

Rovnosť, umelá inteligencia, algoritmy  
a automatizované rozhodnutia:

ako odhaľovať a riešiť diskrimináciu  
v kontexte umelej inteligencie?

Miroslav Chlipala



Financované v rámci programu Európskej komisie „Práva, rovnosť a občianstvo na obdobie rokov 2014-2020“

Časté používanie pojmu

<>

rôzny výklad významu

Neexistuje žiadna oficiálna definícia

Tradícia sci-fi

Všeobecná umelá inteligencia

vs.

Úzka umelá inteligencia

## Všeobecná umelá inteligencia

predstavuje stroj, ktorý dokáže zvládnuť akúkoľvek intelektuálnu úlohu

## Úzka umelá inteligencia

označuje systém, ktorý vykonáva jeden úkon

Všeobecná umelá inteligencia je doménou science-fiction

Všetky súčasné metódy umelej inteligencie patria  
do oblasti úzkej umelej inteligencie!

## Kľúčové pojmy

### Autonómnosť

Schopnosť vykonávať úlohy v zložitom prostredí bez neustáleho usmerňovania zo strany používateľa

### Adaptívnosť

Schopnosť zlepšovať výkonnosť učením sa zo skúseností

# Obrovský potenciál umelej inteligencie

Zdravie

Doprava

Energia

Poľnohospodárstvo

Cestovný ruch

Kybernetická bezpečnosť

Dôveryhodná umelá inteligencia

Dôvera ako nevyhnutný predpoklad  
Etická technológia



Európska koncepcia  
dôveryhodnej umelej inteligencie

Nariadenie o umelej inteligencii

## Koncepcia rizika

Neprijateľné riziko

Vysoké riziko

Obmedzené riziko

Minimálne riziko

Neprijateľné riziko

Zakázané systémy umelej inteligencie

Ohrozenie bezpečnosti, života a práv ľudí

## Vysoké riziko

Systemy aj s anti-diskriminačným dopadom

napr. vzdelávanie (bodovanie testov)

riadenie zamestnancov (triedenie životopisov)

Vysoké riziko

Prísne povinnosti:

Posudzovanie a zmierňovanie rizika

Vysoká kvalita súborov údajov vstupujúcich do systému AI

Zaznamenávanie činnosti (log management)

Podrobná dokumentácia

Jasné a primerané informácie pre používateľa

Primeraný ľudský dohľad

Vysoká úroveň spoľahlivosti, bezpečnosti a správnosti

## Vysoké riziko

Biometrická identifikácia na d'ial'ku

Výnimky (s povolením súdu, iného nezávislého orgánu)

Pátranie po nezvestnom dieťati

Zabránenie konkrétnej a bezprostrednej teroristickej hrozbe

Páchateľ závažného trestného činu

Obmedzené riziko

(napr. chatboty)

Vedomosť používateľa  
pre jeho rozhodnutie či bude používať takýto systém

Minimálne riziko

Videohry s AI  
Spamové filtre

Predstavuje iba minimálne riziko pre práva a bezpečnosť  
(Nariadenie sa tieto systémy nevzťahuje)



## Návrh Komisie

Vnútroštátne orgány dohľadu  
Európska rada pre umelú inteligenciu

# Autonómne vozidlá

(roboty na doručovanie zásielok, drony či autonómne plavidlá)

## Odporúčaný obsah

Zatiaľ čo titulka tlačenej verzie „novín“ je rovnaká pre všetkých čitateľov, úvodná stránka online verzie je pre každého čitateľa  
iná

Algoritmy, ktoré určujú zobrazovaný obsah, sú založené na  
umelej inteligencii

## Spracovanie obrazu a videa

(rozpoznávanie tváre – triedenie fotografií, automatické označovanie v sociálnych médiách)

Generovanie alebo úprava digitálneho obsahu – filtre fotografií

Počítačovo generované postavy vo filmoch  
(napodobovanie ľudských gest)

Ďakujem za pozornosť!

Vaše otázky..?