Możliwość dofinansowania

AI-101 Generatywna AI w praktyce – bezpieczeństwo i ryzyka (Safe Generative AI Fundamentals) - szkolenie

Numer usługi 2026/07/29/18395/3717227

1 900,00 PLN brutto
1 900,00 PLN netto
135,71 PLN brutto/h
135,71 PLN netto/h
261,33 PLN cena rynkowa

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

4,6 / 5

45 ocen

Kraków
Usługa szkoleniowa
stacjonarna
Zajęcia grupowe
14:00 h
29.09.2026 do 30.09.2026

Informacje podstawowe

Kategoria	Informatyka i telekomunikacja / Bezpieczeństwo IT
Grupa docelowa usługi	Pracownicy IT, kadra zarządzająca, liderzy projektów i procesów, specjaliści biznesowi, a także pracownicy naukowo-dydaktyczni i administracyjni uczelni, którzy w codziennej pracy wykorzystują lub planują wykorzystywać narzędzia klasy Generative AI.
Minimalna liczba uczestników	8
Maksymalna liczba uczestników	16
Data zakończenia rekrutacji	24-09-2026
Forma prowadzenia usługi	stacjonarna
Podstawa uzyskania wpisu do BUR	art. 163 ust. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (t. j. Dz. U. z 2024 r. poz. 1571, z późn. zm.)
Zakres uprawnień	kształcenie w innych formachkształcenia prowadzonych przez uczelnie

Cel

Cel edukacyjny

Celem szkolenia jest przybliżenie mechanizmów działania modeli LLM oraz wykształcenie dobrych praktyk w zakresie ich bezpiecznego stosowania. Program koncentruje się na technicznych i organizacyjnych uwarunkowaniach pracy z generatywną sztuczną inteligencją. Szkolenie ma charakter aplikacyjny i zostało stworzone z myślą o użytkownikach,

którzy wykorzystują na co dzień narzędzia takie jak ChatGPT, Copilot czy Gemini i oczekują od siebie rzetelnego i odpowiedzialnego podejścia do zarządzania infor

Efekty uczenia się oraz kryteria weryfikacji ich osiągnięcia i Metody walidacji

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
Charakteryzuje mechanizmy działania modeli językowych (LLM) oraz różnice między publicznymi a firmowymi/lokalnymi wdrożeniami narzędzi GenAI	opisuje ewolucję sztucznej inteligencji od systemów eksperckich do modeli LLM, opisuje proces uczenia i działania modelu językowego, rozróżnia publiczne, korporacyjne i lokalne wdrożenia narzędzi GenAI pod kątem bezpieczeństwa danych	Test teoretyczny
Identyfikuje zagrożenia związane z wykorzystaniem narzędzi GenAI w kontekście wycieku danych, własności intelektualnej i praw autorskich	wskazuje mechanizmy i przykłady wycieku danych (data leakage) przy korzystaniu z publicznych modeli AI, charakteryzuje zagrożenia dla własności intelektualnej i danych badawczych związane z GenAI, ocenia status prawnoautorski treści generowanych przez AI	Test teoretyczny
Ocenia wiarygodność treści generowanych przez modele językowe i rozpoznaje zjawisko halucynacji oraz uprzedzeń (bias)	opisuje przyczyny występowania halucynacji w odpowiedziach modeli językowych, weryfikuje fakty podane przez model poprzez krzyżowe sprawdzenie źródeł, rozpoznaje przejawy uprzedzeń (bias) w wynikach generowanych przez AI	Test teoretyczny
Projektuje bezpieczne zapytania (prompty) do narzędzi GenAI oraz identyfikuje techniki ataków na modele od strony użytkownika	konstruuje zapytania z zachowaniem zasad anonimizacji danych osobowych i poufnych (PII, RODO), redaguje przykładowe prompty usuwając z nich dane wrażliwe, rozróżnia techniki prompt injection i jailbreakingu oraz ryzyka związane z wtyczkami (plugins) i niestandardowymi GPT	Test teoretyczny
Proponuje założenia polityki bezpiecznego korzystania z GenAI w organizacji zgodne z wymogami regulacyjnymi	opisuje kluczowe elementy polityki korzystania z AI dla działu, uczelni lub firmy, wskazuje wymogi wynikające z nadchodzących regulacji (EU AI Act) istotne dla praktyki biznesowej i edukacyjnej, uzasadnia dobór zasad polityki adekwatnie do specyfiki organizacji	Test teoretyczny

Kwalifikacje

Kompetencje

Usługa prowadzi do nabycia kompetencji.

Warunki uznania kompetencji

Pytanie 1. Czy dokument potwierdzający uzyskanie kompetencji lub wyraźnie z nim powiązane inne dokumenty związane ze wsparciem zawierają opis efektów uczenia się?

TAK

Pytanie 2. Czy dokument lub wyraźnie z nim powiązane inne dokumenty związane ze wsparciem potwierdzają, że walidacja została przeprowadzona w oparciu o zdefiniowane w efektach uczenia się kryteria ich weryfikacji i zgodnie z zaplanowanymi metodami walidacji?

TAK

Pytanie 3. Czy dokument lub wyraźnie z nim powiązane inne dokumenty związane ze wsparciem potwierdzają zastosowanie rozwiązań zapewniających rozdzielenie procesów kształcenia i szkolenia od walidacji?

TAK

Program

1. Anatomia sztucznej inteligencji: krótka historia i ewolucja AI, podstawy działania dużych modeli językowych (LLM), różnice między modelami publicznymi a środowiskami korporacyjnymi i lokalnymi.

2. Krajobraz zagrożeń GenAI: ryzyka wycieku danych, zagrożenia dla własności intelektualnej i danych badawczych oraz aspekty praw autorskich związanych z treściami generowanymi przez AI.

3. Krytyczne myślenie i halucynacje modeli: weryfikacja faktów, krzyżowe sprawdzanie źródeł oraz analiza zjawiska bias.

4. Bezpieczny prompt engineering: konstruowanie bezpiecznych i skutecznych zapytań, anonimizacja danych oraz redagowanie promptów z usunięciem PII i danych poufnych.

5. Ataki na modele z perspektywy użytkownika: prompt injection, jailbreaking, ryzyka związane z wtyczkami i niestandardowymi akcjami.

6. AI governance i polityki organizacyjne: podstawy tworzenia polityk korzystania z AI oraz wprowadzenie do regulacji, w szczególności EU AI Act, w kontekście biznesu i edukacji.

Zajęcia odbywają się w jednej grupie szkoleniowej. Na jednego uczestnika przypada jedno stanowisko komputerowe wraz z niezbędnym sprzętem i oprogramowaniem.

Walidacja będzie realizowana na ostatnich zajęciach w postaci testu teoretycznego.

Zajęcia teoretyczne (wykład interaktywny): ok. 40% czasu szkolenia, tj. ok. 6,4 godziny dydaktycznej. Zajęcia praktyczne (warsztaty i analiza przypadków): ok. 60% czasu szkolenia, tj. ok. 9,6 godziny dydaktycznej.

Harmonogram

Liczba pozycji harmonogramu: 9

Przedmiot / temat	Typ aktywności	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
1 z 9 Anatomia Sztucznej Inteligencji – krótka historia i ewolucja AI, zasady działania modeli językowych (LLM), różnice między publicznymi a korporacyjnymi/lokalnymi wersjami modeli. Wykład interaktywny.	Zajęcia	Tomasz Bartel	29-09-2026	08:00	10:15	02:15
2 z 9 Krajobraz Zagrożeń GenAI – ryzyko wycieku danych (data leakage), zagrożenia dla własności intelektualnej i danych badawczych, prawa autorskie a treści generowane przez AI. Wykład interaktywny.	Zajęcia	Tomasz Bartel	29-09-2026	10:15	12:30	02:15
3 z 9 -	Przerwa	-	29-09-2026	12:30	13:30	01:00

Przedmiot / temat	Typ aktywności	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
4 z 9 Krytyczne myślenie i „Halucynacje” – przyczyny halucynacji modeli, warsztat fact-checkingu i krzyżowej weryfikacji źródeł, zjawisko bias. Wykład + warsztat.	Zajęcia	Tomasz Bartel	29-09-2026	13:30	15:00	01:30
5 z 9 Bezpieczny Prompt Engineering – konstruowanie bezpiecznych zapytań, anonimizacja danych, warsztat redagowania promptów z usunięciem PII i danych poufnych. Wykład + warsztat.	Zajęcia	Tomasz Bartel	30-09-2026	08:00	11:00	03:00
6 z 9 -	Przerwa	-	30-09-2026	11:00	12:00	01:00
7 z 9 Ataki na modele (Perspektywa Użytkownika) – prompt injection i jailbreaking, ryzyka związane z wtyczkami (plugins) i niestandardowymi akcjami (Custom GPTs). Wykład interaktywny.	Zajęcia	Tomasz Bartel	30-09-2026	12:00	13:30	01:30

Przedmiot / temat	Typ aktywności	Prowadzący	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
8 z 9 AI Governance i Polityki Organizacyjne – zasady tworzenia polityki korzystania z AI dla działu/uczelni /firmy, wprowadzenie do regulacji EU AI Act. Wykład interaktywny.	Zajęcia	Tomasz Bartel	30-09-2026	13:30	14:30	01:00
9 z 9 -	Walidacja	-	30-09-2026	14:30	15:00	00:30

Podsumowanie

Rodzaj godzin	Liczba godzin
Suma godzin zegarowych usługi	14:00
w tym suma godzin zajęć	11:30
w tym suma godzin walidacji	00:30
w tym suma przerw	02:00
Suma godzin dydaktycznych bez przerw	16:00

Cennik

Cennik

Rodzaj ceny	Cena
Koszt przypadający na 1 uczestnika brutto	1 900,00 PLN
Podmiot uprawniony do zwolnienia z VAT na podstawie art. 43 ust. 1 ustawy o VAT	
Koszt przypadający na 1 uczestnika netto	1 900,00 PLN
Koszt osobogodziny brutto	135,71 PLN

Koszt osobogodziny netto	135,71 PLN

Liczba godzin usługi

Rodzaj godzin	Liczba godzin
Liczba godzin zegarowych usługi	14:00

Prowadzący

Liczba prowadzących: 1

1 z 1

Tomasz Bartel

Założyciel i Prezes Zarządu Vigil Guard PSA, praktyk cyberbezpieczeństwa z ponad dwudziestoletnim doświadczeniem zdobytym w CrowdStrike, Cisco Systems, IBM i Riverbed Technology. Certyfikowany ekspert Cisco (CCIE). Specjalizuje się w bezpieczeństwie aplikacji opartych na dużych modelach językowych i systemach agentowych — twórca rozwiązań z obszaru AIDR (AI Detection and Response). Prelegent konferencji branżowych w kraju i za granicą.

Informacje dodatkowe

Informacje o materiałach dla uczestników usługi

Uczestnicy otrzymują prezentację szkoleniową w formacie PDF obejmującą całość materiału programowego, zestaw praktycznych wskazówek (checklistę) dotyczących bezpiecznego korzystania z narzędzi GenAI oraz przykładowe scenariusze promptów wykorzystywane podczas warsztatów. Po zakończeniu szkolenia uczestnik otrzymuje dokument potwierdzający ukończenie usługi.

Adres

ul. Władysława Reymonta 23/D-8
30-055 Kraków
woj. małopolskie

Udogodnienia w miejscu realizacji usługi

- Klimatyzacja
- Wi-fi
- Laboratorium komputerowe

Kontakt

Agata Rzepka

E-mail szkolenia@informatyka.agh.edu.pl
Telefon (+48) 12 3283 414