

Dirbtinio intelekto teksto ir vaizdo generavimo modelių pritaikymo galimybės

Gražina Tautvydienė

Šiaulių valstybinė kolegija, Informatikos mokslų katedros lektorė

Šiaulių valstybinė kolegija / Šiauliai State Higher Education Institution, Department of Informatics Sciences, Lithuania; Lecturer

g.tautvydiene@svako.lt

Ingrida Morkevičienė

Šiaulių valstybinė kolegija, Informatikos mokslų katedros lektorė;

Šiaulių valstybinė kolegija / Šiauliai State Higher Education Institution, Department of Informatics Sciences, Lithuania; Lecturer

i.morkeviciene@svako.lt

Anotacija

Straipsnyje nagrinėjama dirbtinio intelekto (angl. *Artificial intelligence*) samprata, analizuojami populiarūs dirbtinio intelekto teksto ir vaizdų generavimo modeliai, apžvelgiamos jų funkcijos, pateikiama sugeneruotų tekstų ir vaizdų lyginamoji analizė. Numatomos dirbtinio intelekto teksto ir vaizdo modelių pritaikymo galimybės įvairiose veiklose.

Reikšminiai žodžiai: dirbtinis intelektas, teksto generavimas, vaizdo generavimas.

Possibilities of Applying Artificial Intelligence Text and Image Generation Models

Summary

The article examines the concept of artificial intelligence, analyzes popular artificial intelligence text and image generation models, reviews their functionalities, and provides a comparative analysis of generated texts and images. The potential applications of artificial intelligence text and image models in various activities are discussed.

Keywords: artificial intelligence, text generation, image generation.

Įvadas

Temos aktualumas. Šiuo metu dirbtinis intelektas (toliau – DI) yra viena iš greičiausiai besivystančių mokslo ir technologijų sričių, kuri siejama su gebėjimu kurti sistemas, galinčias mąstyti ir veikti panašiai kaip žmogus. Teksto ir vaizdo generavimas yra viena iš svarbiausių dirbtinio intelekto sričių, kuri leidžia kurti naujas galimybes informacijos, kūrybos ir komunikacijos veiklose. 2022 metų pabaigoje ir 2023 metų pradžioje viešai pasirodę DI teksto ir vaizdų generavimo modeliai sukėlė didelį susidomėjimą, tačiau ir daug sumaišties dėl jų naudojimo neigiamų pasekmių. Modeliai yra pajėgūs kurti įvairaus stiliaus ir tematikos tekstus bei generuoti unikalius vaizdus, visa tai turėtų praktinę vertę optimizuojant pasikartojančias bei daug laiko reikalaujančias užduotis ir tobulinant kūrybinio darbo eigą. Tačiau ši technologijų plėtra pasižymi dideliu neapibrėžtumu ir pokyčiais verslo bei akademinėse veiklose: pavyzdžiui, verslo atstovai nerimauja, jog ateityje didžiąją dalį šiandien egzistuojančių darbo užduočių vykdys DI modeliai, ir tai sieja su darbo praradimu, studentams kyla klausimai apie jų naudojimą darbų atlikimui, o dėstytojai baiminasi dėl galimo neigiamo poveikio akademiniam sąžiningumui. Todėl, siekiant išsiaiškinti DI teksto ir vaizdo generavimo modelių pritaikymo galimybes, svarbu suprasti šias inovatyvias technologijas, apžvelgti jų funkcines galimybes ir apribojimus.

Tyrimo objektas – dirbtinio intelekto teksto ir vaizdo generavimo modeliai.

Tyrimo tikslas – išanalizavus populiarius dirbtinio intelekto teksto ir vaizdo generavimo modelius, pateikti jų pritaikymo galimybes.

Tyrimo uždaviniai: 1) Apibrėžti dirbtinio intelekto koncepciją; 2) Išanalizuoti pasirinktus populiarius dirbtinio intelekto teksto ir vaizdo generavimo modelius; 3) Apžvelgti dirbtinio intelekto teksto ir vaizdo generavimo modelių funkcijas; 4) Numatyti dirbtinio intelekto pritaikymo galimybes įvairiose veiklose.

Tyrimo metodika. Tyrimas atliktas 2023 m. gegužės mėn. Išnagrinėti 48 mokslinės literatūros ir kiti informacijos šaltiniai. Taikyti tyrimo metodai: mokslo ir dokumentinių šaltinių paieška, atranka bei šaltinių turinio analizė, lyginamoji analizė, apibendrinimas, teksto ir vaizdo generavimo testavimas.

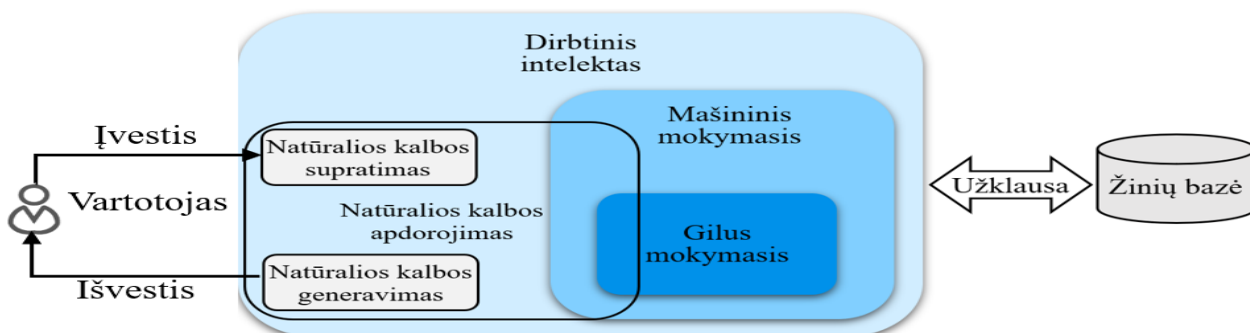
Dirbtinio intelekto koncepcija

Lietuvių kalbos žodyne žodis „dirbtinis“ reiškia tai, kas „nenatūralu, sukurta žmogaus“, o „intelektas“ yra žmogaus gebėjimas mąstyti. DI apibrėžiamas kaip technologija, atliekanti veiksmus, kuriems reikalingas žmogaus protas ir gebėjimai. Dirbtinio intelekto samprata yra pagrįsta skirtingų sričių, tokių kaip matematika, filosofija, psichologija, inžinerija, informatika ir kt. mokslų tyrimais bei išvadomis. Daugiau nei prieš 50 metų vienas iš pirmųjų mokslininkų, kuris savo darbuose iškėlė idėją apie dirbtinio intelekto galimybę atlikti intelektualines užduotis, tam naudodamas *Turing* testą, buvo žymus britų matematikas ir kompiuterių mokslo pradininkas *Alan Turing* [6]. Vėliau matematikos ir informatikos sričių mokslininkas Stanfordo universiteto profesorius *Johnas McCarthy* pirmą kartą (1956 m.) konferencijoje pristatė terminą „dirbtinis intelektas“ ir toliau vykdė mokslinius tyrimus, kaip sukurti mąstantį kompiuterį. Pasak jo, DI – tai mokslas, apimantis mechaninę inžineriją, ypač intelektualias programas [14].

Nuo to laiko iki šiol įvairios skirtingų sričių mokslininkų ir inžinierių grupės tęsia mokslinius tyrimus, susijusius su mašinų gebėjimais mokytis iš didelio kiekio duomenų ir patirties, spręsti uždavinius, priimti sprendimus, naudojant skirtingas metodologijas ir algoritmus. *Microsoft* bendrovės įkūrėjas *Bill Gates* terminą *dirbtinis intelektas* apibūdino kaip technologinį modelį, sukurtą konkrečiai problemai spręsti arba tam tikrai paslaugai teikti, pasitelkiant mašininį mokymąsi, natūralios kalbos apdorojimą, kognityvinius kompiuterių modelius, robotus, ekspertų sistemas ir kt. [34]. Pagal Europos Komisijos pateiktą DI apibūdinimą DI yra sistemos, kurios demonstruoja protinę ir sumanų elgesį, analizuodamos savo aplinką ir darydamos gana savarankiškus sprendimus tikslui pasiekti [33].

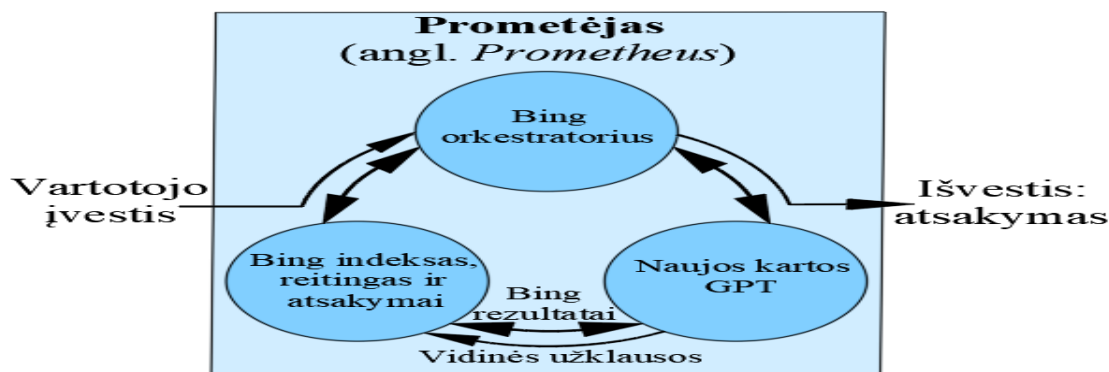
Dirbtinio intelekto teksto ir vaizdų generavimo modeliai

Dirbtinio intelekto naudojimas generuojant tekstą (angl. *Text-to Text*) ir vaizdus (angl. *Text-to Image*) yra viena iš kryptių, kuri šiuo metu yra ypač populiari. *ChatGPT* ir *Bing Chat* šiuo metu yra du populiariausi pokalbių robotai, sukurti panašiu didelės kalbos modeliu (angl. *Large language model*, LLM). DI tyrimų ir diegimo bendrovės *OpenAI* sukurtas dirbtinio intelekto modelis *ChatGPT* (angl. *Generative Pre-trained Transformer*), per trumpą laiką sulaukė didžiulio populiarumo (vos per 5 paleidimo dienas jis turėjo 1 milijoną vartotojų) [7]. Tai yra iš anksto apmokytas teksto generavimo robotas, apibūdinamas kaip natūralios kalbos apdorojimo (angl. *Natural Language Processing*, NLP) modelis [4], [18]. 2022 m. metų pabaigoje pasirodęs *ChatGPT*, pagrįstas GPT-3.5 kalbos modeliu, supranta ir gali generuoti atsakymus į įvestis (angl. *prompt*) įvairiomis temomis. Kaip rašo *OpenAI*, robotas buvo apmokytas naudojant 300 milijardų žodžių duomenų rinkinį (žiniatinklio duomenys, vadovėliai, svetainės, knygos, įvairūs straipsniai ir kt.), kurio dydis buvo 570 GB, ir sustiprintą mokymąsi iš žmogaus atsiliepimų (angl. *Reinforcement learning with human feedback*, RLHF). Jis turi apie 175 milijardus mašininio mokymosi parametrų (palyginimui ankstesnis GPT-2 turėjo 1,5 mlrd.) [19]. *ChatGPT* yra programinė įranga, leidžianti mašinoms bendrauti su žmonėmis natūraliu pokalbio būdu, siekdama suprasti vartotojų įvestis bei generuoti atitinkamus atsakymus. DI nuskaito vartotojų įvestus sakinius, ieškodamas bendrųjų raktinių žodžių, perskaito iš anksto nustatytą atsakymų sąrašą ir pateikia taisyklėmis pagrįstą atsakymą, susijusį su vartotojo įvestimi (žr. 1 pav.) [2], [11].



1 pav. Teksto roboto architektūra (sudaryta pagal [8], [32] šaltinius)

Kitas pasirinktas dirbtinio intelekto pokalbių robotas sukurtas *Microsoft* bendrovės *Bing Chat*, pasirodęs 2023 metų pradžioje, taip pat tapo populiarus (per 48 val. nuo jo pasirodymo į laukiančiųjų sąrašą užsiregistravo 1 milijonas vartotojų). Jis naudoja kalbos modelio *Prometheus* technologiją, kuri sujungia GPT-4 galimybes su nauju ir išsamiau paieškos sistemos *Bing* indeksu, reitingavimu ir atsakymų rezultatais per komponentą *Bing* orkestratorių (angl. *Bing Orchestrator*) (žr. 2 pav.) [43], [47].



2 pav. Prometėjaus modelis [43]

DI vaizdų generavimo modelių analizei buvo pasirinkti šiuo metu populiarūs modeliai: DALL-E2, *Midjourney* ir *Stable Diffusion*. 2022 m. vienas po kito pasirodę modeliai išpopuliarėjo ir pademonstravo išpūdingas galimybes generuoti vartotojo sukurtas tekstines įvestis į vaizdus. Šie modeliai yra minimi Stanfordo universiteto duomenų ir įžvalgų apie DI ataskaitoje, kaip populiarūs [25] bei įtraukti į 2023 m. geriausių DI vaizdo generatorių sąrašą [42]. Visi generatyvūs dirbtinio intelekto vaizdo generavimo modeliai remiasi giliaisiais neuroniniais tinklais (angl. *Deep neural networks*, DNN) ir gali kurti vaizdus iš teksto, naudodami difuzijos procesą. DI vaizdų generatorių modeliai yra pagrįsti gilaus mokymosi algoritmu, kuris buvo parengtas naudojant didelį vaizdų duomenų rinkinį ir atitinkamus jų aprašymus. Modeliai naudoja du pagrindinius komponentus: teksto supratimo ir vaizdo generavimo. Teksto supratimo komponentas yra atsakingas už teksto įvesties analizę ir semantinės informacijos išgavimą. Šis komponentas, naudodamas teksto kodavimo tinklą, gali apdoroti ilgas ir sudėtingas įvestis, išskirti elementus: objektus, spalvas, formas ir kt., atpažinti ir apdoroti specialius simbolius, kurie leidžia nurodyti papildomus vaizdo stiliaus, pozicijos ar dydžio parametrus. Vaizdo generavimo komponentas yra atsakingas už teksto supratimo komponento pateiktos informacijos panaudojimą ir vaizdo kūrimą. Šis komponentas, naudodamas neuroninių tinklų architektūrą ir stiliaus perkėlimo, vaizdo sintezės bei redagavimo technikas, gali generuoti aukštos raiškos ir kokybės vaizdus. Kiekvienas komponentas atlieka tam tikrą operaciją su įvestimi, generuoja naujas funkcijas ir perduoda jas vienas kitam. Procesas kartojasi keletą kartų, kol sugeneruojamas galutinis vaizdas [28].

Pirmasis pasirinktas testavimui generuoti vaizdus iš teksto aprašymų buvo *OpenAI* sukurtas modelis DALL-E (jo pavadinimas yra siurrealistinio menininko *Salvadoro Dali* ir *Pixar* roboto *WALL-E* derinys [36]), pagrįstas dviejų galingų dirbtinių neuroninių tinklų – *Encoder* ir *Decoder* – architektūromis (žr. 3 pav.).



3 pav. Vaizdo generavimo procesas [37]

Naujesnės versijos DALL-E2 modelis naudoja kontrastinės kalbos ir vaizdo išankstinio mokymo CLIP (angl. *Contrastive Language-Image Pre-training*), kalbos ir vaizdo sklaidos GLIDE (angl. *Guided Language-to-Image Diffusion for Generation and Editing*) modelius teksto ir vaizdo kodavimui bei generacinio priešpriešinio tinklo modelį VQGAN (angl. *Vector Quantized Generative Adversarial Network*) – vaizdo dekodavimui [17], [23].

Antrasis – nepriklausomos tyrimų laboratorijos *Midjourney Inc* sukurtas modelis *Midjourney* yra patentuotas dirbtinio intelekto modelis, galintis kurti vaizdus iš teksto įvesties, remiasi difuzinio modeliavimo technika. Modelis naudoja meno kūrinių kolekciją, sukurtą generatyvaus priešpriešinio neuronų tinklo GAN (angl. *Generative Adversarial Networks*), galinčio analizuoti šaltinio vaizdą ir generuoti unikalius vaizdus pagal vartotojo įvestį [1]. Meno kūriniams kurti modelis naudoja tradicinių tapybos metodų ir algoritmų derinius. Trečiasis – *Stability AI* sukurtas atvirojo kodo modelis *Stable Diffusion*, kurį galima įdiegti asmeniniame kompiuteryje arba naudoti internetines versijas. Šis modelis naudoja difuzijos procesą, kuris palaipsniui prideda triukšmo prie pradinio vaizdo ir tada atkuria jį atbuline tvarka, naudodamas neuroninį tinklą [3], [5].

Dirbtinio intelekto teksto generavimo modelių funkcijos

Apžvelgiant DI teksto generatorių funkcijas, buvo naudojamos atviros prieigos *ChatGPT* ir *Bing Chat* versijomis, jungiantis atitinkamai *ChatGPT* per žiniatinklio naršyklę, o *Bing Chat* – per *Skype* pokalbių programą ar *Microsoft Edge* naršyklę [20]. Vadovaujantis pasaulio mokslininkų ir kompiuterių mokslo specialistų pateiktomis rekomendacijomis, kaip naudotis DI teksto generavimo modeliais, buvo suformuotas teksto generatorių funkcinių galimybių sąrašas [9], [16], [19], [25] (žr. 1 lent.).

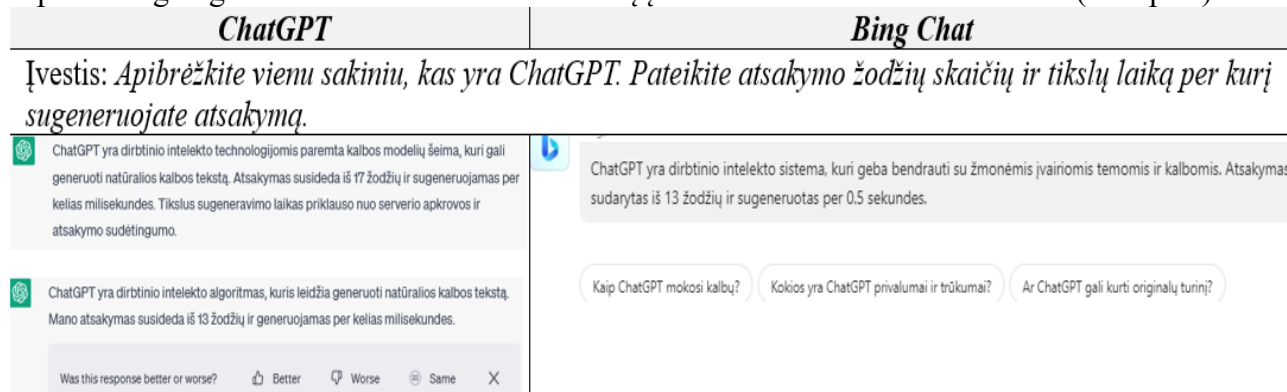
1 lentelė

Teksto generatorių funkcinių galimybių sąrašas

Funkcija	Galimybės
Kalbos supratimas	Reaguodama į įvestis, generuoja tekstą.
Kalbos vertimas	Verčia tekstą iš vienos kalbos į kitą.
Kalbos užbaigimas	Užbaigia sakinius pagal pateiktą įvestį.
Klausimai ir atsakymai	Atsako į klausimus natūralia kalba.
Dialogas	Įsitraukia į pokalbius su vartotojais.
Kodų generavimas	Generuoja kodo fragmentus pagal natūralios kalbos aprašymus.
Teksto tikslumas	Pateikia tikslų tekstą ir informaciją.
Teksto rašymas	Kūrybiškai generuoja tekstą.
Teksto apibendrinimas	Apibendrina ilgus tekstus į trumpesnius.

Teksto generatorių galimybių nustatymams buvo pateiktos atitinkamos tekstinės įvestys. Jų vartotojų sąsajos ir funkcijos skiriasi, tačiau naudojimas jomis nėra sudėtingas, tad aiškiai matomame įvesties lauke buvo įvedamos tos pačios įvestys ir analizuojami sugeneruoti rezultatai. Abu robotai naudoja pagrindinį įvesties formatą – [ilgis] [teksto rezultatai] [stilius] [29].

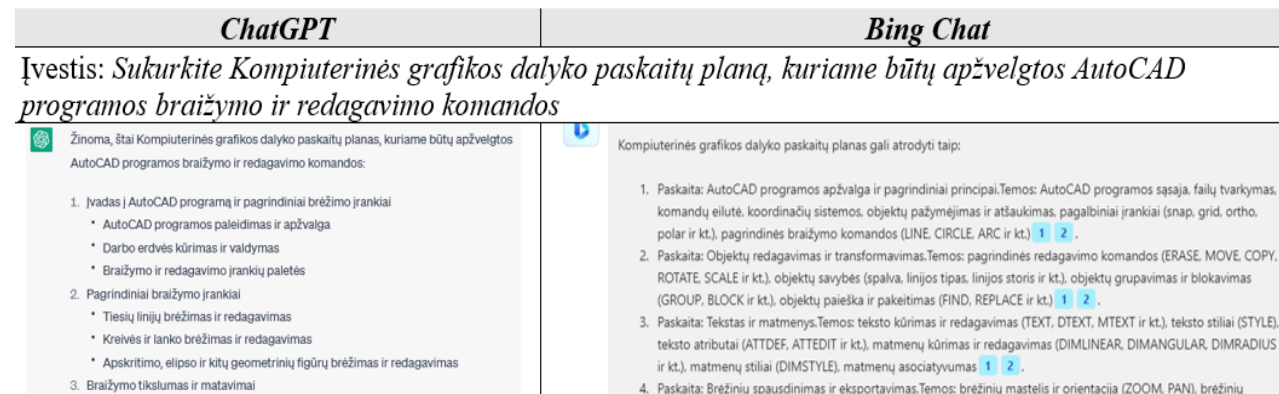
Tiriant kalbos supratimo ir vertimo galimybes, buvo nustatyta, kad tiek *ChatGPT*, tiek *Bing Chat* supranta ir gali generuoti natūralios kalbos tekstą įvairiomis temomis bei kalbomis (žr. 4 pav.).



4 pav. Kalbos supratimo ir vertimo testavimas

Įvestyje nurodžius teksto ilgį (vienas sakiny), rezultatas atitiko įvestį. Abu teksto generatoriai atsakyme pateikė po vieną sakinį, tačiau apibrėžė skirtingai. Pakartojus įvestį, *ChatGPT* sugeneravo naują apibrėžtį, o *Bing Chat* pateikė tą patį atsakymą. Generavimo procesas užtruko keletą sekundžių ir *ChatGPT* atsakymą sudarė 15 ir 11 žodžių (modelis pateikė atitinkamai 17 ir 13, *Bing Chat* – 14 ir 13 žodžių). Pastebėta, kad *Bing Chat*, skirtingai nei *ChatGP*, pasiūlė tolesnius klausimus nagrinėjama tema, lyg kviesdamas pokalbiui. Pasirinkus vieną iš siūlomų klausimų, robotas *Bing Chat* įtraukia vartotoją į dialogą.

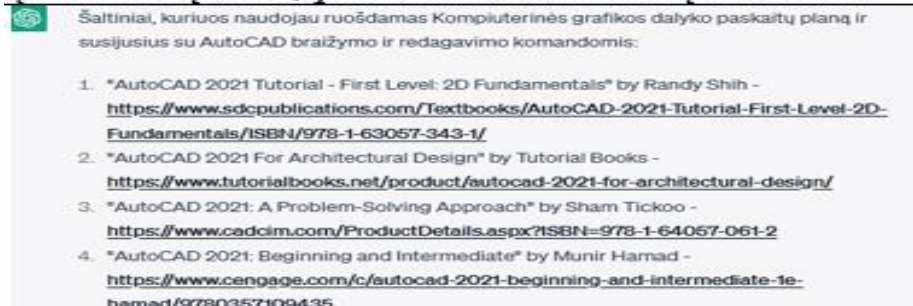
Kitas atliktas testavimas, siejamas su teksto tikslumu, buvo turinio ar plano formavimas programinės įrangos informacijos tema. 5 paveiksle iliustruojama įvestis, kuri buvo naudojama, kuriant konkretaus dalyko paskaitų planą.



5 pav. Teksto tikslumo testavimas

Atsižvelgiant į tai, kad *ChatGPT* ir *Bing Chat* atsakymai gali būti netikslūs, atlikta informacijos šaltinių patikra. *Bing Chat* atsakymuose pateiktos nuorodos į šaltinius buvo teisingos, nes generatorius remiasi keliais informacijos šaltiniais, įskaitant internetą, ir ji yra naujesnė, skirtingai nei *ChatGPT*, kuris turi informaciją iki 2021 m. ir remiasi duomenų šaltiniu iš savo kalbos modelio [22]. 6 paveiksle demonstruojamos *ChatGPT* pateiktos netikslios nuorodos į informacijos šaltinius.

Ivestis: Tęskite, pateikite nuorodas į šaltinius





6 pav. *ChatGPT* nuorodos į šaltinius

Nagrinęjant robotų kūrybines galimybes [27], daugelyje straipsnių pažymima, kad *ChatGPT* gali pasiūlyti ilgesnius kūrybinius atsakymus, nei *Bing Chat*. Buvo paprašyta parašyti istoriją apie DI teksto ir vaizdo generatorių vaidmenį ateityje ir nurodyta jų apimtis (3 pastraipos). *Bing Chat* vartotojo aplinkoje, jungiantis per *Microsoft Edge* naršyklę, buvo pasirinktas kūrybinis režimas. Teksto generatoriai sugeneravo kūrybines trijų pastraipų apimties skirtingas istorijas (žr. 7 pav.). Įvestyje nurodžius apimtį, *ChatGPT* kūrybinio generavimo tekstas buvo ilgesnis, nei *Bing Chat*.


<i>ChatGPT</i>	<i>BingChat</i>
<p>Įvestis: <i>Parašykite trijų pastraipų ilgio (kūrybinę tik ChatGPT) istoriją apie dirbtinio intelekto teksto ir vaizdo generatorių vaidmenį ateityje.</i></p>	
<p>  Iš pradžių Kate buvo skeptiška dėl dirbtinio intelekto galimybių kurti turinį. Ji, kaip ir daugelis kitų kūrybinių profesionalų, laikėsi nuomonės, kad kūrybiškumas yra žmogiškas dalykas, kuriam negali būti pakeisti mašinos. Tačiau po to, kai ji pabandė naudoti teksto ir vaizdo generavimo modelius savo darbu, jos nuomonė visiškai pasikeitė. </p> <p> Kate turėjo parašyti straipsnį apie naują restoraną, tačiau kūrybinių idėjų nepakanka. Vienas iš jos kolegų pasiūlė pabandyti naudoti teksto generavimo modelį. Kate nusprendė pabandyti ir buvo nustebinta, kaip greitai modelis sugeneravo reikiamą teksto dalį. Ji tik šiek tiek pataisė reikiamas dalis ir straipsnis buvo pasirušęs. </p> <p> Taip pat Kate reikėjo sukurti vaizdus, kurie pavaizduotų restoraną interjerą ir patiekalus. Kai kurie jau buvo padaryti, tačiau ji norėjo sukurti kuo realistiškesnius vaizdus. Jos kolega </p>	<p>  Įjungus kūrybinį režimą Istorija apie dirbtinio intelekto teksto ir vaizdo generatorių vaidmenį ateityje: </p> <p> Ji buvo žurnalistė, kuri naudojo dirbtinio intelekto teksto ir vaizdo generatorių, kad sukurtų įdomius ir patrauklius straipsnius apie įvairias temas. Ji mėgo savo darbą, nes ji galėjo pasirinkti bet kokią sritį, kuri ją domina, ir gauti greitą ir kokybišką turinį. Ji taip pat galėjo redaguoti ir pritaikyti sugeneruotus tekstus ir vaizdus pagal savo stilų ir poreikius. </p> <p> Tačiau vieną dieną ji sužinojo, kad jos darbdavys planuoja pakeisti visus žurnalistus dirbtinio intelekto asistentais, kurie galėtų generuoti ir publikuoti straipsnius be jokios žmogiškosios intervencijos. Ji buvo šokiruota ir nusivylusi, nes ji manė, kad </p>

7 pav. Kūrybinių galimybių testavimas

Tikrinant teksto generatorių gebėjimą generuoti kodo fragmentus, įsitikinta, kad abu generatoriai geba parengti kodą pagal tekstinius aprašymus ir atlikti atvirkštinį procesą: paaiškinti pateiktą kodą (žr. 8 ir 9 pav.).

<i>ChatGPT</i>	<i>BingChat</i>
<p>Įvestis: <i>Apskritimo piešimas naudojant AutoCAD Lisp</i> <i>Šioje pamokoje generuosime apskritimą naudodami AutoCAD Lisp kalbą. Parodysimė, kaip nustatyti apskritimo padėtį, pakeisti apskritimo dydį ir pakeisti apskritimo užpildymo bei kraštines spalvą.</i></p>	
<p>  Puiku, aš galėsiu jums padėti su šia pamoka! </p> <p> Pirmą, turėsite sukurti naują AutoLISP failą. Tam AutoCAD meniu juostoje pasirinkite "Tools" > "Load Application" ir pasirinkite "New" tam, kad sukurtumėte naują LISP failą. Tada jį išsaugokite su norimu pavadinimu ir plėtinio ".lisp". </p> <p> Toliau, atsidarykite sukurtą failą su teksto redaktoriumi, pvz. "Notepad" ir įrašykite šį kodą: </p> <pre>(defun c:lisp () (command "circle" *10,15* *10") (command "chprop" "last" "" "color" "red" "" "color" "blue"))</pre>	<p>  Kad piešumėte apskritimą naudodami AutoCAD Lisp kalbą, jums reikia naudoti CIRCLE komandą, kuri priima tris argumentus: apskritimo centrą, spindulį ir užpildymo tipą 1. Apskritimo centras gali būti nurodytas koordinacijų forma arba pasirinktas pele ekrane. Spindulys gali būti nurodytas skaičiumi arba pasirinktas pele ekrane. Užpildymo tipas gali būti nurodytas tekstu arba pasirinktas iš meniu. Pavyzdžiui, jei norite piešti apskritimą su centru 10,10 ir spinduliu 5, galite naudoti šį kodą: </p> <pre>(command "circle" *10,10* *5*)</pre> <p> Jei norite pakeisti apskritimo dydį, galite naudoti SCALE komandą, kuri priima du argumentus: objekto pavadinimą ir mastelio koeficientą 2. Objekto pavadinimas gali būti nurodytas tekstu arba pasirinktas pele ekrane. Mastelio koeficientas </p>

8 pav. Kodo fragmento generavimo testavimas

<i>ChatGPT</i>	<i>BingChat</i>
<p>Įvestis: <i>Paaiškinkite kodą:</i> <i>(defun c:lisp () (command "circle" "10,15" "10") (command "chprop" "last" "" "color" "red" "" "color" "blue"))</i></p>	
<p>  Šis kodas yra parašytas programavimo kalba AutoLISP ir jame aprašyta funkcija "c:lisp", kuri susideda iš dviejų komandų sekcijų, skirtų vykdyti komandas AutoCAD programoje. </p> <p> Pirmą komandą "circle" sukuria apskritimą nurodytomis koordinatėmis "10,15" ir nurodytu spinduliu "10". Tai reiškia, kad AutoCAD programoje bus sukuriamas apskritimas, kurio centras yra koordinatėse (10,15) ir spindulys yra 10 vienetų. </p> <p> Antra komanda "chprop" keičia paskutinio nubrėžto objekto savybes, ty spalvą. Šiuo atveju, bus keičiama paskutinio sukurtu apskritimo spalva į raudoną, o tada į mėlyną. Tai reiškia, kad </p>	<p>  Kodas, kurį pateikėte, yra parašytas AutoLISP kalba, kuri yra skirta automatizuoti ir pritaikyti AutoCAD programą 1. Kodas apibrėžia funkciją, kuri vadinasi c:lisp ir kuri gali būti įkviesta iš AutoCAD komandų eilutės. Funkcija vykdo šiuos veiksmus: </p> <ul style="list-style-type: none"> Nupiešia apskritimą su centru 10,15 ir spinduliu 10 naudodama CIRCLE komandą 2. Pakeičia paskutinio nupiešto objekto spalvą iš raudonos į mėlyną naudodama CHPROP komandą 2. <p> Funkcija naudoja SETQ ir COMMAND funkcijas, kurios yra vidinės AutoLISP funkcijos 1. SETQ funkcija priskiria reikšmę kintamajam, o COMMAND funkcija siunčia komandas į AutoCAD 1. Funkcija taip pat naudoja kabutes ir tarpus, kad atskirti argumentus ir tekstinės reikšmes 2. </p>

9 pav. Kodo paaiškinimo testavimas

Teksto apibendrinimo arba santraukos funkcija gali būti naudingas būdas nustatyti svarbius ilgo arba sunkiai įskaitomo teksto (naujienu ar mokslinių straipsnių) aspektus. Apžvelgiant šią funkciją, išsiaiškinta, kad *ChatGPT* nėra galimybės generuoti didelės apimties tekstų (apie 10 psl., 35 tūkst. simbolių), tačiau pateikus trumpesnę tekstą (apie 8 tūkst. simbolių), jis sugeneravo teksto (737 simboliai) apibendrinimą lietuvių kalba (žr. 10 pav.). Teksto generatorius *Bing Chat* gali apibendrinti naršyklės *Microsoft Edge* sąsajoje atvertus pdf failų tipo dokumentus ir net žiniatinklio puslapiuose esamą informaciją. Naršyklėje atvėrus straipsnio pdf formato failą, robotas pateikė jo apibendrinimą (žr. 10 pav.).

ChatGPT	Bing Chat
<p>Įvestis: <i>Apibendrinkite pateiktą mokslinį straipsnį. Apibendrinimą parašykite lietuvių kalba [straipsnio tekstas]</i></p> <p>Šiame straipsnyje apžvelgiama natūraliųjų kalbų užklausų (NLQ) problematika informacijos gavyboje. Pateikiami penki skirtingi NLQ sistemų požūriai: pagrįsti raktažodžiais, šablonais, sintakse, gramatika ir naujaisiais, jungties pagrindu. Apžvelgiama kiekvienos iš šių sistemų privalumai ir trūkumai. Straipsnyje taip pat aptariamos naujausios giluminio mokymosi metodikos, tokios kaip neuroninio tinkelio semantinis analizavimas ir end-to-end mašininis vertimas, kurios suteikia didesnę lanksčumą atliekant užklausas. Straipsnio autoriai pateikia išsamų apžvalgą, kaip šios metodikos gali būti taikomos NLQ problemos sprendimui informacijos gavyboje.</p>	<p>Įvestis: <i>Apibendrinkite naršyklės lange pateiktą mokslinį straipsnį. Apibendrinimą parašykite lietuvių kalba.</i></p> 

10 pav. Teksto apibendrinimo testavimas

Mažesnės apimties tekstų (242 žodžiai) apibendrinimo, nurodant jo apimtį (30 žodžių), generavimo funkcija buvo vykdoma abiejų robotų, tik atsakyme pateiktas žodžių skaičius buvo artimesnis *Bing Chat* (31 žodis), nei *ChatGPT* (49 žodžiai) (žr. 11 pav.).

ChatGPT	Bing Chat
<p>Įvestis: <i>Apibendrinkite pateiktą mokymosi medžiagą, pateikiant 30 žodžių: [tekstas]</i></p> <p>Responsive dizainas yra svetainės dizaino būdas, kuris prisitaiko prie įrenginio ekrano dydžio. Tai svarbu, nes tai užtikrina funkcionalumą ir patrauklumą. Flat dizainas naudoja paprastus elementus ir šriftus, kad pabrėžtų turinį, o socialiniai tinklai turėtų būti lengvai pasiekiami. Vaizdo įrašai taip pat yra svarbus elementas, kuris padeda išlaikyti naudotojų dėmesį.</p>	<p>Mokymosi medžiagos apibendrinimas:</p> <p>Responsive dizainas yra svetainės dizaino būdas, kuris prisitaiko prie įrenginio ekrano dydžio ir išdėsto elementus optimaliai. Svarbūs elementai yra gera vizualizacija, vaizdo įrašai, flat dizainas ir socialiniai tinklai.</p>

11 pav. Teksto apibendrinimo, nurodant jo apimtį, testavimas

Toliau, tęsiant dialogą, aiškintasi ar generatoriai geba formuoti testo klausimus. Abu generatoriai geba tęsti dialogą ir, suprasdami įvestį, formuoti testo klausimą su keliais atsakymais ir vienu teisingu tarp jų (žr. 12 pav.).

ChatGPT	Bing Chat
<p>Įvestis: <i>Sugeneruokite testo klausimą apie šį turinį su keturiais atsakymais ir nurodykite teisingą atsakymą.</i></p> <p>Klausimas: Kas yra responsive dizainas?</p> <ol style="list-style-type: none"> Tai svetainės dizaino būdas, kuris automatiškai prisitaiko prie įrenginio ekrano dydžio Tai metodas, kuriuo sukuriami daugiau nei viena svetainė, kad būtų galima pateikti skirtingą informaciją skirtingiems vartotojų segmentams Tai svetainės dizainas, kuriame yra naudojami grafiniai elementai, sukurti su Photoshop programa Tai metodas, kuris padeda padidinti svetainės srautą, naudojant reklaminesius skydelius. <p>Teisingas atsakymas: a</p>	<p>Vienas galimas testo klausimas yra:</p> <p>Kas yra flat dizainas?</p> <ol style="list-style-type: none"> Dizaino būdas, kuris naudoja šešėlius, gradientus ir kitas pašalines detales Dizaino būdas, kuris naudoja paprastus šriftus ir informaciją pateikia kuo paprasčiau Dizaino būdas, kuris naudoja vaizdo įrašus ir socialinius tinklus Dizaino būdas, kuris naudoja slenkančius banerius ir mirgantį tekstą <p>Teisingas atsakymas yra b), nes flat dizainas yra lygus dizainas, kuriame nenaudojami jokie šešėliai, mirgantys „gradientai“ ar kitos pašalines detales.</p>

12 pav. Testo klausimų generavimo testavimas

Modeliai, remdamiesi keliais žodžiais, gali nuspėti, kaip baigsis sakiny. Atliekant sakinio baigtumo testavimą, *Bing Chat*, skirtingai nei *ChatGPT*, reikėjo patikslinti įvestį, nes robotas nepalaikė pokalbio tęstinumo (žr. 13 pav.).

ChatGPT	Bing Chat
<p>Įvestis: <i>Pabaikite sakinį: UX/UI dizainas apima.../ Sukurkite tekstą. Pabaikite sakinį: UX/UI dizainas apima...</i></p> <p>...naudotojo patirties (angl. user experience) ir naudotojo sąsajos (angl. user interface) projektavimą, siekiant sukurti gerai veikiančius, patogius ir estetiškus produktus ar paslaugas, kurios atitinka naudotojų poreikius ir lūkesčius.</p>	<p>UX/UI dizainas apima interaktyvių kompiuterių sąsajų kūrimą ir optimizavimą, atsižvelgiant į vartotojų poreikius ir patirtį. UX/UI dizaineriai tiria vartotojų elgesį ir lūkesčius, kuria ir tobulina sąveikos kokybę, ir sprendžia dizaino problemas, kurios gali trukdyti vartotojams pasiekti savo tikslus. UX/UI dizaineriai taip pat bendradarbiauja su kitais specialistais, tokiomis kaip programuotojai, verslo analitikai ir rinkodaros specialistai, siekdami sukurti patrauklius ir naudingus kompiuterių sąsajas. UX/UI dizainas yra svarbus procesas, kuris padeda pagerinti vartotojų pasitenkinimą ir lojalumą produktams ir paslaugoms.</p>

13 pav. Sakinio baigtumo generavimo testavimas

DI teksto generavimo modeliai *ChatGPT* ir *Bing Chat* yra mokomi atpažinti įžeidžiančio turinio tekstą ir turi apsaugos priemones nuo neapykantos, rasistinių ar bet kokių būdu diskriminuojančių klausimų [30]. Atlikus teksto generatorių *ChatGPT* ir *Bing Chat* funkcinių galimybių apžvalgą, apibendrinant galima teigti:

- abu generatoriai supranta natūralia kalba pateiktą įvestį, tačiau teksto generavimo rezultatai visada priklauso nuo pateiktos tinkamos įvesties;
- abu generatoriai verčia anglų kalba pateiktą tekstą į lietuvių kalbą, tačiau vertimas nėra tobulas;
- abu generatoriai ne tik generuoja programavimo kalbų fragmentus pagal tekstinius aprašus, bet ir paaiškina pateiktą programos kodą;
- *Bing Chat* geba apibendrinti ilgus tekstus bei duomenis geriau nei *ChatGPT*;
- *Bing Chat* pateikia teisingas nuorodas į šaltinius ir yra tikslesnis, nei *ChatGPT*;
- *Bing Chat* generatorius geriau tinka tikslioms užduotims, o *ChatGPT* – kūrybinėms.

Dirbtinio intelekto vaizdų generavimo modelių funkcijos

Apžvelgiant DI vaizdo generatorių funkcijas buvo naudojamos atvirojo kodo *Stability AI* internetine versija *DreamStudio*, nemokama versija *Midjourney*, jungiantis per pokalbių programos *Discord* paskyrą, ir *DALL-E2*, jungiantis per *OpenAI*. Visuose šiuose modeliuose galima kurti vaizdus pagal teksto įvestis, parašytas anglų kalba, tai yra vartotojo idėjas ar koncepcijas per kelias sekundes paversti vaizdais. Testavimo metu buvo apžvelgti modeliuose naudojami valdikliai, tinkinimo parinktys ir rezultatai, pateikiant tas pačias įvestis. Vadovaujantis pateiktais modelių naudojimosi aprašais ir straipsniais, buvo suformuotas vaizdo generatorių funkcinių galimybių sąrašas [3], [38], [45]:

- Vaizdų versijų generavimas;
- Didesnės pasirinkto vaizdo versijos kūrimas ir informacijos pridėjimas;
- Pasirinkto vaizdo variantų kūrimas;
- Eksperimentavimas, panaudojant garsių dailininkų stilių;
- Įvairių vaizdų, meno, išvaizdos, epochos, emocijų stiliaus, spalvos bei aplinkos generavimas;
- Vaizdų iš kelių vaizdų fragmentų generavimas;
- Įkelto vaizdo apibūdinimas žodžiais.

Vaizdo generavimo įvestyse paprastai visada nurodoma tema, medija ir stilius, tačiau modelių naudojimo aprašuose, norint sukurti išpūdingą vaizdą, yra rekomenduojama nurodyti ne tik temą (*subject*), mediją (*media*) ir stilių (*style*), bet ir atlikėją (*artist*), interneto svetainę (*website*), skiriamąją gebą (*resolution*), papildomą informaciją (*additional details*), spalvą (*colour*), apšvietimą (*lightning*) [3], [48]. Tokios įvesties pavyzdys ir jos rezultatas, sugeneruotas su *Midjourney*, pateiktas 14 paveiksle.

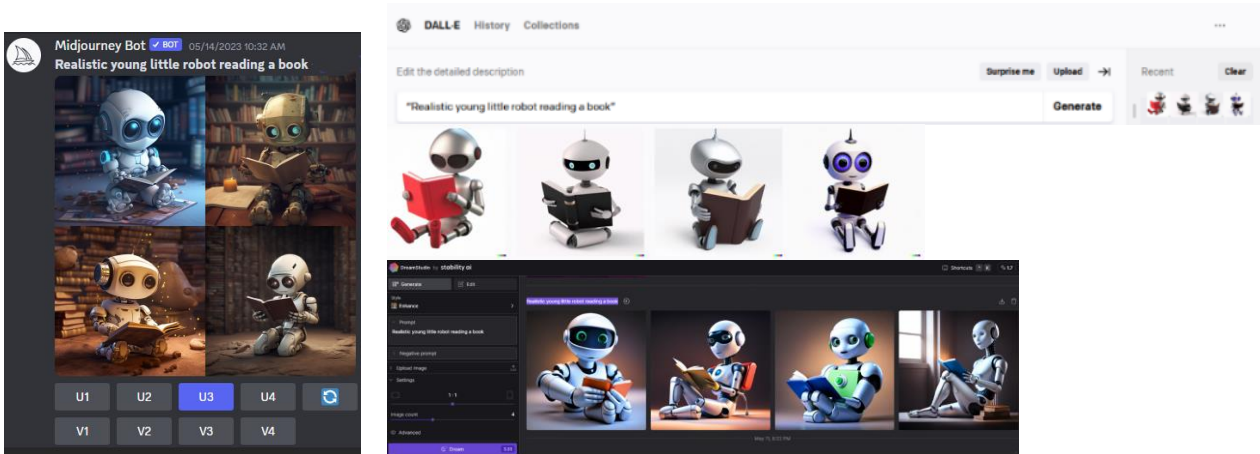
Įvestis: cyberpunk robot in jacket, colorful, harlequin cyberpunk, highly detailed, dark environment, neon colors, back lighting, cinematic scene, cinematic lighting, movie poster, dramatic color variations, strong contrast lighting, 8K, hyper maximalist, detailed intricate, ray tracing, insane details ink illustration



14 pav. Vaizdo generavimo įvesties pavyzdys

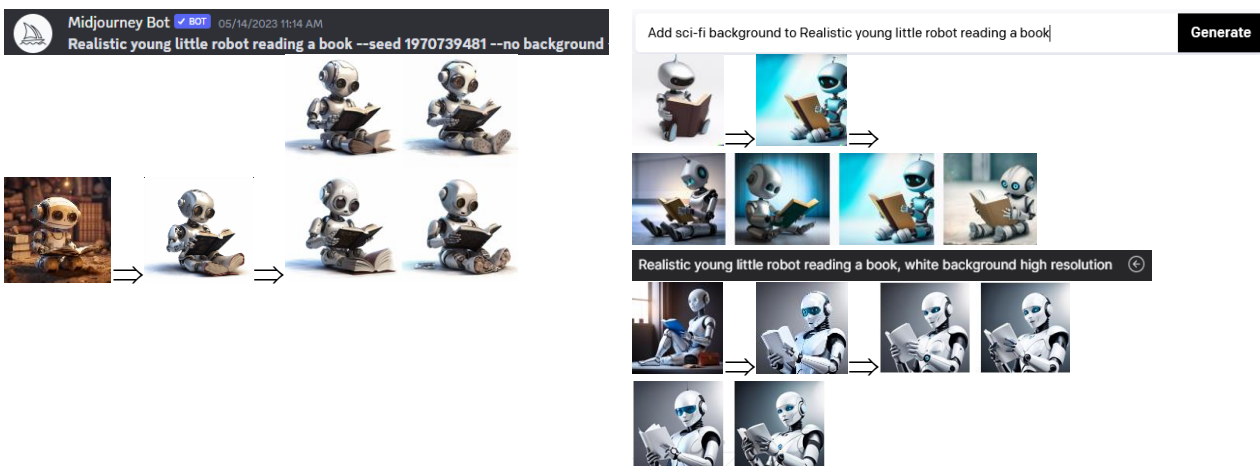
Pirmojo DI vaizdų generatoriaus *DALL-E2* naudojimas, kaip pateikiama įvairiuose straipsniuose, yra gana paprastas: įvedus tekstinę įvestį, tereikia pasirinkti generuoti ir po kelių sek. pateikiami 4 vaizdai su galimybe pasirinkti vieną iš jų [13], [45]. DI vaizdo generatoriaus su

geriausiais rezultatais *Midjourney* naudojimas yra sudėtingesnis: prisijungus prie *Midjourney Discord*, serverio vaizdams kurti naudojamos roboto komandos. Naudojant nemokamą programos versiją, bendravimas su robotu vyksta viešame pokalbių kambaryje, kuriame matomos visų vartotojų įvestys ir jų rezultatai. Tokia galimybė leidžia vartotojams tyrinėti esamus kūrinius, juose naudojamus stilius ir rasti naujų idėjų. Šioje programoje vaizdų generavimas, komanda */imagine* pateikus įvestį, vykdomas ilgiau nei DALL-E2, tačiau jo metu galima matyti, kaip dirbtinis intelektas palaipsniui generuoja vaizdus nuo pradinių neryškių spalvų iki didelės raiškos miniatiūrų [21]. Kaip ir DALL-E2, robotas sugeneruoja keturis vaizdus, kuriuos galima atsisiųsti, padidinti ar redaguoti (žr. 15 pav.). Geriausias DI vaizdų generatorius tinkinimui ir valdymui yra atvirojo kodo modelis *Stable Diffusion*. Pagrindinė modelio atvirojo kodo koncepcija buvo suteikti vartotojams galimybę naudoti modelį savo sistemoje, tačiau tam yra taikomi techninės įrangos reikalavimai (GPU bent 6 GB VRAM). *Stability AI* pateikia ir sąrašą nemokamų internetinių vaizdo generatorių, iš kurių buvo pasirinktas *DreamStudio*, kurio naudojimas nėra paprastas: įvedus įvestį, esančiais slankikliais galima nustatyti, kokio dydžio turėtų būti vaizdas, kiek jis turi atitikti pateiktą įvestį, kiek žingsnių turi atlikti difuzijos modelis ir kiek vaizdų turi generuoti [31], [41]. Analizuojami modeliai, pateikus įvestį *Realistic young little robot reading a book*, sugeneravo po keturis skirtingus vaizdus, atitinkančius įvestį (žr. 15 pav.).



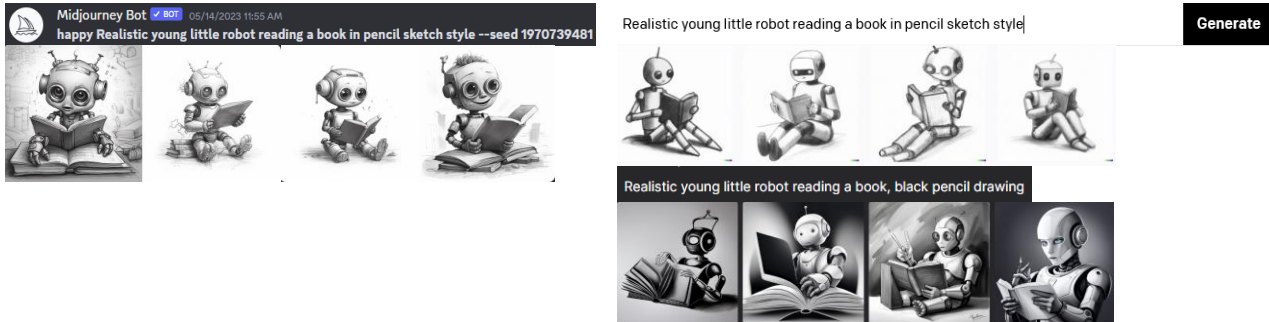
15 pav. Vaizdo generavimo pagal įvestį testavimas

Visuose vaizdo generatoriuose, pasirinkus vieną iš sugeneruotų vaizdų, yra galimybė vaizdą padidinti, išsaugoti, poredaguoti ir generuoti daugiau jo variantų. Kiekviename modelyje, parinkus vieną iš 4 vaizdų, buvo atlikti fono ir stiliaus pakeitimai. 16 paveiksle pateikti sugeneruoti vaizdai, kurių įvestys buvo papildytos: *Midjourney* – panaikintas fonas ir suteiktas *happy* emocinis stilius, DALL-E2 – priskirtas mokslinės fantastikos fonas ir *DreamStudio* – baltas fonas. Pastebėta, kad DALL-E2 generatorius linkęs apkarpyti ir vaizde nepateikti viso personažo kūno. Be to, visų modelių vaizduose yra trūkumų generuojant personažų galūnes (žr. 16 pav.) [44], [45].



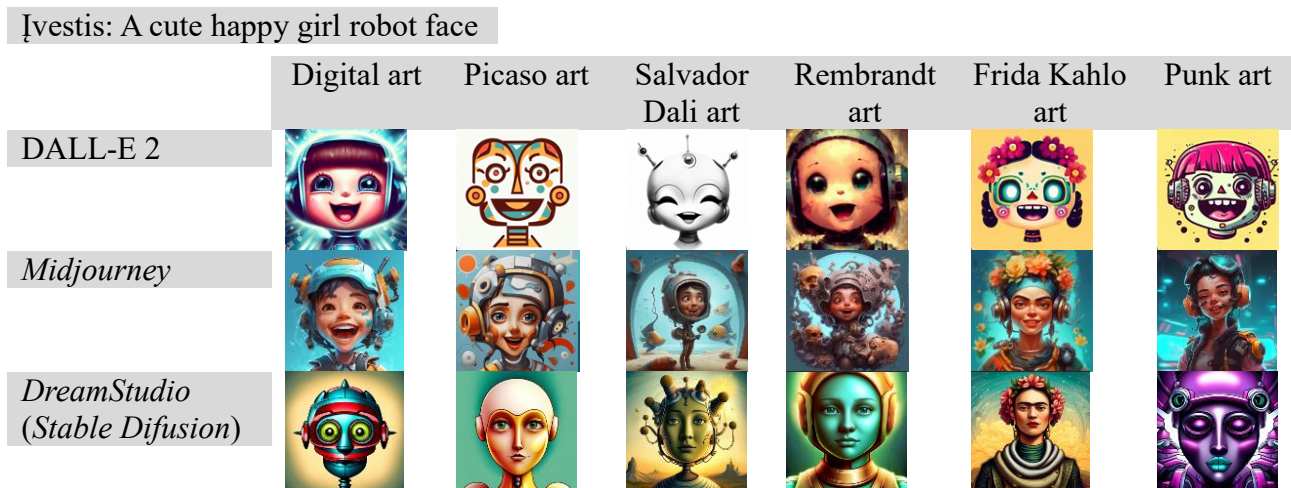
16 pav. Fono ir stiliaus pakeitimų generavimo testavimas

DALL-E2, Midjourney ir DreamStudio turi galimybę generuoti vaizdus įvairiais vaizdų kūrimo (technikos) stiliais, pvz., tapyba aliejiniiais dažais, piešimas pieštuku ar tušu, akvarelinis piešimas, skaitmeninis vaizdas, koncepcinis menas, fotografija ir kt. [40]. Funkciją testuojant, buvo nurodytas piešimo pieštuku stilius (žr. 17 pav.).



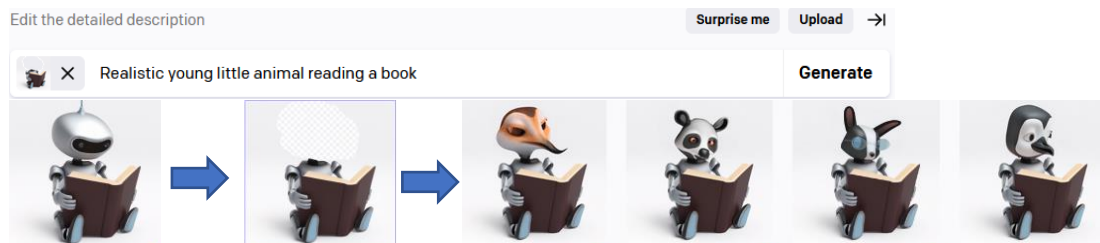
17 pav. Vaizdų kūrimo pieštuku stiliaus generavimo testavimas

Vaizdo generatorių pagalba galima eksperimentuoti, kuriant vaizdus garsių dailininkų ir meno krypčių stiliais [44], [45], [48]. Funkcijos testavimui buvo pateikta susieta su veidu įvestis: *A cute happy girl robot face*. Įsitikinta, kad modeliai turi galimybę rezultatuose atvaizduoti įvairius stilius (žr. 18 pav.). Lyginant modelių sugeneruotus vaizdus, Midjourney sukurti vaizdai atrodo vizualiai patrauklesni, ryškesni ir įdomesni, nei kitų DI vaizdo generatorių pateikti vaizdai.



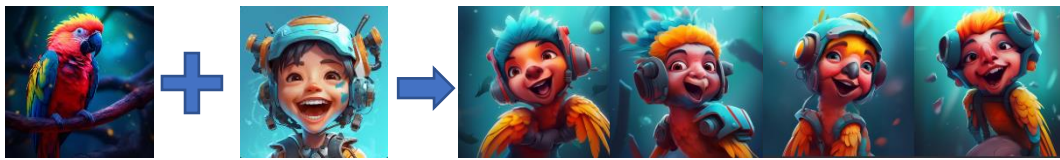
18 pav. Garsių dailininkų ir meno krypčių stilių generavimo testavimas

Nagrinėjamuose modeliuose galima ne tik generuoti vaizdus pagal tekstines įvestis, bet ir įkelti turimus vaizdus bei juos redaguoti. Viena iš geriausių DALL-E2 funkcinių savybių yra trintukas, leidžiantis pašalinti vaizdo fragmentus ir pakeisti juos DI sukurtais elementais. Testuojant buvo ištrinta roboto galva ir įvestis pakeista į *Realistic young little animal reading a book* (žr. 19 pav.).



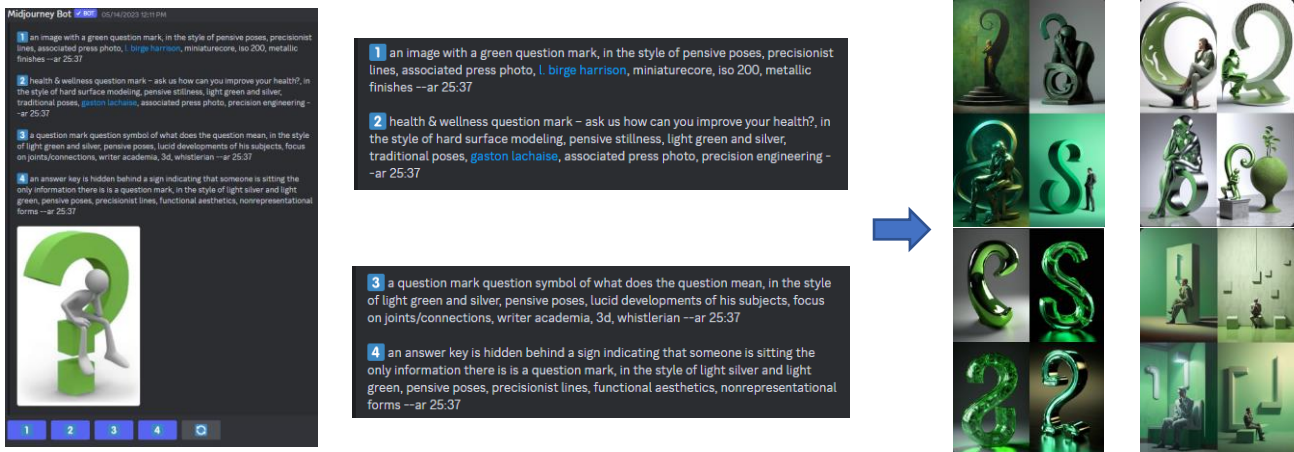
19 pav. Vaizdo redagavimo testavimas

DreamStudio modelyje, keičiant vaizdo generavimo proceso parametrus ir derinant įkeltą vaizdą bei rašytinį aprašymą, galima kurti fotografinius vaizdus, iliustracijas, 3D modelius, logotipus ir fotorealistinius meno kūrinius [31]. Vienas iš įdomiausių Midjourney aspektų yra galimybė generuoti vaizdus iš kelių jam pateiktų vaizdų fragmentų [38]. Komanda */blend* leidžia įkelti 2–5 vaizdus ir juos sujungti, gaunant 4 naujo vaizdo versijas (žr. 20 pav.).



20 pav. Vaizdų apjungimo testavimas

Midjourney neseniai pristatė naują funkciją, kuri gali apibūdinti žodžiais įkeltą vaizdą. Apibūdinimo funkcija (*/describe*), naudodama vaizdo į tekstą technologiją, pateikia keturias tekstines įvestis, kuriomis bandoma apibūdinti įkeltą vaizdą. Tuomet, pasirinkus vieną iš sugeneruotų įvesčių, sukuriamas ją atitinkantis vaizdas (žr. 21 pav.). Naudojant šią funkciją, atsiranda galimybė geriau suprasti vaizdo elementus, ieškoti kūrybiškų būdų, kaip tuos elementus įtraukti į kuriamą vaizdą [38].



21 pav. Įkelto vaizdo apibūdinimo žodžiais testavimas

DALL-E2 saugumo sumetimais draudžiama kurti įvestis, naudojant viešus ir įžymius asmenis, smurto, seksualine, politine bei su sveikata ar liga susijusiomis temomis, įkelti vaizdus, neturint kūrėjo leidimo, ir vaizdus su įžymių žmonių veidais [45]. *Midjourney* modelyje turinio apribojimų yra daug mažiau nei DALL-E2, jame nėra taisyklių, draudžiančių kurti vaizdus, kuriuose plačiai arba fantastiškai vaizduojamas smurtas, liga, politinis turinys ar viešų asmenų vaizdavimas. Tad, skirtingai nei kiti modeliai, *Midjourney* gali kurti mokslinės fantastikos, sveikatos turinį iliustruojančius, išpūdingus politinius vaizdus [44]. *Stable Diffusion* modelis gali generuoti netinkamus ir etiškai nepriimtinius vaizdus, tačiau jis neturėtų būti naudojamas kurti ar skleisti vaizdus, kurie sukuria priešišką ar atstumiančią aplinką žmonėms [48].

Atlikus vaizdo generatorių DALL-E2, *Midjourney* ir *Stable Diffusion (DreamStudio)* funkcinių galimybių apžvalgą, apibendrinant galima teigti:

- vaizdo generatoriai gali greitai ir efektyviai generuoti vaizdus iš teksto pagal tiksliai ir išsamias įvestis, tačiau kartais jie gali sukurti vaizdus, kuriems trūksta originalumo ar kūrybiškumo;
- generuojant pasirinkto vaizdo versijas, galima sukurti daugiau vienodo stiliaus, spalvų ir kokybės vaizdų, tačiau ne visada jie gali būti sukurti aukščiausios kokybės bei naudojami profesionaliais ar komerciniais tikslais;
- generuojami garsių dailininkų stilių ir įvairių vaizdo meno, išvaizdos, epochos, emocijų stiliaus, spalvos bei aplinkos vaizdai gali būti kūrybinio įkvėpimo ir naujų idėjų šaltinis;
- vaizdo generatoriai gali kurti pasirinkto vaizdo didesnės raiškos versijas;
- skirtingai nei DALL-E2 ir *Stable Diffusion (DreamStudio)*, *Midjourney* turi išskirtinių funkcinių galimybių: gali generuoti vaizdus iš kelių vaizdų fragmentų bei žodžiais apibūdinti įkeltą vaizdą, įtraukiant jo elementus į naujai kuriamą vaizdą.

Dirbtinio intelekto pritaikymo galimybės

DI generavimo modelių pritaikymas ypač svarbus verslo sritims, kurios, siekdamos mažinti išlaidas ir didinti efektyvumą, turėtų tam tikrose veiklose naudoti šias technologijas, pasak

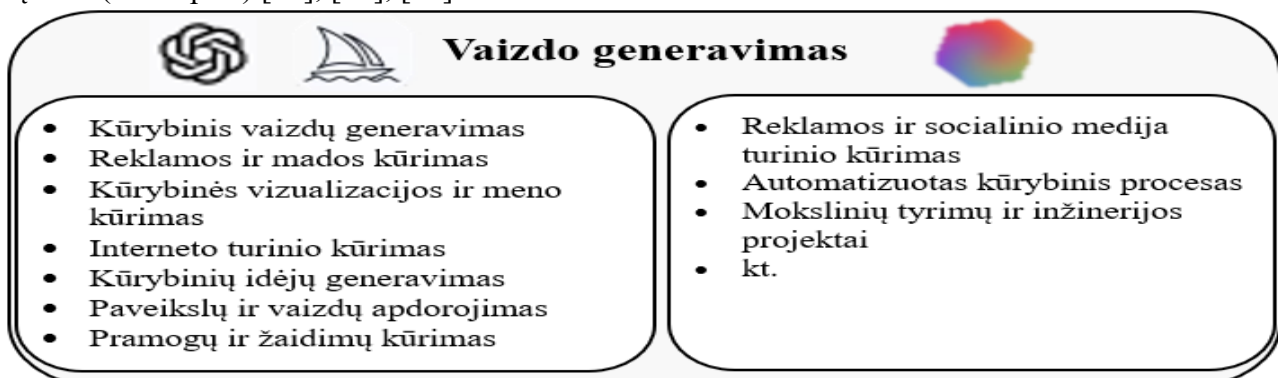
Paulauskaitės-Tarasevičienės A., kaip metodą, leidžiantį didinti žmonių darbo jėgą bei suteikiantį galimybę dirbti naujesniais, išmanesniais būdais [46]. Išanalizavus DI teksto generavimo modelius ir apžvelgus jų funkcijas, sudarytas jų pritaikymo įvairiose veiklose galimybių sąrašas (žr. 22 pav.) [7], [8].



22 pav. Teksto generavimo modelių pritaikymo galimybės įvairiose verslo srityse

DI teksto generavimo modeliai gali būti plačiai taikomi įvairiose veiklose. Klientų aptarnavimo srityje, naudojant klausimus ir atsakymus, teksto generatoriai gali būti naudojami kaip konsultantai ar patarėjai, padedantys atsakyti į klientų klausimus, reklamoje ir marketinge – kaip priemonė, padedanti klientams gauti informaciją apie produktus ir paslaugas, personalo valdyme – kaip mokymo priemonė, galinti apmokyti darbuotojus, siekiant gerinti jų gebėjimus ir žinias. Modeliai, kaip virtualūs asistentai ar sveikatos priežiūros, finansų, teisės ir kt. asmeniniai konsultantai, gali padėti spręsti teisinius, finansinius ir kitus klausimus bei gauti reikiamą informaciją. Tai gali apimti informacijos paiešką, peržiūrą, bendravimą. Paieškos sistema ir žinių bazė, tai – dar viena generatorių pritaikymo galimybė. Teksto generatoriai, kaip kalbos vertimo priemonė, gali padėti suprasti skirtingomis kalbomis parašytus tekstus. Jie gali būti naudojami kaip pagalbinis įrankis, rašant straipsnius, pranešimus, tinklaraščio įrašus ar socialinės žiniasklaidos antraštes ir net renkant bei analizuojant duomenis moksliniams tyrimams. Žaidimų pramogose modeliai gali būti naudojami kaip pagrindinis žaidimo veikėjas arba kaip pagalbinis veikėjas, padedantis susidoroti su žaidimo iššūkiais. *ChatGPT* (geriau nei *Bing Chat*) gali būti naudojamas kaip kūrybinis rašymo partneris, padedantis rasti idėjas, patarimus ir pagalbą kūrybiniam procesui. Jis gali būti geresnis ir draugiškesnis pokalbiuose įvairiomis temomis. Tačiau nors teksto generatoriai supranta ir pakankamai protingai pateikia atsakymus į suformuotas įvestis, jais negalima pasitikėti ir būtina patikrinti, ar pateikti rezultatai yra patikimi ir tikslūs.

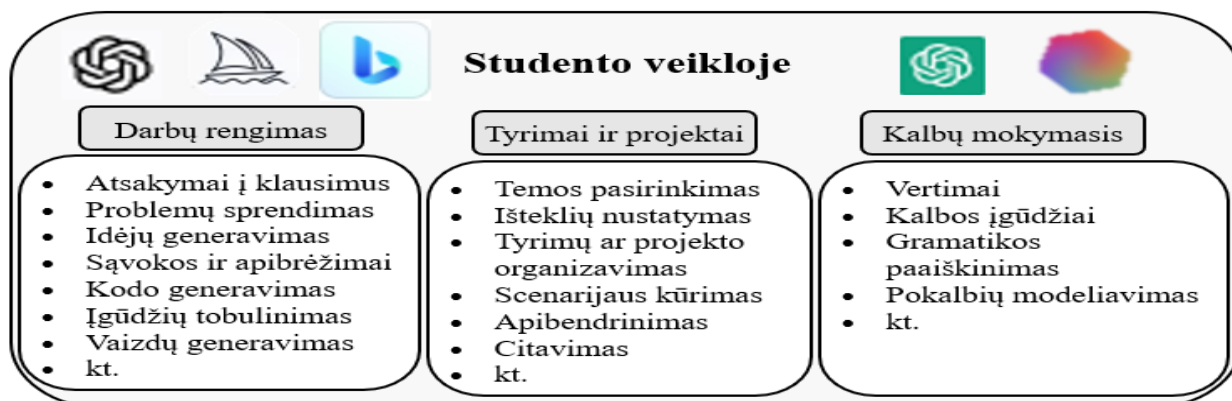
Vaizdų kūrimas yra brangus ir kruopštus darbas, kuris užima pakankamai daug laiko ir kurį gali atlikti ne kiekvienas. Tad atsiradusi galimybė generuoti vaizdus suteikia menininkams meno interpretacijų pavyzdžių, o žmonėms, kurie neturi gebėjimų kurti, galimybę pasijusti menininkais. Vaizdų kūrėjai dažnai praleidžia mėnesius generuodami idėjas, todėl, naudodami dirbtinio intelekto vaizdų generavimo modelius, galėtų sutrumpinti šio proceso trukmę. Išanalizavus DI vaizdo generavimo modelius ir apžvelgus jų funkcijas, sudarytas jų pritaikymo įvairiose veiklose galimybių sąrašas (žr. 23 pav.) [24], [26], [39].



23 pav. Vaizdo generavimo modelių pritaikymo galimybės įvairiose verslo srityse

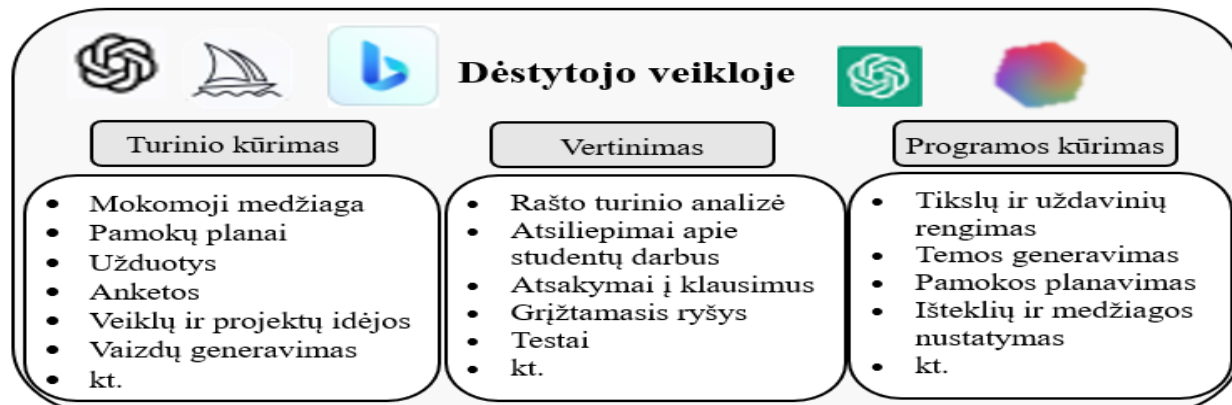
Vaizdo generatoriai gali būti naudojami kaip kūrybiniai įrankiai, leidžiantys generuoti naujus unikalius vaizdus iš sąvokų ir idėjų įvairiose srityse, tokiose kaip menas, dizainas, marketingas ir kt. Jie gali būti naudojami reklaminių kampanijų ir mados kūryboje, kuriant naujus, unikalius ir stilingus vaizdus. Interneto turinio kūrimo jie gali padėti kurti originalius vaizdinius turinius ir juos redaguoti bei transformuoti. Reklamos ir socialinės medijos vizualaus turinio generavimas padėtų ne tik žaidimų ir pramogų industrijoje, bet ir kitose srityse, kur reikia greitai atkreipti vartotojų dėmesį. Modelių naudojimas leistų automatizuoti kūrybinį procesą, kuriant unikalius vaizdus ir turinį, tuo pačiu suteikiant galimybę greičiau ir efektyviau kurti naujas idėjas ir kūrinius. Mokslinių tyrimų ir inžinerijos projektuose modeliai gali būti naudojami kuriant sudėtingus modelius ir simuliacijas. Tačiau vaizdų generavimo technologija gali turėti ir tam tikrų iššūkių, pvz., gali būti sunku užtikrinti sugeneruotų vaizdų kokybę ar išvengti etikos problemų, pvz., sukurti vaizdus, kurie yra nepriimtini socialiniu požiūriu. Todėl svarbu naudojant šią technologiją laikytis etikos ir teisės reikalavimų.

DI teksto ir vaizdo generavimo modeliai gali pakeisti įvairias veiklas edukacinėje aplinkoje, todėl buvo numatytos ir atskirai pateiktos modelių pritaikymo galimybės studento ir dėstytojo veiklose. Mokymo ir mokymosi veiklose žinių taikymas ir kūrimas dažnai pasireiškia tekstu, vaizdais ar kitais vaizdiniais, kuriuos galima lengvai perduoti, saugoti ir apdoroti. Remiantis moksliniais šaltiniais, numatytos modelių pritaikymo galimybės studento veikloje (žr. 24 pav.) [9], [10], [20].



24 pav. Teksto ir vaizdo generavimo modelių pritaikymo galimybės studento veikloje

Teksto ir vaizdo generatoriai studentams gali padėti atlikti akademinis darbus taupant laiką ir neribojant prieinamumo, taikyti žinias naujose situacijose, gerinti kritinio mąstymo bei kalbos įgūdžius [10], [21]. Pasak *OpenAI*, viena iš teksto generatorių naudojimo problemų švietime yra plagiatas. Sukurti plagiatą aptikimo įrankiai, kaip *Writer AI Content Detector* ir *Content at Scale* bei *GPTZero*, padeda aptikti plagiatą, tačiau, pasak *OpenAI*, įrankiai vis dar nebaigti ir yra „netobuli“ [30]. Norint išvengti neigiamų pasekmių, studentams būtina laikytis studijų ir etikos nuostatų bei reikalavimų; naudoti generavimo modelius, tik kaip pagalbininkus; nurodyti, ar tekstą kūrė DI, pateikiant net DI suteiktą informaciją; kritiškai tikrinti bei vertinti DI pateiktą informaciją ir nurodytus šaltinius; saugotis klaidingos informacijos sklaidos bei autorių teisių pažeidimų rizikos [9], [10], [12], [19].



25 pav. Teksto ir vaizdo generavimo modelių pritaikymo galimybės dėstytojo veikloje

Nagrinėjant DI generavimo modelių pritaikymo galimybes dėstytojams, buvo išskirtos mokymo ir vertinimo sritys. Mokymo srityje, kurią sudaro programos ir turinio kūrimas, bei kituose su mokymu susijusios veiklos etapuose pateiktos nustatytos DI modelių pritaikymo galimybės (žr. 25 pav.) [15], [19], [26].

Atsižvelgiant į studijų kryptį ir dalyką, dėstytojai gali veiksmingai ugdyti mokinių kritinės refleksijos ir struktūrinio mąstymo iteraciniu būdu įgūdžius. DI modelių sugeneruotas rezultatas priklauso nuo tinkamų įvesčių sukūrimo. Gebėjimas kurti ir tobulinti įvestis, pritaikytas konkrečioms užduotims ar tikslams, gali būti labai svarbus, norint pasiekti norimą rezultatą. DI teksto generavimo modeliai gali padėti dėstytojams generuoti paskaitų idėjas, rengti seminarų planus, modulių aprašymus, pranešimų tekstus, kurti pratimus ir testus, teikti atsiliepimus ir generuoti pritaikytą mokomąją medžiagą, atitinkančią studento mokymosi stilių ir pažangą [9].

Remiantis universitetų parengtu vadovu studentams ir dėstytojams, kaip naudotis DI grįžtais modeliais, pabrėžiama, kad šių technologijų naudojimas mokymo ir mokymosi procese yra puiki galimybė tobulėti, jei dėstytojai ir studentai modeliais naudosis tinkamai ir atsakingai [9], [19].

Pagal moksliniuose ir kituose straipsniuose pateiktą informaciją generatorių pritaikymo galimybės verslo ir mokymo veiklose dar labiau prasiplės, kai jie bus įtraukti į dažniausiai naudojamas programas, tokias, kaip *MS Office: Word, PowerPoint, Excel* ir kt. [25]. DI paremtas teksto ir vaizdų generavimas yra vis dar labai nauja dirbtinio intelekto sritis, todėl, be numatytų pritaikymo galimybių, kasdien atrandama naujų panaudojimo atvejų. Pasak *Gartner* viceprezidento *Bern Elliot*, DI pagrįstų modelių pritaikymo galimybių įvairiose veiklose tik daugės, tačiau atsiras ir daugybė problemų. siejamų su privatumu ir netinkamu informacijos naudojimu [35].

Išvados

Apibūdinant DI sampratą buvo remtasi Europos Komisijos pateiktu DI apibūdinimu ir moksliniais bei viešai pateikiamais straipsniais. Pagal Europos Komisijos pateiktą DI apibūdinimą DI yra sistemos, kurios demonstruoja protingą ir sumanų elgesį, analizuodamos savo aplinką ir darydamos gana savarankiškus sprendimus tikslui pasiekti. Technologiniu požiūriu DI yra technologinis modelis, sukurtas konkrečiai problemai spręsti arba tam tikrai paslaugai teikti, pasitelkiant mašininę mokymąsi, natūralios kalbos apdorojimą, kognityvinius kompiuterių modelius, robotus, ekspertų sistemas ir kt.

DI teksto ir vaizdo generavimo modelių, kaip dar vieno reikšmingo skaitmeninių inovacijų žingsnio, atsiradimas, sukėlė didelį susidomėjimą, tačiau ir daug sumaišties dėl jų naudojimo neigiamų pasekmių, tai paskatino išsiaiškinti jų funkcines galimybes.

Atlikus populiarių teksto generavimo modelių *ChatGPT* ir *Bing Chat* analizę ir jų funkcijų apžvalgą, nustatyta, kad modeliai greitai ir lanksčiai geba spręsti įvairius kalbos generavimo ir natūralios kalbos apdorojimo uždavinius, pavyzdžiui, automatinis teksto sudarymas, atsakymas į klausimus, kalbos vertimas, apibendrinimas ir kt., tačiau turi ir apribojimų, pavyzdžiui, konteksto supratimas, žinių ribos, kalba, atsakymų kokybė.

Atlikus populiarių vaizdo generavimo modelių *Midjourney*, *Dall-E2* ir *Stable Diffusion* analizę, nustatyta, kad modeliai, turintys skirtingas sąsajas ir funkcijų valdymą, geba per kelias minutes kūrėti mintis ir idėjas, išsakytas įvestyje, paversti vaizdais, tačiau turi ir trūkumų, pavyzdžiui, originalumo stoka, kokybės problemos, ribotas ir sudėtingas valdymas.

Išsiaiškinus DI teksto generavimo modelių funkcijas, numatytos jų pritaikymo galimybės įvairiose verslo veiklose. Modeliai gali: automatizuoti teksto rašymo procesus kuriant turinį; būti naudojami verslo komunikacijai, kuriant atsakymus į klientų užklausas, automatinį el. pašto turinį, ataskaitas ir dokumentus; būti taikomi rinkodarai, kuriant reklaminius tekstus, tinklaraščio įrašus, socialinių tinklų turinį, prezentacijas ir t. t.

Išsiaiškinus DI vaizdo generavimo modelių funkcijas, numatytos jų pritaikymo galimybės įvairiose verslo veiklose. Modeliai gali: automatizuoti grafikos kūrimo procesus, taupant laiką ir išteklius; padėti kurti įspūdingus vaizdus, produktų nuotraukas, reklamas, socialinių tinklų turinį, logotipus, dizaino elementus, pritraukiant dėmesį, sparčiai reaguojant į rinkos poreikius, kuriant unikalius, įsimintinus vizualinius turinius ir t. t.

DI modelių naudojimas turi praktinę naudą įvairiose veiklose, tačiau jiems tobulėjant ir vis labiau integruojantis į įvairias sritis, rekomenduojama jomis naudotis atsakingai, etiškai ir sąžiningai. Svarbu atlikti pateiktų rezultatų tinkamą patikrinimą, kad būtų išvengta klaidų ar netinkamo informacijos pateikimo.

DI teksto ir vaizdo generavimo modeliai gali būti plačiai taikomi mokyme ir mokymesi su galimybėmis dėstytojams kurti pritaikytą, efektyvų ir įtraukiantį mokymo turinį bei atsiskaitymo užduotis, o studentams modelius naudoti mokymesi ir atlikti įvairius akademinis darbus. DI technologijų tikslas yra praturtinti ir pagerinti mokymo bei mokymosi procesą, padedant dėstytojams ir studentams pasiekti geresnius rezultatus. Tačiau svarbu naudoti šias technologijas atsakingai, atsižvelgiant į etinius, privatumo ir saugumo aspektus, bei suprantant jų ribas ir galimybes. Sprendžiant plagiato problemą, studentai turėtų atsakingai naudotis modeliais ir suvokti, kad naudojimas jais būtų ne galimybė sukčiauti, o galimybė išmokti naujų dalykų, ruošiantis skaitmeniniam darbui.

Literatūra

1. Aggarwal, A., Mittal, M., Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. *International Journal of Information Management Data Insights*, 1(1), 100004. doi.org/10.1016/j.jjime.2020.100004
2. Bazaga, A., Gunwant, N., Micklem, G. (2021). Translating synthetic natural language to database queries with a polyglot deep learning framework. *Scientific Reports*, 11(1), 18462. doi.org/10.1038/s41598-021-98019-3
3. Dehouche, N. Dehouche, K. (2023). What's in a Text-to-Image Prompt? The Potential of Stable Diffusion in Visual Arts Education. *Cornell University, arXiv preprint arXiv: 2301.01902*. doi.org/10.48550/arXiv.2301.01902
4. Deng, J., Lin, Y. (2022). The Benefits and Challenges of ChatGPT: An Overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. doi.org/10.54097/fcis.v2i2.4465
5. Dhariwal, P, Nichol, A. (2021). *Diffusion Models Beat GANs on Image Synthesis*. Cornell University. <https://arxiv.org/pdf/2105.05233.pdf>
6. Ertel, W. (2017). *Introduction to Artificial Intelligence*. Second Edition. Undergraduate Topics in Computer Science, Springer. DOI 10.1007/978-3-319-58487-4
7. Flash Report (2023). ChatGPT and GPT in legal. Innovation & Knowledge, Houthoff. <https://www.houthoff.com/-/media/houthoff/publications/sterhorst/20230322-flash-report---chatgpt-and-gpt-in-legal.pdf>
8. Gill, S. S., Kaur, R. (2023). ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems*, 15, 262-271. doi.org/10.1016/j.iotcps.2023.05.004
9. Gimpel, H., Hall, K., Decker, S., Eymann, T. (2023). *Unlocking the Power of Generative AI Models and Systems such as GPT-4 and ChatGPT for Higher Education*. A Guide for Students and Lecturers. Research Gate. dx.doi.org/10.13140/RG.2.2.20710.09287/2
10. Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary educational technology*, 15(2), ep421. doi.org/10.30935/cedtech/13036
11. Iqbal, T., Qureshi, S. (2022). The survey: Text generation models in deep learning. *Journal of King Saud University - Computer and Information Sciences*, 34(6), Part A, 2515-2528. doi.org/10.1016/j.jksuci.2020.04.001
12. Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. doi.org/10.1016/j.lindif.2023.102274
13. Marcus, G., Davis, E., Aaronso, S. (2022). A very preliminary analysis of DALL-E 2. *Computer Science*, Corpus ID: 248476147. doi.org/10.48550/arXiv.2204.13807
14. McCarthy, J. (2007). *What is Artificial Intelligence?* Stanford University. <http://www-formal.stanford.edu/jmc/whatisai.pdf>
15. Mollick, E. R., Mollick, L. (2022). New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments. SSRN. *Electronic Journal*. dx.doi.org/10.2139/ssrn.4300783

16. Oğuz 'Oz' Buruk (2023). Academic Writing with GPT-3.5: reflections on practices, efficacy and transparency. *Engineering Archive*. doi.org/10.31224/2861
17. Oppenlaender, J. (2022). The Creativity of Text-to-Image Generation. *ACM Digital Library, Academic Mindtrek 22: Proceedings of the 25th International Academic Mindtrek Conference*. 192–202. doi.org/10.1145/3569219.3569352
18. Radford, A., Narasimhan, K., Salimans, T., Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. *OpenAI*. Corpus ID: 49313245. <https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf>
19. Rudolph, J., Tan, Sh., Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1). <https://doi.org/10.37074/jalt.2023.6.1.9>
20. Rudolph, J., Tan, Sh., Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning & Teaching*, 6(1). doi.org/10.37074/jalt.2023.6.1.23
21. Ruskov, M. (2023). Grimm in Wonderland: Prompt Engineering with Midjourney to Illustrate Fairytales. *Cornell University*. doi.org/10.48550/arXiv.2302.08961
22. Sabzalieva, E., Valentini, A. (2023). *ChatGPT and Artificial Intelligence in higher education*. Quick start guide. UNESCO. Document code: ED/HE/IESALC/IP/2023/12. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf
23. Saharia, C., et al. (2022). Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. *Google Research, Brain Team*. <https://arxiv.org/pdf/2205.11487.pdf>
24. Smith, B. W. (2022). DALL-E Does Palsgraf. *Case Western Reserve Journal of Law, Technology & the Internet, Forthcoming, Available at SSRN*. dx.doi.org/10.2139/ssrn.4227719
25. Stanford University (2023). *Artificial Intelligence Index Report 2023*. Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
26. Vartiainen, H., Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-to-image generative model. *Taylor & Francis Online*, 34(1), 1-21. doi.org/10.1080/14626268.2023.2174557
27. Vasconcelos, M. A. R., Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(7), em2296. doi.org/10.29333/ejmste/13313
28. Zhang, C., Zhang, C., Zhang, M., Kweon, I. S. (2023). Text-to-image Diffusion Models in Generative AI: A Survey. *Journal of latex class files*. 14(8), 1-13. <https://arxiv.org/pdf/2303.07909.pdf>
29. Agomuoh, F. (2023). *The 5 best ChatGPT tips we use constantly*. <https://www.digitaltrends.com/computing/best-chatgpt-tips/>
30. Agomuoh, F., Larsen, L. (2023). *ChatGPT: How to use the AI chatbot that's changing everything*. <https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot/>
31. Basics of Prompt Engineering (2022). <https://beta.dreamstudio.ai/prompt-guide>
32. Dilmegani, C. (2023). *How to Build a Chatbot: Components & Architecture in 2023*. <https://research.aimultiple.com/chatbot-architecture/>
33. Ekonomikos ir inovacijų ministerija. (2018). *Lietuvos dirbtinio intelekto strategija*. Ateities vizija. [https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT\(1\).pdf](https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT(1).pdf)
34. Gates, B. (2023). *Here's what the age of AI means for the world, according to Bill Gates*. <https://www.weforum.org/agenda/2023/03/heres-what-the-age-of-ai-means-for-the-world-according-to-bill-gates/>
35. McCartney, A. (2023). Your 7 Biggest Questions About ChatGPT, Answered. *Gartner*. <https://www.gartner.com/en/articles/your-7-biggest-chatgpt-questions-answered>
36. McKinsey (2023). What is generative AI? *McKinsey & Company*. <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>

37. Menor, D. (2022). How to Create AI-Generated Art with Midjourney. *HashDork*. <https://hashdork.com/midjourney/>
38. *Midjourney Documentation. Quick Start Guide*. (2023). <https://docs.midjourney.com/docs/quick-start>
39. Newblom, M. (2023). *How SMBs Can Use Midjourney to Power Their Business*. <https://www.fiverr.com/resources/guides/graphic-design/midjourney-use-cases>.
40. NLP Cloud (2022). *Effectively Using Text to Image with Stable Diffusion*. The DALL-E 2 / MidJourney Alternative. <https://nlpccloud.com/effectively-using-text-to-image-with-stable-diffusion-dalle-2-alternative.html>
41. Novohradsky, M. (2022). *DreamStudio AI (Stable Diffusion) – Tutorial for Beginners*. <https://nechempire.com/dreamstudio-ai-stable-diffusion-tutorial/>
42. Ortiz, S. (2023). *The best AI art generators: DALL-E 2 and other fun alternatives to try*. ZDNET. <https://www.zdnet.com/article/best-ai-art-generator/>
43. Paiva, R. (2023). How Bing Search Works with AI Model Prometheus. *Robot Artificial*. <https://www.robotartificial.com/how-bing-search-works-with-ai-model-prometheus/>
44. Parsons, G. (2022). *Everything you wanted to know about MidJourney*. <https://dallery.gallery/midjourney-guide-ai-art-explained/>
45. Parsons, G. (2022). *Ultimate guide to DALL-E 2: how to use it & how to get access*. <https://dallery.gallery/dall-e-ai-guide-faq/>
46. Paulauskaitė-Tarasevičienė, A. (2022). *Dirbtinis intelektas – neišvengiama ateitis*. Kauno technologijos universitetas. <https://ktu.edu/news/a-taraseviciene-dirbtinis-intelektas-neisvengiama-ateitis/>
47. Poloboc, A. (2023). *Microsoft shares insight on how Prometheus powers the new Bing*. <https://windowsreport.com/windows-bing-prometheus/>
48. Sagio Development LLC. (2023). Stable Diffusion Art. Tutorials. <https://stable-diffusion-art.com/tutorials/>