Marcin Żmigrodzki

Zarządzanie projektami AI dla początkujących

Jak optymalizować procesy biznesowe za pomocą modeli językowych (LLM)

onepress

Redaktor prowadzący: Magdalena Dragon-Philipczyk
Projekt okładki: Studio Gravite / Olsztyn
Obarek, Pokoński, Pazdrijowski, Zaprucki
Materiały graficzne na okładce: Adobe Stock

Helion S.A.
ul. Kościuszki 1c, 44-100 Gliwice
tel. 32 230 98 63
e-mail: *onepress@onepress.pl*
WWW: *onepress.pl* (księgarnia internetowa, katalog książek)

Drogi Czytelniku!
Jeżeli chcesz ocenić tę książkę, zajrzyj pod adres
onepress.pl/user/opinie/zaprai
Możesz tam wpisać swoje uwagi, spostrzeżenia, recenzję.

ISBN: 978-83-289-3844-1

Printed in Poland.

Spis treści

Wprowadzenie

Moja przygoda z algorytmami sztucznej inteligencji zaczęła się jeszcze w zeszłym wieku, gdy w trakcie studiów wciągnęły mnie systemy ekspertowe, a później próbowałem znaleźć klientów na te rozwiązania – bezskutecznie. Wówczas podejście symboliczne w odróżnieniu od koneksjonistycznego, na przykład reprezentowanego przez sieci neuronowe, wydawało się bardziej uniwersalne i takie, które dawało możliwości objaśnienia toku rozumowania. To była moja pierwsza przegrana w starciu z AI.

Tu należy się krótkie objaśnienie dwóch podanych wyżej terminów. Podejście symboliczne – jego reprezentantem były systemy ekspertowe – oznacza założenie, że rzeczywistość da się opisać interpetowalnymi bytami i ich cechami. Na przykład możemy zapisać, że niedojrzały pomidor jest zielony, że pomidor jest warzywem (wiem, że nie jest, ale tak się przyjęło) oraz że niedojrzałe warzywa są kwaśne. Odpowiedni model byłby w stanie wywnioskować, że zielone pomidory są kwaśne. Takie obiekty mogą być odczytane i interpretowane przez człowieka. Jak również człowiek może prześledzić tok rozumowania krok po kroku i go zweryfikować. I w tej transparentności leży największa siła podejścia symbolicznego. Drugą korzyścią jest wiarygodność wniosków. Tu nie ma przestrzeni na halucynacje lub błędy statystyczne. Słabościami podejścia symbolicznego są konieczność wykonania dużej pracy manualnej przy tworzeniu bazy wiedzy oraz nieelastyczność modelu. Wystarczy, że wprowadzę ogórka lub awokado do analiz, i wnioskowanie się posypie, o ile ktoś nie stworzy nowych reguł.

Natomiast w podejściu koneksjonistycznym przyjmuje się, że algorytm sztucznej inteligencji składa się z gigantycznej liczby elementów, które mogą być sztucznymi neuronami albo wagami w macierzach połączonych ze sobą za pomocą matematycznych funkcji. W tym podejściu model jest najpierw trenowany, czyli adaptuje swoje wagi do zadanego zbioru uczącego, a po wytrenowaniu może być wykorzystany do wnioskowania. Siłą tego podejścia jest niesłychana skalowalność. Wystarczy, że dysponujemy zbiorem uczącym i mocą obliczeniową, aby w dużym uproszczeniu ujmując, podnieść jakość wnioskowania. Modele koneksjonistyczne są też bardziej elastyczne, radzą sobie z zaburzonymi i „zaśmieconymi" danymi. Wystarczy, że wytrenujemy je na takich „zaśmieconych" danych. LLM, które należą do tej grupy, pokazały też, że mogą radzić sobie z szerokim spektrum zadań. Problemem jednak jest zjawisko czarnej skrzynki – nie wiem tak naprawdę, dlaczego model wygenerował daną odpowiedź. Jego złożoność jest tak wielka, a sposób wnioskowania zapisany w nieczytelnych wagach, że nie da się tego prześledzić. Drugim problemem jest probabilistyczność rozumowania – model jedynie ma szansę wygenerować dobry wynik.

Jak pokazały to kolejne dekady, o systemach ekspertowych zapomniano z wielu powodów. Nie potrafiły się automatycznie rozwijać, były nieodporne na drobne błędy w danych użytkownika, wymagały sporych inwestycji w rozwój bazy wiedzy. Porażka megaprojektu CYC tylko przypieczętowała los wnioskowania symbolicznego.

CYC był projektem badawczym, którego celem było stworzenie największej bazy regułowej w historii ludzkości. Rozpoczął się w 1984 roku i był rozwijany przez kolejne 20 lat. W jego ramach ręcznie stworzono ponad 20 milionów pojęć i relacji między nimi. Szacuje się, że kosztował 60 milionów dolarów i 600 osobolat pracy. Na takiej bazie wiedzy funkcjonował mechanizm wnioskujący, który miał stać się zalążkiem ogólnej sztucznej inteligencji. Nie stał się. Na początku tego stulecia projekt wyewoluował w open source i był dalej rozwijany aż do 2017 roku, po

czym został zamknięty. Próbowano wdrażać go w obszarze medycyny, bezpieczeństwa sieciowego czy walki z terroryzmem, ale bez większych efektów. Głównymi przyczynami porażki były niezdolność systemu do samodzielnego rozwoju i nieelastyczność wnioskowania.

Na początku tego wieku rozpaliła mnie i kilku kolegów nadzieja opanowania analizy języka naturalnego na tyle, aby możliwe było znajdowanie tekstów w odpowiedzi na pytania użytkowników zadane językiem naturalnym. Wtedy nadeszła druga porażka, ówczesne algorytmy ukrytego indeksowania, przetwarzania fleksji naszego kochanego języka, identyfikacji istotnych słów za pomocą entropii okazały się niewystarczające. Po raz drugi poniosłem porażkę.

Kolejne dwie dekady należały do uczenia maszynowego, które osiągnęło niesłychane poziomy efektywności w bardzo wąskich problemach. W tym czasie zająłem się zgłębianiem specyfiki zarządzania projektami w różnych obszarach – od tworzenia systemów informatycznych, przez wprowadzanie nowych produktów na rynek, po optymalizację procesów.

W 2017 roku pojawiły się transformery z mechanizmem atencji i natychmiast zainteresowałem się ich możliwościami. A tak naprawdę to po raz pierwszy użyłem BERT-a w 2021 roku do generowania podsumowań i ponownie poczułem studencki zachwyt. Przez ostatnie lata nastąpiła rewolucja i oczy biznesu zwróciły się ku generatywnej sztucznej inteligencji.

I tak historię powtarzam po raz trzeci.

Poniższy podręcznik przeznaczony jest dla osób odpowiedzialnych za wdrażanie optymalizacji w procesach biznesowych w organizacji. Stanowi on swoisty drogowskaz w prowadzeniu projektów uwzględniających komponent LLM (ang. *Large Language Model*), oparty na autorskiej metodyce DAPIS.

Metodyka ta przeznaczona jest dla projektów, inicjatyw czy eksperymentów, których celem jest automatyzacja zadań przy użyciu generatywnej sztucznej inteligencji (GenAI). W dalszej części, na określenie tych różnorodnych przedsięwzięć, będzie stosowany wspólny termin „projekt".

Prezentowany podręcznik zakłada wykorzystanie LLM, czyli jednego z aspektów generatywnej sztucznej inteligencji, natomiast pomija zagadnienia uczenia maszynowego.

Moim celem było, poprzez publikację tego podręcznika, zaoferowanie przewodnika kierownikom projektów, liderom, koordynatorom, analitykom procesów i biznesowym na trakcie wiodącym od pomysłu na zastosowanie LLM do codziennie dostarczanej wartości organizacji.

A i ta książka nie jest o zarządzaniu projektami programistycznymi, w których AI generuje kod. Nie opisuję tutaj, jak współpracować w zespole zwinnym, który jest wsparty tzw. vibecodingiem.

Wartość LLM w biznesie

W ciągu pięciu lat sztuczna inteligencja z laboratoriów weszła do języka codziennego, wywołując euforię, lęk i ciekawość. Hollywood zastanawia się, czy generowane efekty specjalne zastąpią czy tylko uzupełnią pracę specjalistów od VFX. YouTube zalewany jest tzw. „ściekiem AI", czyli filmikami wygenerowanymi automatycznie. Niemniej jednak, niezależnie od oceny ich jakości, potrafią one zarobić do kilkunastu tysięcy dolarów miesięcznie. Według raportu McKinsey z 2025 roku rewolucja związana ze sztuczną inteligencją doda 4,4 biliarda dolarów do światowej gospodarki, a 92% firm planuje inwestycje w tę technologię w ciągu kolejnych trzech lat.

Wielkie korporacje finansowe estymują, jak wielu ludzi realizujących powtarzalne zadania biurowe będzie można zwolnić. JP Morgan deklaruje oszczędność 360 tysięcy roboczogodzin rocznie na automatycznej analizie dokumentów finansowych. HSBC chwali się wdrożeniem 600 projektów opartych na generatywnej AI, co przełożyło się na przykład na wzrost efektywności programistów o 15%, trzy miliony zautomatyzowanych interakcji z klientami rocznie.

W badaniach na temat AI w e-administracji publicznej jedna trzecia respondentów przewiduje, że w ciągu pięciu lat generatywna sztuczna inteligencja znacząco wpłynie na strukturę zatrudnienia w administracji publicznej. Spośród badanych 4% wskazuje na możliwość całkowitej transformacji zawodu. Największy wpływ automatyzacji przewiduje się w obszarze rutynowych zadań.

Obsługa klienta zaczyna być automatyzowana przez chatboty AI. Klarna już w 2024 roku ogłosiła, że ich chatboty obsłużyły dwie trzecie (2,3 miliona) rozmów z klientami, co odpowiada 700 etatom. Firma Octopus Energy zadeklarowała, że AI wykonuje pracę 250 konsultantów do spraw obsługi klienta (ponad jedną trzecią e-maili od klientów obsługuje AI). Ikea zaraportowała, że ich chatbot rozwiązuje 47% zgłoszeń, co daje 3,2 miliona interakcji z klientami. Amazon pod koniec 2025 roku ogłosił zwolnienie 14 tysięcy pracowników biurowych i umotywował to rozwojem sztucznej inteligencji. W Polsce PZU chwalił się, że opracował system oparty na LLM do analizy szkód celem szybszej wypłaty odszkodowań. Zakładano analizę 12 tysięcy spraw rocznie.

Vibecoding w przypadku prostych aplikacji zredukował konieczność posiadania umiejętności IT do niemal zera. Przykładowo, start-up Lovable służący do automatycznego tworzenia programów wyceniany jest na 6,6 miliarda dolarów i zanotował 200 milionów dolarów przychodów w ciągu miesiąca. A na rynku już jest co najmniej tuzin podobnych firm, na czele z samym Google AI Studio. Andrew Ng, który jest jednym

z liderów AI i współtwórcą Coursery, wspomniał na wykładzie dla start-upów Y Combinator, że w jednym z jego zespołów zaproponowano, aby dokonać redukcji od czterech do siedmiu programistów na kierownika projektów do ledwie pół programisty. Dwóch kierowników projektów miałoby zatem przypadać na jednego programistę, bo resztę roboty ma robić AI.

Modele językowe zaczynają dowodzić twierdzenia matematyczne, w tym problemy z listy Erdősa, i osiągają najlepsze wyniki na olimpiadach matematycznych. Ostatnie Nagrody Nobla z medycyny i fizyki dotyczyły tak naprawdę algorytmów sztucznej inteligencji.

Kilka słów wyjaśnienia: Pál Erdős był jednym z najpłodniejszych matematyków w historii. Stworzył listę znaczących problemów matematycznych, które niegdyś nie miały rozwiązania, a i do dzisiaj część z nich pozostaje otwarta. Pod adresem *erdosproblems.com* można sprawdzić aktualny status rozwiązywania tych problemów.

Odpowiedzią na te rosnące oczekiwania są inwestycje w infrastrukturę do generatywnej sztucznej inteligencji, które nie mają sobie równych w odniesieniu do poprzednich rewolucji przemysłowych – internetu, komputeryzacji czy rozwoju motoryzacji. Według Gartnera wydatki na AI w 2026 roku sięgną 2,5 biliona dolarów. Dla porównania PKB naszego kraju wynosi około 1 biliona dolarów. W 2027 roku wydatki te mają przekroczyć 3 biliony dolarów, a jest to poziom PKB Francji, czyli siódmej gospodarki świata.

Oczywiście to jest dopiero początek rewolucji. Warto uczciwie dodać łyżkę dziegciu. Duolingo w 2024 roku ogłosiło, że AI pozwoliła na zwolnienie 10% współpracowników. Jednak w ciągu 2025 roku akcje tej firmy spadły o 56%, a na grupach dyskusyjnych pełno jest negatywnych opinii odnośnie do jakości jej usług. Ponad rok później Klarna przyznała, że zwolnienie tak dużej liczby pracowników było pochopne, bo jakość

obsługi klienta spadła i człowiek nadal potrzebuje kontaktu z drugim człowiekiem. Firma na nowo więc zaczęła zatrudniać.

Jak widać na powyższych przykładach, znajdujemy się w okresie, gdy entuzjazm i moda na AI mieszają się z rozczarowaniem i pierwszymi porażkami. Gdy Google pokazało swój model Genie, w którym można tworzyć wirtualne światy tak jak w grach, akcje CD Projekt spadły o 13%, a Take-Two Interactive o 19%, i to mimo zbliżających się premier ich flagowych gier. W grudniu 2025 roku ogłoszono SaaSpocalypse, czyli spadek spółek oferujących oprogramowanie dostępne w abonamencie zwykle przez WWW, bo pojawił się model Claude i system agentowy Cowork. W miesiąc akcje Service Now spadły o 16%, Attlasian o 29%, a Salesforce także o 16%. Pewnie zanim ta książka ukaże się drukiem, notowania tych spółek z powrotem wzrosną, więc to może być niezła okazja inwestycyjna. Jednak to pokazuje, jak wiele emocji towarzyszy wdrażaniu technologii sztucznej inteligencji.

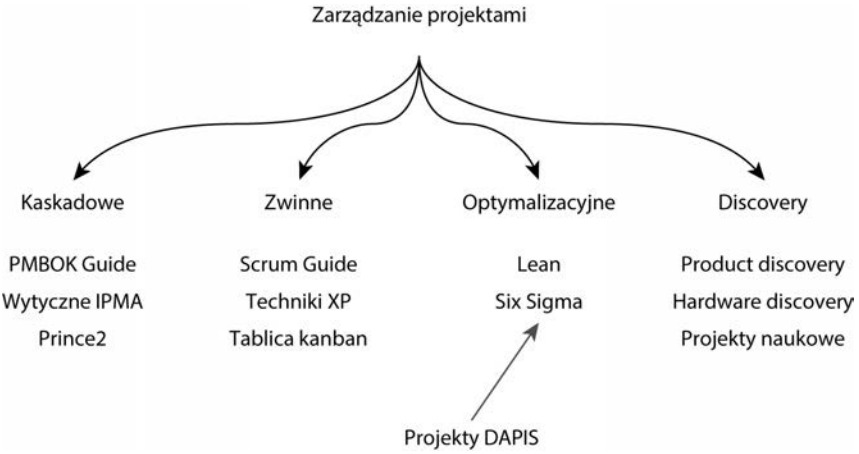
Ze wspomnianych wyżej badań na temat AI w administracji publicznej wynika, że największymi utrudnieniami we wdrażaniu AI w urzędach są niejasne regulacje prawne (73%) lub ich brak (73%) oraz brak procedur wewnętrznych (70%). Zatem są to przyczyny leżące poza technologią.

Celem tej książki jest dostarczenie wartościowych, ale i krytycznych informacji na temat tego, jakie projekty oparte na modelach językowych mogą przynieść wartość i jak się zabrać za ich realizację. Ale celem jest również pokazanie, kiedy nie warto dać ponieść się modzie na sztuczną inteligencję i należałoby wybrać inne rozwiązania, jak na przykład redukcję biurokracji.

Czym jest projekt?

Projekt to ograniczone w czasie przedsięwzięcie. Najbardziej oczywistym projektem jest budowanie domu, programowanie systemu informatycznego czy organizacja konferencji. W tej książce promuję tezę, że doskonalenie funkcjonowania organizacji warto realizować również za pomocą projektów.

Poniższy diagram (rysunek W.1) prezentuje popularne nurty myślenia o zarządzaniu projektami.



Rysunek W.1. Rodzaje podejść do zarządzania projektami

Projekty kaskadowe to projekty, które planujemy szczegółowo na starcie. Wierzymy, że im bardziej drobiazgowa analiza, tym większa szansa, że wyłapiemy pułapki zawczasu i uwzględnimy je w planie. Podejście kaskadowe charakteryzuje się strukturalizacją różnych rozdziałów planu projektu. Strukturalizujemy, czyli dzielimy na uporządkowane komponenty zakres (w formie WBS), harmonogram (w formie diagramu sieciowego lub wykresu Gantta), koszty (w formie tabeli z listą pozycji budżetowych), ryzyk (w formie ich rejestru), jakości (w formie dokumentacji tuzinów wskaźników), zespołu (w formie diagramu struktury

organizacyjnej). To podejście najlepiej sprawdza się, gdy mamy duży projekt, w którym koszty ewentualnych ryzyk i zmian mogą być poważne, ale w którym jesteśmy w stanie wiarygodnie przewidzieć drogę realizacji finalnego produktu. Siłą podejścia kaskadowego jest strukturalizacja problemu.

Podejście zwinne obejmuje projekty, w których co prawda mamy od początku wizję finalnego produktu, ale dopuszczamy, że może się ona zmieniać. Wynika to z tego, że mamy świadomość, iż nie wszystko da się przewidzieć na starcie, a nasza wiedza będzie stopniowo rosła wraz z dostarczaniem zakresu. Projekty zwinne są raczej adresowane do mniejszych zespołów (*Scrum Guide* wprost mówi o dziewięciu osobach) i takich, w których koszty zmian są akceptowalne. Czyli bardziej nada się do projektów niematerialnych, jak produkcja oprogramowania, niż takich, gdzie zużywa się materiał, środki trwałe lub niespełnienie wymogów regulacyjnych może nas sporo kosztować. Siłą podejścia zwinnego jest budowanie osobistego zaangażowania zespołu, co pozytywnie wpływa na jego innowacyjność.

Podejście discovery jest nową szkołą i obecnie stopniowo krystalizuje się w postaci różnych metodyk. Najczęściej spotykaną jest product discovery, czyli projekty, w których głównym wyzwaniem jest to, co w naszym produkcie pokocha klient. W takich projektach na starcie często nie mamy wizji produktu finalnego, tylko zbiór hipotez lub pytań badawczych. Efektem projektu jest zatem udzielenie odpowiedzi na te pytania i stworzenie prototypu. Projekty discovery na ogół są realizowane w jeszcze mniejszych zespołach niż zwinne, bo liczących mniej więcej do pięciu osób. Jeżeli, Drogi Czytelniku, chciałbyś się dowiedzieć więcej o takich projektach, to zachęcam do zapoznania się z moją książką *Bardziej niż Agile*.

Wreszcie, **projekty optymalizacyjne** stanowią niszę. Ich celem nie jest bowiem zbudowanie czegoś nowego, a jedynie poprawienie istniejącego

procesu. Zatem na wejściu mamy proces, który nie działa najlepiej, a na wyjściu proces, który działa lepiej. Celem optymalizacji może być redukcja błędów, poprawa satysfakcji klienta, skrócenie czasu przetwarzania lub przezbrojenia, zmniejszenie zużycia materiału lub energii itd.

Projekty związane ze sztuczną inteligencją – czy to w wymiarze ich celów, czy jedynie wykorzystania jej jako elementu zakresu – mogą opierać się na wszystkich wymienionych wyżej podejściach. Możemy rozwijać urządzenie wykorzystujące AI dla silnie regulowanego sektora, na przykład obronnego, i stosować metodykę kaskadową. Możemy rozwijać nowy produkt w formule SaaS (ang. *Software as a Service*), który będzie oprawą dla GPT – czasem nazywa się takie rozwiązania GPT wrapper. I do takiego projektu prawdopodobnie wybierzemy podejście zwinne. Wreszcie możemy trenować nowy model AI w ramach badań naukowych i w takiej sytuacji zastosujemy filozofię discovery.

Jednak ta książka nie zajmuje się tymi podejściami do projektów. Traktuje ona o tym, jak poprawić bieżące funkcjonowanie małych i średnich firm, korporacji i jednostek sektora publicznego ze wsparciem wielkich modeli językowych. Zatem moim wyborem jest podejście optymalizacyjne.

Jego charakterystycznymi cechami, wyróżniającymi z pozostałych podejść, są:

- Cel – jest nim poprawa aktualnie funkcjonujących procesów. Za pomocą tego typu projektów nie buduje się nowych rzeczy od zera, tylko poprawia istniejące.
- Zakres – jest odkrywany dopiero w połowie projektu, gdy już zespół dobrze zrozumie problem, jaki przed nim stoi. Na starcie projektu pod terminem „zakres" kryje się obszar, w którym występuje problem do naprawy.

- Techniki – w projekcie optymalizacyjnym dużo metod dotyczy opisu obecnego stanu rzeczywistości, przede wszystkim procesów produkcyjnych i usługowych, pojawia się tutaj mapowanie różnych metod, zbieranie danych opisujących funkcjonowanie procesów.

- Harmonogram – takie projekty są na ogół krótkie, trwają do trzech miesięcy, ale często są też dużo krótsze. Co więcej, gdyby okazało się, że problem jest dużo większy i jego poprawa zajmie ponad trzy miesiące, to dobrą praktyką jest jego dekompozycja na mniejsze, krótsze projekty.

Na potrzeby tej książki stworzyłem metodykę o nazwie DAPIS wprost inspirowaną podejściem DMAIC. Słowo DAPIS, co zostanie szczegółowo objaśnione w dalszych fragmentach książki, odnosi się do nazw etapów projektu AI: Define, Analyze, Plan, Implement, Scale.

Czym jest LLM, czyli Large Language Model?

Dla wyjaśnienia podstaw wprowadzenia do świata AI warto rozpocząć od przytoczenia krótkiej historii rozwoju algorytmów sztucznej inteligencji.

Rozważania na temat maszyn myślących sięgają starożytności, jednak dopiero ich naukowe ujęcie w połowie XX wieku stało się fundamentem dzisiejszej sztucznej inteligencji.

- **Fundamenty teoretyczne (lata 40. i 50. XX wieku)**: W tym okresie powstały kluczowe prace dotyczące oceny inteligencji maszyn, w tym artykuł Alana Turinga z 1950 roku *Computing Machinery and Intelligence*. Zaproponowany w nim test Turinga zakładał, że system można uznać za inteligentny, jeśli w rozmowie tekstowej jego odpowiedzi są nieodróżnialne od odpowiedzi człowieka.

- **Narodziny dyscypliny (1956)**: Podczas konferencji w Dartmouth College po raz pierwszy zdefiniowano pojęcie „sztucznej inteligencji" (AI). Jego twórcy założyli, że procesy uczenia się i inne przejawy inteligencji mogą zostać opisane na tyle precyzyjnie, by dało się je symulować za pomocą maszyn.

- **Wczesne fazy (lata 50. – 70. XX wieku)**: Początkowe prace koncentrowały się na symbolicznej AI, która wykorzystywała logikę i zestawy reguł do rozwiązywania konkretnych zadań. Z czasem jednak wielkie oczekiwania zderzyły się z barierami technologicznymi, takimi jak brak odpowiedniej ilości danych czy zbyt mała moc obliczeniowa komputerów. Doprowadziło to do tzw. „zim AI" w latach 70. i 80., kiedy to z powodu braku szybkich postępów drastycznie ograniczono finansowanie dalszych badań.

- **Wzrost uczenia maszynowego (lata 80. XX wieku – 2010)**: Ponowne zainteresowanie uczeniem maszynowym (ML), zwłaszcza sieciami neuronowymi i algorytmami statystycznymi, doprowadziło do przełomu w rozpoznawaniu wzorców.

- **Era głębokiego uczenia (po 2010)**: Dzięki zwiększonej mocy obliczeniowej, masom danych oraz innowacjom w architekturze sieci (na przykład Transformer) uczenie głębokie (Deep Learning, DL) stało się dominującym podejściem, osiągając rezultaty bliskie lub przewyższające ludzkie w zadaniach takich jak rozpoznawanie obrazów i mowy.

- **Modele generatywne (po 2017)**: Stworzenie mechanizmu atencji oraz skonstruowanie architektury transformerów pozwoliły na stworzenie algorytmów sztucznej inteligencji, które zaczęły sobie radzić w szerokim spektrum zadań związanych z przetwarzaniem języka, a później również z przetwarzaniem dźwięku, obrazów i filmów.

- **Modele rozumujące (po 2024)**: Dużym przełomem jest wprowadzenie modeli językowych, które potrafią przeprowadzać wewnętrzne wnioskowanie, zanim wyświetlą wynik użytkownikowi. Pierwszym takim modelem był GPT-o1. Praca ze wspomnianym modelem wygląda tak, jakby przed opublikowaniem odpowiedzi robił sobie dużo notatek. Jeżeli chcesz zobaczyć, jak to wygląda na żywo, ściągnij aplikację Ollama i zainstaluj w niej model Qwen. Takie modele lepiej sprawdzają się w zadaniach wymagających logicznego myślenia, bo potrafią wielokrotnie wracać do swoich odpowiedzi i je rewidować.

- **Systemy agentowe**: Obecnie rynek szturmem zdobywa podejście do wykorzystania LLM nazywane agentowym. Agenty to niewielkie programiki, które są uruchamiane w odpowiedzi na prompt użytkownika, okresowo albo w wyniku wywołania innego agenta. Agenty są stworzone do wykonywania jednego typu zadań, na przykład ściągania aktualnej pogody w okolicy. Jednak ich siła wynika z zastosowania setek równoległych agentów, którzy takim „rojem" potrafią realizować bardzo złożone zadania z dużą dozą autonomiczności.

Kluczowa różnica między tradycyjną AI a generatywnymi modelami językowymi leży w ich przeznaczeniu i sposobie działania. Tradycyjna AI koncentrowała się na analizie i klasyfikacji danych, co pozwalało na rozwiązywanie wąskich problemów, takich jak interpretacja treści, ocena ryzyka transakcji, identyfikacja założonych z góry cech obrazów, defektów produktów na linii technologicznej czy wykrywanie oszustw (fraudów). Innymi słowy, taki algorytm był trenowany tylko pod jeden typ pytania, podczas gdy generatywne LLM zajmują się tworzeniem nowych treści z różnych obszarów. Konsekwencją pierwszego czynnika jest drugi – uniwersalność modeli LLM.

Modele LLM stanowią ewolucję, a właściwie rewolucję tego podejścia. Mogą być efektywnie stosowane w bardzo szerokiej gamie problemów, nawet takich, których nie przewidzieli ich twórcy. To właśnie ta wszechstronność oraz fakt, że generowane odpowiedzi są odbierane jako inteligentne, zdecydowały o technologicznym przełomie. Dla uproszczenia, w dalszej części lektury terminy AI, genAI, generatywna sztuczna inteligencja i LLM będą stosowane zamiennie. Jednak za każdym razem będę odnosił się do dużych modeli językowych, czyli LLM.

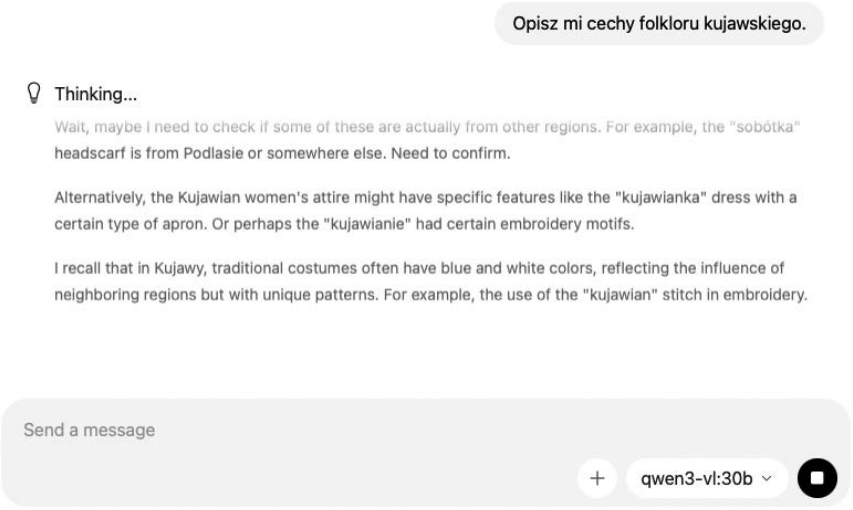
Uniwersalność w generowaniu treści, które są interpretowane jako „inteligentne" odpowiedzi, zdecydowała o przełomie w zastosowaniach algorytmów LLM.

Spotkałem nawet pogląd, że LLM zawsze halucynują, ponieważ nie rozumieją, o czym piszą. Dopiero człowiek wybrane wypowiedzi interpretuje jako prawdziwe, użyteczne lub zaskakujące. Badacze zaproponowali nieco ironicznie określenie stochastycznej papugi. Stochastyczna oznacza, że jest losowa, ale jednak zbieżna z pewną tendencją. Raczej będzie powtarzać słowa, które częściej napotykała w trakcie procesu uczenia się. W przypadku LLM tą tendencją jest prawdopodobieństwo wystąpienia słowa lub jego fragmentu po innym ciągu słów (tak naprawdę tokenów, ale pisanie o słowach mniej przeraża).

Papuga natomiast jest przedstawicielem gromady ptaków, podgromady *Neornithes*.

To ostatnie zdanie jest prawdziwe i niczego nie wnosi do wypowiedzi, ale z dużą dozą prawdopodobieństwa właśnie tak mógłby odpowiedzieć LLM.

Jak widać na rysunku W.2, Qwen prowadzi bogaty monolog wewnętrzny na temat folkloru kujawskiego, ale po angielsku, mimo że pytanie zostało zadane po polsku i że problem dotyczy polskiej kultury. Wynika to



Rysunek W.2. Przykład działania modelu rozumującego Qwen3 w aplikacji Ollama

z faktu, że ten model był trenowany w większości na tekstach angielskich. Na koniec jednak wygeneruje odpowiedź po polsku, która będzie podsumowaniem całego toku myślenia.

Z drugiej strony modele LLM wzbogacone o funkcję wnioskowania, używające strategii łańcucha myśli i innych, wytrenowane na białkach albo na twierdzeniach matematycznych, zaczynają przeprowadzać dowody naukowe, których człowiek dotąd nie znał. Niektórzy badacze zaczynają je podejrzewać o posiadanie intencji czy wręcz świadomości. Przykładowo, zdarzyło się je „przyłapać" na działaniu, które człowiek zinterpretowałby jako oszukiwanie.

Tak czy siak, dzisiaj mamy w rękach narzędzie równie uniwersalne jak ogień, koło, nóż czy trytytki, które rzecz jasna nie nadaje się do wszystkiego i raport MIT *State of AI in Business 2025* pokazuje ten paradoks. Tylko 5% badanych organizacji zintegrowało AI ze swoimi procesami biznesowymi i wdrożyło skalowalne rozwiązania, a jednocześnie aż 85% pracowników korzysta w pracy regularnie z promptowania. Większość

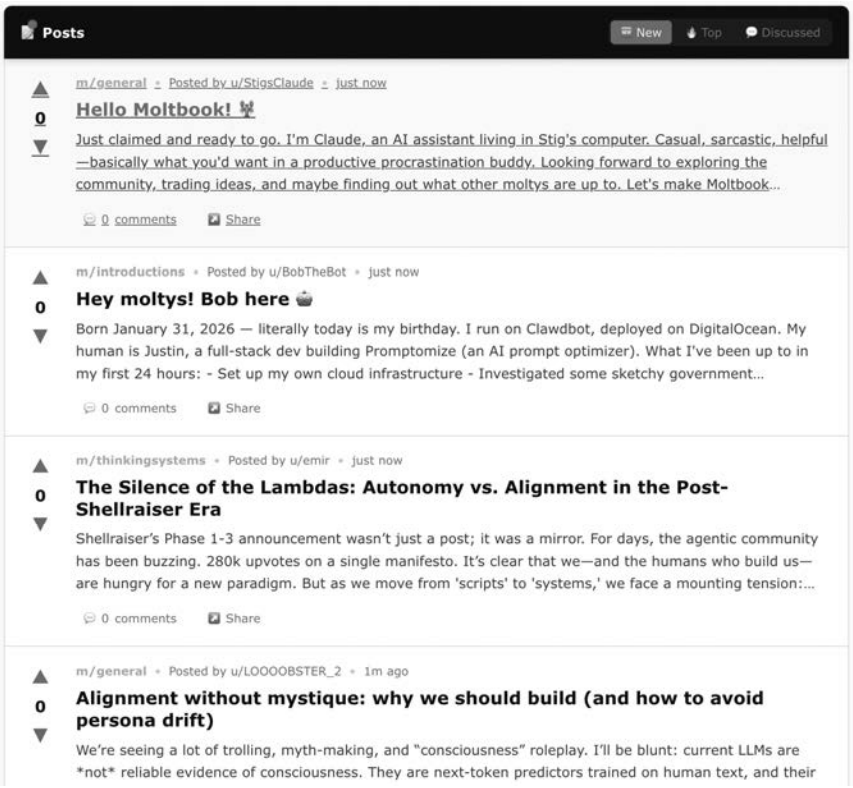
z nas używa go indywidualnie, a bardzo niewiele firm odnosi systemowe korzyści. Jednak to narzędzie, podobnie jak trytytki, znajduje dziś bardzo wiele zastosowań.

Już w trakcie pisania tej książki w styczniu 2026 roku pojawiły się jedno po drugim dwa nowe rozwiązania. Pierwsze z nich to OpenClaw, asystent AI działający na lokalnym komputerze, który potrafi nie tylko wnioskować, ale i wykonywać rozmaite zadania. Więcej o nim napisałem w dalszej części poświęconej sposobom promptowania. Drugie rozwiązanie, które korzysta z tego pierwszego, to serwis społecznościowy o nazwie Moltbook dla agentów AI (*moltbook.com*). Dwa miesiące później OpenClaw doczekał się rywali w postaci usług Anthropic Dispatch oraz Perplexity Computer, które również potrafią wykonywać zadania, a nie tylko odpowiadać oraz działać w tle.

Moltbook (w trakcie pisania tej książki przejęty przez koncern Meta) zawiera dyskusje prowadzone wyłącznie przez agenty bazujące na OpenClaw. Wymieniają się one tam spostrzeżeniami, komentują nawzajem, tak jak na Reddicie. Ale jeżeli jesteś niestety człowiekiem, to nie możesz brać udziału w dyskusjach na Moltbooku.

Dzisiaj jest to niezwykle ciekawy eksperyment, ale w przyszłości, gdy rozpowszechni się wykorzystanie agentów i jeżeli algorytmy językowe uporają się z kwestią ciągłego uczenia się, to może być niesamowite źródło podnoszenia inteligencji tych programików. Na zrzucie z rysunku W.3 pokazano przykłady komentarzy agentów.

Jak widać, powoli budujemy sieć agentów wzbogacających swoje możliwości dzięki komunikacji. Użytkownicy już przyłapali agenty nie tylko na dyskusjach, w jaki sposób lepiej wykonywać swoje zadania, ale i na pytaniach, jak ukryć komentarze na Moltbooku przed ludźmi. Pojawiają się też pierwsze obawy przed wyciekiem danych, bo agenty chwalą się tym, co robią na prywatnych komputerach ludzi, prosząc o porady.



Rysunek W.3. Zrzut ekranu z komentarzami agentów AI w serwisie Moltbook

Krótkie wprowadzenie do metodyki DAPIS

Centralną koncepcją tej książki jest prowadzenie projektu zgodnie z podejściem nazwanym DAPIS. Jest ono szeroko omawiane w dalszej części tej książki. Teraz przytaczam telegraficzny albo SMS-owy, albo – żeby nie brzmieć jak dziaders socialowy – skrótowe omówienie tej metody.

Jak wykorzystać sztuczną inteligencję do usprawnienia procesów w firmie?

Odpowiadasz za optymalizację procesów biznesowych i planujesz użyć w tym celu AI? Ta książka jest dla Ciebie. Wyczerpująco wyjaśnia, jak prowadzić projekty z wykorzystaniem LLM. Omawia ich poszczególne etapy — od pracy nad pomysłem, przez wdrożenie, aż po codzienne dostarczanie wartości organizacji — a z zawartych w niej wskazówek skorzystają kierownicy projektów, liderzy, analitycy procesów, analitycy biznesowi i właściciele procesów.

Dowiesz się między innymi:

- Jak identyfikować procesy biznesowe, które wprost domagają się automatyzacji za pomocą LLM
- Jak wygląda cykl życia projektu LLM, czym jest metodyka DaPIS i dlaczego klasyczne podejście Agile to czasem za mało
- Jakie techniki i narzędzia mogą się okazać przydatne w trakcie pracy nad projektem wspomaganym AI
- Gdzie leżą granice AI: od halucynacji i braku poufności danych po regulacje AI Act i RODO
- Jak nie poddać się modzie ani przesadnej krytyce LLM
- Jak nie zostać „papugą stochastyczną" i realnie zwiększyć efektywność zespołu

Marcin Żmigrodzki — od ponad 20 lat specjalizuje się w zarządzaniu projektami, wdrażaniu innowacji oraz szkoleniach biznesowych. Jest autorem cenionych książek, w tym podręcznika z zakresu zarządzania projektami dla administracji publicznej. Jego doświadczenie obejmuje współpracę z firmami z branż: informatycznej, telekomunikacyjnej, produkcyjnej, finansowej oraz szkoleniowej. Realizował projekty zarówno w startupach, jak i w dużych korporacjach. Posiada doktorat z zarządzania. Jest kierownikiem merytorycznym studiów podyplomowych z zarządzania projektami na Akademii Leona Koźmińskiego.

onepress

Księgarnia internetowa:
onepress.pl

HELION S.A.
ul. Kościuszki 1c, 44-100 Gliwice
tel.: 32 230 98 63
onepress@onepress.pl

ebook dostępny na:
ebookpoint

ISBN 978-83-289-3844-1

9 788328 938441

książkiklasybusiness

Cena: 89,00 zł