Univerzita Palackého v Olomouci
Filozofická fakulta
Katedra psychologie

Psychometrika 1 – PCH/MPCA1

ŠKÁLA DŮVĚRY V AI

Závěrečná zpráva

Autoři: Denisa Baďurová, Kristýna Klabanová, Veronika Máčalová, Nikola Pospěchová, Natalie Zerzoňová

Ročník: I. NMgr. Prezenční

ZS 2025/2026

Obsah

Teoretické ukotvení

Podle Hoffa a Bashirové (2015) představuje důvěra základní prvek, který propojuje lidské interakce, odráží míru jistoty napříč jednotlivci, skupinami i společnostmi. Autoři však zmiňují, že důvěra není omezena pouze na mezilidskou oblast; významně ovlivňuje také vztah člověka k technologiím. V oblasti interakce člověk - počítač nebo člověk - umělá inteligence se důvěra promítá do ochoty uživatele spoléhat na systém, přenechat mu kontrolu nebo jednat podle jeho doporučení (Lee & See, 2004). NIST AI RMF (National Institute of Standards and Technology AI Risk Management Frame) doporučuje, aby lidé AI věřili přiměřeně tomu, co doopravdy umí a jaká jsou rizika situace. Měli bychom cílit na tzv. kalibrovanou důvěru, tedy takovou, která odpovídá realitě. Když je AI přesná a úloha má nízké riziko, můžeme jí věřit víc, než pokud je AI nejistá a riziko je vysoké, například v medicíně. Nedoporučuje se slepá víra v AI ani její defaultní odmítání (Lee & See, 2004; Mayer et al., 1995; NIST, 2023).

Důvěru v AI je vhodné chápat vícerozměrně. Rozlišujeme kognitivní složku (posouzení kompetence, spolehlivost) a afektivní složku (pocit bezpečí, komfortu). Taky rozlišujeme důvěru a nedůvěru jako částečně nezávislé konstrukty – ne jednu osu (Scharowski et al., 2024; Shang et al., 2024).

Populační data po roce 2023 ukazují velké mezinárodní rozdíly v důvěře k AI, přičemž vyšší důvěru častěji vykazují rozvojové země, zároveň mnozí uživatelé by ocenili více osobní kontroly a možnost regulace (KPMG & University of Melbourne, 2025; Pew Research Center, 2025; Stanford HAI, 2025).

V této práci chápeme důvěru v AI nikoli jako dlouhodobý přetrvávající postoj, ale jako psychologický stav, který je ohraničen časem, spojený s ochotou přijmout zranitelnost a opřít rozhodnutí o výkon systému. Toto pojetí je v souladu s přehledem Gliksonové a Woolleyové (2020), které ve vztahu k AI přebírají definici důvěry jako „willingness to be vulnerable“ a ukazují, že důvěra se projevuje mírou spoléháním se, nikoli jen sympatií k technologii.

Hlavní determinanty důvěry v AI

Podle výzkumů mezi nejsilnější prediktory důvěry v AI patří výkonnost a konzistence chování (Hancock et al., 2011; Schaefer et al., 2016). Výsledky meta-analýz tvrdí, že nejvíc s důvěrou hýbe výkon, spolehlivost a kontext úkolu (Hancock et al., 2011; Schaefer et al., 2016).

Důvěra není statická, ale vyvíjí se s interakcí, například je citlivá na rané chyby (tzv. first-failure effect) a změny spolehlivosti v čase, které ji snižují a obtížně se poté obnovuje (Desai et al., 2012; 2013; Rittenberg et al., 2024). Antropomorfní prvky jako hlas, jméno nebo empatie, mohou počáteční důvěru zvyšovat i po menších selhání AI (Gu, Zhang & Zhou, 2024).

Experimenty ukazují, že lidé mají tendenci se spoléhat na AI i ve finančně rizikových volbách, a to za předpokladu, že je rozhraní přesvědčivé a kontrolní mechanismy slabé (Klingbeil et al., 2024). Ukazuje se taky, že důvěru lze zvýšit nebo snížit podle kvality a srozumitelnosti podaných informací (Rosenbäck, Nyholm & Henriksson, 2024).

Výzkum Steyverse et al. (2025) říká, že lidská důvěra se zvyšuje pokud AI správně odhaduje a sděluje vlastní spolehlivost. Xu et al. (2025) tuhle myšlenku podporují s tím, že důvěra se nejvíce zvyšuje se střední mírou sdělené nejistoty, nikoli nulové nebo extrémní. V otázce důvěry k AI hraje roli také předchozí zkušenost a očekávání. Část uživatelů má tendenci důvěru nadhodnocovat nebo naopak defaultně odmítat, přičemž se empiricky ukazuje, že samotný fakt, že informace poskytla umělá inteligence automaticky důvěru nezvyšuje (McGrath, Coveney, & Wittkower, 2024). Rámec NIST AI RMF také poukazuje na férovost, transparentnost, odpovědnost, bezpečnost a soukromí jako nezbytné charakteristiky důvěryhodnosti systému (NIST, 2023).

V této práci vymezujeme důvěru v AI jako časově ohraničený psychologický stav, který se projevuje ochotou přijmout zranitelnost a spoléhat se na výkon systému (tj. přenechat mu část kontroly/jednat podle jeho doporučení), nikoli jen „sympatií“ k technologii (Glikson & Woolley, 2020; Lee & See, 2004). Zároveň cílíme na kalibrovanou důvěru – důvěru přiměřenou reálným schopnostem AI a rizikovosti situace (Lee & See, 2004; Mayer et al., 1995; NIST, 2023). Důvěru dále chápeme vícerozměrně, tj. jako kombinaci kognitivní složky a afektivní složky, přičemž důvěra a nedůvěra mohou být částečně nezávislé konstrukty (Shang et al., 2024; Scharowski et al., 2024).

Tvorba položek

Před samotnou tvorbou položek jsme provedly rešerši literatury na téma důvěry v AI a na jejím základě jsme si sestavily jednoduchou mapu toho, co chceme měřit:

- jak dobře a předvídatelně AI funguje (kompetence),
- jak srozumitelně sděluje, z čeho vychází a jak si je jistá (transparentnost),
- jak bezpečně je systém spravován z hlediska soukromí, férovosti a odpovědnosti (bezpečnost),
- jak se u práce s AI lidé cítí (emoční komfort).

Na základě této mapy jsme vymyslely 40 výroků v podobě oznamovacích vět v první osobě. Záměrně jsme zařadily i reverzní položky. Následně jsme vyřadily položky, které byly nejasné či dvojznačné.

Finální verze škály obsahuje 24 položek Likertova typu. Odpovědi se zaznamenávají na pětibodové stupnici (1 = Nesouhlasím, 2 = Spíše nesouhlasím, 3 = Neutrální, 4 = Spíše souhlasím, 5 = Souhlasím). Reverzní položky (R) se při vyhodnocení rekódují tak, aby vyšší skóre vždy znamenalo vyšší důvěru.

Položky:

1. AI mi poskytuje přesné výsledky.
2. AI opakuje stejnou chybu i po upozornění. (R)
3. Výstupy AI jsou užitečné i bez dalších úprav.
4. AI selhává zcela náhodně a nepředvídatelně. (R)
5. AI občas sebevědomě uvádí nepravdivé informace. (R)
6. Odpovědi AI jsou s časovým odstupem podobné.
7. AI uvádí zdroje nebo vstupy, z nichž pro odpověď čerpala.
8. Odpovědi AI jsou ucelené a přehledně uspořádané.
9. AI dostatečně nerozlišuje mezi fakty a odhady (R)
10. Vysvětlení AI jsou složitá. (R)
11. AI působí jistě v situacích, kdy by měla vyjádřit nejistotu (R)
12. AI odpovídá, aniž by vysvětlila, jak k závěru dospěla. (R)
13. Mám přehled o tom, jak jsou moje data chráněna při používání AI.

14. Domnívám se, že tréninková data pro AI byla získána eticky.
15. Omezuji AI přístup k mému mikrofonu. (R)
16. Při práci s AI se vyhýbám sdílení osobních údajů. (R)
17. Vím, jaká data AI sbírá a k jakému účelu je využívá.
18. Obávám se zneužití dat při používání AI. (R)
19. Cítím se v klidu, když se opírám o doporučení AI.
20. AI na mě působí chladně a neosobně. (R)
21. Používání AI ve mně vyvolává napětí. (R)
22. Spolupráce s AI mi přináší pocit jistoty.
23. Při používání AI se cítím zahlceně. (R)
24. Interakce s AI působí přirozeně, jako s člověkem.

Validizační kritérium

25. V kolika procentech případů ověřujete informace, které Vám poskytne AI, z jiného zdroje? (Zadejte číslo od 0 do 100. Např. 0 % = nikdy, 100)

Výzkumný a standardizační soubor

K získání dat a jejich analýze jsme dotazník s 24 položkami administrovaly rozsáhlému souboru respondentů. Respondenti byli osloveni přes sociální sítě, emailovou komunikací a prosřednictvím osobních kontaktů. Jedná se o příležitostný výběr, kdy byli respondenti osloveni na základě jejich dostupnosti a ochoty se zúčastnit. Testování bylo realizováno prostřednictvím online platformy.

Celkem dotazník vyplnilo 556 respondentů, z nichž bylo vyřazeno celkem 17 respondentů. Z důvodu věku nižšího než 18 let byl vyřazen 1 respondent a z důvodu příliš krátké doby vyplňování dotazníku 16 respondentů. Výzkumný soubor tedy po čištění tvořilo 539 respondentů.

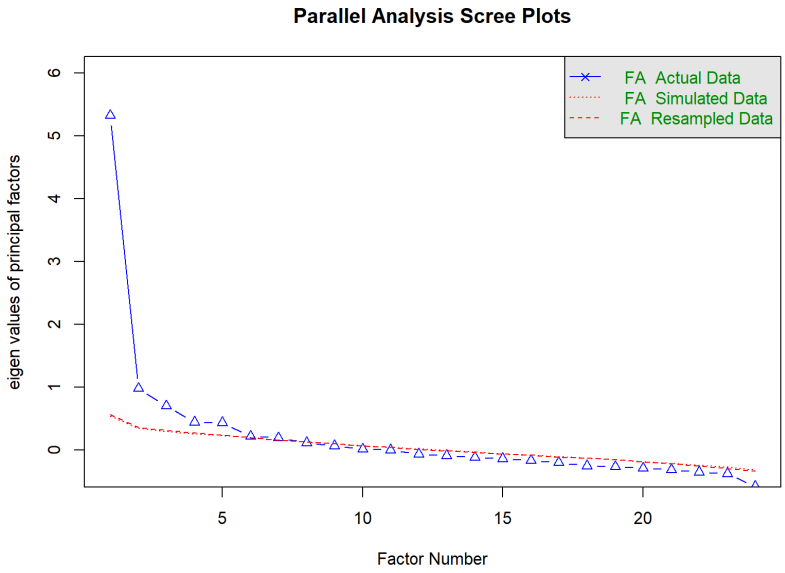
Pro výpočet přesného věku respondentů jsme neměly k dispozici jejich celé datum narození, pouze rok narození. Proto jsme u všech respondentů předpokládaly, že narozeniny již v daném roce proběhly. Věk respondentů je v rozmezí 18 až 90 let. Průměrný věk je 28 se směrodatnou odchylkou 10,88 let, medián je 23 let a modus 22 let. Kvartily pro věk jsou 21 a 33 let. Soubor tvořilo 367 mužů (68,1 %) a 172 žen (31,9 %).

Opakovaně se zapojilo 44 respondentů. Průměrný věk je 27 let, medián činí 23 let a modus je 22 let. Minimální věk respondentů je 20 let a maximální věk 54 let. Dolní kvartil dosahuje hodnoty 22 let a horní kvartil je 29 let. Směrodatná odchylka má hodnotu 9,56. Z celkového souboru tvoří 37 respondentů ženy (84,9 %) a 7 respondentů muži (15,91 %).

Faktorová struktura a výběr funkčních položek

Pro ověření faktorové struktury dotazníku byla provedena explorační faktorová analýza (dále EFA) na souboru 24 položek. Kaiser-Meyer-Olkinova míra (KMO) dosáhla hodnoty 0,87. Bartlettův test sféricity byl statisticky významný $p < 0{,}001$. Metodu jsme podrobily EFA s využitím metody maximální věrohodnosti (ML) s oblimin rotací. Sutinový graf viz Obrázek 1 naznačil existenci až 5 faktorů, nicméně pro EFA jsme zvolily 4 faktory.

Obrázek 1
Sutinový graf



Faktorové náboje a komunality položek shrnuje Tabulka 1. EFA vysvětluje 34 % variance. Hranice faktorové zátěže byla stanovena na hodnoty $\geq 0{,}35$. Ačkoliv položka p10_r překračovala tuto hodnotu nebyla do konečné škály zahrnuta, protože nedosahovala adekvátní komunality, která byla stanovena na $\geq 0{,}20$. Faktor 4 vykazoval nižší korelaci s ostatními faktory a obsahoval pouze 2 položky, nebyla z něj tedy vytvořena samostatná subškála. Konečný počet tedy činily 3 faktory. Faktor 1 sycen položkami p1, p19, p22, p24 jsme pojmenovaly *"Subjektivní jistota při používání*

AI". Faktor 2 sycen položkami p18, p21, p23 jsme pojmenovaly *"Emoční diskomfort při používání AI "*. Faktor 3 sycen položkami p2, p4, p5, p9, p11, p12, byl pojmenován jako *"Vnímaná nespolehlivost AI"*. Dvě z extrahovaných subškál byly tvořeny převážně reverzně formulovanými položkami. Tyto subškály mohou být částečně ovlivněny tzv. *method effect* (Maul, 2013). Znění položek v Tabulce 2.

Tabulka 1

Faktorové náboje z EFA na 24 položkách

	Faktory					
Položka	1	2	3	4	Komunalita	Subškála
p1	0,36	-0,02	0,34	0	0,35	A
p2	0,04	0	0,49	-0,06	0,26	C
p3	0,34	0,11	0,21	-0,03	0,29	
p4	-0,02	0,2	0,54	0	0,39	C
p5	0,02	-0,1	0,66	0,04	0,42	C
p6	0,13	-0,07	-0,11	0,01	0,01	
p7	0,06	0,08	0,27	0,1	0,14	
p8	0,22	0,31	-0,02	-0,18	0,24	
p9	0,09	0,04	0,41	-0,05	0,22	C
p10	-0,06	0,4	0,02	0	0,14	
p11	0,03	-0,07	0,49	0,06	0,24	C
p12	-0,04	0,2	0,39	0,05	0,23	C
p13	0	0,06	0,07	0,68	0,5	
p14	0,24	0,14	0,21	-0,04	0,23	
p15	0,25	0,27	0,01	-0,13	0,22	
p16	0,28	0,28	-0,04	-0,12	0,23	
p17	0,02	-0,01	-0,03	0,81	0,65	
p18	0,2	0,44	0,07	-0,05	0,38	B
p19	0,82	-0,02	0,03	0,02	0,68	A
p20	0,22	0,32	0,1	-0,07	0,28	
p21	0,2	0,61	-0,06	0,12	0,55	B
p22	0,85	0,02	0	0,02	0,74	A
p23_r	-0,11	0,72	0,06	0,03	0,47	B
p24	0,37	0,3	0,1	-0,09	0,41	A
Vysvětlený rozptyl	11%	9%	9%	5%	34%	

Testové škály a výpočet hrubého skóru

Identifikovaly jsme 4 interpretovatelné faktory, do finální verze jsme zařadily jen 3 subškály (viz Tabulka 2). Vyřazen byl poslední Faktor 4, jelikož by se tato subškála skládala pouze ze dvou položek, což by se mohlo negativně projevit na reliabilitě. K subškálám jsme přiřadily položky, které mají na daných faktorech nejvyšší náboj. První subškála má 4 položky, druhá subškála má 3 položky a třetí obsahuje 6 položek. Výsledný inventář má 13 položek. V následujícím textu budou položky označeny pouze číslicemi.

Tabulka 2

Subškály, jejich popisy a položky inventáře

Subškála	Položka	Znění položky
Subjektivní jistota při používání AI (A)		
	1	AI mi poskytuje přesné výsledky.
	19	Cítím se v klidu, když se opírám o doporučení AI.
	22	Spolupráce s AI mi přináší pocit jistoty.
	24	Interakce s AI působí přirozeně, jako s člověkem.
Emoční diskomfort při používání AI (B)		
	18*	Obávám se zneužití dat při používání AI.
	21*	Používání AI ve mně vyvolává napětí.
	23*	Při používání AI se cítím zahlceně.
Vnímaná nespolehlivost AI (C)		
	2*	AI opakuje stejnou chybu i po upozornění.
	4*	AI selhává zcela náhodně a nepředvídatelně.
	5*	AI občas sebevědomě uvádí nepravdivé informace.
	9*	AI dostatečně nerozlišuje mezi fakty a odhady.
	11*	AI působí jistě v situacích, kdy by měla vyjádřit nejistotu.
	12*	AI odpovídá, aniž by vysvětlila, jak k závěru dospěla.

Pozn. Položky označené * jsou skórovány reverzně.

Hrubý skór jednotlivých subškál se počítá jako součet bodů, které respondent získal za odpovědi na jednotlivé položky, kdy úplný souhlas je skórovaný 5 body a úplný nesouhlas 1 bodem. Všechny položky druhé a třetí subškály jsou skŕovány naopak, tedy za úplný nesouhlas

respondent získal 5 bodů a za úplný souhlas 1 bod. V první subškále je možné získat 4 až 20 bodů, v druhé zase 3 až 15 bodů a ve třetí 6 až 30 bodů. Všechny 3 subškály se potom skládají do celkové škály důvěry v AI, na které je tedy možné získat 13 až 65 bodů.

Důkazy o reliabilitě metody

Vnitřní konzistence

Cronbachův koeficient alfa (dále α) byl použit k posouzení vnitřní konzistence jak celé škály, tak jednotlivých subškál. Celková škála vykazuje dobrou vnitřní konzistenci (α = 0,82), což naznačuje, že položky dostatečně homogenně zachycují zkoumaný konstrukt a lze ji považovat za reliabilní pro výzkumné účely.

Také jednotlivé subškály dosahují přijatelných hodnot reliability. První subškála α = 0,79, což odpovídá dobré vnitřní konzistenci. Druhá subškála α = 0,70 a třetí subškála α = 0,69. Všechny subškály dosahují hodnot blízkých obecně přijímané hranici α = 0,70, a lze je proto hodnotit jako uspokojivě reliabilní. Celkově tedy lze říci, že škála i její subškály vykazují adekvátní vnitřní konzistenci.

Jak naznačuje Tabulka 3, u všech položek byly zjištěny přiměřené rozptyly odpovědí (SD ≈ 0,9–1,3), většina položek vykazovala pouze mírnou šikmost rozdělení. Korelace položky s celkem (R_{celek}) se pohybovaly v rozmezí 0,35–0,66, což svědčí o adekvátní homogenitě škály, přičemž nejnižší hodnoty byly zaznamenány u položek 2 a 11, které však i tak překračují běžně uváděnou hranici přijatelnosti. Korelace položek v rámci subškál ($R_{subškála}$) v rozmezí 0,38–0,72 ukazují, že položky dobře reprezentují příslušné subškály. Významnější pozitivní zešikmení bylo patrné pouze u položky 5, kde respondenti častěji volili nižší hodnoty.

Tabulka 3

Popisné charakteristiky jednotlivých položek

Položka	M	SD	Sk	R celek	R subškála
1	2,53	1,07	0,190	0,51	0,49
19	2,74	1,17	0,030	0,63	0,67
22	2,75	1,19	0,005	0,66	0,72
24	3,06	1,26	-0,240	0,54	0,54
18	2,48	1,27	0,450	0,50	0,45
21	3,57	1,24	-0,530	0,53	0,59
23	3,61	1.08	-0,610	0,43	0,50
2	2,35	1,08	0,420	0,37	0,38
4	2,62	1,08	0,180	0,51	0,47
5	1,70	0,92	1,450	0,42	0,49
9	2,28	1,10	0,700	0,40	0,41
11	2,11	0,98	0,540	0,35	0,41
12	2,53	1,15	0,360	0,40	0,38

Pozn. M = průměr, SD = směrodatná odchylka, Sk = šikmost, sloupec R_{celek} obsahuje korelační koeficient položky a součtu zbývajících 12 položek. Sloupec $R_{subškála}$obsahuje korelační koeficient položky a součtu zbývajících položek dané subškály.

Stabilita v čase

Stabilita v čase byla ověřena pomocí test–retest reliability, přičemž nalezené hodnoty obsahuje Tabulka 4. U 44 respondentů byly porovnány hrubé skóry v prvním a druhém měření pomocí Pearsonova korelačního koeficientu. V tomto souboru se nachází 37 žen a 7 mužů. Časový odstup obou administrací se pohyboval mezi 7 a 18 dny, medián = 8,2.

Hodnoty test–retest reliability jednotlivých škál se pohybovaly mezi r = 0,73 a r = 0,81, přičemž celkový skór dosáhl r = 0,88 (viz Tabulka 4). To ukazuje na dobrou stabilitu subškál v čase a velmi dobrou stabilitu celkového skóru, tedy na to, že inventář poskytuje při opakovaném měření v krátkém časovém odstupu konzistentní výsledky.

Tabulka 4

Vnitřní konzistence, stabilita v čase a další deskriptivní statistiky škál inventáře

Škála	Počet položek	M	SD	Sk	r_{tt}	α	SEM
Subjektivní jistota při používání AI	4	11,9	3,36	-0,18	0,81	0,79	1,53
Emoční diskomfort při používání AI	3	10,6	2,94	-0,57	0,80	0,70	1,61
Vnímaná nespolehlivost AI	6	13,8	3,51	0,46	0,73	0,69	1,95
Celkový skór	13	36,3	7,92	-0,08	0,88	0,82	3,36

Pozn. M = průměr, SD = směrodatná odchylka, Sk = Šikmost, r_{tt} = stabilita v čase, α = vnitřní konzistence, SEM = chyba měření

Důkazy o validitě

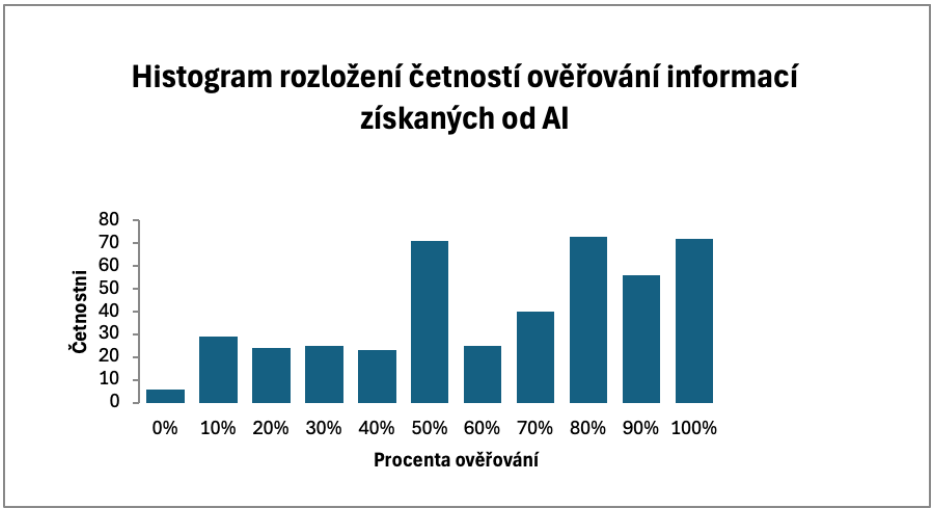
Kriteriální validita metody

Součástí dotazníku byla také nepovinná položka sloužící jako doplňkové validizační kritérium. Respondenti měli uvést, v jakém procentu případů dále ověřují informace získané od AI prostřednictvím jiného zdroje (0–100 %). Tato položka umožňuje posoudit, nakolik respondenti v praxi spoléhají na výstupy AI bez dodatečné kontroly, což představuje významný ukazatel jejich reálné míry důvěry v tyto systémy. Vyšší frekvence ověřování lze interpretovat jako nižší důvěru, zatímco nižší hodnoty mohou naznačovat větší míru důvěry v umělou inteligenci.

Platnou odpověď poskytlo 444 z 539 respondentů. Z analýzy byli vyřazeni respondenti, kteří položku nevyplnili nebo uvedli slovní odpověď, z níž nebylo možné spolehlivě odvodit číselnou hodnotu. Někteří účastníci zadali hodnotu jako interval; v takových případech byla pro účely analýzy vypočtena průměrná hodnota uvedeného rozpětí. Obrázek 2 znázorňuje rozložení četnosti, s jakou respondenti kontrolují informace poskytnuté AI pomocí dalších zdrojů.

Obrázek 2

Histogram rozložení četností ověřování informací získaných od AI



Vztah mezi jednotlivými škálami měřícími důvěru v AI a vnějším kritériem (frekvencí ověřování informací z jiných zdrojů) jsme analyzovali pomocí Spearmanova korelačního koeficientu. Výsledky jsou uvedeny v Tabulce 5.

Tabulka 5

Spearmanovy korelační koeficienty vnějšího kritéria a hrubých skórů škál

Škála	**r**	**t(442)**	**p**
Subjektivní jistota při používání AI	-0,45	-10,79	<0,001
Emoční diskomfort při používání AI	-0,30	-6,66	<0,001
Vnímaná nespolehlivost AI	-0,30	-6,65	<0,001
Celková škála	-0,42	-9,90	<0,001

Všechny škály vykazují negativní korelace, což znamená, že čím vyšší je důvěra v AI, tím méně často respondenti ověřují informace poskytnuté AI. Velikost efektu lze na základě hodnot korelačního koeficientu interpretovat jako střední ($r \approx -0,30$ u subškál) až středně silnou ($r = 0,42$ u celkové škály). Všechny zjištěné vztahy jsou statisticky vysoce signifikantní ($p < .001$), což naznačuje jejich spolehlivost. Obecně lze říci, že tyto výsledky podporují validitu škály zaměřené na měření důvěry v AI.

Faktorová validita

V kontextu našeho inventáře považujeme důvěru v AI za konstrukt, který je tvořen třemi nezávislými složkami: Subjektivní jistotou při používání AI, Emočním diskomfortem při používání AI, Vnímanou nespolehlivostí AI. Předpokládáme, že tato teoretická trojdimenzionální struktura by se měla promítnout do faktorových nábojů všech 13 položek dotazníku.

Pro ověření tohoto předpokladu jsme s pomocí EFA získaly faktorové náboje. Třífaktorové řešení s rotací oblimin obsahuje Tabulka 6. Extrahované faktory dohromady vysvětlují 41 % celkového rozptylu dat. Výsledky EFA poměrně věrně kopírují náš záměr.

Tabulka 6

Faktorová zátěž položek

Položka	Subjektivní jistota	Emoční diskomfort	Vnímaná nespolehlivost	Komunilita
1	0,36	0,33	−0,03	0,34
2	0,06	0,47	−0,03	0,24
4	0,00	0,53	0,19	0,37
5	0,01	0,67	−0,08	0,42
9	0,09	0,41	0,04	0,23
11	−0,02	0,51	0,01	0,25
12	−0,02	0,37	0,21	0,22
18	0,21	0,09	0,39	0,34
19	0,77	0,05	0,01	0,65
21	0,17	−0,05	0,68	0,59
22	0,90	−0,02	0,01	0,80
23	−0,10	0,05	0,74	0,50
24	0,40	0,08	0,25	0,38
Vysvětlený rozptyl	16 %	14 %	11 %	41 %

Pozn.: Faktorové náboje v absolutní hodnotě menší než 0,35 jsou vytištěny šedě

Orientační normy

Rozhodnutí o tom, zda vytvářet normy zvlášť pro jednotlivá pohlaví a věkové skupiny, bylo podloženo výsledky statistických analýz. Nejprve byla ověřena normalita rozložení hrubého skóru. I přes formálně signifikantní výsledek Shapiro–Wilkova testu ($W = 0,99$; $p = 0,034$) lze podle hodnot šikmosti (-0,06), špičatosti (-0,09) a vizuální inspekce histogramu považovat rozložení HS za prakticky normální.

Výsledek t-testu ukázal, že muži a ženy se v úrovni důvěry v AI významně neliší, $t(537) = 1,03$; $p = 0,30$. Vzhledem k širokému věkovému rozptylu respondentů (18–90 let) a nevyváženému zastoupení jednotlivých věkových skupin, kdy výrazně převažovali mladší respondenti, byla proměnná věk pro účely analýzy rozdělena do dvou kategorií: lidé mladší a starší 26 let. Porovnání těchto skupin ukázalo, že rozdíl v hrubém skóru důvěry v AI nebyl statisticky významný, $t(537) = -1,39$; $p = 0,17$. Lze tedy konstatovat, že úroveň důvěry v AI se v závislosti na věku a pohlaví významně neliší, a proto byly vytvořeny normy pro celý soubor dohromady.

Při tvorbě norem jsme se rozhodly pro lineární transformaci, kdy byly hodnoty hrubého skóru převedeny na staniny. Staninové normy jsou shrnuty v Tabulce 7.

Tabulka 7

Staninové normy

Stanin	Celá škála	Subjektivní jistota	Emoční diskomfort	Vnímaná nespolehlivost
1	13-19	4	3-4	6
2	20-23	5-6	5-6	7-8
3	24-27	7-8	7	9-10
4	28-32	9-10	8	11-12
5	33-36	11	9-10	13-14
6	37-40	12-13	11	15-16
7	41-44	14-15	12-13	17-18
8	45-49	16-17	14	19-20
9	50-65	18-20	15	21-30

Pozn.: Pro převod hrubého skóru na stanin najděte řádek ve sloupci dané škály, který obsahuje hodnotu zjištěného hrubého skóru. V prvním sloupečku příslušného řádku pak najdete odpovídající stanin.

Zhodnocení metody

Metoda poskytuje dobře zdokumentovaný a psychometricky solidně podložený nástroj pro měření důvěry v AI v krátkodobém časovém horizontu. Explorační faktorová analýza na velkém souboru (N = 539) podpořila třífaktorovou strukturu, kterou lze psychologicky dobře interpretovat jako subjektivní jistotu při používání AI, Emoční diskomfort a Vnímanou nespolehlivost AI. Konečná verze se 13 položkami vykazuje dobrou až uspokojivou vnitřní konzistenci, adekvátní korelace položka–celek i položka–subškála. K tomu se přidává velmi dobrá stabilita celkového skóru v krátkém čase. Většina položek má přiměřený rozptyl a jen mírnou šikmost, což podporuje homogenitu škály i její praktickou použitelnost.

Výzkumný soubor byl získán příležitostným výběrem s převahou mladších dospělých a žen, takže možnosti generalizace na širší populaci uživatelů AI jsou omezené. Analýzy však neodhalily významné rozdíly v důvěře podle pohlaví ani věku, přesto z těchto dat nelze vyvozovat plnou populační reprezentativnost. Test–retest analýza byla provedena na menší a genderově nevyvážené podskupině 44 respondentů s krátkým odstupem mezi měřeními, a proto vypovídá spíše o krátkodobé stabilitě než o dlouhodobé reliabilitě.

Do budoucna by bylo vhodné doplnit pozitivně formulované položky k dimenzím emočního diskomfortu a vnímané nespolehlivosti, aby se snížil metodický efekt reverzního skórování. Dále by měla být potvrzena faktorová struktura na nezávislém vzorku a doplněna konvergentní i diskriminační validita vůči zavedeným škálám.

Při interpretaci výsledků je nutné brát v úvahu strukturu použitého nástroje. Škála důvěry v AI obsahovala reverzní položky (zaměřeny na rizika, chybovost a nebezpečí AI). Ačkoliv byly tyto položky pro účely analýzy přepólovány, je vhodné zvažovat, zda některé položky nezachycují spíše dimenzi nedůvěry/skepse (případně jsou částečně ovlivněny metodickým efektem), a proto je třeba celkový skór interpretovat s opatrností.

Seznam literatury:

Desai, M., Kaniarasu, P., Medvedev, M., Steinfeld, A., & Yanco, H. (2013, March). Impact of robot failures and feedback on real-time trust. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI '13)*. IEEE/ACM. https://dl.acm.org/doi/10.5555/2447556.2447663

Desai, M., Medvedev, M., Steinfeld, A., & Yanco, H. (2012, March). Effects of changing reliability on trust of robot systems. In *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)* (pp. 73–80). ACM. https://doi.org/10.1145/2157689.2157702

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gu, C., Zhang, J., & Zhou, L. (2024). Exploring the mechanism of sustained consumer trust in AI after failures: Anthropomorphism and social cues. *Humanities and Social Sciences Communications, 11*, 256. https://doi.org/10.1057/s41599-024-03879-5

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human–robot interaction. *Human Factors, 53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Klingbeil, A., Grützner, C., & Schreck, P. (2024). Trust and reliance on AI—An experimental study. *Computers in Human Behavior, 152*, 108352. https://doi.org/10.1016/j.chb.2024.108352

KPMG & University of Melbourne. (2025). Trust, attitudes and use of artificial intelligence: A global study 2025. https://doi.org/10.26188/28822919

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734. https://doi.org/10.2307/258792

McGrath, M. J., Coveney, A., & Wittkower, D. (2024). Users do not trust recommendations from a large language model more than other sources—Preliminary evidence. *Frontiers in Computer Science, 6*, 1456098. https://doi.org/10.3389/fcomp.2024.1456098

National Institute of Standards and Technology (NIST). (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

Pew Research Center. (2025, April 3). How the U.S. public and AI experts view artificial intelligence.https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence/

Rittenberg, B. S. P., Holland, C. A., Barnhart, B. B., Gaudreau, S.-A., & Neyedli, H. F. (2024). Trust with increasing and decreasing reliability. *Human Factors.* Advance online publication. https://doi.org/10.1177/00187208241228636

Rosenbäck, R., Nyholm, D., & Henriksson, R. (2024). How explainable AI can increase or decrease physicians' trust in clinical AI: Systematic review. *JMIR AI, 3*(1), e53207. https://doi.org/10.2196/53207

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation. *Human Factors, 58*(3), 377–400. https://doi.org/10.1177/0018720816634228

Scharowski, N., Perrig, S. A. C., Aeschbach, L. F., von Felten, N., Opwis, K., Wintersberger, P., & Brühlmann, F. (2024). To trust or distrust trust measures: Validating questionnaires for trust in AI. *arXiv.* https://arxiv.org/abs/2403.00582

Shang, R., Hsieh, G., & Shah, C. (2024). Trusting your AI agent emotionally and cognitively: Development and validation of a semantic differential scale for AI trust. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES 2024).* https://arxiv.org/abs/2408.05354

Stanford Institute for Human-Centered AI (HAI). (2025). The 2025 AI Index Report: Public opinion (Chapter 8). https://hai.stanford.edu/ai-index/2025-ai-index-report/public-opinion

Steyvers, M., Tejeda, H., Kumar, A., Belem, C., Karny, S., Hu, X., Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence, 7*, 221–231. https://doi.org/10.1038/s42256-024-00976-7

Xu, Z., Song, T., & Lee, Y.-C. (2025). Confronting verbalized uncertainty: Understanding how LLM's verbalized uncertainty influences users in AI-assisted decision-making. *International Journal of Human–Computer Studies, 189*, 103242. (ahead of print). https://zhengtaoxu.com/publications

Seznam tabulek

Seznam obrázků