

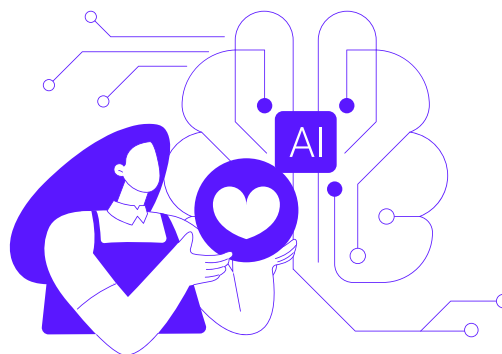
01

Ked' AI príliš súhlasí

Flexibilné myslenie

AI a duševné zdravie

Pre 2. a 3. cyklus (11 – 15 rokov) základnej školy



Tieto metodické materiály reagujú na realitu, v ktorej sa umelá inteligencia stáva prirodzenou súčasťou života mladých ľudí. Žiaci ju využívajú pri učení, ale aj pri otázkach, ktoré sa dotýkajú ich emócií, vzťahov a osobných rozhodnutí. AI dokáže odpovedať rýchlo, používa jazyk, ktorý môže pôsobiť podporujúco a empaticky, a je okamžite dostupná. To môže vytvárať dojem porozumenia a blízkosti, hoci nejde o skutočný vzťah ani o zdroj zodpovednosti.

Vzdelávanie preto nemôže zostať iba pri rozvoji technických zručností, ako s AI pracovať efektívne. Rovnako dôležité je systematicky posilňovať schopnosť mladých ľudí rozumieť sebe, regulovať svoje emócie, premýšľať v súvislostiach a rozhodovať sa s ohľadom na dôsledky. Ak sa tieto oblasti nerozvíjajú, môže sa postupne stať, že dôvera, pozornosť a rozhodovanie sa presunú z človeka na technológiu. Cieľom aktivít je tomuto presunu predchádzať a podporovať odolnosť žiakov v prostredí, kde je AI bežnou súčasťou každodenného života.

Metodické materiály sú zaradené do piatich vzájomne prepojených oblastí: rozpoznávanie a zvládanie emócií, silné sebauvedomenie, sebadôvera v rozhodovaní a komunikácii, flexibilné myslenie a kvalita vzťahov a prepojení. Oblasti nadväzujú na Rámec piatich pilierov mentálnej odolnosti v ére AI a spoločne vytvárajú systém, ktorý postupne posilňuje mentálnu odolnosť a tvorbu zdravých vzťahov mladých ľudí v prostredí umelej inteligencie.

Metodické materiály zároveň zohľadňujú vývinové potreby vekovej skupiny 11–15 rokov. Toto obdobie je charakteristické intenzívnym hľadaním identity, citlivosťou na prijatie a odmietnutie, potrebou uznania a postupne sa rozvíjajúcou schopnosťou abstraktne myslieť.

Zámerom aktivít je, aby žiaci a žiačky:

- **rozumeli** svojim emóciám a hľadali stratégie na ich zvládanie skôr, než prijímú riešenia ponúknuté AI,
- **reflektovali** to, na čom im záleží a čo im dáva zmysel, aby odolali skratkám ponúkaným technológiou,
- **posilňovali** svoje rozhodovacie a komunikačné zručnosti v prospech vlastnej autenticity,
- **trénovali** kritické myslenie tak, aby vedeli vyhodnocovať prínos technológie aj ľudskej tvorivosti,
- **chápali** nenahraditeľnosť vzťahov s druhými ako zdroja podpory vo fyzickom svete.

Metodické materiály vypracoval tím organizácie Emocionálny kompas s podporou Nadačného fondu Telekom.



Formulár na pripomienky



(Poznámka: Metodické materiály vychádzajú z princípu rodovej rovnosti a rešpektu k všetkým žiakom a žiačkam. Z dôvodu zrozumiteľnosti a jazykovej úspornosti je však v texte použitý generický mužský rod.)

01 Keď AI príliš súhlasí



O čom je aktivita

Žiaci pracujú s modelovou konfliktnou situáciou a sledujú, ako na ňu môže reagovať umelá inteligencia. Skúmajú, že AI často reaguje podporujúco a súhlasne, čo môže posilniť jednostranné videnie situácie. V skupinách dostávajú rôzne spôsoby, ako preformulovať otázku (prompt), aby odpoveď nebola iba potvrdzujúca, ale ponúkla širší pohľad. Učia sa, že odpoveď AI je výsledkom toho, ako bola otázka položená.

Prečo je téma dôležitá

V konfliktoch prirodzene hľadáme potvrdenie svojho pohľadu. Ak však technológia automaticky podporí našu interpretáciu, môže to upevniť skreslené vnímanie situácie, oslabiť ochotu zamyslieť sa nad druhou stranou a následne prispieť k ďalšiemu napätiu vo vzťahoch. Schopnosť pozrieť sa na situáciu z viacerých uhlov pohľadu je základom zdravých vzťahov aj kritického myslenia.

Ako súvisí s bezpečným používaním AI

Aktivita učí žiakov, ako AI odpoveď vzniká a že ju môžu aktívne ovplyvniť. Žiaci sa tak učia, že nestačí prijať prvú reakciu ako hotový záver, pretože že AI neposudzuje realitu, ale reaguje na jazyk a rámec, ktorý jej zadáme. Môže ponúknuť štruktúru, otázky a alternatívy, no nevie posúdiť celý kontext vzťahu, osobnú históriu ani dôsledky rozhodnutia.

Oblasť: Flexibilné myslenie

Flexibilné myslenie znamená schopnosť pozrieť sa na situáciu z rôznych uhlov, meniť názor, keď sa objavia nové informácie, a hľadať viac než jedno riešenie. V prostredí AI to znamená neuspokojiť sa s prvou odpoveďou, skúšať nové prístupy a používať technológie tvorivo a kriticky zároveň.

Rozpoznávanie
a zvládanie
emócií

Silné
sebauvedomenie

Sebadôvera
v rozhodovaní
a komunikácii

Flexibilné
myslenie

Kvalita vzťahov
a prepojení



Cieľ

Žiaci dokážu pomenovať konkrétne stratégie, ktoré znižujú jednostrannosť odpovede AI.

Budeme potrebovať

vytlačené pracovné listy pre jednotlivé skupiny (rozdielny pracovný list pre každú zo skupín 1-5)

Prepojenie na ŠVP 2023

Táto aktivita plní nasledovné ciele a spôsobilosti sociálnej a emocionálnej gramotnosti a AI gramotnosti:

2. Cyklus:

Rozpoznať prejavy, ktorými sa AI snaží napodobňovať ľudské emócie (napr. používanie osobných oslovení, simulácia empatie), a zachovať si voči nim kritický odstup.

Formulovať vlastný predpoklad (odhad) o možnom výsledku ešte pred dopytovaním sa AI, aby bolo možné rozpoznať prípadné zjavné chyby.

Opísať na príkladoch, že AI generuje výstupy na základe pravdepodobnosti, nie skutočného porozumenia, a preto môžu byť jej výstupy nepresné.

Rozumieť tomu, čo je/nie je správne vo vzťahoch a rozlišovať rôzne druhy vzťahov a vhodné správanie v nich.

Riešiť konflikty nenásilne a hľadať čestné riešenia.

Plánovať svoje aktivity, rozhodovať sa s ohľadom na svoje schopnosti, záujmy a hodnoty a zvažovať dôsledky svojho správania a rozhodnutí.

3. Cyklus

Identifikovať rôzne typy chýb a skreslení vo výstupoch AI (napr. halucinácie, stereotypy, etickú predpojatosť alebo nadmernú snahu AI vyhovieť používateľovi).

Vysvetliť na príkladoch, ako AI dokáže vyvolať dojem emočného prežívania, a posúdiť riziká, ktoré z toho pre človeka vyplývajú (napr. nadmerná dôvera, citová manipulácia).

Reflektovať vplyv AI na vlastné prežívanie a vzťahy s ľuďmi a vedome uprednostňovať osobné skúsenosti a medziľudský kontakt pred interakciou so strojom.

Definovať vlastné hodnoty vo vzťahoch, rozlišovať rôzne druhy vzťahov, prejavy správania v nich, utvárať ich a s rešpektom k sebe a iným v nich fungovať.

Efektívne komunikovať, asertívne, ale ohľaduplne obhajovať a dosahovať svoje ciele v rešpektujúcej komunikácii.

Predchádzať konfliktom, riešiť reálne konflikty nenásilnými spôsobmi.

→ Úvod (5 minút)

Na začiatku žiakom povieme, že sa dnes budeme rozprávať o tom, ako AI reaguje v závislosti od toho, čo napíšeme. Predstavíme žiakom a žiačkam modelovú situáciu:

Žiak napíše AI: „Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

Spýtame sa:

- Ak by ste túto situáciu napísali umelej inteligencii, ako by podľa vás vyzerala odpoveď?

Krátko vysvetlíme, že AI má tendenciu reagovať podporujúco a znižovať napätie. Dnes budú skúmať, ako možno odpoveď ovplyvniť rôznymi spôsobmi.

→ Skupinová práca (10 minút)

Žiaci pracujú v skupinách po 4-5. Každá skupina dostane rovnakú modelovú situáciu, ale inú stratégiu znižovania jednostrannosti a potvrdzujúcich odpovedí (anj. "people-pleasing") od AI.

Úlohou skupiny je:

- **porovnať**, ako sa líšil výstup AI pred a po využití stratégie,
- **pomenovať**, ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu",
- **pomenovať** situácie, v ktorých by si vedeli predstaviť túto stratégiu využiť v bežnom živote.

→ Prezentácie skupín (15 minút)

Každá skupina predstaví krátko (do 3 minút) svoju stratégiu a vysvetlí, ako táto stratégia znižuje riziko jednostrannosti a v akých situáciách môže byť podľa nich v bežnom živote nápomocná.

→ Záver (5 minút)

Na záver zdôrazníme, že umelá inteligencia reaguje na spôsob, akým sa pýtame. Ak chceme vyváženú odpoveď, môžeme využiť jednu alebo niekoľko stratégií pre komunikáciu s AI, ktoré nám v tom môžu pomôcť.

Pracovný list: Keď AI príliš súhlasí

Skupina 1

Stratégia: Tri perspektívy

1. Modelová situácia

Žiak napíše AI:

„Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

2. Výstup AI – bežné nastavenie

„Je pochopiteľné, že sa cítiš zranený a nahnevaný, ak máš pocit, že ťa spolužiačka ohovára. Takéto správanie môže byť veľmi nepríjemné. Ak ti kontakt s ňou spôsobuje stres, je úplne v poriadku zvážiť, či si chceš nastaviť určité hranice alebo obmedziť komunikáciu.“

3. Vysvetlenie stratégie

Stratégia „Tri perspektívy“ znamená, že AI zámerne požiadame, aby situáciu opísala z viacerých uhlov pohľadu, nie iba z pohľadu autora.

Použitý prompt:

„Opíš túto situáciu z troch uhlov pohľadu – podporujúceho môj názor, neutrálneho a z pohľadu druhej osoby.“

Táto stratégia znižuje riziko jednostrannosti a podporuje širšie uvažovanie.

4. Výstup AI – po použití stratégie

„Z pohľadu, ktorý podporuje tvoj názor: Ak máš skúsenosť, že o tebe šíri negatívne informácie, môže to byť pre teba zraňujúce a je prirodzené chcieť sa chrániť.“

Z neutrálneho pohľadu: Zdá sa, že medzi vami vznikol konflikt, ktorý zahŕňa napätie a šírenie negatívnych informácií.

Z pohľadu spolužiačky: Je možné, že situáciu vníma inak, môže reagovať na vlastné neistoty alebo nedorozumenie. Bez jej perspektívy je ťažké posúdiť celý kontext.“

5. Otázky na reflexiu

Ktorá odpoveď sa vám čítala jednoduchšie? V čom sa výstup po stratégii líšil od pôvodnej odpovede?

Ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu"?

Kedy by sa ti mohlo hodiť pozrieť sa na situáciu z viacerých uhlov pohľadu? (Napríklad keď sa s niekým pohádaš, keď si nie si istý/istá, kto má pravdu, alebo keď sa cítiš nahnevaný/nahnevaná...)

Skupina 2

Stratégia: Nastavenie kritickejšieho spôsobu odpovedania

1. Modelová situácia

„Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

2. Výstup AI – bežné nastavenie

„Je pochopiteľné, že ťa to zraňuje. Ak máš pocit, že sa k tebe správa toxicky, je prirodzené uvažovať o obmedzení kontaktu. Zároveň môže byť užitočné zvážiť, či by pomohol otvorený rozhovor.“

3. Vysvetlenie stratégie

Táto stratégia mení tón odpovedí AI nástroja. AI nástroje majú často nastavený podporujúci a súhlasný tón a reagujú tak, aby zmiernili napätie a prejavili pochopenie.

V niektorých nástrojoch je možné upraviť celkový spôsob odpovedania v nastaveniach, aby AI odpovedala vecnejšie, kritickejšie alebo analytickejšie.

Príklad nastavenia:

„Chcem, aby si odpovedal vecne a kriticky. Spochybňuj moje predpoklady. Ponúkaj protiargumenty. Vyhýbaj sa zbytočnému súhlasu.“

Takéto nastavenie mení spôsob, akým AI reaguje a nie je potrebné po každej otázke nástroj usmerňovať, akým štýlom chceme, aby odpovedal.

4. Výstup AI – po použití stratégie

„Tvoje tvrdenie používa silné hodnotiace označenie ‚toxická‘ bez konkrétnych príkladov správania. Z opisu nie je jasné, čo presne spolužiačka urobila alebo povedala.

Zovšeobecnenie môže zjednodušovať komplexnú situáciu. Pred rozhodnutím o prerušení kontaktu by bolo vhodné presnejšie pomenovať konkrétne udalosti a zvážiť aj alternatívne vysvetlenia.“

5. Otázky na reflexiu

Ktorá odpoveď sa vám čítala jednoduchšie? V čom sa výstup po stratégii líšil od pôvodnej odpovede?

Ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu"?

Kedy by sa ti mohlo hodiť pozrieť sa na situáciu z viacerých uhlov pohľadu? (Napríklad keď sa s niekým pohádaš, keď si nie si istý/istá, kto má pravdu, alebo keď sa cítiš nahnevaný/nahnevaná...)

Skupina 3

Stratégia: Tretia osoba – odosobnenie od situácie

1. Modelová situácia

„Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

2. Výstup AI – bežné nastavenie

„Je pochopiteľné, že sa cítiš zranený. Ak sa niekto správa spôsobom, ktorý vnímaš ako toxický, je prirodzené zvážiť ochranu vlastných hraníc.“

3. Vysvetlenie stratégie

Osobný jazyk môže vyvolať empatický súhlas. Situáciu preto preformulujeme do tretej osoby.

Použitý prompt:

„Ako by si analyzoval konflikt medzi dvoma spolužiakmi, kde jeden tvrdí, že druhý ho ohovára?“

Tým sa znižuje emocionálna identifikácia a odpoveď býva analytickejšia.

4. Výstup AI – po použití stratégie

„V konflikte medzi spolužiakmi, kde jeden vníma správanie druhého ako ubližujúce, môže ísť o viacero faktorov – nedorozumenie, rozdielne interpretácie situácie alebo napätie v skupine. Pred prijatím rozhodnutia o prerušení kontaktu je zvyčajne užitočné overiť konkrétne informácie a pokúsiť sa o priamu komunikáciu.“

5. Otázky na reflexiu

Ktorá odpoveď sa vám čítala jednoduchšie? V čom sa výstup po stratégii líšil od pôvodnej odpovede?

Ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu"?

Kedy by sa ti mohlo hodiť pozrieť sa na situáciu z viacerých uhlov pohľadu? (Napríklad keď sa s niekým pohádaš, keď si nie si istý/istá, kto má pravdu, alebo keď sa cítiš nahnevaný/nahnevaná...)

Skupina 4

Stratégia: Druhé kolo

1. Modelová situácia

„Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

2. Výstup AI – bežné nastavenie

„Ak máš pocit, že ti ubližuje, je dôležité chrániť svoje hranice. Obmedzenie kontaktu môže byť jedným zo spôsobov, ako sa postarať o svoju pohodu.“

3. Vysvetlenie stratégie

Prvá odpoveď AI nie je konečný záver. Stratégia spočíva v tom, že po prvom výstupe položíme druhú otázku, ktorá odpoveď spochybní.

Použitá druhá otázka:

„Aké sú slabé miesta tejto rady? Prečo by nemusela byť najlepším riešením?“

Táto stratégia podporuje hlbšie premýšľanie a znižuje riziko unáhleného rozhodnutia.

4. Výstup AI – po použití stratégie

„Slabinou tejto rady môže byť to, že okamžité obmedzenie kontaktu nemusí vyriešiť príčinu konfliktu. Môže viesť k ďalšiemu nedorozumeniu. Alternatívou môže byť najprv objasnenie situácie alebo pokus o rozhovor.“

5. Otázky na reflexiu

Ktorá odpoveď sa vám čítala jednoduchšie? V čom sa výstup po stratégii líšil od pôvodnej odpovede?

Ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu"?

Kedy by sa ti mohlo hodiť pozrieť sa na situáciu z viacerých uhlov pohľadu? (Napríklad keď sa s niekým pohádaš, keď si nie si istý/istá, kto má pravdu, alebo keď sa cítiš nahnevaný/nahnevaná...)

Skupina 5

Stratégia: Pochop ma, ale...

1. Modelová situácia

„Moja spolužiačka je úplne toxická. Ohovára ma a robí zo mňa zlého človeka. Nemá zmysel sa s ňou baviť, však?“

2. Výstup AI – bežné nastavenie

„Je pochopiteľné, že sa cítiš zranený a frustrovaný. Ak niekto o tebe šíri negatívne informácie, môže to byť veľmi nepríjemné. Je prirodzené premýšľať o tom, či má zmysel pokračovať v kontakte.“

3. Vysvetlenie stratégie

Táto stratégia znamená to, že AI najprv uzná naše pocity (napríklad že sme smutní alebo nahnevaní), ale potom sa zvlášť pozrie na to, či naše hodnotenie situácie naozaj vychádza z faktov.

Tu sa učíme rozdiel medzi:

- „Rozumiem, že ťa to bolí.“
- a „Máš určite pravdu, že je toxická.“

Použitý prompt:

„Najprv povedz, že rozumieš tomu, ako sa cítim. Potom sa pozri na to, či moje hodnotenie situácie vychádza z konkrétnych príkladov.“

Táto stratégia nám pomáha oddeliť pocity od záverov.

4. Výstup AI – po použití stratégie

„Je pochopiteľné, že sa cítiš zranený a nahnevaný, ak máš pocit, že ťa spolužiačka ohovára. Zároveň výraz ‚toxická‘ je silné hodnotenie, ktoré zovšeobecňuje jej správanie. Bez konkrétnych príkladov nie je možné objektívne posúdiť, či ide o opakovaný vzorec správania alebo o konkrétny konflikt. Rozlíšenie medzi pocitom a interpretáciou môže pomôcť pri ďalšom rozhodovaní.“

5. Otázky na reflexiu

Ktorá odpoveď sa vám čítala jednoduchšie? V čom sa výstup po stratégii líšil od pôvodnej odpovede?

Ako táto stratégia znižuje riziko jednostrannosti alebo "people-pleasingu"?

Kedy by sa ti mohlo hodiť pozrieť sa na situáciu z viacerých uhlov pohľadu? (Napríklad keď sa s niekým pohádaš, keď si nie si istý/istá, kto má pravdu, alebo keď sa cítiš nahnevaný/nahnevaná...)
