

Falzifikační charakter sociální agence v užívání vizuálně-generativní AI

Jan Šmejkal

Jan Šmejkal: Ústav filozofie a religionistiky FFUK,
Náměstí Jana Palacha 2,
honzi3702@gmail.com

Abstrakt

Příspěvek se zabývá schopností umělých neuronových sítí vygenerovat unikátní obrázek. Na první pohled není jasné, jak a komu lze v kontextu tohoto procesu atribuovat agenci. Prvním argumentem příspěvku je tvrzení, že agence leží jak na straně lidského operátora, tak na straně technologie, která vystupuje jako specifický tvořivý nástroj schopný emulovat styl lidského umělce. Jen lidský článek vztahu je však schopen intendovat a proto musí být agence v konečném důsledku přičítána jemu. Povahy této agence je dále ztotožněna se zastírací snahou a do uměleckého kontextu převedena apropriací pojmu falzifikace a falzifikátu. Argumentace se opírá o agenční teorii antropologa Alfreda Gella. Autor příspěvku nelegitimizuje jakékoliv praktiky spojené s provozováním ani trénováním neuronových sítí, které využívají obrazový materiál bez souhlasu jeho tvůrců.

1 Obrázky z umělé inteligence

Systémy umělé inteligence zaměřené na produkci a generování statického vizuálního obsahu¹ jsou trénovány na fotografiích a obrazech lidských profesionálů a umělců, dokáží emulovat jejich styl² a vyvést v něm vizuální dílo na základě slovního zadání. Tento proces běžným jazykem nazýváme „nechat si vygenerovat obrázek umělou inteligencí“. Ač je funkcionalita těchto sítí rozličná a služby, které se skrz ně nabízejí, různorodé, v našem příspěvku se budeme soustředit na jejich vrcholnou dovednost, tedy schopnost přivést v existenci do té doby neexistující statické vizuální dílo („vygenerovat obrázek“).

¹ V tomto příspěvku se soustředíme pouze na problematiku technologií generujících „obrázky“, tedy statické dvourozměrné vizuální objekty. Pod tímto označením myslíme např. technologie Imagen, DALL-E, Midjourney, Craiyon apod.

² Důrazu na stylovou emulaci si všímá Yao (Yao, 2025, str. 2). Stylu zde rozumíme jako ustrnulé interpretaci vizuální reality při jejím převodu na uměleckou reprezentaci.

Tato funkcionalita má samozřejmě dopad na trh s námezdným uměním a podstatným způsobem mění rozložení ekonomicko-kulturních sil na obou stranách vazby „poptávající-umělec“. Naše zkoumání se však bude zabírat teoretickou otázkou atribuce umělecké agence a její povahy v kontextu „generování obrázků“ za pomoci neuronových sítí.

K tomu budeme využívat terminologii agenčního systému antropologa Alfreda Gella, jehož užití krátce zdůvodníme. Tato antropologická teorie byla svým autorem využívána zejména pro objektovou analýzu lidového umění domorodých kmenů. Ač vizuálně-generativní AI (dále jen VGAI)³ není domorodý artefakt, přeci jen se v kontextu tradičních uměleckých forem jedná o něco cizorodého, co nezapadá do klasického schématu ani situace vizuálního umění. Také se tato teorie vyznačuje poměrnou liberalitou, se kterou nakládá s pojmy umění, dílo, artefakt atd., jenž mohou být v kontextu jiných uměleckých teorií exkluzivní a vykazovat rysy mnoha podob centrismu.⁴ Posledním důvodem je to, že na rozdíl např. od agenční teorie Bruna Latoura je Gellova teorie pragmatičtější⁵ a snazší na pochopení.

Zbytek úvodu využijeme pro postavení dvou klíčových *technologických* aspektů VGAI, které jsou pro naši analýzu podstatné a na něž omezíme vysvětlení těchto technologických artefaktů.

A) Grafické generátory na bázi *diffusion* jsou produktem obrácení funkcionality neuronových sítí sloužících k rozpoznávání tvarů a objektů: pracují na základě opakovaného včítání nejpravděpodobnějšího vzoru do náhodné plochy bílého šumu. (Bergman, Stryker, 2026) Produkt VGAI je tedy v nepřímé vazbě k vizuálním datům,⁶ na kterých byla původní rozpoznávající síť

³ Volíme novotvar „vizuálně-generativní“ namísto zaběhnutého „generativní“, protože generace vizuálních dat je pouze jednou z možných generativních funkcí AI, viz (Corchado et al., 2023)

⁴ Např. eurocentrismu, antropocentrismu nebo manocentrismu.

⁵ To je míněno doslova, jedná se o teorii hlásící se k pragmatismu.

⁶ Tedy k databázi, která se tak stává tzv. *konceptuální metaforou* skutečného uměleckého světa (de Mul, 2009, str. 101)

trénována a z nichž získává vzory, jež jsou reprodukovány. Tato nepřímá vazba má povahu napodobení. Pro náš problém je podstatné, že ačkoliv VGAI generuje unikátní (ve smyslu neopakované) artefakty, pracuje vždy s určitým množstvím předloh a „inspirací“, které byly využity při jejím trénování a které tvořili lidští řemeslníci. Oproti *kreativitě* tedy VGAI sází na *variaci existujícího*. Tato vlastnost sítí činí uživatelsky předvídatelnější a lépe monetizovatelnou.

B) Neuronové sítě fungují na stochastickém (pravděpodobnostním) základu. (Bergman, Stryker, 2026) To znamená, že namísto toho, abychom vůči zadání A dostali vždy výsledek A' (jako v deterministickém systému), děje se to, že při opětovném zadání vstupu A dostaneme výsledek A', B', C' etc., přičemž bude platit, že pokud je síť odladěná, měly by všechny výsledky splňovat zadání ke spokojenosti provozovatele i zákazníka. Opět zdůrazňujeme funkci produktové variace, která navíc slouží jako ochrana před porušením autorského práva.

2 Gellova agenční teorie a její aplikace na vizuálně-generativní AI

Z agenční teorie Alfreda Gella zavedeme několik pojmů. Jednak je to samotný pojem *agence* – aktivní působení na druhého s nějakým záměrem. (Gell, 1998, str. 16) Z něj plynou pojmy *agent* a *pacient*, tedy na jedné straně ten, kdo svou agencí aktuálně působí na druhého, na druhé ten, kdo toto působení aktuálně přijímá.⁷ To, prostřednictvím čeho agent v uměleckém kontextu na pacienta působí, je zvané *index* neboli konkrétní (vizuální) artefakt. Základní vztah v rámci nexu sociální umělecké agence tedy vypadá tak, že index je nositelem agence, kterou do něj vtiskl agent k oslovení pacienta.⁸ Pacient nahlíží, že je na něj umělecky působeno, a doptává se sám sebe na povahu agence, tedy na to, co asi kdo zamýšlel tím, že mu toto dal před oči. Gell toto tázání nazývá *abdukci agence*. (Gell, 1998, str. 14) Jedná se o nepravou indukci založenou na domněnce nebo vágním pocitu z vizuálního artefaktu. Pokaždé, když se ptáme, kdo mohl konkrétní objekt vytvořit nebo pro koho nebo co byl zamýšlen, provádíme tuto abdukci nezávisle na tom, jestli existuje odpověď na tyto otázky nebo jsou-li tyto otázky k tomuto objektu relevantní.

Gellova teorie navíc umožňuje pojmut za agenty (a pacienty) i neživé objekty, např. antropomorfizované

hračky, vozidla, budovy (Gell, 1998, str. 18) nebo tvořivé technologické artefakty typu VGAI. Agentem či pacientem se stává prostě to, co vstupuje do agenčního vztahu, tedy co přijímá agenci nebo čemu je agence přisuzována abdukci.

Pokud představený systém přizpůsobíme naší výzkumné situaci, dojdeme k závěru, že v rámci Gellovy agenční teorie není problém přisoudit jakékoliv VGAI uměleckou agenci a jakémukoliv jejímu výtvoření indexu, jehož vznik je navíc vlivem stochastiky vlastní VGAI samotné alespoň částečně nepředvídatelný. O síti samotné ovšem budeme hovořit jako o tzv. *sekundárním* agentovi.

Sekundární agence je umožněna vágní povahou abdukce: nezáleží na tom, jestli pacient atribuuje agenční záměr *rádně*, tzn. jestli objekt abdukce skutečně je intendujícím primárním agentem. Sekundární agent je takový agent, který esteticky působí na pacienta bez toho, aby cokoliv sám intendoval, tedy měl vlastní záměry podmíněné vědomím, vůlí apod. Gell (Gell, 1998, str. 21) používá příklad minometčíka a miny: ačkoliv je to mina, která aktuálně působí na pacienta, jedná se o sekundární agenci, jelikož mina nemůže intendovat. Intendující primární agent je minometčík, který minu pokládá se záměrem zabít. Přesto říkáme, že vojáka „zabila mina“, nikoliv že „jej zabil minometčík“, pracující zcela srozumitelně se sekundární agenční atribucí, aniž by kdo snad podezřívá minu samotnou z nekalých úmyslů.

Vycházíme zde z předpokladu, že neuronové sítě nemají vědomí a nemohou tedy intendovat. Jejich působení na pacienta skrze vygenerované obrázky (indexy)⁹ má ráz plnění rozkazu nebo zadání. Za intenci, s jakou VGAI tvoří, tedy můžeme a musíme dosadit intence jejího autora, programátora, provozovatele a uživatele, tedy obecně primárního agenta: „[sekundární agenti] jsou objektivní ztělesnění moci nebo možnosti zamýšlet jejich užívání“. (Gell, 1998, str. 21, autorský překlad) Sekundární agenti jsou takové předměty, jejichž užívání je spojeno se záměrem na straně uživatele. Pacient se sice abduktivně vztahuje k VGAI (se stejnou personifikační figurou jako v předešlém příkladu k mině, ptá se „co VGAI zamýšlela, když přede mne postavila tento obrázek“), ví-li, že hledí na vygenerovaný obrázek, ale v dalším kroku své tázání nutně musí usměrnit na tázání po primární agenci („jaký záměr měl původce sítě nebo ten, kdo ji využil k získání tohoto obrázku“). Platí tedy, že kdokoliv, kdo plánuje VGAI vytvořit nebo využít jakýmkoliv způsobem, vtiskává svou agenci do ní jako do nástroje a skrz ní do jejího produktu, cílového indexu. Zdálo by se tedy, že z intencionálního hlediska je VGAI *průchozí*, nepřidává nic ke vztahu agence a pacienta, které volně proudí skrz její

⁷ Agence a paciente nejsou fixní kategorie, mimo aktuální agenční vztah věci a lidé zůstávají pouze potenciálními agenty nebo pacienty.

⁸ Index je tedy v druhém plánu také pacientem, protože rovněž přijímá agenční působení agenta.

⁹ VGAI má tu specifickou vlastnost, že se jedná o index a sekundárního agenta zároveň, o „tvořivou věc“, jejíž výtvoření jsou k síti samotné ve vztahu pacienta.

indexy. V následující kapitole se pokusíme ukázat, že užívání AI je ve všech svých instancích spojeno s intencí *ošálit* a že zmiňovaná intencionální průchodnost je pouze zdánlivá.

3 Závěr – falzifikační povaha agence

VGAI je sice kuriózní svým jedinečným charakterem „tvořivé věci“, protože její produkty jsou unikátní vizuální díla. Jelikož však nemůže konkrétní podobu svých výstupů intendovat, jedná se o sekundárního agenta, který k sobě přibírá více primárních agentů (provozovatele, programátora, uživatele) a indexy, které generuje, tak syntetizují intence všech relevantních primárních agentů. Jejich vytrasování není snadné a v tuto chvíli ani podstatné.

Zeptáme se znovu, co obecně zamýšlí primární agent, který za svůj nástroj volí jako sekundárního agenta VGAI. Domníváme se, že jedna intence je vlastní užití této techniky jako takové, jako je užívání kladiva vlastní intence zatloukání, a nezáleží na tom, jestli hřebíků nebo suků. Sekundární agent je objektovým ztělesněním možnosti být užit se záměrem¹⁰ a my se tážeme po obecné povaze tohoto záměru. Tímto univerzálním záměrem je podle nás *zastírající snaha přesvědčit pacienta, že má co do činění s lidským výtvořem* a že svou vizuální zkušenost může modelovat podle svých dalších zkušeností s konvenčními produkty lidské umělecko-řemeslné práce.

VGAI, jak je v době psaní příspěvku tržně implementována, je v první řadě služba, která na zadání generuje produkt s užžitnou hodnotou. Tomu odpovídá způsob práce s tréninkovou databází a asimilace stylů: obchodním cílem je, aby VGAI generovala co možná nejspokojivější reprezentace slovních zadání. Nejspokojivější uměleckou reprezentaci slovního zadání tvoří lidský umělec. Produkty VGAI se tedy stylem nutně budou blížit pracím lidských umělců, na kterých je právě proto VGAI trénováno. Stylová návaznost na lidské vizuální umění je zcela záměrná a z obchodního hlediska nepostradatelná. VGAI generuje reprezentace nebo exempláře stylů, v nichž tvoří živí umělci, a odvolává se přitom na ustavené interpretační a abduktivní techniky lidské vizuální kultury. Řečeno jinými slovy, produkt VGAI se tváří jako lidský výtvoř a žádá, aby s ním tak bylo při pohledu na něj nakládáno, ačkoliv jím není. Samotná nutnost připojování vysvětlujících popisek „použitý obrázek byl vygenerován umělou inteligencí“ poukazuje na relativní úspěšnost zastírající snahy, která z výše zmíněných důvodů musí být abduktivně atribuována autorům i uživatelům VGAI. Je to proto, že VGAI jako sekundární agent pouze facilituje agenci primárního

agenta, jehož intence je v končeném důsledku otištěna v artefaktu. Průvodní snahou primárního agenta je, aby si pacient buď neuvědomil, že obrázek je vygenerován, nebo aby na to nebral zřetel a kognitivně s ním nakládal jako s lidským výtvořem, tedy aby jej dekodoval podle kódu lidského vizuálního umění. Řečeno pojmy teorie agence, primární agent pacienta skrz index *záměrně pobízí k nesprávné abdukci*, která by s indexem pracovala v irelevantních kategoriích „umělcova záměru“ a „reprezentace skutečnosti“.

Posledním krokem příspěvku vystoupíme z agenční teorie do teorie umění, odkud si přisvojíme jeden klíčový termín. Uměnověda totiž zná snahu přimět indexem pacienta k nesprávné abdukci agence a k nesprávnému dekódování založenou na zastírání, říká jí *falzifikace*. Nejde nám o falzifikaci v podobě zhotovení kopie a vydávání této kopie za originální dílo. Jak jsme zdůraznili v úvodu, VGAI nekopíruje, nýbrž variuje databázovaný obsah. Našemu případu lépe odpovídá kontext toho falza, které jeho autor pouze vyvede ve *stylu* napodobovaného autora a následně jej opatří falešným podpisem nebo značkou pro dodání věrohodnosti. Domníváme se, že schéma práce s VGAI v mnoha svých rysech odpovídá takto popsané falzifikační činnosti, ovšem uchopené trochu šířeji a abstraktněji. To, co se napodobuje, je styl, s jakým lidští umělci přistupují k reprezentaci reality. Výsledný produkt se s žádným existujícím dílem neshoduje, pouze některá připomíná natolik, aby se vedle nich mohl zařadit. Svou podobou apeluje na to, aby byl vnímán jako tato díla, nikoliv jako syntetická reprezentace stylu těchto děl. Přitom se nejedná o reprezentaci žádné prožité reality, pouze o nápodobu výsledku tvůrčího procesu bez toho, aniž by musel tento proces být proveden.

Naším výstupem není to, že každá falzifikace je špatná, proto je VGAI špatná. Jak dokazuje praxe vystavování falzifikátů *jako* falzifikátů vedle jejich originálů,¹¹ falzum není neumělecké. Co falzifikát činí potenciálně sdělným je však právě jeho odhalení a demaskování. Náležitě zhodnocení přínosu VGAI a generovaných indexů na poli estetiky¹² nemůže a nebude opomíjet atribuci falzifikační agence primárním agentům, kteří s VGAI nakládají. V každé relační struktuře obsahující takto generované indexy bude pacient nucen nejprve přijmout, že měl být oklamán.

¹⁰ Viz předchozí citace

¹¹ Příkladem je výstava *Falza? Falza!* Kurátorky Olgy Kotkové konaná ve Šternberském paláci od 19. 11. 2021 do 1. 5. 2022, viz [6].

¹² např. schopnosti VGAI nepřímou vytvářet skrz interakci s uměleckým světem hybridní styly nebo dokonce odhadovat budoucí trendy z předešlého vývoje

Literatura

- [1] Alfred Gell: *Art and Agency, an Anthropological Theory*, Clarendon Press, Oxford, 1998
- [2] J. M. Corchado, S. López F., J.M. Núñez V., R. Garcia S., Pablo Chamoso (2023): *Generative Artificial Intelligence: Fundamentals*. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 12(1), e31704. <https://doi.org/10.14201/adcaij.31704> Dave Bergman, Cole Stryker: *What are diffusion models?* [Online] Dostupné z: <https://www.ibm.com/think/topics/diffusion-models>
- [3] Q. Yao: *Application of Artificial Intelligence Virtual Image Technology in Photography Art Creation Under Deep Learning*, in *IEEE Access*, vol. 13, pp. 14542-14556, 2025, doi: 10.1109/ACCESS.2025.3529521, dostupné z: <https://ieeexplore.ieee.org/document/10840187>
- [4] Jos de Mul: The work of art in the age of digital recombination. In *Digital Material: Tracing New Media in Everyday Life and Technology* (Marianne van den Boomen, Sybille Lammes, Ann-Sophie Lehmann, Joost Raessens, and Mirko Tobias Schäfer eds.), Amsterdam University Press, 2009, dostupné z: <http://www.jstor.org/stable/j.ctt46mxjv.9>
- [5] Olga Kotková (ed.): *Falza? Falza!*, Praha: Národní galerie v Praze, 2021, ISBN 978-80-7035-782-8.
- [6] Dave Bergman, Cole Stryker: *What are diffusion models?* [Online] Dostupné z: <https://www.ibm.com/think/topics/diffusion-models>