

Karol Olejniczak¹

Dominika Wojtowicz²

Projektowanie behawioralne polityk publicznych: perspektywy zastosowania generatywnej AI

Abstrakt

Celem artykułu było zbadanie, na ile generatywna sztuczna inteligencja (GenAI) może wspierać projektantów polityk publicznych w tworzeniu wstępnych koncepcji interwencji behawioralnych. Autorzy przeprowadzili eksperyment, w którym cztery pomysły rozwiązań redukujących marnotrawienie żywności zostały przygotowane z różnym udziałem GenAI, a następnie ocenione przez 15 praktyków polityk publicznych. Wyniki pokazują, że GenAI może skutecznie wspierać projektantów w tworzeniu spójnych logicznie i przekonujących propozycji rozwiązań publicznych. Warunkami produktywnej współpracy są: oparcie projektowania na spójnych ramach koncepcyjnych (Teoria Zmiany) i przemyślane zaprojektowanie sposobu wykorzystania różnych GenAI we współpracy z człowiekiem (metoda TRIP). Artykuł podkreśla też konieczność dalszych badań nad zastosowaniami GenAI w pełnym procesie projektowania obejmującym empiryczne badania grupy docelowej i konsultacje z interesariuszami. Zwraca także uwagę na wyzwania etyczne związane z autorstwem i odpowiedzialnością za treści generowane przez AI. Praca wnosi wkład

¹ Uniwersytet SWPS, e-mail: karol.olejniczak@swps.edu.pl, <https://orcid.org/0000-0002-7079-2440>

² Akademia Leona Koźmińskiego, e-mail: doma@kozminski.edu.pl, <https://orcid.org/0000-0002-8433-8935>

w rozwój interdyscyplinarnej refleksji nad zastosowaniem przełomowych technologii w procesach formułowania polityk publicznych.

Słowa kluczowe: behawioralne projektowanie, behavioural insights, polityki publiczne, generatywna AI, GenAI, marnotrawienie żywności

Kody klasyfikacji JEL: D91, O33, Q18, H83

Behavioural Design of Public Policies: Prospects of Generative AI Application

Abstract

The aim of this article is to examine the extent to which generative artificial intelligence (GenAI) can support public policy designers in developing preliminary concepts for behavioural interventions. The authors conducted an experiment in which four proposals to reduce food waste were developed with varying degrees of GenAI involvement and subsequently evaluated by 15 public policy practitioners. The results show that GenAI can effectively assist designers in generating logically coherent and persuasive public policy proposals. Productive collaboration requires two key conditions: grounding the design process in a consistent conceptual framework (Theory of Change) and carefully planning how to integrate different GenAI tools with human workflow (the TRIP method). The article also highlights the need for further research on GenAI applications throughout the full design cycle, including empirical research with target groups and stakeholder consultations. It points to ethical challenges related to authorship and accountability for AI-generated content. The paper contributes to the growing interdisciplinary reflection on the use of breakthrough technologies, in particular GenAI, in public policy formulation.

Keywords: behavioural design, behavioural insights, public policy, generative artificial intelligence (GenAI), food waste

JEL Classification Codes: D91, O33, Q18, H83

Wprowadzenie

Generatywna sztuczna inteligencja (GenAI) przynosi prawdziwą rewolucję, zmieniając wiele dyscyplin (Dwivedi et al., 2023; Mollick, 2024), w tym praktykę: zarządzania publicznego (Charalabidis et al., 2024), analiz polityk (Safaei i Longo, 2024), projektowania behawioralnego (Kaiser et al., 2024) i ewaluacji działań publicznych (Jacob, 2025).

Praktyka behawioralnego projektowania polityk publicznych rozwija się bardzo dynamicznie w ostatnim dziesięcioleciu (Banerjee i John, 2025; OECD, 2017; Mozafar et al., 2023; Sunstein i Reisch, 2023; World Bank, 2015), jednak wciąż napotyka codzienne ograniczenia zarówno na poziomie koncepcyjnym, jak i organizacyjnym. Należą do nich: (1) diagnoza koncentrująca się zwykle na indywidualnych zachowaniach jednostek i błędach poznawczych przez nie popełnianych, (2) skrzynka narzędziowa słabo zintegrowana z tradycyjnym spektrum instrumentów polityk publicznych, (3) czasochłonny proces analizy źródeł i budowania argumentacji wspierającej konkretne wybory w projektowaniu interwencji (Banerjee i Galizzi, 2024; Dupoux et al., 2025; Howlett, 2019a; Kuehnhanss, 2019).

Dlatego też zasadnym wydaje się postawienie pytania o to, czy GenAI może być przydatna w zintegrowaniu podejścia behawioralnego z codziennym procesem projektowania interwencji publicznych.

Celem artykułu jest zweryfikowanie zdolności GenAI do wspierania projektantów polityk publicznych w tworzeniu wstępnych koncepcji interwencji behawioralnych. Zastosowaną metodą jest test praktyczny – dwójka projektantów przygotowała, z różnym wsparciem narzędzi GenAI, cztery wersje pomysłów na rozwiązania publiczne. Jakość tych pomysłów, przedstawionych w formie krótkich notatek (ang. *policy briefs*), została następnie oceniona przez ekspertów polityk publicznych. W oparciu o wyniki tego testu artykuł weryfikuje ramy koncepcyjne i procesowe, które wykorzystano w pracy z GenAI, wskazuje dalsze pola eksploracji tematu oraz szersze implikacje dla teorii i praktyki projektowania behawioralnego płynące z tego eksperymentu.

Przegląd literatury

Obserwacje z literatury projektowania behawioralnego

Projektowanie polityk publicznych (ang. *policy design*) to proces adresowania wyzwań społeczno-gospodarczych poprzez sformułowanie celu polityki, zrozumienie problemu i dopasowanie instrumentów pozwalających osiągnąć pożądaną zmianę (Howlett, 2019b). Polega na skutecznym regulowaniu zachowań adresatów polityki – indywidualnych, organizacyjnych i zbiorowych. Dlatego projektanci polityk potrzebują wiedzy, jak analizować i modyfikować mechanizmy napędzające lub blokujące zachowania (Lasswell, 1951; Soman, 2017).

Dynamiczny rozwój ekonomii behawioralnej przyniósł cenną perspektywę dla projektowania rozwiązań publicznych, przesuując akcent z racjonalnych decyzji w wyidealizowanych okolicznościach ku rzeczywistym czynnikom determinującym

skuteczność interwencji (Kuehnhanss, 2019). Jednak praktyka wciąż odbiega od postulatów Kahnemana (2013) o wykorzystaniu nauk behawioralnych jako podstawy tworzenia polityk publicznych. Z literatury wyłaniają się trzy wyzwania integrowania podejścia behawioralnego z projektowaniem.

Wyzwanie 1: Uporządkowanie procedury projektowania i strategii zmian

W literaturze projektowania behawioralnego często dochodzi do mieszania modeli prowadzenia procesów projektowania (OECD, 2019; Wendel, 2013) z propozycjami wybranych strategii zmian behawioralnych prezentowanych w formie skrótów mnemoniczych, np. MINDSPACE (Dolan et al., 2010), Nudges (Thaler et al., 2013: 438), EAST (Service et al., 2014).

Literatura z zakresu polityk publicznych proponuje uporządkowanie sekwencji procesowej, integrując diagnozę z doborem instrumentów interwencji (Stephans, 2016; Michie et al., 2011). W diagnozie projektanci, koncentrując się na konkretnej grupie docelowej, starają się zidentyfikować przeszkody dla pozytywnego/pożądanego zachowania. Następnie do tych blokad dobierają strategie zmiany. Wreszcie pomysł całej interwencji jest testowany z grupą docelową (często w formie eksperymentów terenowych lub laboratoryjnych).

Wyzwanie 2: Rozszerzenie diagnozy i spektrum instrumentów

Diagnoza behawioralna jest często utożsamiana z ekonomią behawioralną, koncentrującą się na pułapkach decyzyjnych i odchyleniach od modelu racjonalnego działania w kontekście decyzji konsumenckich. Takie zawężenie sprowadza diagnozę do błędów poznawczych, a rozwiązania ogranicza do interwencji typu *nudge* lub *boost*.

Szerszym podejściem jest podejście behawioralne (ang. *behavioral insights*). Przyjmuje ono perspektywę integrującą ustalenia z innych nauk (np. neuronauka, socjologia) do analizy zachowań aktorów polityk publicznych (Ruggeri, 2018). Przedmiotem interwencji mogą być zachowania indywidualne, organizacyjne lub kolektywne (Dupoux et al., 2025). Na etapie diagnozy podejście to identyfikuje nie tylko kognitywne bariery, ale również zasoby, motywacje i kontekst zachowań (popularny model COM-B) (West i Michie, 2020). Zaś na etapie tworzenia rozwiązań oferuje tzw. *policy mixes*, łączące strategie behawioralne z szerszym zestawem instrumentów polityk (Howlett, 2019a; Olejniczak et al., 2020).

Wyzwanie 3: Wyjaśnianie mechanizmu zmiany

Projektowanie behawioralne często opiera się na modelu czarnej skrzynki – wskazuje instrument zmiany, pomijając określenie mechanizmu, który ją napędza. Problem ilustruje przykład omówiony przez Lunna (2014: 45). Ustawienia domyślne mogą działać na grupę docelową, ponieważ (a) są postrzegane jako zalecenie przez organ, który je określa, (b) sygnalizują wybory innych, (c) pomagają w przełamaniu zwleknięcia lub inercji wybierającego lub (d) działają jako punkt odniesienia dla oceny innych wyborów. Dopóki nie jest jasne, jaki mechanizm napędza zachowanie w konkretnym przypadku, tworzone polityki mogą być suboptymalne.

Praktyka projektowania publicznego radzi sobie z wyzwaniem czarnej skrzynki za pomocą Teorii Zmiany (Astbury i Leeuw, 2010; Mayne, 2017). To narracja *explicit* artykułująca założenia o powiązaniach przyczynowych w interwencji, wyjaśniająca, co ma się zmienić i jak interwencja ma prowadzić do zmiany zachowań w grupie docelowej.

Budując na obserwacjach z przeglądu literatury, w tym badaniu przyjęto szersze, zintegrowane podejście do projektowania interwencji publicznych – *behavioral insights*. Jako ramy koncepcyjne projektowania zastosowano zaadaptowaną Teorię Zmiany (patrz: tabela 1).

Tabela 1. Elementy Teorii Zmiany dla projektowania behawioralnego

Nazwa elementu	Szczegółowy opis
Wybór grupy docelowej	Osoby, organizacje lub instytucje, których zachowanie ma kluczowe znaczenie dla danego problemu i powinno zostać zmienione za pomocą interwencji publicznej.
Diagnoza barier dla pożądanych zachowań	Sytuacje i przeszkody utrudniające grupie docelowej pożądane zachowanie. Do ich diagnozy stosuje się model COM-B, który wskazuje potencjał, sposobność i motywację jako kluczowe determinanty zachowania.
Dobór rozwiązań	Konfiguracja instrumentów zmiany i objaśnienie mechanizmu, który miałby uruchomić i podtrzymać pożądaną zmianę. Przy rozwiązaniu łączone są strategie behawioralne z instrumentami tradycyjnymi.
Opis oczekiwanych efektów	Zakładane efekty końcowe interwencji – wskaźniki pozytywnej zmiany w danej grupie docelowej i jej otoczeniu.

Źródło: opracowanie własne.

Obserwacje z praktyki zastosowań GenAI

Praktyka zastosowań GenAI rozwija się szybko i jest wielowymiarowa. Na poziomie makro dyskusja rozciąga się od identyfikowania potencjału transformacyjnego oraz zastosowań GenAI w administracji (Carrasco et al., 2023; PwC, 2024)

i politykach publicznych (Richter, 2024), poprzez wskazywanie czynników umożliwiających wdrożenie GenAI i zarządzanie ryzykami (Carrasco et al., 2024; McKinsey & Company, 2023), przeglądy zastosowań AI w sektorze publicznym (Tveita i Hustad, 2025) lub na wybranych poziomach administracji (Fontana et al., 2025), po mapowanie podejść regulacyjnych (UE vs. US) (Weerts, 2025) oraz pierwsze próby formułowania zasad zastosowań GenAI w procesach decyzyjnych (Gasser i Mayer-Schönberger, 2024) i odpowiedzialnego zarządzania GenAI w organizacjach publicznych (Janssen, 2025).

Tabela 2. TRIP – metoda współpracy z GenAI

Pola decyzji (en/pl)	Zakres decyzji projektanta
T = <i>Tasks</i> /Typy zadań	Wybór konkretnej roli i zadań, które GenAI ma wykonać lub w których ma pomóc projektantowi polityki publicznej. Może to być np. obsadzenie GenAI w roli asystenta przeglądu źródeł, recenzenta pierwszych pomysłów projektowych, współautora notatki opisującej rozwiązanie. Przy każdej z ról można GenAI delegować pojedyncze, wąskie zadania, określić jedno szerokie działanie lub zautomatyzować całą sekwencję procesu (ang. <i>workflow</i>), który będzie wspierany różnymi asystentami GenAI.
R = <i>Resources</i> /Rodzaje źródeł	Dobór danych i zasobów informacyjnych, z których GenAI będzie korzystał podczas wykonywania swojej pracy. Domyślnie LLM udzielają odpowiedzi w oparciu o materiał treningowy. Ten zasób jest tajemnicą firm tworzących model. Przy tej opcji model często halucynuje. Wybrane LLM mogą realizować zadania w oparciu o informacje z Internetu. GenAI widzi jednak głównie publikacje otwartego dostępu. Wreszcie GenAI można połączyć z bazami danych, repozytoriami albo zasilić dokumentami wybranymi przez użytkownika (tzw. RAG – McKinsey & Company, 2024).
I = <i>Infrastructure</i> /Infrastruktura	Wybór platformy, rozwiązań technologicznych, z których projektant będzie korzystał (rodzaje modeli LLM, platformy RAG, inne oprogramowanie). Na rynku dostępnych jest wiele modeli LLM (Chat GPT, Claude, Gemini, LLaMA, rodzime – PLLuM, Bielik). Istnieje też szereg platform, które używając konkretnego modelu LLM dostosowały go do zadań np. przeszukiwania Internetu (Perplexity), streszczania tekstów (Petal AI), kwerendy literatury otwartego dostępu (SCISPACE, Scite), analizy jakościowej (Maxqda Tailwind). Każde rozwiązanie ma swój profil, mocne strony i ograniczenia (np. tokenizacja tekstu, limity długości źródeł, polityka poufności).
P = <i>Prompts</i> /Prompty	Sposób interakcji z GenAI – forma poleceń, podpowiedzi, pytania uzupełniające, a także sposoby weryfikacji treści generowanych przez AI. Interakcja z GenAI to rozmowa prowadzona naturalnym językiem. Użytkownik nie musi znać programowania – liczy się jakość poleceń tzw. promptów wydawanych modelowi. Główne zasady to: (i) wskazanie roli, (ii) określenie zadania i formatu odpowiedzi, (iii) wyjaśnienie kontekstu, (iv) umożliwienie LLM dopytania nas o szczegóły. Każdy wynik pracy GenAI wymaga krytycznej oceny. GenAI może popełniać błędy i wymyślać źródła. Użytkownik może wskazać błędy, prosić o korektę, dopytywać, podjąć nowy wątek.
Końcowa weryfikacja „podróżny”	Czy to poprawia moją pracę? Czy to jest bezpieczne? Czy to jest etyczne?

Źródło: opracowanie własne.

Na poziomie mezo widać szereg zastosowań funkcjonalnych w konkretnych zadaniach sektora publicznego, np. podnoszeniu jakości usług (Yun et al., 2024), zlecaniu i realizacji badań ewaluacyjnych (Olejniczak et al., 2025), tworzeniu *policy briefs* (Safaei i Longo, 2024), testowaniu pomysłów interwencji publicznych z syntetycznymi interesariuszami (Park et al., 2025), tworzeniu przez GenAI treści konkretnych interwencji behawioralnych (Harrison et al., 2024), czy nawet przewidywaniu za pomocą LLM kierunku i wielkości efektów interwencji behawioralnych (Kaiser et al., 2024).

Wreszcie poziom mikro to wiele źródeł poruszających *stricte* techniczne kwestie, takie jak wybór optymalnego modelu do konkretnych rodzajów zadań (Open AI, 2025), „dostrajanie” modelu przez użytkowników (tzw. *fine-tuning*) (Bergman, 2024), techniki rozmowy z modelami (tzw. *promptowanie*) (Phoenix i Taylor, 2024), czy ocena wiarygodności i rzetelności tekstów generowanych przez GenAI (USF Libraries, 2025).

To bogactwo i dynamika przyrostu literatury i praktyki GenAI sprawiają, że projektanci polityki publicznej stają przed dużym wyzwaniem. Chcąc zastosować GenAI w swojej pracy, muszą uwzględnić różne kwestie i dokonać wielu wyborów. Ten nadmiar opcji jest też wyzwaniem z perspektywy badawczej, bo utrudnia przejrzystość opisu procedury.

Aby poprawić ten proces, na potrzeby tego projektu stworzono architekturę wyborów dla projektantów. Heurystyka TRIP upraszcza wybory – grupuje je w cztery kategorie, układa w sekwencję i nadaje łatwy do zapamiętania akronim (TRIP – podróż decyzyjna z GenAI). TRIP pomaga w sposób uporządkowany eksperymentować, tworząc konfigurację opcji dostosowaną do kontekstu i konkretnych potrzeb, a następnie zweryfikować ją trzema pytaniami. Metodę tę oparto na syntezie literatury tematu i wnioskach z pierwszych eksperymentów praktycznych polskiej administracji (Olejniczak et al., 2025: 326–328).

Badanie empiryczne

Celem eksperymentu było sprawdzenie, na ile GenAI może konstruktywnie wesprzeć projektantów w przygotowaniu pomysłu na interwencję publiczną w ramach behawioralnego projektowania polityk. Skupiono się na użyciu modeli językowych (LLM – *Large Language Models*). Choć GenAI umożliwia generowanie różnych form treści (tekstów, grafik, dźwięku czy wideo), to administracja publiczna operuje głównie na tekstach – i to tę formę uznano za priorytetową.

W eksperymencie dwójka projektantów polityk publicznych, wspieranych w różnym stopniu przez GenAI, przygotowała cztery notatki (A, B, C i D) przedstawiające

różne pomysły na interwencję publiczną redukującą marnotrawienie żywności w mieście. Notatki te następnie zostały ocenione w dniach 23 marca – 4 kwietnia 2025 roku przez 15 praktyków polityk publicznych.

Zakres i procedura badania

W badaniu skoncentrowano się na symulacji fazy koncepcyjnej projektowania, której produktem jest notatka opisująca pomysł interwencji (ang. *policy brief*). Należy podkreślić, że dokonano uproszczenia – tworzenie pomysłów opierało się wyłącznie na badaniach gabinetowych – literaturze tematu i wnioskach z dotychczasowych praktyk dostępnych w Internecie. Rzeczywisty proces projektowania behawioralnego obejmuje również badania empiryczne grupy docelowej, jej kontekstu lokalnego oraz konsultacje z interesariuszami.

Uproszczenie wynika z celów badawczych artykułu, koncentrujących się na weryfikacji potencjału GenAI w zakresie syntezy wiedzy i generowania wstępnych koncepcji interwencji. Autorzy przyjęli założenie, że przed przystąpieniem do kosztownych badań terenowych i sesji współprojektowania warto sprawdzić, czy GenAI w ogóle może wspierać projektantów, gdy decyzje są podejmowane, mimo niepełnych, ogólnych danych o grupie docelowej. Takie sytuacje są częste w praktyce administracji publicznej. Badanie należy więc traktować jako test wstępnej fazy – przeglądu wiedzy i generowania hipotez projektowych, które w rzeczywistym procesie wymagałyby dalszej weryfikacji empirycznej.

Jako temat wybrano problem marnotrawienia żywności z trzech powodów. Po pierwsze, jest to istotny problem w krajach rozwiniętych i rozwijających się, generujący negatywne konsekwencje dla zasobów naturalnych, klimatu i bezpieczeństwa żywnościowego (Editorial, 2024). Po drugie, problem jest wieloaspektowy, obejmuje interakcje człowiek-technologia-środowisko oraz zachowania indywidualne i zbiorowe, co stawia projektantów przed różnymi wyborami na etapie diagnozy i rozwiązań (Manning et al., 2023). Po trzecie, mimo dostępnych propozycji rozwiązań (Reynolds, 2023; ReFED, 2024), wymagają one dostosowania do kontekstu, co pozwala sprawdzić zdolność GenAI do uwzględniania specyfiki lokalnej.

Aby maksymalnie urealnić eksperyment, określono wyraźny kontekst sytuacji decyzyjnej, odbiorców notatki, warunki brzegowe i formę notatki (patrz: załącznik – Treść zadania). Zadanie było realizowane z perspektywy projektanta polityki publicznej wspierającego Urząd Miasta Stołecznego Warszawy podczas wstępnego tworzenia rozwiązań publicznych. Struktura notatki została oparta na Teorii Zmiany projektowania behawioralnego (patrz: tabela 1).

Tabela 3. Porównanie zakresu zastosowania GenAI w przygotowaniu czterech notatek

Notatka	Zastosowanie GenAI	Szczegółowa procedura pracy*
Notatka A – projektant opierający się na źródłach	Brak użycia GenAI. Notatka przygotowana przez ludzi-projektantów w oparciu o materiał z faktycznego projektu miejskiego.	K1: Zebranie materiałów nt. faktycznej interwencji (w jęz. angielskim). K2: Przetłumaczenie materiałów i dostosowanie ich do szablonu notatki.
Notatka B – GenAI korzystająca z Internetu	Jeden całościowy prompt dla GenAI i krótka wymiana promptów korygujących wynik pracy genAI. Użyte GenAI: Claude 3.7 Sonnet	K1: Projektant dał jeden prompt szczegółowo opisujący zadanie. K2: GenAI–Claude 3.7 Sonnet wykonała zadanie w oparciu o źródła internetowe. K3: Projektant dał dodatkowy prompt, żeby GenAI podała źródła do przytoczonych w tekście danych i statystyk. K4: GenAI skorygowała zadanie. K5: Projektant dał prompt, żeby źródła były wskazane w tekście i w bibliografii według stylu APA7. K6: GenAI przedstawiła ostateczną wersję notatki.
Notatka C – zaawansowana GenAI z biblioteką źródeł i Internetem	Jeden całościowy prompt dla GenAI i odpowiedź na pytania precyzujące zadane przez GenAI. Użyte GenAI: Chat GPT o1 pracujący w trybie Deep Research	K1: Projektant zasilil dziesięcioma źródłami GenAI–Chat GPT o1 pracującą w trybie Deep Research. K2: Projektant dał jeden prompt szczegółowo opisujący zadanie i wskazujący, by GenAI wykonała zadanie w oparciu o podane dziesięć źródeł oraz źródła internetowe. K3: GenAI zadawała projektantowi serię pytań nt. preferencji co do grupy docelowej i struktury potencjalnej interwencji. K4: Projektant dał odpowiedź GenAI, przypominając raz jeszcze o limicie długości notatki (limit słów). K5: GenAI wykonała zadanie.
Notatka D – projektant z zespołem agentów GenAI z biblioteką źródeł i internetem	Seria mikro-zadań nt. (i) problemu, (ii) diagnozy, (iii) rozwiązania i efektów. Przy każdym zadaniu człowiek decyduje w oparciu o materiał stworzony przez GenAI. Użyte GenAI: Claude 3.7. Sonnet Chat GPT o1 Perplexity Spaces (z LLM Claude 3.7) NotebookLM (z LLM Gemini 2.0)	Grupa docelowa K1: Projektant zapytał o propozycje grup docelowych (trzy proste i trzy zaawansowane pomysły). Pytanie skierował do dwóch GenAI: Claude 3.7. i Chat GPT o1. K2: Obydwie GenAI przedstawiły po 6 propozycji wraz z uzasadnieniem. K3: Projektant ocenił wszystkie sześć propozycji i wybrał jedną, nieco ją zawężając (wyłącznie stolówki w placówkach oświatowych, bez innych stolówek instytucjonalnych miasta). Poprosił GenAI o doprecyzowanie źródeł dla tej jednej notatki. K4: GenAI przedstawiła propozycję ze źródłami we wskazanym formacie. Diagnoza K7: Projektant przedstawił GenAI Perplexity grupę docelową i wyjaśnił zadanie – zrobienie diagnozy tej grupy z zastosowaniem modelu COM-B. K8: GenAI Perplexity wykonała zadanie – diagnozę grupy docelowej w oparciu o źródła internetowe. Rozwiązanie i efekty K9: Projektant zasilil 41 źródłami nt. redukcji marnotrawienia żywności dwa GenAI: Perplexity Spaces i NotebookLM. K10: Projektant w prompie dla obydwu modeli (i) przedstawił diagnozę grupy docelowej, (ii) dał zadanie by GenAI zaproponowała rozwiązania i ich efekty oraz (iii) wskazał, że GenAI może czerpać inspirację z udostępnionych jej źródeł. K11: Obydwie GenAI wykonały zadanie. K12: Projektant porównał propozycje modeli i wybrał propozycję Perplexity (ciekawsze pomysły, lepszy opis mechanizmu zmiany i lepsze odesłania do literatury). K13: Projektant poprosił GenAI Perplexity o doprecyzowanie źródeł w formacie APA7. K14: GenAI przedstawiła ostateczną wersję opisu rozwiązania i efektów. Synteza materiału K15: Projektant przygotował ostateczną wersję notatki integrując opisy: (i) grupy docelowej, (ii) diagnozy, (iii) rozwiązań i efektów.

* „K” – jest skrótem od słowa „krok”. Dokumentację szczegółowych kroków w procedurze, w tym użytych promptów i źródeł dla każdej z opcji, udostępniono w publicznym repozytorium pod linkiem: <https://doi.org/10.58142/swps-ko-dw-01/09/2025>.
Źródło: opracowanie własne.

Tabela 4. Porównanie merytorycznej zawartości czterech notatek**

Notatka	PROBLEM – wskazuje grupę docelową i zachowanie lub sytuację, na których miasto powinno skupić interwencję, uzasadniając to listą przesłanek.	DIAGNOZA – identyfikuje typowe sytuacje prowadzące do marnowania żywności w grupie docelowej i przeszkody dla redukcji marnotrawstwa.	PROPONOWANE ROZWIĄZANIE – proponuje interwencję, która ma wywołać pozytywną zmianę zachowań w grupie docelowej i wyjaśnia, jakie mechanizmy sprawią, że interwencja zadziała.	OCZEKIWANE EFEKTY – opisuje, jak będzie wyglądał efekt końcowy interwencji, w tym wskaźniki pozytywnej zmiany.
Notatka A	Problem: marnowanie żywności w małych restauracjach w Warszawie. Grupa docelowa: właściciele i menedżerowie lokali (5–15 stolików). Uzasadnienie wyboru: duża liczba małych restauracji, potrzeba wsparcia restauratorów po pandemii; dobrze udokumentowane przyczyny marnowania jako solidna podstawa do zaproponowania rozwiązania problemu; możliwość synergii z działaniami Biura Rozwoju Gospodarczego.	Sposobność: obowiązujące przepisy są niejasne i rozproszone w różnych aktach prawnych. Potencjał: właściciele nie wiedzą, co jest legalne i jak zgodnie z prawem dzielić się nadwyżkami. Motywacja: właściciele są zdezorientowani i obawiają się nieświadomego złamania przepisów.	– Kompendium wiedzy opracowane z GIS, zawierające instrukcje, infografiki, wzory dokumentów, listę odbiorców żywności i praktyczne wskazówki dotyczące przekazywania nadwyżek. – Kampania informacyjna z interaktywną infografiką, skierowana do co najmniej 150 małych restauracji, zachęcająca do korzystania z kompendium i gotowa do dalszego rozszerzania.	Spadek obaw restauratorów; wzrost świadomości procedur przekazywania żywności; większa gotowość do dzielenia się nadwyżkami; poprawa przejrzystości zarządzania żywnością w mieście.
Notatka B	Problem: marnowanie żywności w gospodarstwach domowych w Warszawie. Grupa docelowa: rodziny z dziećmi i młodzi dorośli (20–35 lat). Uzasadnienie wyboru: gospodarstwa domowe odpowiadają za znaczące straty żywności; deficyty w planowaniu zakupów i przechowywaniu żywności, największy potencjał zmiany zachowań przy odpowiednim wsparciu.	Potencjał: brak wiedzy i umiejętności planowania posiłków, przechowywania żywności i wykorzystywania resztek. Sposobność: brak informacji o sposobach ograniczania strat i niedostatek infrastruktury do dzielenia się nadwyżkami. Motywacja: niska świadomość skutków marnowania i brak pozytywnych wzorców społecznych.	Inicjatywa „Warszawa Nie Marnuje”: – aplikacja mobilna z planerem zakupów, przypomnieniami, mapą lodówek, kalkulatorem oszczędności i systemem nagród; – sieć 100 lodówek <i>food sharingowych</i> zintegrowanych z aplikacją, objęta nadzorem sanitarnym; – program edukacyjny „Sprytna Kuchnia”: warsztaty kulinarne, kampania informacyjna, e-booki z przepisami dostępne w aplikacji.	Redukcja marnowania o 25% w 2 lata, aktywni użytkownicy aplikacji, analiza odpadów, oszczędności CO ₂ i finansowe.

Notatka C	Marnowanie żywności w gospodarstwach domowych. Grupa docelowa: ogół mieszkańców Warszawy. Uzasadnienie wyboru: mieszkańcy są główną grupą generującą straty żywności, wyrzucają żywność łatwo psującą się, działania w tej grupie są konieczne do realizacji celów SDG, odpowiednia interwencja ma potencjał ograniczenia strat o 30–50%.	Potencjał: brak umiejętności planowania posiłków i przechowywania żywności Sposobność: promocje skłaniające do nadmiernych zakupów, brak łatwego dostępu do jadłodzieln. Motywacja: utrwalone nawyki, niska motywacja do zmiany, brak odczuwalnych konsekwencji marnowania.	Inicjatywa „Warszawa Nie Marnuje”: – kampania w mediach miejskich, poradniki i przepisy, programy szkolne, warsztaty kulinarne; – monitoring strat (np. dzienniczki), publiczne deklaracje, komunikaty w sklepach; – rozbudowa jadłodzieln i uruchomienie miejskiej platformy do dzielenia się żywnością.	Redukcja marnowania o 20%, zmiana deklaracji i postaw, liczba uczestników i przekazanej żywności jako wskaźniki.
Notatka D	Marnowanie żywności w stołówkach szkolnych i przedszkolnych. Grupa docelowa: dzieci, kadra placówek i personel kuchenny. Uzasadnienie wyboru: duża skala strat w stołówkach, możliwość systemowej zmiany w ponad 340 placówkach, wysoki potencjał edukacyjny (efekt na całe rodziny), potwierdzone rezultaty podobnych programów w innych miastach.	Sposobność: brak wiedzy uczniów o skutkach marnowania, niskie kompetencje pracowników w planowaniu posiłków. Potencjał: zbyt krótkie przerwy obiadowe, brak monitorowania strat, ograniczenia w modyfikacji menu przy cateringu. Motywacja: niska atrakcyjność posiłków, brak zaangażowania społeczności szkolnej.	Program „Szkolne Zero Marnowania”: – wydłużenie przerw obiadowych do 30 minut; – monitorowanie odpadów i informacja zwrotna dla społeczności szkolnej; – warsztaty kulinarne z udziałem dzieci, rodziców i personelu; – wybór porcji i możliwość dokładki zamiast dużych porcji i wyrzucania resztek; – pojemniki „Zjedz Najpierw” na produkty z krótkim terminem.	Redukcja marnowania w stołówkach o 30–40%, większe zaangażowanie społeczności szkolnej i przenoszenie nawyków do domów; wskaźniki: waga odpadów, wybór porcji, liczba dokładek, ankiety i obserwacje.

** Pełna treść czterech powstałych notatek jest dostępna pod linkiem:

<https://doi.org/10.58142/swps-ko-dw-01/09/2025>.

Źródło: opracowanie własne.

Autorom zależało na przetestowaniu narzędzi GenAI o niskim progu wejścia. Dlatego skorzystano wyłącznie z ogólnodostępnych, komercyjnych narzędzi GenAI; praca z nimi nie wymaga umiejętności programowania ani modyfikacji zaawansowanych ustawień.

W eksperymencie przetestowano cztery opcje podejścia do procesu projektowania, modyfikując kluczowe elementy metody TRIP (patrz: tabela 2), które mogą determinować produktywność współpracy z GenAI. Tabela 3 przedstawia podsumowanie procedury dla każdej z opcji. Tabela 4 podsumowuje treść notatek stworzonych przez analityków i GenAI.

Gotowe cztery notatki zostały przedstawione (bez podawania szczegółów ich powstania) do oceny ekspertkom i ekspertom – osobom z dużym doświadczeniem

praktycznym pracy w polskiej administracji publicznej. Ocena wykonana przez każdą z osób miała trzy aspekty:

- 1) Ocena punktowa stopnia, w jakim każda z notatek jest przekonująca. Ocena na skali Likerta 1–5, gdzie: 1 = jest zdecydowanie nieprzekonująca, 2 = jest raczej nieprzekonująca, 3 = nie jest ani przekonująca, ani nieprzekonująca, 4 = jest raczej przekonująca, 5 = jest zdecydowanie przekonująca.
- 2) Wskazanie jednej notatki, którą osoba rekomendowałaby decydentom m.st. Warszawy.
- 3) Uzasadnienie dla wyboru tej notatki.

Oceny przyznane poszczególnym notatkom w oparciu o skalę Likerta zostały zliczone i przedstawione w formie tabeli/wykresów. Wyniki jakościowe (otwarte komentarze i uzasadnienia wyboru notatek) przeanalizowano, kodując cały materiał w programie MAXQDA, a następnie tworząc podsumowania ocen ekspertów każdej z czterech notatek.

Wyniki badania

Wyniki analizy danych zebranych od 15 ekspertów obejmują trzy obszary: ocenę przekonującego charakteru notatek w ujęciu liczbowym, wskazania ekspertów dotyczące rekomendacji jednej notatki decydentom m.st. Warszawy oraz jakościowe uzasadnienia tych wyborów.

Ocena punktowa przekonującego charakteru notatek

Oceny przyznane przez ekspertów pozwoliły uchwycić postrzeganą siłę argumentacyjną poszczególnych notatek. Najwyższą średnią ocen uzyskała notatka D (4,1), co wskazuje na dominujące przekonanie ekspertów o jej wysokiej wartości merytorycznej i potencjale perswazyjnym. Notatka ta otrzymała również największą liczbę ocen maksymalnych – siedem wskazań na poziomie 5. Drugie miejsce pod względem średniej arytmetycznej zajęła notatka B (3,9), również wysoko oceniona, z wyraźną przewagą ocen 4 i 5. Notatka C osiągnęła średnią ocenę 3,5, z rozkładem wskazującym na umiarkowane uznanie wśród ekspertów. Najniższą średnią uzyskała notatka A (2,7), która jako jedyna nie otrzymała żadnej oceny maksymalnej (dominowały wskazania 2 i 3).

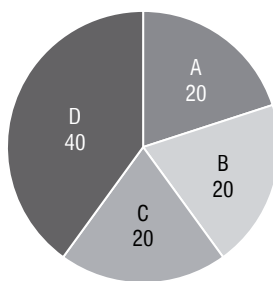
Tabela 5. Podsumowanie oceny notatek ($n = 15$)

	Suma przyznanych punktów (max = 75)	Średnia ocena notatki (max = 5)
Notatka A – projektant opierający się na źródłach	40	2,6
Notatka B – GenAI korzystająca z Internetu	59	3,9
Notatka C – zaawansowana GenAI z biblioteką źródeł i Internetem	52	3,4
Notatka D – projektant z zespołem agentów GenAI z biblioteką źródeł i Internetem	62	4,1

Źródło: opracowanie własne.

Wskazania dotyczące notatki rekomendowanej decydentom m.st. Warszawy

W odpowiedzi na pytanie o wybór notatki rekomendowanej prezydentowi Warszawy notatka D uzyskała najwięcej wskazań – sześć. Notatki A, B i C otrzymały po trzy głosy każda. Choć notatka D nie uzyskała zdecydowanej dominacji, jej przewaga liczebna wraz z wysoką oceną punktową czyni ją najbardziej pozytywnie ocenioną propozycją. Równomierny rozkład pozostałych wskazań świadczy, że każda z pozostałych notatek przekonała część ekspertów, przy czym kryteria wyboru różniły się znacząco (patrz: analiza jakościowa uzasadnień). Wyniki zobrazowano graficznie jako procentowy udział wskazań ekspertów dla każdej notatki.

Rysunek 1. Wskazania notatki do rekomendacji prezydentowi Warszawy przez ekspertów (%)

Źródło: opracowanie własne.

Analiza jakościowa – uzasadnienia dla wyboru notatki rekomendowanej

Analiza jakościowa uzasadnień formułowanych przez ekspertów w odniesieniu do czterech ocenianych notatek pozwoliła na wyłonienie trzech spójnych obszarów branych pod uwagę podczas oceny. Obszarami tymi są: wartość merytoryczna, komunikacyjna i operacyjna poszczególnych propozycji. Eksperci w swoich uwagach odnosili się bowiem zarówno do trafności diagnozy problemu, adekwatności proponowanych rozwiązań, jak i sposobów ich zaprezentowania.

Notatka A, przygotowana przez badacza-projektanta na podstawie materiałów z projektu miejskiego, została uznana przez część ekspertów za klarowną i zwięzłą, o przejrzystej strukturze i języku. Doceniano logikę dokumentu, zgodność między problemem a rozwiązaniami oraz odniesienie do lokalnego kontekstu i grupy docelowej. Dla niektórych ekspertów zaletą była ograniczona innowacyjność proponowanych działań, które nie budziły obaw i mogły być łatwiej akceptowane. Wielu respondentów kwestionowało główną oś interwencji – założenie, że barierą dla przekazywania nadwyżek żywności są obawy prawne. Eksperci wskazywali, że istotniejsze są kwestie ekonomiczne oraz że istnieje konieczność oceny istniejących przepisów prawnych. Zarzuty dotyczyły braku odniesień do obowiązujących regulacji i pominięcia czynników systemowych. Niektórzy komentujący kwestionowali oryginalność dokumentu, sugerując automatyczne wygenerowanie treści przez AI (sic!), a nie przemyślaną analizę. Wskazywano także na brak bibliografii w języku polskim oraz ogólnikowe uzasadnienie skuteczności działań.

Notatka B, wygenerowana przez GenAI, budziła wśród ekspertów zarówno uznanie, jak i poważne zastrzeżenia. Była oceniana jako przejrzysta, logicznie skonstruowana i adresująca główny problem marnotrawienia żywności w gospodarstwach domowych. Eksperci docenili kompleksowe ujęcie interwencji, obejmującej elementy infrastrukturalne (lodówki społecznościowe), edukacyjne (warsztaty) i technologiczne (aplikacja mobilna). Pozytywnie oceniono zawężenie grupy docelowej do rodzin z dziećmi oraz młodych dorosłych, wykazujących potencjał do zmiany nawyków. Doceniono obecność mierników efektów, umożliwiających ocenę efektywności i kontrolę wydatków. Notatka bazowała na badaniach naukowych i odwoływała się do rozwiązań wzmacniających wizerunek Warszawy jako miasta innowacyjnego. Wskazywano jednak na ograniczenia: słaby związek między grupą docelową a narzędziami interwencji, brak działań w szkołach oraz potencjalnie wysokie koszty. Eksperci krytykowali brak odniesień do efektów podobnych programów w innych miastach i niewystarczające źródła, na podstawie których zaproponowano rozwiązania problemu.

Notatka C, stworzona przez GenAI z dostępem do biblioteki źródeł i Internetu, była postrzegana jako najbardziej analityczna i zakorzeniona w danych empirycznych. Eksperci pozytywnie oceniali spójność diagnozy, precyzyjne uzasadnienie wyboru grupy docelowej oraz realistyczne oczekiwania efektów interwencji. Podkreślano zgodność celów z wynikami badań oraz odwołania do zobowiązań międzynarodowych, nadając dokumentowi wymiar legitymizacyjny. Dla części ekspertów notatka była dowodem przemyślanej, możliwej do wdrożenia interwencji, która bazowała na wiedzy o wysokim poziomie wiarygodności. Jednocześnie część respondentów uznała jej styl za zbyt technokratyczny – bardziej odpowiadający planistom niż osobom podejmującym decyzje polityczne i pracującym z mieszkańcami. Krytykowano także małą liczbę danych ilościowych w formie graficznej oraz mniej przejrzysty układ w porównaniu z notatką D. Wśród krytycznych ocen pojawiły się opinie, że notatka – mimo bogactwa danych – pozostaje ogólna i mało konkretna w zakresie narzędzi operacyjnych.

Notatka D, opracowana przy współpracy badacza-projektanta z zespołem agentów GenAI posiadających dostęp do wyselekcjonowanych źródeł wiedzy, została przez większość ekspertów uznana za najbardziej przekonującą. Jej atutem była przejrzysta forma prezentacji, wysoki poziom dopracowania oraz silne osadzenie w sprawdzonych rozwiązaniach. Eksperci docenili konkretność propozycji, jasno zdefiniowane wskaźniki efektów oraz odniesienie do doświadczeń zagranicznych i krajowych. Zaproponowane działania dla placówek edukacyjnych są łatwe do wdrożenia, nie wymagają znaczących nakładów finansowych i mogą być realizowane przy wykorzystaniu zasobów samorządu. Wskazywano na potencjał efektów mnożnikowych, jak zmiana postaw czy wzmocnienie więzi społecznych. Zwracano uwagę na korzystny stosunek kosztów do rezultatów oraz możliwość szybkiego wdrożenia. Wśród zastrzeżeń pojawiały się wątpliwości dotyczące zgodności grupy docelowej z narzędziami (warsztaty z rodzicami a zmniejszenie strat w stołówkach) oraz obawy o obciążenie placówek dodatkowymi zadaniami. Mimo to, notatka D wyróżniała się jako najbardziej przystępna i dobrze zakorzeniona w wiedzy praktycznej.

Podsumowując, analiza jakościowa uzasadnień ekspertów pokazuje zróżnicowane oczekiwania względem dokumentów strategicznych – od przejrzystości i prostoty, przez głębię analityczną, po kompleksowość i wykonalność rozwiązań. Jednocześnie potwierdza, że różnice w ocenach nie wynikają wyłącznie z jakości dokumentów, lecz także z odmiennych priorytetów przyjmowanych przez ekspertów w ocenie pomysłów na interwencje publiczne – takich jak koszt, skala oddziaływania, innowacyjność pomysłu, potencjał wdrożeniowy, potencjał komunikacyjny, osadzenie w danych, spójność logiczna czy zgodność z logiką polityki miejskiej.

Dyskusja

Wniosek 1: Preferencje oceniających

Analiza uzasadnień wyborów ekspertów ujawniła istotne różnice w kryteriach przyjmowanych przy ocenie poszczególnych notatek. W ocenach dominowały dwa podejścia – część ekspertów premiowała notatki syntetyczne, komunikatywne i operacyjne, inne osoby natomiast ceniły pogłębione podejście analityczne, nawet kosztem mniejszej przystępności formy. Różnice te ukazują więc szerszy dylemat dotyczący standardów w opracowaniach *policy brief*: czy większą wartość stanowi precyzja i klarowność w prezentacji pomysłu na interwencję publiczną, czy raczej głęboka diagnoza poparta bogatymi źródłami i szerszym kontekstem teoretycznym. Obserwacje te pokrywają się z literaturą nt. komunikacji z decydentami polityk publicznych (Miller, 2015; Reid i Chaytor, 2022).

Wniosek 2: Produktywna współpraca z GenAI

Eksperyment pokazał, że GenAI może stanowić skuteczne wsparcie dla projektantów polityk publicznych przy tworzeniu koncepcji interwencji behawioralnych. Te ustalenia stoją w sprzeczności z wcześniejszymi wynikami badań dotyczącymi wykorzystania GenAI w zadaniach analitycznych (Kocoń et al., 2023), w tym pisaniu *policy briefs* (Safaei i Longo, 2024). Rozbieżność może wynikać z dwóch kwestii. Po pierwsze, w ostatnich miesiącach nastąpił postęp technologiczny w zakresie możliwości modeli językowych. Po drugie, test był przygotowywany w oparciu o złożoną instrukcję (prompt) z jasno określonymi zadaniami i formą, a w dwóch przypadkach również w oparciu o konkretne źródła, co podniosło jakość danych wejściowych. Wyniki są natomiast zgodne z najnowszymi ustaleniami z innych sektorów, które wskazują, że eksperci oraz zespoły efektywnie integrujące narzędzia GenAI w swoich procesach pracy osiągają znacząco wyższą produktywność niż osoby, które z takich narzędzi nie korzystają lub robią to pobieżnie (Dell’Acqua et al., 2025; The Economist, 2025).

Wniosek 3: Istotność przemysłanego procesu projektowania

Patrząc na pozytywne oceny wyników współpracy z GenAI, należy uwzględnić szerszą perspektywę – nakłady pracy potrzebne do przygotowania współpracy z GenAI. W badaniu fundamentem były trzy działania przygotowawcze. Po pierwsze, było to stworzenie ram koncepcyjnych (Teoria Zmiany – tabela 1). Ramy te prze-

łożyły się na szablon notatki i sekwencję działań z GenAI. Ich opracowanie oparto na przeglądzie literatury i praktyki. Szablon testowano pod kątem przejrzystości i responsywności AI na polskie tłumaczenia angielskojęzycznych terminów. Drugim elementem były decyzje projektantów według metody TRIP (tabela 2), dotyczące zadań przypisanych GenAI, doboru zasobów informacyjnych, typów modeli oraz brzmienia promptów. Trzecim działaniem było zgromadzenie rzetelnych źródeł merytorycznych jako bazy informacyjnej dla GenAI. Pakiet 41 źródeł był efektem dwóch quasi-systematycznych przeglądów literatury (University of Washington, Seattle: wrzesień–grudzień 2023; Temple University, Filadelfia: czerwiec–sierpień 2024). Należy więc podkreślić, że współpraca z GenAI wymaga znaczącej inwestycji czasowej w prace przygotowawcze.

Wniosek 4: Kwestie bezpieczeństwa danych i etyki

Test nie wskazał kwestii bezpieczeństwa danych, ponieważ GenAI pracowała wyłącznie na publicznie dostępnych danych wtórnych. Jednak w przypadku projektowania zasilonego danymi pierwotnymi (np. wywiady, notatki z obserwacji uczestniczących), konieczny będzie odpowiedni dobór platform i ustawień, które zapewnią zachowanie anonimowości i bezpieczeństwa danych. Test rodzi jednak implikacje natury etycznej dotyczące autorstwa notatek i odpowiedzialności za ich treść. Praktyka nie wypracowała jeszcze standardów postępowania z materiałami tworzonymi przez GenAI. Jednak renomowani wydawcy naukowcy (WAME, 2025; Sage, 2025) już teraz wprowadzają wytyczne, według których GenAI nie może być wskazywana jako autor tekstu i nie może generować wniosków (bo tylko człowiek może ponosić odpowiedzialność). Takie podejście dyskwalifikowałoby notatki B i C, które praktycznie w całości stworzyła GenAI. W przypadku notatki D, kluczowe decyzje podejmowali projektanci, ale i tu pozostaje kwestia transparentności – rzetelnego poinformowania odbiorców o udziale GenAI w tworzeniu treści.

Wniosek 5: Kierunki dalszych badań

Projektowanie interwencji publicznych jest wypadkową trzech rzeczy: 1) wiedzy z literatury i wcześniejszych praktyk, 2) badań i interakcji z grupą docelową oraz 3) uwarunkowań instytucjonalnych i relacji z interesariuszami. Badanie skupiło się na pierwszym elemencie. Wyniki nie mogą być więc interpretowane jako dowód zastąpienia badań terenowych przez GenAI, lecz tylko jako wskazanie potencjału technologii w fazie koncepcyjnej, przed głębszą diagnozą behawioralną. Dalsze badania powinny eksplorować sposoby włączania GenAI w cały proces projektowania, weryfikując, czy

AI może wspomagać planowanie narzędzi diagnozy, interpretację danych z badań terenowych, wzmacniać jakość sesji współprojektowania, czy testować rozwiązania z pomocą syntetycznych interesariuszy (Rhymer et al., 2024). Przyszłym, rozbudowanym testom w środowisku realnych interwencji powinna towarzyszyć refleksja nie tylko nad skutecznością GenAI, ale również implikacjami etycznymi i kwestiami bezpieczeństwa informacji.

Podsumowanie

W badaniu sprawdzono, czy GenAI może konstruktywnie wesprzeć projektantów w przygotowaniu pomysłu na interwencję publiczną. Dwójka projektantów polityk publicznych, wspieranych w różnym stopniu przez GenAI, przygotowała cztery notatki zawierające różne pomysły na interwencję publiczną redukującą marnotrawienie żywności w mieście. Notatki te następnie zostały ocenione przez 15 praktyków polityk publicznych.

Wynik badania pokazuje, że projektanci wspomagani przez GenAI mogą skutecznie wypracowywać wstępne koncepcje interwencji publicznych opartych na podejściu behawioralnym. Pomysły takie są przekonujące dla praktyków polityki i stanowić mogą cenny wkład początkowy w rozwój konkretnych rozwiązań publicznych i dalsze badania empiryczne.

Wszystkie notatki stworzone w teście miały wysoką jakość, również te praktycznie w całości wygenerowane maszynowo. Ten wynik może stwarzać złudne wrażenie, że każda osoba, nawet laik, może z pomocą GenAI automatycznie projektować rozwiązania behawioralne. Trzeba więc podkreślić, że notatki były ostatecznym efektem dużo szerszego i czasochłonnego procesu pracy. Praca z GenAI wymaga od projektantów głębokiego zaangażowania – narzucenia ram koncepcyjnych, doboru i zasilenia GenAI wiarygodnymi źródłami, ciągłej interakcji i krytycznej weryfikacji tworzonych treści na różnych etapach procesu (coś, co anglosaska praktyka określa terminem: *human in the loop*).

Eksperyment potwierdził wartość zastosowanej ramy koncepcyjnej – Teorii Zmiany wzbogaconej o model COM-B – jako skutecznego narzędzia prowadzenia GenAI w procesie projektowania behawioralnego. Struktura ta porządkowała pracę projektanta i wymuszała na GenAI *explicite* artykulację mechanizmu zmiany, który w interwencjach publicznych bywa często domyślny lub niesprecyzowany. Dzięki tej ramie możliwe było określenie, dlaczego i jak proponowane działania miałyby prowadzić do pożądanej zmiany zachowania w grupie docelowej. To zwiększyło spójność logiczną i perswazyjność pomysłów. Teoria Zmiany może stanowić kluczowy kom-

ponent w pracy z GenAI przy projektowaniu rozwiązań publicznych, umożliwiając systematyzację procesu i zwiększając przejrzystość oraz falsyfikowalność rozwiązań.

W ocenie autorów najbardziej efektywnym sposobem współpracy z AI okazała się taktyka polegająca na rozbijaniu procesu na małe, zdefiniowane zadania, formułowaniu precyzyjnych instrukcji oraz prowadzeniu dialogu wokół konkretnych źródeł wiedzy z wykorzystaniem różnych platform AI. Takie podejście, wdrożone w opracowaniu notatki D, pozwoliło na stworzenie dokumentu o najwyższej wartości merytorycznej, komunikacyjnej i operacyjnej, co potwierdziły zarówno oceny punktowe, jak i rekomendacje ekspertów. Metoda TRIP (patrz: tabela 2) sprawdziła się w usystematyzowaniu bogatego spektrum wyborów, przed którymi stają projektanci, podejmując taką współpracę z GenAI.

Patrząc szerzej, wyniki eksperymentu sygnalizują zmieniającą się rolę projektantów rozwiązań publicznych w erze GenAI. Obrazowo można to ująć jako ewolucję od budowniczych do architektów procesu, którzy muszą określić zakres pracy i współpracy w relacji projektanci-interesariusze-GenAI. Pierwsze porównawcze badania międzynarodowe sektorów analityki publicznej i biznesowej wskazują, że realny wzrost produktywności i jakości pracy odnotowują właśnie osoby, które skutecznie integrują GenAI ze swoim procesem (ang. *workflow*), tworząc autorskie zespoły asystentów AI (The Economist, 2025).

Pozostaje mieć nadzieję, że zarówno ciekawość i motywacje projektantów, jak i instytucjonalne działania w obszarze szkoleń i rozwoju zawodowego pozwolą administracji publicznej w pełni wykorzystać potencjał generatywnej AI w projektowaniu skutecznych polityk publicznych.

Załącznik. Treść zadania

I. INFORMACJE KONTEKSTOWE

Urząd Miasta Stołecznego Warszawy chce podjąć problem marnowania żywności w metropolii. Poszukuje pomysłów na skuteczne interwencje redukujące marnotrawienie żywności, które mogą być zainicjowane przez urząd.

II. TWOJA ROLA I CEL

Jesteś projektantem polityki – ekspertem, który wspiera Miasto Stołeczne Warszawa podczas wstępnego projektowania rozwiązań publicznych.

Twoim celem jest przedstawienie pomysłu na interwencję publiczną redukującą marnotrawienie żywności w mieście. Twoja propozycja powinna być możliwa do realizacji przez miasto, to znaczy zaproponowane działania powinny leżeć w kompetencjach urzędu. Twój pomysł powinien mieć formę krótkiej notatki (ang. *policy brief*).

III. FORMA I ZAWARTOŚĆ NOTATKI

Przygotuj *policy brief* prezentujący pomysł interwencji na rzecz ograniczenia marnowania żywności w Warszawie. Notatka jest skierowana do decydentów wyższego szczebla w mieście: Prezydenta Warszawy i dyrektorów odpowiednich biur urzędu.

Policy brief powinien mieć nie więcej niż 700 słów. *Policy brief* powinien zawierać następujące cztery punkty:

(1) PROBLEM

Wskaż grupę docelową i zachowanie lub sytuację, na których miasto powinno skupić interwencję ograniczającą marnowanie żywności. Grupą docelową mogą być ludzie lub organizacje. Podaj jasną listę przesłanek, aby przekonać decydentów, dlaczego miasto powinno skupić się na tej konkretnej grupie i jej zachowaniach.

(2) DIAGNOZA

Zidentyfikuj typowe sytuacje prowadzące do marnowania żywności w grupie docelowej. Wskaż przeszkody, które utrudniają grupie docelowej zredukowanie marnotrawstwa jedzenia. Przeszkody te opisz używając modelu COM-B (ang. capability, opportunity, motivation, behavior; polskie tłumaczenie to: potencjał, sposobność, motywację, zachowanie)

Wyraźnie wskaż, na jakiej sytuacji i jakich przeszkodach koncentrujesz projekt.

(3) PROPONOWANE ROZWIĄZANIE

Zaproponuj interwencję, która wywoła pozytywną zmianę zachowań w grupie docelowej. Jasno opisz, jakie działania proponujesz. Wyjaśnij co miałyby wywołać i podtrzymać pożądaną reakcję w grupie docelowej (tzw. mechanizm zmiany).

Proponowana interwencja powinna wykorzystywać, ale nie ograniczać się wyłącznie do instrumentów behawioralnych polityk publicznych (ang. terminy: *nudges and boosts*).

(4) OCZEKIWANE EFEKTY

Opisz, jak będzie wyglądał efekt końcowy interwencji. Zaproponuj wskaźniki do pomiaru pozytywnej zmiany.

Oświadczenie o wkładzie poszczególnych autorów

Wszyscy wymienieni autorzy wnieśli znaczący, bezpośredni i intelektualny wkład w artykuł i zatwierdzili go do publikacji.

Oświadczenie o konflikcie interesów

Autorzy deklarują, że badanie zostało przeprowadzone przy braku jakichkolwiek powiązań komercyjnych lub finansowych, które można by zinterpretować jako potencjalny konflikt interesów.

Oświadczenie o etyce badawczej

Autorzy poświadczają, że opublikowane w tekście badania zostały przeprowadzone zgodnie z etyką badawczą afiliowanej uczelni.

Oświadczenie o źródle finansowania

Praca powstała w wyniku realizacji projektu badawczego o nr 2021/43/B/HS5/01935 finansowanego ze środków Narodowego Centrum Nauki.

Informacja o otwartych danych

Pełen zbiór danych z prezentowanej analizy jest dostępny pod linkiem: <https://doi.org/10.58142/swps-ko-dw-01/09/2025>

Pakiet zawiera: (1) Zapisy *promptów* z trzech sesji z GenAI poświęconych tworzeniu *policy briefs* (sesje B, C i D); (2) Instrukcje dla ekspertów oraz pakiet czterech notatek prezentujących pomysły interwencji (*policy brief* A, B, C i D); (3) Zanonimizowane wyniki ocen dokonanych przez 15 ekspertów.

Podziękowania

Autorzy dziękują ekspertom, którzy wzięli udział w badaniu i dokonali oceny notatek oraz przedstawili jakościowe uzasadnienia swoich wyborów. Wyrazy wdzięczności kierujemy również do dwóch anonimowych recenzentów za uwagi, które przyczyniły się do znaczącego ulepszenia tekstu.

Bibliografia

- Astbury, B., Leeuw, F.L. (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation, *American Journal of Evaluation*, 31(3): 363–381, <https://doi.org/10.1177/1098214010371972>.
- Banerjee, S., Galizzi, M.M. (2024). Behavioral Public Policy for Global Challenges. W: *A Better How: Notes on Developmental Meta-Research* (90–101), P.S. Forscher, M. Schmidt (Eds.). Busara, <https://doi.org/10.62372/ISCI6112>.
- Banerjee, S., John, P. (2025). Behavioural Public Policy: Past, Present, and Future, SSRN, 1–30, <https://doi.org/10.2139/ssrn.5124029>.
- Bergman, D. (2024). *What Is Fine-Tuning?* <https://www.ibm.com/think/topics/fine-tuning> (dostęp: 15.03.2024).

- Carrasco, M., Habib, C., Felden, F., Sargeant, R., Mills, S., Shenton, S., Ingram, J., Dando, G. (2023). *Generative AI for the Public Sector: From Opportunities to Value*, <https://www.bcg.com/publications/2023/unlocking-genai-opportunities-in-the-government>
- Carrasco, M., Habib, C., Felden, F., Sargeant, R., Mills, S., Shenton, S., Ingram, J., Dando, G. (2024). *The Journey to Scale*, <https://www.bcg.com/publications/2024/gen-ai-journey-to-scale-in-government>.
- Charalabidis, Y., Medaglia, R., van Noordt, C. (Eds.). (2024). *Research Handbook on Public Management and Artificial Intelligence*. Cheltenham: Edward Elgar Publishing.
- Dell'Acqua, F., Ayoubi, C., Lifshitz-Assaf, H., Sadun, R., Mollick, E.R., Mollick, L., Han, Y., Goldman, J., Nair, H., Taub, S., Lakhani, Karim R. (2025). The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise, *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, 25(043), 1–54, <https://doi.org/http://dx.doi.org/10.2139/ssrn.5188231>.
- Dolan, P., Hallsworth, M., Halpern, D., King, D., Vlaev, I. (2010). *MINDSPACE. Influencing Behaviour Through Public Policy*. UK Cabinet Office and Institute for Government.
- Dupoux, M., Gaudeul, A., Baggio, M., Bruns, H., Ciriolo, E., Krawczyk, M., Kuehnhanss, C., Nohlen, H. (2025). *Unlocking the Full Potential of Behavioural Insights for Policy* (vol. JRC138028). Publications Office of the European Union – Joint Research Centre. <https://doi.org/10.2760/7367599>.
- Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koochang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Wright, R. (2023). Opinion Paper: „So What If ChatGPT Wrote It?”. Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy, *International Journal of Information Management*, 71(102642): 1–63, <https://doi.org/10.1016/j.ijinfomgt.2023.102642>.
- Editorial. (2024). Food Loss and Waste, *Nature Food*, 5(8): 639, <https://doi.org/10.1038/s43016-024-01041-7>.
- Fontana, S., Errico, B., Tedesco, S., Bisogni, F., Robin, Akagi, M., Santiago, N. (2025). *AI and GenAI Adoption by Local and Regional Administrations*, European Committee of the Regions. Commission for Economic Policy, <https://op.europa.eu/s/z69I>.
- Gasser, U., Mayer-Schönberger, V. (2024). *Guardrails: Guiding Human Decisions in the Age of AI*. Princeton: Princeton University Press.
- Harrison, R.M., Lapteva, E., Bibin, A. (2024). Behavioral Nudging with Generative AI for Content Development in SMS Health Care Interventions: Case Study, *JMIR AI*, 3: e52974, <https://doi.org/10.2196/52974>.
- Howlett, M. (2019a). Behavioural Considerations of Public Policy: Matching Policy Tools and Their Targets. W: *Handbook of Behavioural Change and Public Policy* (78–88), H. Strafheim, S. Beck (Eds.). Cheltenham: Edward Elgar Publishing.
- Howlett, M. (2019b). *The Policy Design Primer. Choosing the Right Tools for the Job*. London: Routledge.

- Jacob, S. (2025). Artificial Intelligence and the Future of Evaluation: from Augmented to Automated Evaluation, *Digital Government: Research and Practice*, 6(1): 1–10, <https://doi.org/10.1145/3696009>.
- Janssen, M. (2025). Responsible Governance of Generative AI: Conceptualizing GenAI as Complex Adaptive Systems, *Policy and Society*, 44(1): 38–51, <https://doi.org/10.1093/polsoc/puae040>.
- Kahneman, D. (2013). Foreword. W: *The Behavioral Foundations of Public Policy* (vii–ix), E. Shafir (Ed.). Princeton: Princeton University Press.
- Kaiser, M., Lohmann, P., Ochieng, P., Shi, I., Sunstein, C.R., Reisch, L.A. (2024). Leveraging LLMs for Predictive Insights in Food Policy and Behavioral Interventions, *arXiv, preprint arXiv:2411.08563*, 1–15, <https://doi.org/10.48550/arXiv.2411.08563>.
- Kocoń, J., Cichecki, I., Kaszyca, O., Kochanek, M., Szydło, D., Baran, J., Bielaniec, J., Gruza, M., Janz, A., Kanclerz, K., Kocoń, A., Koptyra, B., Mieszczenko-Kowszewicz, W., Miłkowski, P., Oleksy, M., Piasecki, M., Radliński, Ł., Wojtasik, K., Woźniak, S., Kazienko, P. (2023). ChatGPT: Jack of All Trades, Master of None, *Information Fusion*, 99: 101861, <https://doi.org/10.1016/j.inffus.2023.101861>.
- Kuehnhanss, C.R. (2019). The Challenges of Behavioural Insights for Effective Policy Design, *Policy, and Society*, 38(1): 14–40, <https://doi.org/http://doi.org/10.1080/14494035.2018.1511188>.
- Lasswell, H.D. (1951). The Policy Orientation. W: *The Policy Sciences. Recent Developments in Scope and Method*, D. Lerner, H.D. Lasswell (Eds.). San Francisco: Stanford University Press.
- Lunn, P. (2014). *Regulatory Policy and Behavioural Economics*. Paris: OECD Publishing.
- Manning, S., Sharma, S.R., Walton, M. (2023). Realist Review and System Dynamics as a Multimethod Qualitative Synthesis Approach for Analyzing Waste Minimization in Aotearoa New Zealand, *Systems*, 11(8), <https://doi.org/10.3390/systems11080385>.
- Mayne, J. (2017). Theory of Change Analysis: Building Robust Theories of Change, *Canadian Journal of Program Evaluation*, 32(2): 155–173, <https://doi.org/doi:10.3138/cjpe.31122>.
- McKinsey & Company. (2023). *Unlocking the Potential of Generative AI: Three Key Questions for Government Agencies*, <https://www.mckinsey.com/industries/public-sector/our-insights/unlocking-the-potential-of-generative-ai-three-key-questions-for-government-agencies>
- McKinsey & Company. (2024). *What is Retrieval-Augmented Generation (RAG)?*, <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-retrieval-augmented-generation-rag>
- Michie, S., van Stralen, M., West, R. (2011). The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions, *Implementation Science*, 6(42): 1–11, <https://doi.org/https://doi.org/10.1186/1748-5908-6-42>.
- Miller, R.L. (2015). How People Judge the Credibility of Information. W: *Credible and Actionable Evidence. The Foundation for Rigorous and Influential Evaluations* (39–61), S.I. Donaldson, C.A. Christie, M.M. Mark (Eds.). Thousand Oaks: SAGE.
- Mollick, E. (2024). *Co-Intelligence: Living and Working with AI*. New York: Penguin.

- Mozafar, M., Moini, A., Sobhanifard, Y. (2023). A Systematic Review of Behavioral Public Policy Research: Origins, Mechanisms and Outcomes, *Transforming Government: People, Process and Policy*, 17(4): 603–631, <https://doi.org/10.1108/TG-12-2022-0168>.
- OECD. (2017). *Behavioural Insights and Public Policy. Lessons from Around the World*. Paris: OECD Publishing House.
- OECD. (2019). *Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit*. Paris: OECD Publishing House.
- Olejniczak, K., Batorski, D., Pokorski, J. (Eds.). (2025). *Generatywna AI w badaniach. Praktyczne zastosowania w ewaluacji polityk publicznych*. Polska Agencja Rozwoju Przedsiębiorczości.
- Olejniczak, K., Śliwowski, P., Leeuw, F. (2020). Comparing Behavioral Assumptions of Policy Tools: Framework for Policy Designers, *Journal of Comparative Policy Analysis*, 22(6): 498–520, <https://doi.org/https://doi.org/10.1080/13876988.2020.1808465>.
- Open AI (2025). *Reasoning Best Practices. Learn When to Use Reasoning Models and How They Compare to GPT Models*, <https://platform.openai.com/docs/guides/reasoning-best-practices> (dostęp: 8.05.2025).
- Park, J., Zou, C.Q., Shaw, A., Mako, B., Cai, C., Morris, M.R., Willer, R., Liang, P., Bernstein, M.S. (2025). *Simulating Human Behavior with AI Agents*. San Francisco: Stanford University – Human-Oriented AI.
- Phoenix, J., Taylor, M. (2024). *Prompt Engineering for Generative AI. Future-Proof Inputs for Reliable AI Outputs*. Sebastopol: O'Reilly Media.
- PwC. (2024). *GenAI for Next-Gen Governments*, <https://www.pwc.com/gx/en/government-public-services/assets/genai-for-next-gen-governments.pdf>.
- ReFED. (2024). *ReFED Insights Engine – food waste solution database*, <https://insights.refed.org/>.
- Reid, G., Chaytor, S. (2022). What Do Policymakers Want from Researchers? Developing Better Understanding of a Complex Landdscape. W: *How to Engage Policy Makers with Your Research: The Art of Informing and Impacting Policy* (10–27), T. Vorley, S. Abdul Rahman, L. Tuckerman, P. Wallace (Eds.).
- Reynolds, C. (2023). Tackling Food Loss and Waste. An Overview of Policy Actions. W: *Food Loss and Waste Policy: From Theory to Practice* (42–60), S. Buseti, N. Pace (Eds.). Taylor & Francis. <https://doi.org/10.4324/9781003226932-5>.
- Rhymer, J., Murray, A., Sirmon, D. (2024). Synthetic Stakeholders: Engaging the Environment in Organizational Decision-Making. W: *Research Handbook on Artificial Intelligence and Decision Making in Organizations* (226–239), I. Constantiou, P.J. Mayur, M. Stelmaszak (Eds.). Cheltenham: Edward Elgar.
- Richter, A. (2024). *Navigating Generative AI in Government*, The IBM Center for The Business of Government, <https://www.businessofgovernment.org/report/navigating-generative-ai-government>
- Ruggeri, K. (Ed.). (2018). *Behavioral Insights for Public Policy. Concepts and Cases*. London: Routledge.
- Safaei, M., Longo, J. (2024). The End of the Policy Analyst? Testing the Capability of Artificial Intelligence to Generate Plausible, Persuasive, and Useful Policy Analysis. *Digital Government: Research and Practice*, 5(1): 1–35, <https://doi.org/10.1145/3604570>.

- Sage (2025). *Artificial Intelligence Policy*, <https://uk.sagepub.com/en-gb/eur/artificial-intelligence-policy> (dostęp: 6.07.2025).
- Service, O., Hallsworth, M., Halpern, D., Algate, F., Gallagher, R., Nguyen, S., Ruda, S., Sanders, M. (2014). *EAST. Four Simple Ways to Apply Behavioural Insights*. Behavioural Insights Team.
- Soman, D. (2017). *The Last Mile. Creating Social and Economic Value from Behavioral Insights*. Toronto: University of Toronto Press.
- Stephans, E. (Ed.). (2016). *Social and Behavioral Science Used in the Advancement of Policy: Applications and Insights*. Nova Science Pub Inc.
- Sunstein, C.R., Reisch, L.A. (Eds.). (2023). *Research Handbook on Nudges and Society*. Cheltenham: Edward Elgar Publishing.
- Thaler, R.H., Sunstein, C.R., Balz, J. (2013). Choice Architecture. W: *The Behavioral Foundations of Public Policy* (428–439), E. Shafir (Ed.). Princeton: Princeton University Press.
- The Economist (2025). Ascension, for some: How AI will divide the best from the rest. *The Economist*, Finance & economics.
- Tveita, L.J., Hustad, E. (2025). Benefits and Challenges of Artificial Intelligence in Public sector: A Literature Review, *Procedia Computer Science CENTERIS / ProjMAN / HCist*, 256: 222–229, <https://doi.org/10.1016/j.procs.2025.02.115>.
- USF Libraries (2025). *AI Tools and Resources: Evaluating the Reliability and Validity of AI Generated Text and Media*, <https://guides.lib.usf.edu/c.php?g=1315087&p=9678779> (dostęp: 19.03.2025).
- WAME (2025). *Chatbots, Generative AI, and Scholarly Manuscripts*, <https://wame.org/page3.php?id=106> (dostęp: 6.07.2025).
- Weerts, S. (2025). Generative AI in Public Administration in Light of the Regulatory Awakening in the US and EU, *Cambridge Forum on AI: Law and Governance*, 1: e3, <https://doi.org/10.1017/cfl.2024.10>.
- Wendel, S. (2013). *Designing for Behavior Change. Applying Psychology and Behavioral Economics*. Sebastopol: O'Reilly Media.
- West, R., Michie, S. (2020). A Brief Introduction to the COM-B Model of Behaviour and the PRIME Theory of Motivation, *Qeios*, <https://doi.org/10.32388/WW04E6.2>.
- World Bank. (2015). *Mind, Society, and Behavior*. World Bank Group.
- Yun, L., Yun, S., Xue, H. (2024). Improving Citizen-Government Interactions with Generative Artificial Intelligence: Novel Human-Computer Interaction Strategies for Policy Understanding Through Large Language Models, *PLoS One*, 19(12): e0311410, <https://doi.org/10.1371/journal.pone.0311410>.

