

MATRYCA WALIDACJI  
DLA KWALIFIKACJI WOLNORYNKOWEJ  
„WYKORZYSTANIE DUŻYCH MODELI JĘZYKOWYCH”

INSTYTUCJA CERTYFIKUJĄCA:  
POLSKI INSTYTUT KOMPETENCJI PRZYSZŁOŚCI

ul. Ligocka 103, bud. 8, 40-568 Katowice

KRS: 0001097397, NIP: 6343037402

## ZESTAW 1. Podstawy wykorzystywania dużych modeli językowych

### Efekt 1.1 Wyjaśnia pojęcie dużego modelu językowego

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 1.1.1 Definiuje pojęcie dużego modelu językowego (LLM).	Część teoretyczna – test wiedzy; pytanie jednokrotnego wyboru lub dopasowanie pojęcie–definicja.	Arkusz testu wiedzy / system testowy.	„Które z poniższych zdań najlepiej definiuje duży model językowy (LLM)?” (4 odpowiedzi, 1 poprawna).
K 1.1.2 Omawia zasady i zastosowanie dużych modeli językowych.	Test wiedzy; pytania wielokrotnego wyboru, prawda/fałsz, porządkowanie etapów.	Arkusz testu, schemat do uzupełnienia.	Zadanie (test): 1) Uporządkuj etapy działania LLM: „wejście tekstowe -> tokenizacja -> predykcja kolejnych tokenów -> generowanie odpowiedzi”. 2) Wskaż poprawny przykład zastosowania LLM w praktyce (np. podsumowanie dokumentu, klasyfikacja zgłoszeń, generowanie odpowiedzi), i krótko uzasadnij wybór.

### Efekt 1.2 Omawia korzyści płynące ze stosowania dużych modeli językowych w różnych dziedzinach

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 1.2.1 Wskazuje dziedziny, w których duże modele językowe są stosowane do osiągnięcia lepszych wyników.	Test wiedzy; pytanie wielokrotnego wyboru (zaznacz wszystkie poprawne branże).	Arkusz testu / system testowy.	„W których z poniższych obszarów LLM mogą podnieść efektywność pracy? (wybierz wszystkie poprawne odpowiedzi)”.
K 1.2.2 Omawia zastosowanie dużych modeli językowych w praktyce, ilustrując korzyści płynące z ich wykorzystania.	Test wiedzy; pytanie scenariuszowe jednokrotnego wyboru.	Arkusz testu.	Case: opis pracy zespołu obsługi klienta → „Która z odpowiedzi najlepiej opisuje korzyści z zastosowania LLM w opisanej sytuacji?”.

### Efekt 1.3 Opisuje proces trenowania dużych modeli językowych

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 1.3.1 Opisuje pojęcia związane z tworzeniem dużych modeli językowych (np. tokenizacja, uczenie nienadzorowane, uczenie nadzorowane, modele sekwencyjne).	Test wiedzy; dopasowanie pojęcie–definicja, prawda/fałsz.	Arkusz testu.	„Dopasuj pojęcia (tokenizacja, uczenie nadzorowane, uczenie nienadzorowane) do ich opisów.”

K 1.3.2 Opisuje etapy trenowania dużych modeli językowych (np. zbieranie i przygotowywanie danych, wybór architektury modelu, trenowanie modelu, walidacja i testowanie, optymalizacja, wdrażanie, monitorowanie i aktualizacja).	Test wiedzy; zadanie porządkowania etapów, uzupełnianie luk w schemacie.	Arkusze testu ze schematem.	„Ułóż etapy cyklu życia dużego modelu językowego w poprawnej kolejności.”
K 1.3.3 Omawia proces strojenia hiperparametrów dużych modeli językowych.	Test wiedzy; pytania jednokrotnego / wielokrotnego wyboru.	Arkusze testu.	„Które z poniższych przykładów są hiperparametrami dużego modelu językowego?” (np. temperatura, rozmiar warstwy, długość kontekstu).

#### Efekt 1.4 Wykorzystuje gotowe duże modele językowe do generowania tekstu na podstawie zestawu danych

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 1.4.1 Prezentuje sposób instalacji i konfiguracji najczęściej stosowanych bibliotek (np. GPT - Generative Pre-trained Transformer, BART - Bidirectional and Auto-Regressive Transformers) oraz związanych z nimi narzędzi.	Test wiedzy; pytania wielokrotnego wyboru i dopasowanie (model ↔ sposób dostępu).	Arkusze testu; opis modeli dostępnych w IC.	„Dopasuj bibliotekę/narzędzie do jego instalacji i konfiguracji (np. biblioteka do pracy z modelami GPT/BART, narzędzie do uruchomienia modelu open source)”.
K 1.4.2 Przygotowuje dane wejściowe do generowania tekstu w języku naturalnym za pomocą dużych modeli językowych.	Część praktyczna – zadanie praktyczne z obserwacją.	Stanowisko z dostępem do min. 2 modeli (komercyjny + open-source); arkusz obserwacji.	Zadanie: „Przygotuj prompt, który pozwoli wygenerować streszczenie raportu dla zarządu w max. 10 zdaniach, w tonie formalnym.”
K 1.4.3 Uruchamia duży model językowy.	Część praktyczna – zadanie praktyczne z obserwacją; możliwość krótkiego wywiadu.	Środowisko LLM; arkusz oceny zadania.	Kandydat loguje się / uruchamia model, wprowadza przygotowany prompt i zapisuje wygenerowany tekst zgodnie z instrukcją.

## ZESTAW 2. Wykorzystanie dużych modeli językowych w ramach zadania lub domeny

### Efekt 2.1 Wykorzystuje wstępnie trenowane duże modele językowe

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
----------------------	---------------------	---------------------	---------------------------------

K 2.1.1 Przedstawia proces dostosowania dużych modeli językowych do konkretnego zadania lub konkretnej domeny.	Test wiedzy; zadanie porządkowania kroków, wielokrotnego wyboru.	Arkusz testu.	„Ułóż kroki dostosowania LLM do domeny (np. finansowej) w poprawnej kolejności.”
K 2.1.2 Dostosowuje duży model językowy do wskazanego zadania lub wskazanej domeny.	Część praktyczna – zadanie praktyczne z obserwacją.	Środowisko z min. 2 modelami; arkusz oceny zadania.	Zadanie: „Wybierz model i sposób pracy z nim, aby przygotować roboczy raport z danych tekstowych klienta z branży X.”
K 2.1.3 Generuje tekst przy użyciu dużego modelu językowego dostosowanego do tego zadania lub do tej domeny.	Część praktyczna – zadanie praktyczne.	Środowisko LLM; arkusz oceny.	Kandydat generuje tekst (np. opis usługi, streszczenie dokumentu) zgodny z domeną i celem zadania.

## Efekt 2.2 Wykorzystuje techniki inżynierii instrukcji (prompt engineering) do efektywnego zastosowania LLM

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 2.2.1 Wyjaśnia pojęcie inżynierii instrukcji (prompt engineering).	Test wiedzy; pytanie jednokrotnego wyboru + dopasowanie pojęć-opis.	Arkusz testu / system testowy.	„Co oznacza pojęcie 'prompt engineering' w kontekście dużych modeli językowych?”
K 2.2.2 Omawia, jak instrukcja (prompt) wpływa na wyniki generowane przez duże modele językowe.	Test wiedzy; wielokrotnego wyboru, prawda/fałsz.	Arkusz testu.	Stwierdzenia typu: „Im mniej informacji w prompt, tym lepsza odpowiedź” - kandydat ocenia P/F i wskazuje uzasadnienie z listy.
K 2.2.3 Tworzy efektywną instrukcję (prompt).	Część praktyczna - zadanie praktyczne z obserwacją.	Środowisko LLM; arkusz oceny promptu.	Zadanie: „Przygotuj prompt, który wymusi odpowiedź w formie tabeli porównującej 3 rozwiązania, z jasno określonymi kryteriami kolumn.”
K 2.2.4 Analizuje jakość i użyteczność wyników generowanych przez duży model językowy przy użyciu różnych instrukcji (promptów) i strategii.	Test wiedzy; zadanie analityczne (ocena wyników) lub ranking/dopasowanie.	Arkusz testu z przykładami promptów i wyników.	„Porównaj wyniki wygenerowane dla 2-3 różnych promptów/strategii (np. z rolą, z przykładami, z ograniczeniami) i wskaż, który wynik jest najbardziej użyteczny względem kryteriów: adekwatność, kompletność, spójność, format.”
K 2.2.5 Wykorzystuje inżynierię instrukcji (prompt engineering) w kontekście zastosowań dużych modeli językowych (np. analiza tekstu, generowanie treści,	Część praktyczna - zadanie praktyczne z obserwacją; opcjonalnie wywiad swobodny (rozmowa z komisją).	Środowisko LLM; zestaw krótkich materiałów wejściowych (np. tekst/FAQ); arkusz oceny.	Zadanie: „Wybierz kontekst (analiza tekstu / generowanie treści / automatyzacja). Zastosuj prompt engineering (rola, kontekst, format, ograniczenia) i zaprezentuj wynik wraz z krótkim uzasadnieniem przyjętej strategii.”

automatyzacja zadań).			
-----------------------	--	--	--

### Efekt 2.3 Ocenia jakość dużego modelu językowego na podstawie wygenerowanych tekstów

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 2.3.1 Wyjaśnia, jakie metryki automatyczne i manualne są stosowane do oceny jakości wygenerowanych tekstów.	Część teoretyczna – test wiedzy; pytania wielokrotnego wyboru, dopasowanie metryka–opis.	Arkusz testu (tabela metryk, definicje).	„Dopasuj metrykę / kryterium (np. adekwatność, spójność, poprawność) do jej opisu.”
K 2.3.2 Wyjaśnia, jak interpretować wyniki metryk automatycznych i manualnych stosowanych do oceny jakości wygenerowanych tekstów.	Test wiedzy; interpretacja krótkiej tabeli, pytanie jednokrotnego wyboru.	Arkusz testu z przykładową tabelą wyników.	„Który z modeli (A/B/C) daje najlepszy wynik według przedstawionych metryk i dlaczego?” (wybór + uzasadnienie z listy).
K 2.3.3 Stosuje metryki automatyczne i manualne do oceny jakości wygenerowanych tekstów.	Część praktyczna – zadanie praktyczne z obserwacją.	Srodowisko LLM; arkusz oceny tekstu (lista kryteriów) + ewentualnie prosty kalkulator / formularz.	Zadanie: kandydat generuje tekst przy użyciu LLM (np. streszczenie maili), a następnie wypełnia arkusz oceny (np. skala 1–5 dla: adekwatności, kompletności, spójności, stylu), zaznaczając poziom spełnienia każdego kryterium.
K 2.3.4 Interpretuje wyniki zastosowania metryk automatycznych i manualnych do oceny jakości wygenerowanych tekstów.	Część praktyczna – zadanie praktyczne (ciąg dalszy powyższego).	Arkusz oceny z miejscem na krótką konkluzję.	Kandydat formułuje 2–3 wnioski na podstawie własnej oceny: np. „tekst wymaga poprawy w obszarze X, Y; nadaje się / nie nadaje się do użycia bez modyfikacji”.

### Efekt 2.4 Dostosowuje parametry dużych modeli językowych w celu uzyskania lepszych wyników

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 2.4.1 Omawia kluczowe parametry dużego modelu językowego (np. temperatura, wielkość okna kontekstowego).	Test wiedzy; pytania jednokrotnego / wielokrotnego wyboru.	Arkusz testu.	„Które parametry mają największy wpływ na kreatywność odpowiedzi modelu?”
K 2.4.2 Konfiguruje parametry dużych modeli językowych w celu poprawy	Część praktyczna - zadanie praktyczne z obserwacją;	Srodowisko LLM z możliwością zmiany parametrów; arkusz obserwacji/oceny.	Zadanie: kandydat konfiguruje parametry (np. temperatura, długość kontekstu) dla dwóch celów (np. precyzyjne streszczenie vs kreatywne propozycje), generuje

jakości generowanych tekstów.	możliwość krótkiego wywiadu swobodnego.		wyniki, porównuje je i wskazuje, które ustawienia poprawiają jakość (np. zgodność z celem, spójność, zwięzłość).
-------------------------------	---	--	--

### ZESTAW 3. Wykorzystywanie dużych modeli językowych w środowisku pracy

#### Efekt 3.1 Wyjaśnia, w jaki sposób LLM mogą wspomagać proces analizy zbiorów danych tekstowych i formułowanie wniosków

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 3.1.1 Omawia, jak duże modele językowe mogą ułatwić analizę zbiorów danych tekstowych i pozyskiwanie istotnych informacji.	Test wiedzy; pytania wielokrotnego wyboru, dopasowanie zadanie – zastosowanie.	Arkusze testu.	„Dopasuj zadanie (streszczenie, ekstrakcja informacji, analiza sentymentu) do sposobu wykorzystania LLM.”
K 3.1.2 Omawia zastosowanie dużych modeli językowych w analizie zbiorów danych tekstowych (np. w badaniu sentymentu, ekstrakcji informacji, klasyfikacji tekstów).	Test wiedzy; scenariusze + jednokrotny wybór.	Arkusze testu.	„Które zastosowanie LLM jest najbardziej adekwatne w opisanym przypadku (badanie opinii klientów)?”
K 3.1.3 Demonstruje, jak wykorzystać duże modele językowe do analizy zbiorów danych tekstowych i formułowania wniosków.	Część praktyczna – zadanie praktyczne z obserwacją.	Zestaw przykładowych tekstów; środowisko LLM; arkusz oceny.	Zadanie: kandydat z pomocą LLM analizuje zestaw zgłoszeń klientów, identyfikuje główne problemy i formułuje wnioski w arkuszu.

#### Efekt 3.2 Wykorzystuje LLM do automatyzacji zadań związanych z analizą tekstu i przetwarzaniem języka naturalnego

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 3.2.1 Omawia proces integracji dużych modeli językowych z istniejącymi narzędziami w celu automatyzacji zadań związanych z analizą tekstu.	Test wiedzy; pytania wielokrotnego wyboru, dopasowanie „bloczek procesu → opis”.	Arkusze testu.	„Ułóż proste kroki procesu: dane → LLM → weryfikacja → rezultat.”
K 3.2.2 Demonstruje, jak używać dużych modeli językowych do automatyzacji różnych zadań przetwarzania języka naturalnego (np. tłumaczenie, podsumowywanie,	Część praktyczna – zadanie praktyczne z obserwacją.	Środowisko LLM; arkusz oceny zadania.	Zadanie: kandydat przygotowuje z pomocą LLM odpowiedzi na serię prostych zapytań klientów, zapisuje wyniki w wskazanej formie.

generowanie odpowiedzi na pytania).			
-------------------------------------	--	--	--

### Efekt 3.3 Rozpoznaje zastosowania LLM w różnych branżach

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 3.3.1 Opisuje zastosowanie dużych modeli językowych w różnych branżach.	Test wiedzy; zadania dopasowania „branża → zastosowanie”.	Arkusze testu.	„Dopasuj branżę (marketing, obsługa klienta, HR) do przykładowego zastosowania LLM.”
K 3.3.2 Omawia korzyści i zagrożenia wynikające z zastosowania dużych modeli językowych w różnych branżach.	Test wiedzy; scenariusze, pytania wielokrotnego wyboru.	Arkusze testu.	Case: opis wdrożenia LLM w dziale HR → pytanie: „Jakie są główne korzyści i ryzyka w tej sytuacji?” (z listy).

## ZESTAW 4. Etyka i odpowiedzialność w wykorzystywaniu dużych modeli językowych

### Efekt 4.1 Charakteryzuje zagrożenia związane z wykorzystaniem LLM i możliwe działania minimalizujące te zagrożenia

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 4.1.1 Identyfikuje potencjalne zagrożenia związane z wykorzystaniem dużych modeli językowych (np. naruszenie prywatności, nieodpowiednie generowanie treści, dezinformacja, halucynacje).	Test wiedzy; wielokrotnego wyboru, dopasowanie „scenariusz → rodzaj ryzyka”.	Arkusze testu z krótkimi case'ami.	„Jakie ryzyko dominuje w opisanym przypadku użycia LLM?” (np. halucynacje, utrata prywatności).
K 4.1.2 Wskazuje metody minimalizowania zagrożeń związanych z wykorzystaniem dużych modeli językowych.	Test wiedzy; dopasowanie „ryzyko → działanie zapobiegawcze”.	Arkusze testu.	„Które działania najlepiej ograniczą ryzyko halucynacji / dezinformacji w opisywanym zastosowaniu LLM?”

### Efekt 4.2 Charakteryzuje zasady ochrony danych osobowych w kontekście LLM

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 4.2.1 Identyfikuje potencjalne zagrożenia związane z wykorzystaniem dużych modeli językowych w kontekście ochrony danych osobowych.	Test wiedzy; pytania wielokrotnego wyboru i prawda/fałsz.	Arkusze testu.	„Które z poniższych danych NIE powinny być wprowadzane do publicznego LLM?” (lista danych).

K 4.2.2 Opisuje, jak duże modele językowe mogą wpłynąć na prywatność danych osobowych.	Test wiedzy; scenariusz MCQ.	Arkusze testu.	Case: pracownik wkleja fragment dokumentu z danymi osobowymi do LLM → „Jakie przepisy/zasady mogą zostać naruszone?”.
--	------------------------------	----------------	---

#### Efekt 4.3 Ocenia wpływ LLM na bezpieczeństwo danych instytucjonalnych

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 4.3.1 Identyfikuje możliwości naruszeń zasad ochrony danych instytucjonalnych.	Test wiedzy; case study, jedno- i wielokrotny wybór.	Arkusze testu.	„Czy w opisanej sytuacji (wklejenie strategii firmy do LLM) doszło do naruszenia zasad ochrony danych instytucjonalnych? Jakiego rodzaju?”.
K 4.3.2 Wyjaśnia, jak zasady ochrony danych instytucjonalnych wpływają na wykorzystywanie dużych modeli językowych.	Test wiedzy; dopasowanie „zasada bezpieczeństwa → przykład jej zastosowania przy LLM”.	Arkusze testu.	„Dopasuj zasadę (np. minimalizacja danych, ograniczanie dostępu) do opisanego przykładu użycia LLM.”

#### Efekt 4.4 Identyfikuje zagrożenia związane z dyskryminacją i uprzedzeniami możliwe przy wykorzystywaniu LLM

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 4.4.1 Wskazuje, w jaki sposób dane treningowe wpływają na generowanie treści dyskryminujących lub wyrażających uprzedzenia.	Test wiedzy; pytania wielokrotnego wyboru, prawda/fałsz.	Arkusze testu.	„Które stwierdzenia dotyczące wpływu danych treningowych na bias modeli są prawdziwe?”.
K 4.4.2 Wyjaśnia, jak zidentyfikować treści dyskryminujące lub wyrażające uprzedzenia i redukować występowanie tych treści w dużych modelach językowych.	Test wiedzy; analiza krótkich przykładów tekstów (MCQ).	Arkusze testu z przykładami odpowiedzi LLM.	„Czy poniższy fragment tekstu jest neutralny, czy zawiera potencjalnie dyskryminujące elementy? Co można zrobić, aby ograniczyć takie treści?”.

#### Efekt 4.5 Opisuje zasady przestrzegania praw autorskich w kontekście wykorzystywania LLM

Kryteria weryfikacji	Metoda(y) walidacji	Narzędzia walidacji	Pytanie/zadanie/polecenie/dowód
K 4.5.1 Wskazuje, w jaki sposób wykorzystywanie dużych modeli językowych może prowadzić do naruszenia praw autorskich.	Test wiedzy; scenariusze wielokrotnego wyboru.	Arkusze testu.	„W której z poniższych sytuacji istnieje największe ryzyko naruszenia praw autorskich przy korzystaniu z LLM?”.

K 4.5.2 Rozpoznaje sytuacje, w których wynik działania dużego modelu językowego może naruszyć prawa autorskie.	Test wiedzy; dopasowanie „opis sytuacji → ocena ryzyka naruszenia”.	Arkusze testu.	Przykłady: „tekst bardzo podobny do znanego artykułu”, „krótka parafraza” itd. Kandydat wybiera poziom ryzyka.
K 4.5.3 Wskazuje metody zapobiegania naruszeniom praw autorskich podczas wykorzystywania dużych modeli językowych.	Test wiedzy; pytania wielokrotnego wyboru.	Arkusze testu.	„Które praktyki pracy z LLM ograniczają ryzyko naruszeń praw autorskich?” (oznacz wszystkie poprawne odpowiedzi).