

Program studiów podyplomowych

**Kierunek: Big Data - Analytics & Society**

rok akademicki 2026/2027

LP.	NAZWA MODUŁU/ NAZWA KURSU	LICZBA GODZIN ZAJĘĆ
1	Wprowadzenie do technologii Big Data	6
2	Podstawy programowania (Python, R)	46
3	Metody przetwarzania i analizy danych	20
4	Algorytmy i struktury danych	20
5	Bazy danych	30
6	Machine Learning	18
7	Analiza sieci społecznych	15
8	Wizualizacja danych	18
9	Internetowe narzędzia analityczne	15
10	Badania Marketingowe i społeczne	12
11	Psychologia Big Data	6
12	Prawne i etyczne aspekty gromadzenia i przetwarzania danych	6
13	Seminarium	9
<b>LICZBA GODZIN DYDAKTYCZNYCH 221 h</b>		

1. Wprowadzenie do technologii Big Data (6h)

- Czym jest, a czym nie jest “Big Data”? Filozofia, technologia, analityka.
- Historia procesów informacyjnych: od “small data” do “Big Data”;
- Definicje: Big Words, czyli “nowy” język analityka danych;
- Wybrane techniki gromadzenia, weryfikacji i przetwarzania danych wielkiej skali;
- Rozwój modeli analizy danych “nV” - ewolucja modeli biznesowych.
- Wybrane aspekty rozwiązywania problemów w skali Big Data (algorytm MapReduce, koncepcja BigTable, NoSQL, rozproszone systemy plików HDFS);

- Problemy integracji narzędzi Big Data z oprogramowaniem użytkowym..
- Big Data i Artificial Intelligence. “Nowoczesna” analityka danych: ewolucja czy zmiana?
- O co nam właściwie chodzi, czyli: nauka i biznes w świecie “wielkich zbiorów danych”. Aspekty poznawcze i praktyczne. Wybrane zastosowania, szanse i zagrożenia.

## 2. Podstawy programowania 46h (23h Python/ 23h R)

### Język programowania R

- Wprowadzenie do środowiska R i podstaw składni
- Struktury danych (wektor, macierz, czynnik/factor, lista, ramka danych)
- Podstawowe techniki przetwarzania danych (pakiet tidyverse)
- Wizualizacja danych (pakiet ggplot2)
- Elementy programowania w R (tworzenie funkcji, pętle, instrukcje warunkowe)
- Wprowadzenie do wybranych metod statystycznych

### Język programowania Python

- Wprowadzenie do języka Python
- Składnia języka Python
- Wyrażenia regularne w języku Python
- Klasy, obiekty, moduły
- Skrypty i kontrola przepływu programu
- Rozwiązywanie praktycznych problemów Big Data z wykorzystaniem języka Python, biblioteki Pandas, Numpy
- Wykorzystanie API do stworzenia aplikacji webowej opartej na frameworku Flask
- Przykłady zastosowania sztucznej inteligencji w oparciu o najnowsze biblioteki opensource (chatgpt, gemini, claude)

## 3. Metody przetwarzania i analizy danych (20h)

- Eksploracyjna analiza danych
- Metody regresji (różne algorytmy, ocena jakości modeli)
- Metody klasyfikacji (modele white-box i black-box)
- Wyjaśnialna sztuczna inteligencja (XAI)
- Klasteryzacja i segmentacja danych
- Redukcja wymiarowości

## 4. Algorytmy i struktury danych (20h)

- 
- Podstawowe definicje, schematy zapisów algorytmów
- Podstawowe struktury danych: zmienne, wskaźniki i referencje, rekordy, tablice i macierze.
- Struktury listowe, słownikowe i programistyczne
- Algorytmy iteracyjne i rekurencyjne. Sortowanie: bąbelkowe, przez selekcję i wstawianie, quicksort.
- Wyszukiwanie liniowe i binarne.
- Algorytmy we współczesnym świecie

## 5. Bazy danych (30h)

- Relacyjny model baz danych
- podstawowe elementy modelu relacyjnego, ograniczenia integralnościowe
- zasady tworzenia relacyjnych baz danych, normalizacja

- modelowanie konceptualne baz danych, diagramy ER - zbiory encji, atrybuty, związki, klucze, krotność i obligatoryjność
  - Wprowadzenie do języka SQL:
  - składnia SQL : DML, DDL, DCL, DQL
  - wykorzystanie języka SQL: dodawanie, usuwanie, modyfikowanie, wyświetlanie danych, praca ze schematem bazy danych
  - procedury, funkcje, widoki
  - praktyczne zastosowanie SQL – zapytania analityczne
  - indeksy w bazach danych
  - Bazy transakcyjne a analityczne; definicja hurtowni danych wg. B. Inmona, R. Kimballa, architektura i projektowanie hurtowni danych, proces ETL
  - Rodzaje baz NoSQL, architektura rozproszona
6. Machine Learning (18h)
- Charakterystyka sztucznej inteligencji
  - Metody i zastosowania uczenia maszynowego
  - Sieć neuronowa jako klasyczny przykład SI w zakresie ML
  - Reprezentacja wiedzy i jej konsekwencje
  - Błędy ludzi, które powielają lub których unikają maszyny
  - Dobór metody z zakresu Si i ML do klasy zadania
  - Optymalizacja i walidacja rozwiązań
  - Uczenie głębokie (deep learning)
  - Praca z frameworkami TensorFlow oraz PyTorch
  - Praktyczne zastosowanie sieci neuronowych w analizie Big Data
7. Analiza sieci społecznych (15h)
- Treści kształcenia w ramach realizowanego kursu zgromadzone są w następujących blokach:
  - SNA jako metoda i narzędzie badawcze
  - Analiza sieci społecznych jako narzędzie badawcze w naukach społecznych i marketingowych; znaczenie miar i relacji w badaniach; wykorzystanie SNA do interpretacji struktur i zachowań w sieciach społecznych.
  - Podstawowe pojęcia i elementy sieci
  - Definicja sieci społecznych, aktora sieci społecznej, połączenia i zbioru relacji;
  - różnice między sieciami społecznymi a wirtualnymi sieciami społecznymi; poziomy analizy sieci społecznych (makro, mezo, mikro).
  - Podstawy analizy sieci społecznych
  - Matematyczna teoria grafów, algebra macierzowa w opisie relacji, podstawy statystyki w analizie sieci; miary wykorzystywane w analizie sieci społecznych: centralność, prestiż, pośrednictwo, bliskość, gęstość, kapitał społeczny.
  - Zastosowania SNA w biznesie i analizie społecznej
  - Obserwowanie zmian społeczno-gospodarczych; analiza aktywności użytkowników Internetu; wyszukiwanie kluczowych użytkowników; marketing skierowany do grup docelowych; systemy rekomendacyjne; tworzenie grup projektowych w sieciach organizacyjnych.
  - Oprogramowanie i praktyka analityczna
  - Oprogramowanie badawcze do analizy sieci; pozyskiwanie danych, tworzenie sieci, obliczanie miar, wizualizacja i interpretacja; praktyczne wykorzystanie oprogramowania na przykładach (case study).

- Stosowane metody dydaktyczne w trakcie kursu kształtują się następująco:
- Wykład interaktywny - Prezentacja pojęć, technik i zastosowań analizy sieci społecznych; omówienie różnic między sieciami rzeczywistymi i wirtualnymi.
- Studia przypadków (case study) - przykłady zastosowań SNA w biznesie, analizach społeczno - gospodarczych oraz w aktywności użytkowników Internetu.
- Ćwiczenia praktyczne z oprogramowaniem badawczym tworzenie sieci, wyznaczanie relacji, obliczanie miar, wizualizacja i interpretacja danych.
- Ćwiczenia problemowe identyfikacja kluczowych użytkowników sieci, wyznaczanie grup docelowych, projektowanie systemów rekomendacyjnych, tworzenie grup projektowych.
- Dyskusja dydaktyczna omówienie różnic i podobieństw sieci społecznych i wirtualnych; znaczenie miar w praktyce biznesowej.

#### 8. Wizualizacja danych (18h)

W ramach zajęć z wizualizacji danych (część teoretyczna) student zdobywa wiedzę dotyczącą:

- zasad percepcji wizualnej i ich znaczenia w procesie projektowania wizualizacji,
- teoretycznych podstaw doboru odpowiednich typów wykresów i form prezentacji danych do określonych kategorii informacji,
- reguł projektowania czytelnych, spójnych i efektywnych wizualnie przekazów danych,
- zasad stosowania koloru, kompozycji i hierarchii informacji w wizualizacjach,
- standardów i zasad etycznych w prezentacji danych, w tym identyfikacji i unikania najczęstszych błędów i manipulacji wizualnych,
- podstaw tworzenia narracji opartej na danych (data storytelling) oraz sposobów wspierania procesu decyzyjnego poprzez odpowiednią prezentację wyników,
- wymogów dostępności oraz dobrych praktyk projektowych zapewniających zrozumiałość wizualizacji dla różnych grup odbiorców,
- przeglądu i omówienia wybranych narzędzi wspierających proces wizualizacji danych, takich jak Power BI, MicroStrategy oraz Tableau, wraz z ich praktycznymi zastosowaniami w organizacjach.

Zajęcia praktyczne koncentrują się na rozwijaniu umiejętności technicznych i projektowych niezbędnych do efektywnego tworzenia wizualizacji danych oraz interaktywnych raportów wspierających proces decyzyjny. W trakcie zajęć studenci:

- przygotowują dane do wizualizacji, obejmując procesy importowania, czyszczenia i transformowania danych w celu zapewnienia ich jakości oraz gotowości do dalszej analizy,
- projektują i tworzą interaktywne dashboardy, umożliwiające użytkownikom eksplorowanie danych, analizowanie zależności oraz odkrywanie ukrytych wzorców,
- wykorzystują narzędzia Business Intelligence, takie jak Power BI, Tableau do budowania zaawansowanych wizualizacji, analizy trendów, identyfikacji kluczowych wskaźników efektywności (KPI) oraz generowania raportów biznesowych,
- analizują rzeczywiste przypadki biznesowe, aby zrozumieć, w jaki sposób organizacje wykorzystują wizualizacje danych do usprawniania procesów operacyjnych i podejmowania decyzji strategicznych,

## 9. Internetowe narzędzia analityczne (15h)

- Narzędzia służące do: monitoringu mediów, zarządzania social media, influencer marketingu, badania zaangażowania użytkowników, content marketingu itp.
- Analityka internetowa: wprowadzenie, konfiguracja narzędzi analitycznych, budowa kompetencji w zespole, tworzenie ciągłych procesów analizy danych
- Przegląd wybranych narzędzi służących do: monitoringu mediów, zarządzania social media, influencer marketingu, badania zaangażowania użytkowników, content marketingu.
- Analityka webowa: dane o ruchu na stronie, analiza jego źródeł oraz zachowania użytkowników, analiza konwersji (Google Analytics)
- Narzędzia SEO: analityka widoczności strony, ruchu organicznego, słów kluczowych oraz backlinków
- Analityka mediów społecznościowych i social listening: monitorowanie i analizowanie aktywności w mediach społecznościowych, raportowanie, analiza sentymentu, trendów, reputacji marki (Sotrender i Brand24)

## 10. Badania Marketingowe i społeczne (12h)

Treści kształcenia w ramach realizowanego kursu zgromadzone są w następujących blokach:

- Wprowadzenie do badań marketingowych i społecznych
- Zakres i funkcje badań w procesie decyzyjnym; badania eksploracyjne, opisowe, przyczynowe; badania ilościowe vs jakościowe; badania satysfakcji, opinii, wizerunku, UX; różnice między badaniami rynkowymi a społecznymi.
- Projektowanie procesu badawczego
- Definiowanie problemu i celów badawczych; formułowanie pytań badawczych i hipotez; dobór podejścia metodologicznego; etyka badań i zasady RODO w badaniach; planowanie próby, harmonogramu i kosztów.
- Metody i narzędzia badań ilościowych
- Projektowanie kwestionariusza (rodzaje pytań, zasady konstrukcji, skale pomiarowe); metody przeprowadzania badań (CAWI, CATI, CAPI, PAPI); techniki doboru próby (probabilistyczne i nieprobabilistyczne); pilotaż i walidacja narzędzi; podstawy przygotowania danych do analizy (czyszczenie, kodowanie).
- Metody i narzędzia badań jakościowych
- Wywiady indywidualne, grupowe (fokusy), obserwacje; konstrukcja scenariusza badawczego; techniki moderowania i sondowania; rekrutacja uczestników; metody analizy jakościowej (kodowanie, kategorie tematyczne, triangulacja); wprowadzenie do metod UX (testy użyteczności, eye-tracking jako technika wspierająca).
- Analiza wyników, wnioski i raportowanie
- Zasady interpretacji danych ilościowych i jakościowych; łączenie metod (triangulacja, insighty komplementarne); struktura raportu badawczego; zasady prezentacji wyników (Executive Summary, rekomendacje, ograniczenia badań); dobre praktyki wizualizacji danych i przedstawiania wniosków interesariuszom.
- Błędy i ryzyka w badaniach
- Błędy projektowe (nieprecyzyjne cele, niewłaściwa próba, brak pilotażu), błędy pomiaru (błędy ankiety, influencja moderatora), błędy analityczne i interpretacyjne; checklisty kontrolne jakości procesu badawczego.
- Stosowane metody dydaktyczne w trakcie kursu kształtują się następująco:
- Wykład problemowy i interaktywny (omówienie zasad metodologii, analiza

- przypadków).
- Studia przypadków (case studies) - interpretacja rzeczywistych projektów badawczych.
  - Warsztaty projektowe — formułowanie pytań i hipotez, projektowanie narzędzi (kwestionariusz, scenariusz IDI/fokusa).
  - Ćwiczenia praktyczne — krytyka narzędzi badawczych, pilotaż, kodowanie danych jakościowych, analiza fragmentów wyników.
  - Dyskusja moderowana — ocenianie decyzji metodologicznych, interpretacja wniosków.
  - Praca projektowa / mini-projekt badawczy — opracowanie skróconego planu badania oraz szkicu raportu z wynikami.

#### 11. Psychologia Big Data (6h)

- Jak nasz dominujący styl komunikacji determinuje przekaz i odbiór informacji?
- Psychologiczny model OCEAN i jego wykorzystanie w tworzeniu psychograficznych profili konsumentów/wyborców w powiązaniu z danymi z FB.
- Przewidywanie cech i atrybutów na podstawie cyfrowej rejestracji ludzkich zachowań. Rozwój segmentacji psychograficznej.
- Wybrane badania wykorzystujące Big Data i AI w analizie ludzkich zachowań i ich konfrontacja z dotychczasowymi ustaleniami psychologii.
- Laboratorium psychologiczne oparte na Big Data: The Psychometric Centre, University of Cambridge
- Multitasking i efektywność w zakresie przetwarzania i zapamiętywania informacji.
- Duże zbiory danych a procesy poznawcze i emocjonalne - podsumowanie.

#### 12. Prawne i etyczne aspekty gromadzenia i przetwarzania danych (6h)

- Big Data a prawo do prywatności
- Ustawa o ochronie danych osobowych versus praktyka przetwarzania danych
- Tajemnice sektorowe
- Implikacje etyczne gromadzenia i przetwarzania informacji
- Etyczne aspekty sztucznej inteligencji oraz aktualne zmiany prawne w tym zakresie
- Zastosowanie Big Data w cyberbezpieczeństwie, detekcji anomalii i ochronie danych

#### Seminarium (9h)

Seminarium i towarzysząca tej formule metoda pracy prowadzi do finalizacji studiów podyplomowych, które kończą się egzaminem dyplomowym (oceniana jest wykonana praca i wynik egzaminu). To jedyna, a zarazem najważniejsza część studiów, związana z uzyskaniem w efekcie wymiernej, całościowej ich oceny (skala pozytywna od 3,0 / dostateczny do 5,5 / celujący - finalizacja studiów z wyróżnieniem). Dokumentem potwierdzającym jest dyplom ukończenia studiów podyplomowych.

#### Model seminarium uwzględnia:

- Cel poznawczy: Podstawy prowadzenia badań naukowych i projektów biznesowych z użyciem narzędzi Big Data. Seminarium weryfikuje jakość przebytej w trakcie studiów "podróży": umiejętność łączenia myślenia analitycznego, tj. wiedzy i znajomości narzędzi analizy danych z kompetencjami praktycznymi ich realnego zastosowania.
- Cel użytkarny: Indywidualna lub zespołowa realizacja przedsięwzięcia (projektu

dyplomowego) obejmującego propozycję rozwiązania wybranego problemu praktycznego (biznes, nauka, relacje społeczne) metodami adekwatnymi dla zakresu przedmiotowego owego problemu, wykorzystując poznane w ramach studiów techniki i narzędzia analizy danych.

Dokumentacja: Praca dyplomowa (dokument nazywany potocznie projektem), zawierający zweryfikowaną analitycznie propozycję rozwiązania konkretnego problemu praktycznego (wskazujący produkt końcowy: co nim jest?).

Sugerowana forma: struktura artykułu naukowego. Nie chodzi zatem o kolejną w życiu pracę pisaną do archiwum, ale o jej sens. Kluczem nie jest jej objętość (ilość stron), ale spójność merytoryczna. Chodzi o zwięzłe przedstawienie realnego problemu, poprzez określenie celu, źródeł danych, metody i sposobu uzyskania odpowiedzi na stawiane przez analityka pytania, a także zestawienie wyników i wskazanie odpowiedzi i ich ocenę. Ujawnia się wówczas realna: poznawcza i praktyczna wartość pracy. Potwierdzają ją: samodzielne wnioskowanie i stanowiące konsekwencję sformułowanych wniosków - rekomendacje autorskie.

Przygoda w ramach seminarium pozwala sprawdzić się w praktyce: weryfikując zdobytą wiedzę i poznane narzędzia Big Data z konkretnym problemem, który wymaga faktycznego rozwiązania. Jeśli świadomie się ową aktywność podejmie - rozwija pasję i zainteresowania, motywuje do dalszego działania. W konsekwencji inspiruje do kolejnych wyzwań w przestrzeni informacji, tworzonej przez wciąż nowe, aktualizowane, a nawet nieograniczone zbiory danych. Serdecznie zapraszamy!

Czas trwania: 2 semestry, 12 zjazdów, 221 godzin