

# Svet slik v besedah: Kako umetna inteligenca približuje slike slepim in slabovidnim

*Ali lahko AI opisuje slike in fotografije bolje kot človek?*

Marino Kačič

Objavljeno v reviji: [Rikoss 2024 / 1](#)

V svetu sodobne informacijske in komunikacijske tehnologije smo priča napredku, ki slepim in slabovidnim omogoča dostop do raznolikih vsebin. Od komunikacije do branja tekstov tehnologija znatno zmanjšuje nekatere ovire. Kljub temu je ostal izpostavljen ključni izziv: omogočanje dostopa do fotografij in slik, ki predstavljajo nestrukturirano in kompleksno vsebino. Na tem področju umetna inteligenca odpira nova vrata, saj ponuja napredne rešitve za opisovanje in interpretacijo vizualnih informacij, ki so bile do zdaj nedostopne.

## Moč umetne inteligence v dostopnosti vizualnih vsebin

Umetna inteligenca (AI) je revolucionirala številne sektorje in zdaj nosi potencial za korenite spremembe v dostopnosti vizualnih vsebin za slepe in slabovidne. Uporaba naprednih jezikovnih modelov, kot je ChatGPT-4, omogoča umetni inteligenci, da ne le razlaga in opisuje slike, temveč služi kot most med vizualnim svetom in tistimi, ki ga ne morejo neposredno doživeti (Crompton in Burke, 2023).

## Opisovanje slik z ChatGPT-4

Ko posnamete fotografijo in jo želite deliti s slepim ali slabovidnim prijateljem, jo lahko sedaj obogatite z natančnim opisom. S pomočjo ChatGPT-4 lahko sliko preobrazite iz nevidne v celovito pripoved, dostopno slepemu ali slabovidnemu posamezniku. Naložite sliko v ChatGPT-4 in mu podajte navodila, naj opiše vsebino, in tako omogočite globlje razumevanje in občutenje opisane scene.

Oseba, ki prejme opis s pomočjo ChatGPT-4, bi lahko dožela slikovito plažo ob sončnem zahodu v naslednjih besedah: "Zlato sonce se potaplja za obzorje, prelivajoč mirno plažo v topli sijaj. Valovi tiho šume ob obali, ustvarjajoč melodično simfonijo narave. Galebi elegantno jadrajo po nebu, njihova krila se rišejo kot vzorci proti večerni svetlobi, ki se počasi umika noči."

## Dodajanje kontekstualnih informacij k opisom

Vendar pa ChatGPT-4 omogoča preseganje zgolj osnovnih opisov in zagotavlja dodaten kontekst z dodajanjem specifičnih informacij. Z vključitvijo, na primer, imena lokacije, posebne zgradbe ali naravnega čuda, lahko AI-model obogati interpretacijo

s podatki, pridobljenimi iz različnih virov. To zagotavlja, da prejemnik fotografije ne le razume, kaj je na njej, ampak pridobi tudi poglobljeno znanje o lokaciji ali posebnostih predstavljene scene. S tem pristopom AI ne samo da ponuja vizualne opise, temveč vključuje tudi bogate, izobraževalne informacije, ki širijo znanje in razumevanje slepega ali slabovidnega posameznika. Pravzaprav je to uporabno za vse ljudi.

### **Opisovanje vizualnega skozi besede: Konkretni primeri opisov ChatGPT-4**

V nadaljevanju članka bom predstavil konkretno uporabo ChatGPT-4 za opisovanje in razumevanje slik ter fotografij. Predstavil bom serijo testov (opravljenih 5. 3. 2024), kjer sem isto fotografijo predložil AI s petimi različnimi navodili. Rezultati, predstavljeni kot citati, ponujajo vpogled v sposobnost AI, da prilagodi svoje opise glede na različne potrebe in perspektive uporabnikov.



Slika (foto Marjan Leskovar): Obala jezera Manasarovar

### **Primer opisa 1:**

*(Za ta opis sem ChatGPT-ju dal poziv: opiši priloženo sliko)*

»Na fotografiji je prikazano obalno območje ob velikem jezeru. V ospredju vidimo suho, puščavnato pokrajino, s posameznimi suhimi travami, ki prehaja v peščeno obalo. Voda ima prelepe odtenke modre in zelene, ki kažejo na čistočo in globino. Obzorje krasi dramatično nebo s kombinacijo modre, bele in sivih oblakov, ki napovedujejo spremenljivo vreme. V daljavi se vidijo gorski vrhovi, nekateri so morda celo pokriti s snegom, kar daje sceni občutek mirovanja in divjine. Celotna slika oddaja občutek miru, odmaknjenosti in naravne lepote.« (ChatGPT-4, 5.3.2024).

### **Primer opisa 2:**

*(Za ta opis sem ChatGPT-ju dal poziv: opiši sliko človeku, ki ne vidi)*

»Predstavlajte si obsežno in mirno jezersko površino, ki se razteza do obzorja, kjer se nebo in voda srečata v popolni harmoniji. Voda je mirna, z rahlimi odsevi sončne svetlobe, ki se lomijo skozi oblake. Obala je nežna mešanica rumene in rjave, kjer naravna vegetacija nežno prehaja v peščene obale. Začutite lahko suhost zemlje pod nogami, prekinjeno s posameznimi trdimi grmički in travami, ki kljubujejo vetru. Nad vami se razprostira prostrano nebo, deloma jasno, deloma prekrito z debelimi, vlažnimi oblaki, ki napovedujejo prihod dežja ali nevihte. V daljavi, tam, kjer nebo objema zemljo, se dvigajo mogočni gorski vršaci, obdani s snežnimi kapami, ki odsevajo mogočnost in večnost narave. Vsak element na sliki – od vode, zemlje, neba do gora – pripoveduje zgodbo o miru, samoti in neokrnjeni lepoti narave.« (ChatGPT-4, 5.3.2024).

### **Primer opisa 3:**

*(Za ta opis sem ChatGPT-ju dal poziv: napiši podrobnejši opis slike in poleg izstopajočih delov na sliki se osredotoči tudi na opis manjših obrobni delov tako, da jih v opisu proti koncu izpostaviš)*

»Na sliki je upodobljen veličasten prizor jezera. V ospredju slike opazimo nežno rumenkasto-rjavo obalo, prepleteno z suho, nizko, razpršeno vegetacijo, ki daje občutek divje, nedotaknjene narave. Ta del zemlje postopoma prehaja v peščeno obalo, ki obdaja modro-zelene vode jezera, ustvarjajoč mejo med suho zemljo in mokro modrino.

Voda v jezeru deluje mirno in nepremično, kar omogoča, da se nebo z vsemi svojimi barvnimi prelivi odraža na njeni površini. Odražanje oblačnega neba doda jezeru

dodatno globino in dinamiko. Oblaki nad jezerom so mešanica bele in sive, nakazujejo na spremenljivo vreme, morda celo na bližajočo se nevihto, ki pa še vedno dopušča, da se skozi oblake prebije modrina neba.

V daljavi opazimo mogočne gorske verige, ki obkrožajo jezero, s svojimi vrhovi, ki se zdi, da se dotikajo samega neba. Te gore dodajo dramatičen občutek pokrajini, hkrati pa poudarjajo odmaknjenost in izoliranost jezera od preostalega sveta. Sneg, ki pokriva vrhove gora, poudarja njihovo višino in divjost, ter ustvarja popolno nasprotje s toplimi toni zemlje v ospredju.

Ko pogledamo proti robovom slike, lahko opazimo nekoliko bolj subtilne podrobnosti, kot so manjši kamni in grušč na obali, ki pričajo o naravnih procesih erozije in nanosa, ki so oblikovali to pokrajino skozi tisočletja. Manjši valovi ob obali nežno lomijo monotonijo peščenih in travnatih površin ter dodajajo občutek gibanja in življenja v celotno sceno.

V tem umirjenem, a hkrati divjem okolju je mogoče začutiti duhovno energijo in tišino, ki omogočata refleksijo in meditacijo. Ta prizor je več kot le vizualna podoba; je vabilo k razmišljanju o večnosti, lepoti in miru, ki ga narava ponuja v svoji najčistejši obliki.« (ChatGPT-4, 5.3.2024).

#### **Primer opisa 4:**

*(Za ta opis sem ChatGPT-ju dal poziv: opiši sliko in kaj več o jezeru na sliki, ki nosi ime Manasarovar)*

»Na sliki je prikazano jezero Manasarovar, ki leži v odmaknjeni in mirni pokrajini Tibeta, obdano z mogočnimi vrhovi Himalaje. To sveto jezero, ki se nahaja na nadmorski višini približno 4,590 metrov, predstavlja enega izmed najvišjih svežih vodnih jezer na svetu in je znano po svojih čistih, prozornih vodah. Jezero je v obliki skoraj popolnega kroga z obsegom okoli 88 kilometrov in globino, ki doseže do 90 metrov. Njegova površina meri približno 320 kvadratnih kilometrov. Manasarovar je prehranjen z ledeniškimi vodami, kar zagotavlja nenehen dotok sveže in čiste vode, ki dodatno poudarja njegovo svetost in čistost.

Duhovni pomen jezera Manasarovar presega njegovo naravno lepoto, saj je kraj globokih verskih tradicij in mitologije. Jezero je sveto v hinduizmu, budizmu, džainizmu in bonu, kar ga naredi za pomembno romarsko destinacijo. Hindujska mitologija ga omenja kot mesto, ustvarjeno s strani boga Brahme kot prostor za duhovno čiščenje.« (ChatGPT-4, 5.3.2024).

## **Primer opisa 5:**

*(Za ta opis sem ChatGPT-ju dal poziv: napiši še več informacij o tem jezeru)*

»Na fotografiji je upodobljeno jezero Manasarovar, eno najvišjih svežih vodnih jezer na svetu, ki se nahaja v tibetanski avtonomni regiji, blizu gore Kailash. Jezero, znano po svoji izjemni lepoti in duhovnem pomenu, leži na višini 4,590 metrov nadmorske višine in obsega približno 320 kvadratnih kilometrov. Je simbol čistosti in se šteje za sveto v štirih religijah: hinduizmu, budizmu, džainizmu in bonu. Ime "Manasarovar" izhaja iz sanskrtskih besed "Mānas" (um ali zavest) in "sarovara" (jezero), kar skupaj pomeni jezero uma ali zavesti. Jezero, ki je znano po svojih kristalno čistih vodah, zagotavlja osupljivo zrcalno sliko okoliških snežnih gora, kar ustvarja prizor, ki spominja na mirno in sveto območje.

Jezero Manasarovar ni samo naravna lepota, ampak tudi kraj duhovnega iskanja in čiščenja. Za hindujce, jezero predstavlja mesto, kjer se lahko po opravljenem romarskem potovanju okoli gore Kailash očistijo grehov. Budisti verjamejo, da je bila mati Bude prenesena na to sveto mesto, kjer se je okopala v njegovih vodah, da bi očistila svoje telo in duha. Voda jezera Manasarovar se meni, da je tako čista, da lahko vsak obiskovalec jasno vidi do njegovih globljih delov. Ob sončnem vzhodu in zahodu jezero spreminja svoje barve, kar doda dodatno razsežnost njegovi lepoti in svetosti.« (ChatGPT-4, 5.3.2024).

## **Možnosti umetne inteligence za slepe in slabovidne**

Umetna inteligenca odpira nove dimenzije v življenjih slepih in slabovidnih. Z razvojem naprednih jezikovnih modelov, kot je ChatGPT-4, se slepim in slabovidnim posameznikom odpirajo vrata do sveta, ki so ga prej lahko le slutili ali poslušali prek opisov videčih ljudi. Zdaj lahko slike doživljajo skozi podrobne opise, ki jim jih zagotavlja umetna inteligenca. Ta prelom v dostopnosti razkriva obilje možnosti, omogoča vključevanje z vizualno vsebino na načine, ki so bili prej nedostopni (Zhou in Izzo Hunter, 2023).

Vloga umetne inteligence pri podpiranju slepih in slabovidnih bo s stalnim razvojem tehnologije postajala vse večja. Z izkoriščanjem njenih zmožnosti lahko vsaj delno ublažimo razkorak med vizualnim svetom in tistimi, ki ga ne morejo neposredno doživeti. ChatGPT-4 ne samo da interpretira in opisuje slike, temveč jih lahko obogati tudi s kontekstualnimi informacijami, kar slepim in slabovidnim omogoča, da se približajo dojemanju vizualnih vsebin. Poleg slik, ChatGPT-4 izkazuje izjemne

sposobnosti pri opisovanju tabel, grafov in drugih matematičnih predstavitev, kar še bolj širi obzorje njihovega razumevanja in upajmo da lažjega učenja.

Prihodnost, ki jo oblikuje sodelovanje z umetno inteligenco, nosi v sebi veliko upanja. Z nadaljevanjem tehnološkega napredka lahko pričakujemo bolj vključujoč in dostopen svet za vse. Za uresničitev tega cilja je potrebno vztrajati pri razvoju in izboljšanju orodij, ki omogočajo, da bodo postali opisi slik široko dostopni. Čeprav je pot do popolne dostopnosti še vedno pred nami, so trenutni dosežki vzpodbudni in napovedujejo svetlejšo prihodnost, kjer lahko vsi delijo in cenijo bogastvo vizualnega sveta. Ob tem se poraja vprašanje, kako nas bo nadaljnji razvoj umetne inteligence, zlasti v smeri splošne umetne inteligence (GAI), še presenetil z nepredstavljenimi možnostmi. Čeprav se to morda zdi preveč optimistično, ni neverjetno, če upoštevamo, da so mnoge danes samoumevne tehnologije še do nedavnega spadale na področje znanstvene fantastike.

## Viri

Crompton, H. in Burke, D. Artificial intelligence in higher education: the state of the field. *Int J Educ Technol High Educ* 20, 22 (2023).

<https://doi.org/10.1186/s41239-023-00392-8>

The Future of Inclusion: How AI is Revolutionizing Accessibility (2023). Obiskano, 5.3.2024): <https://accessibilityspark.com/ai-and-accessibility/>

Zhou, H. in Izzo Hunter, S. (2023). Guest Post — Accessibility Powered by AI: How Artificial Intelligence Can Help Universalize Access to Digital Content. Obiskano, 5.3.2024: <https://scholarlykitchen.sspnet.org/2023/06/05/guest-post-accessibility-powered-by-ai-how-artificial-intelligence-can-help-universalize-access-to-digital-content/>