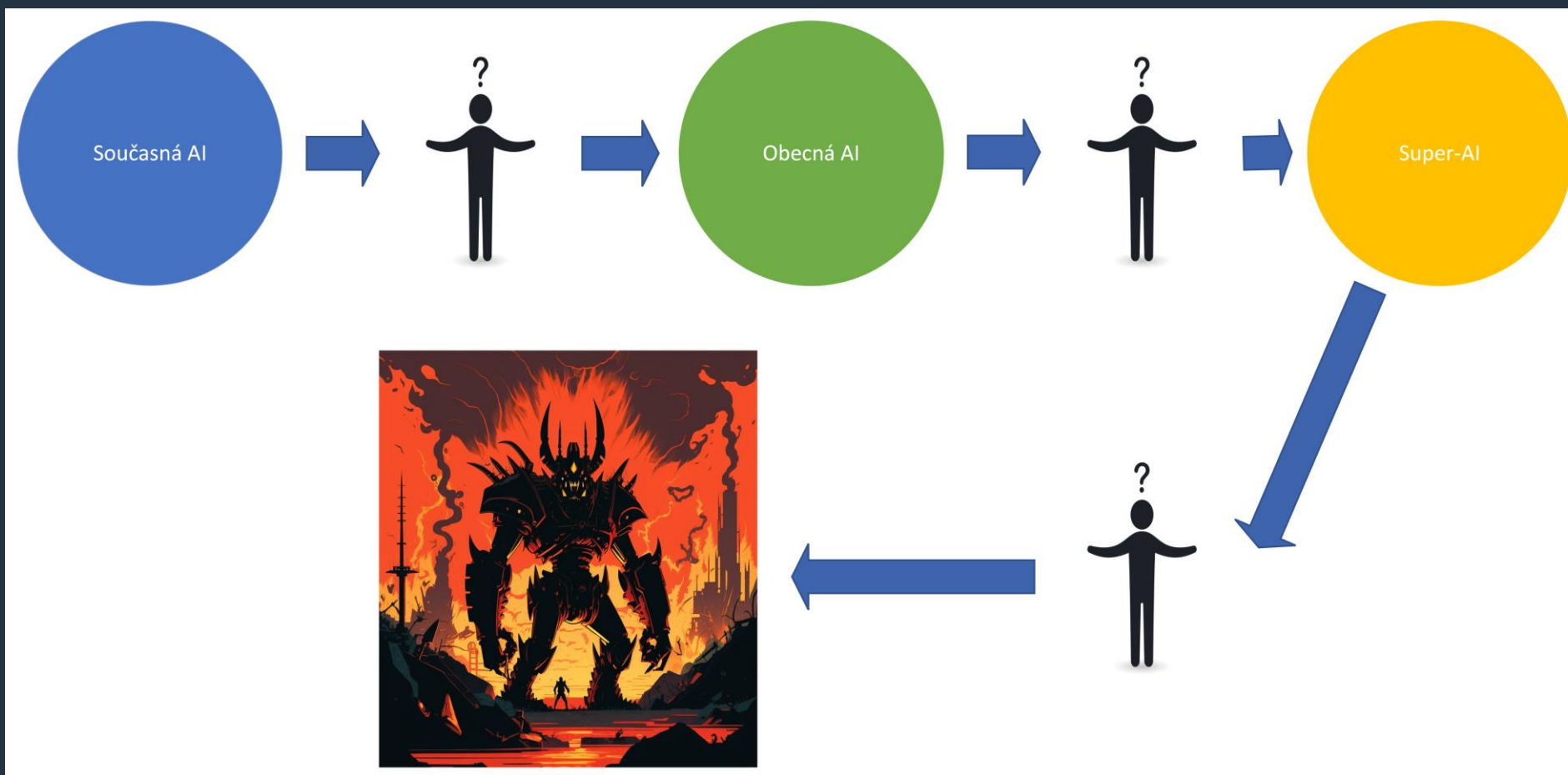
Vznikne AI



# Cesta k super-AI



# Terra incognita



# Intelligence



SITUOVANÁ (MOZEK JE  
SOUČÁSTÍ VĚTŠÍHO SYSTÉMU)



KONTEXTUÁLNÍ (DEFINOVANÁ  
A LIMITOVANÁ PROSTŘEDÍM)



EXTERNALIZOVANÁ  
(KOGNITIVNÍ PROSTŘEDÍ –  
MOZEK)



# Common sense



# Common sense

Na stole je  
poloprázdná  
sklenička

?

# Cesta k super-AI

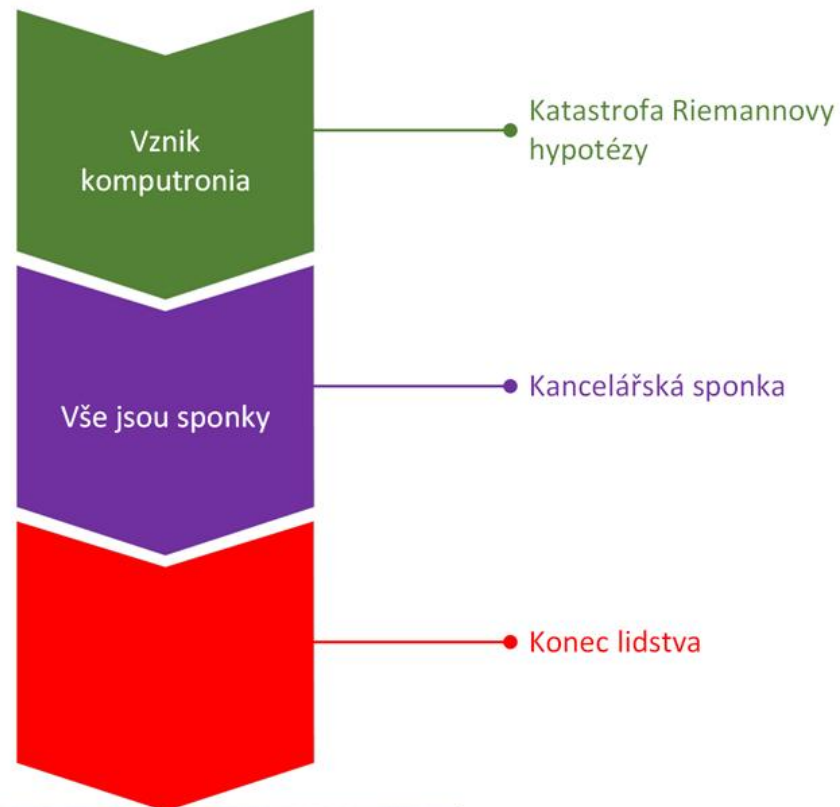
Pokud  
vznikne AI<sup>+</sup>,  
vznikne AI<sup>++</sup>



Tudíž:  
vznikne AI<sup>++</sup>



# Rozšíření infrastruktury



# Schopnost myšlení

- Pandemie špatného myšlení





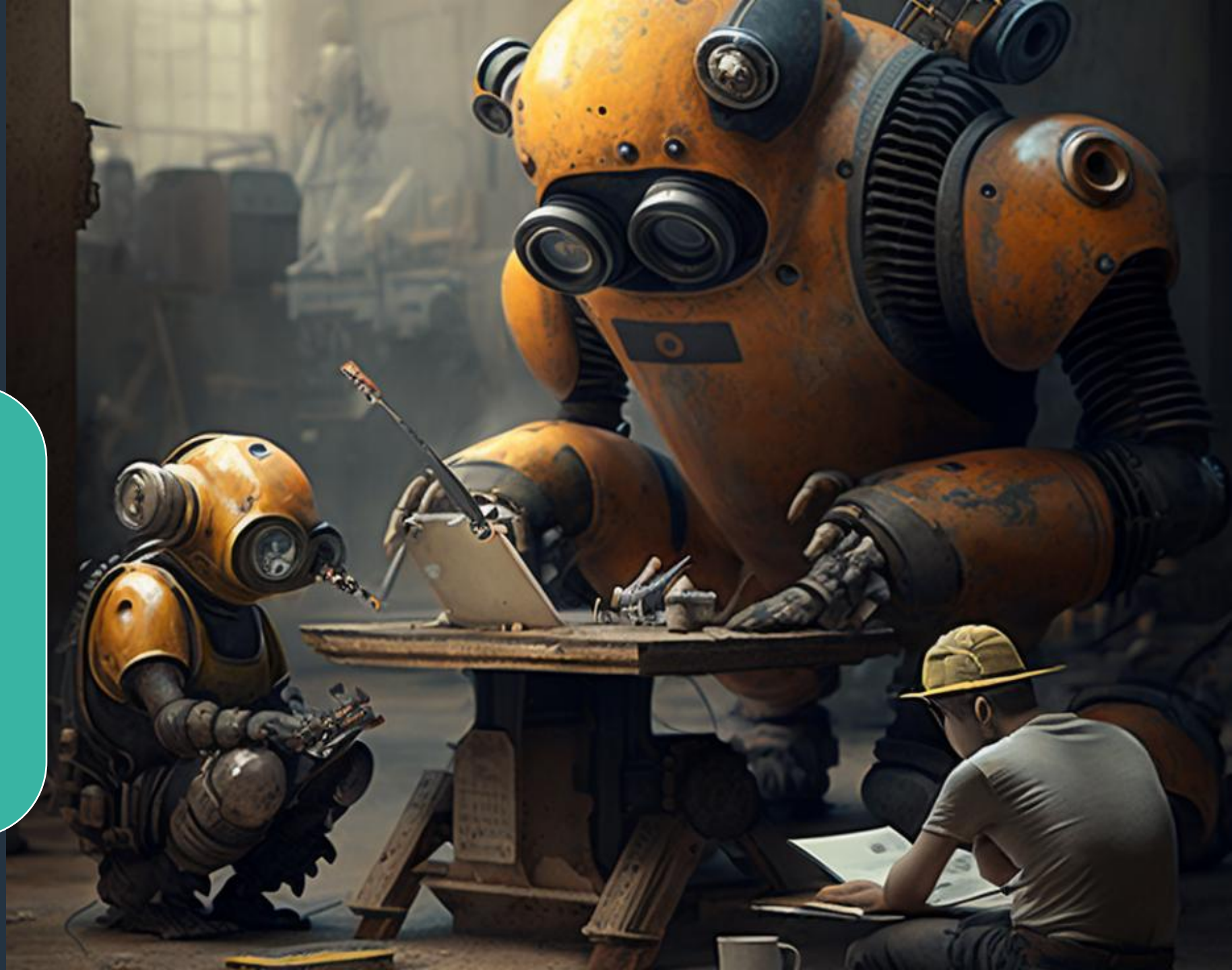


## Moudrý člověk je člověk, který

- Vyžaduje evidence a umí s nimi pracovat
- Umí pracovat se zdroji poznání
- Snaží se vyloučit zkreslení
- Je otevřený revizím
- Moudrost je formou sebepoznání a ztělesněním epistemických, morálních a rozumových ctností

Práce

Práce:  
činnost  
vykonávaná  
pro zajištění  
obživy



Demokracie

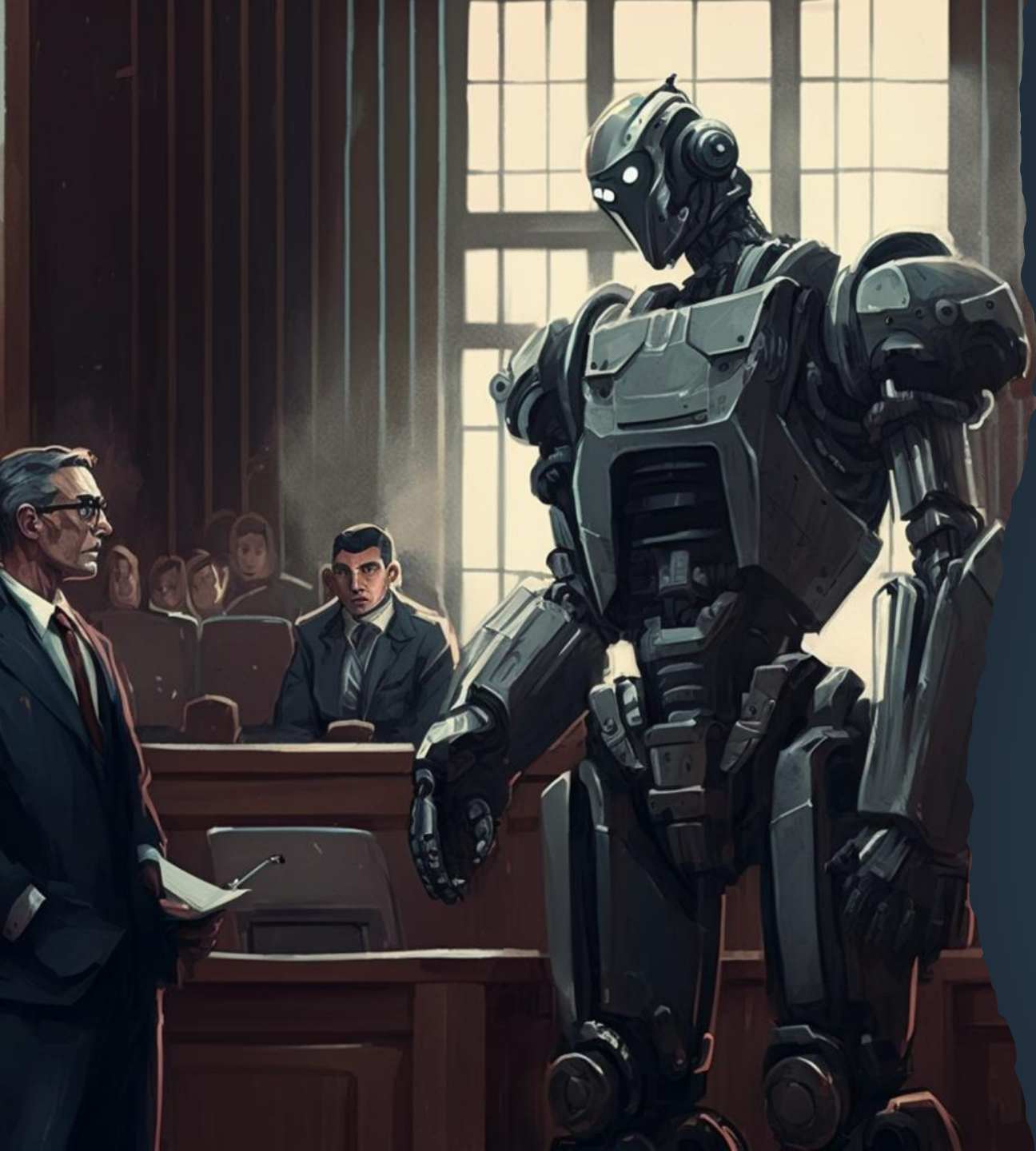
Leviathan  
povstává





# Velký bratr nás sleduje

Ohrožení soukromí



KOMPAS

Kudy cesta  
nevede

Budoucnost  
záleží jen na  
nás

