

CHECKLISTA: PROJEKT GEN AI W TWOJEJ ORGANIZACJI

Michał Jaroń
Adam Gabriel Dobrakowski,
COGITA



WSTĘP

Jesteśmy obecnie świadkami rewolucji technologicznej, która zmienia oblicze wielu branż, przynosząc ze sobą niewyobrażalne wcześniej możliwości. Na czele tych zmian stoi generatywna sztuczna inteligencja (Gen AI), która poprzez takie systemy jak chatboty, personalizowane systemy do obsługi klienta, czy możliwość budowania firmowej bazy wiedzy, może oszczędzić mnóstwo pracy osobom na różnych szczeblach w organizacji, zwiększyć ich efektywność i ograniczyć koszty.

Zrozumienie i wdrożenie tych innowacji jest kluczowe dla przetrwania i rozwoju w coraz bardziej konkurencyjnym środowisku biznesowym. Dzięki wykorzystaniu tej checklisty, będziesz mógł krok po kroku zrozumieć, jak wdrażać te technologie, maksymalizując ich potencjał i minimalizując ryzyko.

Zyskasz dostęp do najlepszych praktyk, które pomagają w identyfikacji najbardziej wartościowych obszarów zastosowania Gen AI, planowaniu strategicznym i wykonaniu skutecznych projektów.

Życzymy Ci powodzenia we wdrażaniu AI w Twojej organizacji i zostania liderem dokonujących się zmian technologicznych!

Michał Jaroń
Adam G. Dobrakowski



O NAS



MICHAŁ JARON

AI expert, Lead AI Strategist, Senior Data Scientist w TVN WBD Discovery. Specjalista od sztucznej inteligencji z doświadczeniem w edukacji oraz wdrażaniu technologii generatywnych w jednym z największych polskich ekosystemów dziennikarskich.

Wdrażał algorytmy AI w różnych branżach, pracując w m.in. Legii Warszawa, gdzie wdrażał AI w unikalnym środowisku piłkarskim, czy też w Gemiusie, zgłębiając polski rynek wydawniczy.

ADAM G. DOBRAKOWSKI

Właściciel i CEO firmy COGITA, specjalizującej się w budowie szytych na miarę rozwiązań AI, m. in. dla branży produkcyjnej, e-commerce, czy medycznej.

Współpracował z takimi firmami jak Oracle, Polkomtel, RTB House, LekSeek czy 7bulls.

Autor publikacji naukowych i prelegent konferencji z Data Science.

ETAPY PROJEKTU GEN AI

Zrozumienie i analiza potrzeb oraz formowanie zespołu

W pierwszych etapach projektu kluczowe jest dokładne zrozumienie wymagań biznesowych oraz oczekiwań odbiorców usług. Na początek otwórz się na jak najszersze spektrum perspektyw.

1 Ważne jest, by tworzyć zespół złożony z osób o różnych specjalizacjach - nie tylko ekspertów od sztucznej inteligencji, ale także przedstawicieli sektora biznesowego, specjalistów od UX/UI, prawników, inżynierów oraz analityków danych. Możesz wtedy dokonać wszechstronnej oceny stawianych wyzwań i razem opracować optymalne rozwiązanie.

Metodologia *double diamond* może pomóc w eksploracyjnej „burzy mózgów” i eksploracji *value propositions*. Przykładowo przy projektowaniu *chatbota* zespół powinien sprawdzić, w jakich obszarach najczęściej zadawane są pytania i jakie procesy mogą być zautomatyzowane.

Definiowanie celów i strategii

Następnym krokiem jest synteza zebranych pomysłów i wymagań oraz precyzyjne zdefiniowanie celów projektu. Cele powinny być mierzalne, osiągalne (SMART) i skorelowane z biznesowymi KPI.

2 Specyfika projektów GenAI sprawia, że raz stworzony fundament może zostać ponownie użyty w innych zastosowaniach, tym samym zwiększając potencjalny ROI (np. ten sam kod *chatbota* dla BOK i HR).

Zadbaj o odpowiednie zarządzanie oczekiwaniami - szum medialny wokół AI, sprawia, że oczekiwania interesariuszy są często nierealistyczne. Upewnij się, że definicje są rozumiane przez wszystkich uczestników projektu w ten sam sposób. Może Ci w tym pomóc stworzenie „słowniczka” pojęć. Zadbaj o jednoznaczne potwierdzenia (np. drogą mailową), pozwoli to zminimalizować szansę na nieporozumienia.

3

Opracowanie i testowanie prototypu

W prowadzenie projektów GenAI nierozdzielnie wpisane jest eksperymentowanie. Dlatego warto pracować w zwinnej metodyce, pozwalającej na szybkie iteracje i na błyskawiczną naukę.

Tempo zmian w GenAI, może spowodować, że nawet w trakcie krótkiego projektu pojawi się nowy model, który wyrzuci do góry nogami wcześniejsze założenia. Zanim zespół programistyczny zacznie cokolwiek implementować - dokonaj analizy rynku, sprawdź istniejące narzędzia pozwalające na łatwą integrację z Twoim ekosystemem.

Tworzenie rozwiązania zacznij od czegoś małego. Nie próbuj od razu budować produktu na pełnych danych. Zacznij od pracy na małej próbce (*dummy data*).

Zanim użyjesz bardziej zaawansowanych metod optymalizacji modelu (RAG, fine tuning), sprawdź czy wystarczające okażą się techniki *prompt engineeringu*.

Zadbaj o stworzenie sposobu obiektywnej analizy kolejnych wersji - stwórz *benchmark*, opisz zasady oceny przez człowieka. W tej fazie skup się na jak szybkim prototypowaniu - pamiętaj jednak o produkcyjnych ograniczeniach.

4

Analiza i wybór modelu bazowego

W fazie MVP możesz testować wiele tzw. modeli fundamentalnych (*foundation models*), które są wytrenowane do ogólnych zastosowań na wielkich korpusach danych. Modele zamknięte, których kod źródłowy nie jest udostępniony (np. *GPT4*, *Gemini*) ciągle oferują najlepszą jakość, jednak nasze dane są wysyłane do zewnętrznego podmiotu. Ze względu na charakter danych, prawne ograniczenia nie zawsze na to pozwalają.

W wielu sytuacjach warto rozważyć modele otwartoźródłowe - *open-source* (*Mistral*, *LLaMA*). Przy odpowiednim *promptowaniu* i dostosowaniu do naszego przypadku pozwalają one osiągnąć zadowalającą jakość.

Czy Twój system będzie działać tylko w języku polskim? W zależności od modelu - jakość dla języka polskiego może być gorsza. Jeżeli zdecydujesz się na model *open-source*, zastanów się, czy posiadasz w organizacji odpowiednie kompetencje i jesteś gotowy utrzymywać rozwiązanie w *CLOUD/on-premise*. W przypadku projektów wymagających rzadkich i szybkich odpowiedzi, hostowanie modelu *open-source* na kartach graficznych (GPU) może być nieefektywne kosztowo.

5

Implementacja i wprowadzenie na rynek

W trakcie tworzenia MVP/PoC, część początkowych założeń może ulec weryfikacji i wymagać dostosowania. Na tym etapie musisz podjąć ostateczną decyzję m.in. odnośnie do modelu bazowego, którego użyjesz (zamknięty model jak OpenAI, open-source na własnym infrastrukturze, model na cloud?).

Musisz także zdecydować, które obliczenia mogą odbywać się *offline (batchowo)*, tzn. nie w momencie uruchamiania modelu przez użytkownika, a które *online* - wówczas ograniczeniem może być wymaganie krótkiego czasu odpowiedzi (np. 400 ms).

W fazie MVP praca odbywała się na próbce danych, dlatego teraz musisz podjąć decyzję związane ze skalowaniem rozwiązania - jak poradzisz sobie ze zwiększonym ruchem, ile to będziesz kosztować?

W tej fazie współpracuj intensywnie z działem IT, szukając najefektywniejszych wdrożenia rozwiązania. Stwórz dokładną *roadmapę*, uwzględniającą testy oraz fazy odbiorów. Regularnie spotykaj się z zespołem w ramach statusów - wprowadź metryczkę opisu postępów - pozwoli to lepiej koordynować działania oraz reagować na potencjalne opóźnienia.

6

Monitorowanie, optymalizacja i skalowanie

Gratulacje! Uruchomiłeś swój projekt! Jednak to jeszcze nie koniec pracy. Niezależnie od tego, czy rozwiązanie zostało stworzone z myślą o optymalizacji pracy wewnątrz organizacji, czy trafiło do użytkownika zewnętrznego - musisz zadbać o monitoring jego działania.

Oprócz klasycznych metryk stosowanych we wszystkich projektach IT (*Average Handling Time AHT, Response Time*), w projektach GenAI kluczowe jest również usystematyzowane zbieranie *feedbacku* od użytkowników dotyczącego jakości odpowiedzi AI.

Kciuki w górę/w dół widoczne w *ChatGPT? OpenAI* zbiera do użytkowników oceny odpowiedzi modelu. Tak prosty *feedback* może stanowić wkład do trenowania kolejnych generacji modeli (metoda *human in the loop*).

Monitoruj również procesy na które wpływa Twój projekt - przeprowadź ankiety wśród pracowników/użytkowników sprawdzające ich satysfakcję. W przypadku chatbota BOK możesz przykładowo sprawdzić liczbę pozytywnie załatwionych spraw.

✓	Zadanie	Data planowana	Data realizacji
	Zidentyfikowanie wszystkich zainteresowanych obszarów w organizacji Stworzenie szerokiej, interdyscyplinarnej grupy roboczej		
	Przeprowadzono warsztaty eksploracyjne Wykorzystaj metodologię <i>double diamond</i> do identyfikacji <i>value proposition</i>		
	Ustalenie terminu i agendy regularnych statusów Oprócz koordynacji działań, statusy mogą pełnić rolę informowania o najnowszych doniesieniach ze świata AI. W tak dynamicznej się branży, warto trzymać rękę na pulsie!		
	Potwierdzenie zrozumienia celów i zakresu projektu przez interesariuszy Zorganizuj sesję Q&A, zapisz i dystrybuuj FAQ.		
	Określono mierzalne KPI Upewnij się, że dokument zawiera specyficzne, mierzalne, osiągalne, realistyczne i czasowo określone cele (metoda SMART).		
	KPIs zatwierdzone przez interesariuszy Uzyskaj potwierdzenie e-mailem.		
	Zaakceptowano SLA (Service Level Agreement) z interesariuszami Ustal krytyczność ciągłości biznesowej. Opracuj plan na wypadek awarii technicznych.		
	Zgromadzono istotne regulaminy usług, z których zamierzasz skorzystać Ułatwi to pracę zespołom prawnym.		
	Akceptacja zespołu prawnego pod kątem praw autorskich i licencji Uzyskaj potwierdzenie e-mailem.		
	Akceptacja zespołu prawnego pod kątem RODO/GDPR Uzyskaj potwierdzenie e-mailem.		
	Ustalony cykl przeglądów Terms & Conditions Regulaminy usług dynamicznie się zmieniają, dlatego ustalenie obiegu informacji i kalendarza przeglądów pozwoli uniknąć potencjalnych problemów.		
	Zatwierdzona ocena wpływu społecznego i etycznego Projekty GenAI wywołują często niepokoje wśród pracowników - warsztaty z działem P&C/HR		
	Przeprowadzono warsztaty z UX/UI Wypracuj wstępny projekt interfejsu użytkownika oraz ustal aspekty specyficzne dla projektów GenAI - takie jak ton wypowiedzi chatabota, płęć, treść komunikatów o halucynacjach itp.		
	Opracowano sposoby mitygacji ograniczeń obecnych modeli Opracuj strategię zarządzania ryzykiem stereotypów zawartych w modelu, halucynacji, braków w wiedzy. Przygotuj plan reakcji na ewentualne incydenty.		
	Przeprowadzono przegląd rynkowych rozwiązań Dokonaj przeglądu gotowych rozwiązań rynkowych, nie opieraj się na sloganach sprzedażowych - poszukaj niezależnych benchmarków.		
	Podjęta decyzja o wariacie wdrożenia Wybierz model (open-source/zamknięty), sposób korzystania z modelu - SaaS (software as a service), wykorzystanie chmury (AWS, Azure GCP), on-premise (własne serwery) - pamiętaj o ograniczeniach związanych z prywatnością danych oraz potencjalnych trade-offach.		

✓	Zadanie	Data planowana	Data realizacji
	Architektura rozwiązania zaakceptowana przez zespół Infosec Skonsultuj pomysł na rozwiązanie pod kątem cyberbezpieczeństwa.		
	Stworzono benchmark/jasne kryteria testowania Idealnie, jeśli mechanizm może być w pełni zautomatyzowany oraz powtarzalny. Jeśli ocena wymaga jednak np. subiektywnej pracy ludzkiej, zadбай o klarowne zasady oceniania.		
	Opracowano przybliżoną roadmapę Zadbaj o odpowiedni bufor - projekty GenAI nierozzerwalnie wiążą się z eksperymentowaniem, Określ funkcjonalności i kamienie milowe.		
	Dokonana analiza i wybór modelu fundamentalnego Rozważ najpopularniejsze modele (GPT4, Gemini, LLaMA, Mistral itp.). Regularnie przeglądaj HuggingFace .		
	Stworzenie pierwszego MVP Pracuj iteracyjnie, najpierw na małych, „zabawkowych” danych. Pracuj wg metodologii „fail fast”. Zaczynaj od najprostszego rozwiązania		
	System zoptymalizowany przy użyciu technik prompt engineeringu Po stworzeniu początkowego MVP, ulepsz jakość odpowiedzi stosując techniki prompt engineeringu (few-shot prompting, CoT, Priming itp.)		
	Dodano RAG (Retrieval Augmented Generation) Jeśli potrzebujesz dostarczyć dodatkową wiedzę do modelu oraz zminimalizować halucynacje - użyj embeddingów trzymanyh w bazie wektorowej. Dodaj kontekst do swoich promptów.		
	Przeprowadzono fine-tuning modelu bazowego Jeśli domena Twojego rozwiązania nadal jest nieobstugiwana w sposób zadowalający - rozważ dotrenowanie modelu na Twoich danych.		
	Zweryfikowano zmiany w specyfice projektu z infosec i zespołem prawnym Po zbudowaniu MVP, wcześniejsze założenia mogą ulec weryfikacji i wymagać dostosowania.		
	Zaplanowano docelową roadmapę Ustal szczegółowy harmonogram z IT.		
	Przeprowadzono testy Red Team Red Team w kontekście rozwiązań Generative AI to grupa ekspertów testujących bezpieczeństwo systemów AI przez symulowanie ataków, aby wykryć i naprawić potencjalne słabości.		
	Zaimplementowano bezpieczniki Wprowadź zmiany na podstawie testów Red Team.		
	Stworzono mechanizm zbierania feedbacku od użytkowników Kciuki w górę/dół, ocena odpowiedzi (optymalizacja modeli tzw. human in the loop)		
	Stworzono raport końcowy Zmapuj początkowe założenia z faktycznie osiągniętymi celami, feedbackiem użytkowników oraz konkretnymi liczbami (spełnione KPI?)		
	Iteracyjnie ulepszaj działanie systemu Feedback zbierany od użytkowników oraz postęp w dziedzinie AI wymagają szybkiej reakcji i gotowości do działania.		

AI PROJECT MANAGER

Uzbrojona/y w wiedzę na temat specyfiki prowadzenia projektów GenAI jesteś gotowa/y, aby odnieść sukces w koordynacji przedsięwzięć opartych na najgorętszej obecnie technologii.

Oczywiście każdy projekt jest inny i kluczowe szczegóły mogą odgrywać niebagatelną rolę. Dlatego warto mieć asystenta, który pomoże nam w decyzjach, zada przydatne pytania i poprowadzi za rękę po meandrach nawet najbardziej zawitych projektów.

Jako bonus do lektury oddajemy w Twoje ręce ChatBOT, który od dzisiaj pomoże w realizacji Twoich projektów!

Co potrafi AI Project Manager?

- Przygotuje roadmapę projektu
- Zaplanuje podział obowiązków
- Przygotuje i wypełni metryczkę ze statusem projektu
- Pomoże stworzyć szyte na miarę KPI
- i wiele, wiele więcej - przetestuj i przekonaj się sam!

**LINK NA NASTĘPNEJ
STRONIE**



TWÓJ OSOBISTY AI PROJECT MANAGER



Custom ChatGPT

(wymaga subskrypcji GPT4)

[<https://cogita.ai/gpt>]

PODSUMOWANIE

Dziękujemy, że jesteś w tym miejscu dokumentu, bo to znaczy, że jego zawartość Cię zainteresowała. To, o czym napisaliśmy, jest oczywiście wycinkiem zakresu wiedzy o tworzeniu projektu opartego o Gen AI. Mamy nadzieję, że jest to interesujący wstęp oraz inspiracja do czegoś nowego!

W naszej internetowej aktywności pogłębiany te oraz inne tematy związane z budową modeli AI.

Jeśli chcesz być na bieżąco z darmową wiedzą, którą udostępniamy, to zachęcamy do śledzenia naszej aktywności w mediach społecznościowych.

MICHAŁ JAROŃ



ADAM G. DOBRAKOWSKI



COGITA

