

Architektura 4.0: proces projektowania wspierany przez sztuczną inteligencję.

Potencjał wykorzystania skryptu generatywnego MidJourney w procesie tworzenia koncepcji architektonicznej



mgr inż. arch.
ANNA JARUGA-ROZDOLSKA
Interdyscyplinarna Szkoła Doktorska
Politechniki Łódzkiej
Dyscyplina Architektura i Urbanistyka
ORCID: 0000-0003-4196-6437

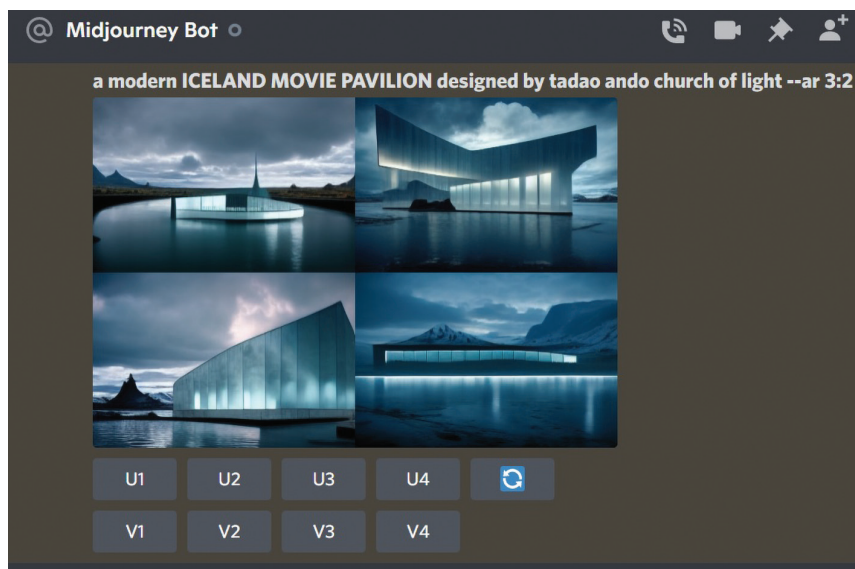
W artykule został omówiony sposób działania nowego narzędzia MidJourney – skryptu generatywnego opartego na sztucznej inteligencji.

Architektura od zawsze jest odzwierciedleniem rozwoju kultury i technologii, namacalnym manifestem określonego etosu, trójwymiarowym zapisem obowiązujących ówczesnie idei [1]. Współczesny świat ulega transformacji – jej katalizatorami są wykładniczy wzrost wydajności obliczeniowej komputerów, rosnąca pojemność pamięci wirtualnej, internet rzeczy (IoT) oraz rozwój sztucznej inteligencji (AI). Zmiany te mają bezpośredni wpływ na proces projektowania i wznoszenia budynków. Najsilniejszy dotychczas współczesny nurt – architektura parametryczna, powstała bezpośrednio w wyniku udostępnienia projektantom oprogramowania służącego modelowaniu, prototypowaniu

i optymalizacji konstrukcji, może niedługo ustąpić nowemu, będącego wynikiem kolejnej automatyzacji – wynikającej z wdrożenia w proces twórczy narzędzi opartych na zaawansowanych skryptach generatywnych, wspieranych przez technologię sztucznej inteligencji. Technologia ta znacząco zmienia charakter procesów twórczych, a komputery już teraz odgrywają kluczową rolę w niezliczonych dziedzinach nauki i sztuki, również w architekturze [2].

W artykule został zaprezentowany przebieg pracy z nowym narzędziem – skryptem generatywnym MidJourney – w kontekście badań nad kreatywnym procesem tworzenia formy architektonicznej. Omawiane narzędzie zdol-

ne jest do tworzenia atrakcyjnych wizualnie cyfrowych obrazów, bazując jedynie na wprowadzonym przez użytkownika ciągu słów kluczowych. Odpowiednio dobrane, mogą wymusić wygenerowanie szkicu koncepcji architektonicznej w określonym stylu, osadzonej w danym kontekście, z użyciem narzuconych materiałów. Renderowanie obrazów na podstawie fraz trwa zaledwie kilkadziesiąt sekund, a widoczny na nich budynek może stanowić gotową koncepcję. Rezultat prac widoczny na wizualizacjach, opublikowanych wraz z wprowadzonym ciągiem słów kluczowych umożliwi lepsze zrozumienie zasady działania. Analiza wygenerowanych cyfrowych obrazów stanowiła podstawę rozważań dotyczących perspektyw implementacji narzędzia w praktyce architektonicznej.



Rys. 1. Zrzut ekranu ilustrujący interfejs skryptu MidJourney. Widoczne cztery wstępne koncepcje mogą zostać decyzją użytkownika rozwinięte (U1, U2...), mogą stanowić podstawę tworzenia wariacji (V1, V2...); jeśli żaden rezultat nie odpowiada oczekiwaniom, można wymusić ponowne wygenerowanie obrazu w oparciu o ten sam zestaw słów kluczowych (symbol dwóch strzałek)

MidJourney

Skrypt MidJourney został udostępniony na platformie społecznościowej Discord [3]. Po fazie testów z udziałem ochotników zakończonej w marcu 2022 r. został on udostępniony do komercyjnego użycia w czerwcu bieżącego roku [4]. Po zalogowaniu i wybraniu jednej z opcji subskrypcji użytkownik zyskuje dostęp do narzędzia będącego botem zdolnym reagować na wprowadzone komendy i frazy – zdolność ich przetwarzania opiera się o technologię wykorzystującą AI. Po wpisaniu komendy /imagine rozpoczynającej proces tworzenia należy wprowadzić ciąg słów kluczowych stanowiących podstawę wygenerowania przez skrypt czterech cyfrowych obrazów o niskiej rozdzielczości, stanowiących propozycję ich wizualnej interpretacji. Użytkownik ocenia ich trafność, podejmując następnie decyzję o ich rozwinięciu – wygenerowaniu obrazu wyższej rozdzielczości



Rys. 2. Zestaw obrazów wygenerowanych w oparciu o słowa kluczowe: Iceland movie pavilion –ar 3:2 [6]



Rys. 3. Wizualizacja wygenerowana w oparciu o słowa kluczowe: concrete modern Iceland movie pavilion designed by Mies Van Der Rohe Barcelona pavilion –ar 3:2 [6]

i z większą ilością szczegółów lub wymuszeniu tworzenia wariacji w oparciu o pierwotnie wygenerowany obraz; jeżeli żadna z propozycji nie jest odpowiednia, użytkownik może podjąć decyzję o ponownym wygenerowaniu kolejnego zestawu propozycji w oparciu o ten sam zestaw słów kluczowych (rys. 1.). Co istotne, wprowadzenie fraz w niezmienionej formie i kolejności za każdym razem da inny rezultat – program losowo dobiera dostępne w internecie grafiki pasujące do wprowadzonych podpowiedzi, dodatkowo zmieniając ich udział procentowy w wygenerowanym obrazie – nie jest więc możliwe ponowne uzyskanie tego samego rezultatu. W sytuacji gdy efekt prac ponownie nie spełnia oczekiwań, należy rozważyć uzupełnienie lub zmianę ciągu słów kluczowych. Pełne zrozumienie mechaniki działania skryptu możliwe jest dzięki ujawnieniu fraz będących podstawą procesu tworzenia.

/Imagine

Założeniem projektowym było stworzenie koncepcji pawilonu filmowego na Islandii, będącego przedmiotem międzynarodowego konkursu architektonicznego ogłoszonego na platformie Buildner [5]. Po wpisaniu komendy /imagine podano ciąg słów kluczowych będących oficjalnym tematem konkursu: Iceland movie pavilion. Skrypt domyślnie generuje obrazy o proporcji 1:1 – można jednak zmienić ich rozdzielczość, dopisując po słowach kluczowych dodatkową komendę – w omawianym przykładzie –ar 3:2, gdzie –ar to współczynnik proporcji (ang. *aspect ratio*), a 3:2 określa stosunek długości do szerokości. Żadna z pierwszych wygenerowanych propozycji nie spełniła oczekiwań; do słów kluczowych dopisano więc frazy concrete – sugerując materiał budulcowy oraz modern – przymiotnik będący sugestią odnośnie do oczekiwanego stylu architekto-

nicznego. Chcąc uzyskać stylizację rezultatu na dzieło wskazanego twórcy, do zestawu słów kluczowych dodano frazę designed by, po której dopisywane zostało nazwisko wybranego architekta. Uzupełnienie ciągu słów kluczowych o nazwę konkretnego dzieła spowodowało wymuszenie stylizacji w oparciu o cechy wybranego obiektu. Uzyskany w ten sposób rezultat, będący jedną z czterech wstępnych propozycji, rozwinęto, a następnie wymuszono jego wariację.

Wpisywane frazy należy rozumieć tym samym jako podpowiedzi, które – właściwie wprowadzone – pomogą narzędziu wygenerować obraz zgodny z oczekiwanym rezultatem. Możliwości są w zasadzie nieskończone; dzięki słowom można określić zarówno parametry budynku, jak i otoczenia – w tym porę dnia, warunki atmosferyczne oraz lokalizację. Ciekawe rezultaty można uzyskać, wpisując pozornie niepasujące do siebie, z punktu widzenia logiki, ciągi słów. Dzięki odpowiednio dobranym wyrazom można także wymusić generowanie abstrakcyjnych obrazów, uzyskując zaskakujące rezultaty. Wszystko, co użytkownik jest w stanie określić słowami, skrypt oparty na sztucznej inteligencji jest sobie w stanie „wyobrazić”.

Perspektywa wdrożenia

Wizualizacje obiektów architektonicznych wygenerowane za pomocą skryptu, choć stanowią ich atrakcyjne wizualnie przedstawienie, nie są gotową koncepcją – mogą być traktowane jedynie jako wstępny zarys budynku, wymagający ewaluacji i rozwinięcia, by móc stać się architekturą. Korekcie należy poddać każdy jego aspekt, również formę. Cyfrowe ilustracje tworzone przez MidJourney są bowiem kolażem złożonym z wielu obrazów i zdjęć, mających często bardzo różne punkty zbiegu lub ogniskowe; przy próbie przełożenia koncepcji budynku na trójwymiarowy model może okazać się, że przeniesienie jego formy w kształcie widocznym na wizualizacji nie jest możliwe [2]. Fakt ten powodować może szereg trudności również na etapie rysowania planu funkcjonalnego oraz rozwiązań konstrukcyjno-materiałowych. Zadaniem projektanta pozostaje więc ocena możliwości realizacji obiektu oraz wykonanie niezbędnych elementów projektu architektonicznego. Wygenerowany obraz stanowi jedynie wstępny zarys, źródło inspiracji.

Należy zauważyć, że automatyzacja procesów kreatywnych możliwa dzięki zastosowaniu narzędzi opartych na sztucznej inteligencji może nieść za sobą szereg negatywnych konsekwencji, które na obecnym etapie badań są niemożliwe do właściwego rozpoznania. W nakreśleniu możliwego scenariusza pomocy okazać się mogą badania dotyczące zmian w sposobie funkcjonowania ludzkiego mózgu na przestrzeni lat wynikające z wdrożenia internetu do powszechnego

użytku. W ciągu ostatnich dwóch dekad znaczna część opracowań wskazuje na istotny wpływ środowiska internetowego na zachowania i zdolności poznawcze człowieka [7]. Dla przykładu: w zakresie przetwarzania informacji wykazano zmianę w kierunku płytkiego trybu przetwarzania danych, charakteryzującą się szybkimi i nieliniowymi przesunięciami uwagi, zmniejszoną zdolnością koncentracji i zapamiętywania [8]. Łatwość wyszukiwania informacji online zmniejszyła w człowieku potrzebę głębokiego przetwarzania w celu utrwalenia informacji. Istnieją dowody na to, że doświadczenie związane z wyszukiwaniem informacji w internecie może powodować trwałe zmiany neurologiczne [9]. Nie jest więc wykluczone, że projektant decydujący się na wdrożenie narzędzia w proces twórczy utraci do pewnego stopnia zdolność samodzielnej kreacji, stając się uzależnionym od generatywnych skryptów opartych na działaniu sztucznej inteligencji – podobnie jak współczesny człowiek uzależniony jest od oprogramowania bazującego na dostępie do internetu. Dostępność i szybkość działania stają się więc – w tym kontekście – paradoksalnie wadą, a nie zaletą nowych technologii.

Podsumowanie

W artykule zaprezentowano schemat pracy z nowym narzędziem MidJourney, bazującym na sztucznej inteligencji – eksperyment przeprowadzono w oparciu o zadanie projektowe będące tematem konkursu architektonicznego. Rezultatem są wizualizacje, które mogłyby posłużyć zarówno jako źródło inspiracji, jak i stać się pierwszym cyfrowym szkicem budynku. Należy zaznaczyć, że wygenerowany obraz przedstawia jedynie wstępny zarys formy obiektu, widocznego z określonej perspektywy. Rozwinięcie go do etapu koncepcji architektonicznej wymaga zaangażowania projektanta, posiadającego odpowiednią wiedzę i umiejętności. Obecnie, z uwagi na brak wdrożenia narzędzia w praktyce, nie jest możliwe przeprowadzenie oceny takich działań – sytuacji, w której szkic stworzony za pomocą skryptu zostałby na dalszych etapach projektowania rozwinięty przez architekta. MidJourney nieustannie się rozwija – dzięki rosnącej popularności i coraz większej liczbie użytkowników [10], analizując ich wybory i gromadząc coraz większą ilość danych, uczy się generować coraz doskonalsze obrazy. Mając na uwadze obecnie możliwe do osiągnięcia efekty, perspektywa wdrożenia skryptów generatywnych w proces kreatywnego projektowania jest wysoce realna, a narzędzia takie jak MidJourney mogą w przyszłości stanowić część standardowej praktyki architektonicznej, redefiniując rolę architekta i być początkiem nowego, silnego nurtu stylistycznego – będącego manifestem przemian zachodzących we współczesnym



Rys. 4. Wariacje uzyskane na podstawie wizualizacji wygenerowane w oparciu o słowa kluczowe: concrete modern Iceland movie pavilion designed by Mies Van Der Rohe Barcelona pavilion –ar 3:2 [6]



Rys. 5. Wizualizacja wygenerowana w oparciu o słowa kluczowe: green parametric organic Iceland movie pavilion –ar 3:2 [6]



Rys. 6. Wizualizacja wygenerowana w oparciu o słowa kluczowe: green parametric organic Iceland movie pavilion –ar 3:2 [6]

świecie w niespotykanym dotychczas tempie i zakresie [11]. Pamiętając o tym, warto zwrócić uwagę na możliwe negatywne rezultaty takich zmian – zwłaszcza w zakresie zmniejszenia potencjału kreatywnego twórcy w dłuższej perspektywie czasowej.

Rozumiejąc zasadę działania narzędzia – zdolnego do tworzenia dzięki dostępowi do ogromnych zasobów cyfrowych obrazów, możliwe jest rozwinięcie własnych zdolności twórczych – dzięki intencjonalnemu zapamiętywaniu i późniejszemu przetwarzaniu form, faktur i barw, obserwowanych w otaczającym krajobrazie. Budowana w ten sposób własna biblioteka zasobów z pewnością wzbogaci wrażliwość i rozwiniętość kreatywność projektanta; obca sztucznej inteligencji zdolność odczuwania emocji towarzyszących przetwarzaniu obrazów w ludzkim mózgu warunkuje bezterminową wyższość człowieka nad narzędziem.

Artykuł jest pierwszym z cyklu publikacji będących rezultatem badań nad potencjałem wykorzystania skryptu w projektowaniu architektonicznym.

Bibliografia

- [1] R. Ivey, Architecture INTL: Celebrating the Past, Designing the Future, Visual Reference Publications, 2008.
- [2] K. Barandy, Interview with Kory Bieg on text-to-image generators & the future of AI in design, <https://www.designboom.com/design/interview-kory-bieg-text-to-image-generators-future-ai-design-08-17-2022> [dostęp: 27.08.2022].
- [3] S. Kaczmarczyk, Komunikacja za pomocą aplikacji Discord w opinii polskich użytkowników, 2021.

[4] A. Walker, MidJourney, <https://knowyourmeme.com/memes/sites/midjourney> [dostęp: 2.08.2022].

[5] Źródło: <https://architecturecompetitions.com/icelandmovie-pavilion/> [dostęp: 10.08.2022].

[6] Cyfrowy obraz wygenerowany za pomocą narzędzia MidJourney; opracowanie: Anna Jaruga-Rozdolska.

[7] K.K. Loh, K. Ryota, How has the Internet reshaped human cognition?. *The Neuroscientist* 22.5 (2016): s. 506–520.

[8] N. Carr, *The Shallows: What the Internet Is Doing to Our Brains*, 2011.

[9] G. Small, G. Vorgan, Meet your iBrain. *Sci Am Mind* 19:42–9, 2008.

[10] W oparciu o wykres GoogleTrends: <https://trends.google.pl/trends/explore?gprop=images&q=midjourney> [dostęp: 8.08.2022].

[11] K. Schwab, *The fourth industrial revolution*. Currency, 2017, s. 17–18.

DOI: 10.5604/01.3001.0015.9893

PRAWIDŁOWY SPOSÓB CYTOWANIA

Jaruga-Rozdolska Anna, 2022, *Architektura 4.0: proces projektowania wspierany przez sztuczną inteligencję. Potencjał wykorzystania skryptu generatywnego MidJourney w procesie tworzenia koncepcji architektonicznej*, „Builder” 10 (303). DOI: 10.5604/01.3001.0015.9893

Streszczenie: W artykule został omówiony sposób działania nowego narzędzia MidJourney – skryptu generatywnego opartego na sztucznej inteligencji. Celem badań jest prezentacja potencjału narzędzia w kontekście tworzenia wstępnego szkicu koncepcji architektonicznej. W oparciu o treść zadania projektowego postawionego przed uczestnikami konkursu, którego przedmiot stanowi projekt koncepcyjny pawilonu filmowego zlokalizowanego na Islandii, wykonano serię obrazów cyfrowych materializujących ideę w oparciu o wprowadzone słowa kluczowe. Wygenerowane za pomocą skryptu wizualizacje są

estetyczną interpretacją wprowadzonych fraz. Artykuł jest zapisem stanu spraw na początkowym etapie komercjalizacji narzędzia.

Słowa kluczowe: sztuczna inteligencja, AI, MidJourney, skrypt generatywny, projektowanie architektoniczne

Abstract: ARCHITECTURE 4.0: AI SUPPORTED DESIGN PROCESS. THE POTENTIAL OF USING MIDJOURNEY GENERATIVE SCRIPT IN THE PROCESS OF CREATING AN ARCHITECTURAL CONCEPT. THE ARTICLE EXPLORES THE PERFORMANCE OF THE MIDJOURNEY TOOL – AN AI-BASED GENERATIVE SCRIPT. The purpose of the research is to present the potential of the tool in the framework of creating a draft of an architectural concept. On the basis of the subject matter of the design task set for participants of the competition, the purpose of which is the conceptual design of a film pavilion located in Iceland, a series of digital images materializing the idea based on the entered keywords were made. The visualizations generated with the script are an aesthetic interpretation of the phrases entered; tools based on similar mechanics of operation may in the future be part of standard architectural practice, forming the basis for the definition of a strong new stylistic trend. The article is a record of the state of affairs in the early stages of the tool's commercialization.

Keywords: artificial intelligence, AI, MidJourney, generative script, architectural design

2022-2023
KONKURS
dla MŁODYCH
INŻYNIERÓW
EDYCJA 6
2022
2023
WYZWANIE MŁODEGO INŻYNIERA
BUILDER
FOR THE
YOUNG
ENGINEERS
REKLAMA