

## Opis przedmiotów kierunku Zarządzanie Big Data

### Program studiów

**Zarządzanie Big Data** - przekazanie wiedzy obejmującej przedmiot studiów w zakresie zarządzania Big Data (BD). Podstawowe własności BD - metody wykorzystania w naukach społecznych i ekonomicznych.

**Propedeutyka technologii informacyjnych** - podstawy informatyki i internetu w zakresie związanym z BD

**Podstawy wizualizacji danych** - przedmiot zapoznaje z metodami wizualizacji danych, szczególnie danych statystycznych. Studenci zapoznają się z graficznymi interpretacjami zjawisk i procesów statystycznych a szczególnie z doбором odpowiedniej grafiki do opisywanego modelu opisywanego wielkościami i zależnościami statystycznymi.

**Organizacja informacji i wiedzy** - wybrane teorie, koncepcje, metody organizowania wiedzy w sztucznych (projektowanych) systemach informacyjnych. Między innymi: pojęcie informacji i wiedzy; kategorie wiedzy i jej własności; organizacja informacji i wiedzy a reprezentacja informacji i wiedzy; organizacja i reprezentacja wiedzy w systemach inteligentnych, organizacja zbiorów wyszukiwawczych a modele i metody wyszukiwania.

**Systemy operacyjne z podstawami programowania** - przedmiot wykładu dotyczy tematyki współczesnych systemów operacyjnych, ich roli i zasad działania. Omawiane są elementy składowe systemów operacyjnych, w szczególności : budowa jądra systemu, procesy, wątki, systemy plików wraz z organizacją fizyczną i logiczną, zarządzanie pamięcią, uprawnienia, pliki konfiguracyjne, usługi sieciowe, interfejsy wejścia wyjścia. Tematyka wykładu dotyczy także administrowania systemami operacyjnym z uwzględnieniem aspektów bezpieczeństwa i zapewnieniem ciągłości działania systemu. Omawiane są również praktyczne aspekty zarządzania systemem operacyjnym, powłoki, skrypty, programowanie systemowe. W ramach wykładu omawiane są podstawowe narzędzia systemów operacyjnych niezbędne do przetwarzania dużych zbiorów danych Big Data, w tym podstawy programowania w powłokach systemowych.

**Statystyka** - celem zajęć jest zaznajomienie studentów metodami analizy danych przy pomocy aparatu statystycznego. Przedstawione zostaną podstawowe miary statystyczne służące do analizy danych. Omówione zostaną metody badania danych przy zastosowaniu algorytmów odkrywania asocjacji, wzorców sekwencji, klasyfikacji oraz grupowania. Przeanalizowane zostaną podstawowe pojęcia statystyczne, rozkłady zmiennych, hipotezy, testy, badania, obróbka danych, prezentacje.

**Ilościowa ocena jakości informacji** - celem zajęć jest zapoznanie studentów z:

- metodami oceny jakości informacji dostępnej w formie elektronicznej.
- dziedzinowymi źródłami informacji.
- gromadzeniem danych drogą białego wywiadu (OSINT).
- analityką internetową
- narzędziami do analizy i wizualizacji informacji.

**Podstawy programowania R** - przedmiot obejmuje podstawy implementacji algorytmów w języku programowania R. Zajęcia mają charakter praktyczny. Podczas ich trwania zostaną omówione zagadnienia dotyczące: składni języka, struktur danych, instrukcji sterujących, procedur, funkcji, operacji wejścia wyjścia, zarządzania pamięcią, kontroli błędów i ich obsługi. Dodatkowo przedmiotem zajęć jest zapoznanie studentów z programowaniem strukturalnym i obiektowym (abstrakcja, enkapsulacja, dziedziczenie, polimorfizm). Ponadto w ramach ćwiczeń zostaną omówione narzędzia programistyczne do tworzenia oprogramowania począwszy od edytora tekstowego po zaawansowane oprogramowanie do testowania, debugowania, wersjonowania i tworzenia dokumentacji. Każdy ze studentów w ramach ćwiczeń stworzy kilka programów realizujących wskazane przez prowadzącego zadania.

**Rafinacja Informacji** - celem przedmiotu jest zapoznanie studentów z ideą rafinacji informacji oraz z praktycznymi aplikacjami tej metody.

**Semantyczne bazy wiedzy** - celem przedmiotu jest z zaznajomienie studentów z założeniami Sieci Semantycznej oraz konstrukcją semantycznych baz wiedzy wykorzystujących ontologie jako narzędzia reprezentacji wiedzy. W ramach kursu zostaną omówione założenia grafowego modelu reprezentacji wiedzy oraz metodyka tworzenia baz wiedzy z wykorzystaniem koncepcji Linked Data. Treści programowe obejmują również wyszukiwanie i pobieranie informacji z semantycznych baz wiedzy z wykorzystaniem języka SPARQL.

**Identyfikacja cyfrowych zasobów informacji online i offline** - celem przedmiotu jest zaznajomienie studentów z wiedzą z zakresu identyfikacji źródeł informacji oraz ich klasyfikacji z punktu widzenia dokumentów cyfrowych oraz tradycyjnych. Pod uwagę zostaną wzięte źródła licencjonowane oraz otwarte. W trakcie zajęć studenci poznają także możliwości dostępu do źródeł ze szczególnym uwzględnieniem zasobów naukowych, specjalistycznych, branżowych i społecznościowych.

**Wykorzystywanie pakietów statystycznych Gretl i Statistica** - studenci zdobędą umiejętność posługiwania się pakietem oprogramowania Gretl i Statistica. Szczególna uwaga zostanie poświęcona analizie danych tekstowych.

**Machine learning** - w ramach przedmiotu studenci zapoznają się z praktycznym zastosowaniem machine learningu rafinacji informacji z dużych zasobów danych. Szczególna uwaga zostanie poświęcona klasyfikacji danych z wykorzystaniem algorytmów machine learningu.

**Elementy przetwarzania wsadowego zbiorów Big Data** - przedmiot, będący zajęciami laboratoryjnymi, obejmuje zagadnienia praktycznego przetwarzania dużych zbiorów danych, w tym:

- wsadowe przetwarzanie danych z wykorzystaniem systemu plików, brokera informacji i baz danych,
- problematykę replikacji, podziału na partycje, transakcje i przetwarzanie strumieniowe,
- rozproszone systemy plikowe i baz danych,
- przechowywanie, kolekcjonowanie i kodowanie danych wraz z realizacją ich przepływu,
- wsadową konwersję różnych formatów danych wejściowych na postać znormalizowaną, będącą surowcem dla procesów rafinacji danych,
- rafinacja dużych zbiorów danych w tym model MapReduce.

**Infometria, webometria, bibliometria** - celem przedmiotu jest zaznajomienie studentów z problematyką wykorzystania metod ilościowych w badaniach dotyczących obiegu informacji w nauce (bibliometria, infometria) oraz w środowisku sieciowym (webometria). Studenci zapoznają się z podstawowymi miarami oraz technikami ich obliczania jak również narzędziami wykorzystywanymi do prowadzenia tego typu badań (gromadzenie, przetwarzanie i analiza).

**Systemy baz danych** - celem przedmiotu jest zapoznanie studentów z problematyką baz danych i projektowaniem schematów baz danych w podejściu strukturalnym i obiektowym. W ramach zajęć studenci poznają narzędzia służące do tworzenia modeli baz danych, optymalizacji schematów, tworzenia podstawowych struktur danych i wykonywania operacji na danych. Zakres przedmiotu obejmuje: projektowanie baz danych - podejście strukturalne i obiektowe, normalizację danych, język SQL, tworzenie i modyfikację struktur danych, operacje na danych, zapytania bazodanowe, metody optymalizacji zapytań, przetwarzanie transakcji, programowanie baz danych, procedury i funkcje w bazie danych.

**Dokumentacja Rafinacji** - zajęcia mają na celu zaznajomienie z metodologią rafinacji danych. Omówione zostaną poszczególne etapy analizy danych: import danych, czyszczenie, modelowanie, wizualizacja oraz prezentacja wyników. Szczegółowo omówiona zostanie prezentacja wyników. Przedstawione będą najlepsze praktyki komunikacji z odbiorcą; informatyka techniczna i telekomunikacja.

**Proces zasilania informacyjnego baz danych** - proces pobierania, czyszczenia i zapisywania materiałów źródłowych w bazie BD do uniwersalnego przedmiotu badań.

**Zarządzanie metadanymi w bazach danych** - celem zajęć jest zapoznanie studentów z:

- pojęciem, funkcjami i typologią metadanych,
- różnymi standardami metadanych stosowanymi w systemach zarządzania treścią,
- standardami kodowania metadanych w systemach zarządzania treścią.
- wymianą metadanych między systemami zarządzania informacją.

**Spoleczne aspekty Big Data** - coraz więcej elementów rzeczywistości i doświadczenia człowieka jest kwantyfikowanych, rejestrowanych i znajduje swój ekwiwalent w potoku danych. Ich namnażaniu i samoczynnemu wychwytywaniu sprzyja naszpikowanie rzeczywistości zespolonymi z internetem czujnikami i miernikami oraz usieciowienie i wirtualizacja świata. Współczesna analityka (biznesowa, społeczna, kulturowa) opiera się na przetwarzaniu owych danych przez złożone algorytmy. Procesy te mają być kluczem do znajdowania wzorów oraz prognozowania stanów rzeczywistości i ludzkich zachowań. Fenomen pozyskiwania, gromadzenia i przetwarzania wielkich zbiorów danych w celu pozyskania z nich wiedzy nosi nazwę Big Data. W trakcie zajęć dyskutowane będą społeczne aspekty Big Data, w szczególności wpływ tego zjawiska na wolność, bezpieczeństwo i dystrybucję władzy.

**Systemy informatyczne zarządzania** - celem przedmiotu jest zapoznanie studentów z funkcjonalnością Systemów Informatycznych Zarządzania (SIZ) i kluczowymi procesami w działalności przedsiębiorstw wspieranymi przez narzędzia informatyczne. Studenci w ramach zajęć poznają przykładowy zintegrowany system do zarządzania przedsiębiorstwem. Zakres przedmiotu obejmuje: rodzaje systemów informatycznych wspierających działalność przedsiębiorstw, korzyści dla przedsiębiorstw wynikające z zastosowania SIZ, generacje

zintegrowanych systemów informatycznych zarządzania, analizę funkcjonalną narzędzia klasy ERP, zagadnienia BI – wsparcie procesów podejmowania decyzji.

**Ekonomia informacji** - przedmiot ma zapoznać studenta z ekonomicznymi aspektami prowadzenia działalności w zakresie szeroko pojętej informacji. W szczególności na identyfikacji relacji między informacją a zjawiskami ekonomicznymi takimi jak konsumpcja, popyt, skłonność do oszczędzania i inwestycje. Przedmiot uwzględnia również zasady finansowania i organizacji działalności informacyjnej w warunkach gospodarki rynkowej, elementy ekonomii behawioralnej.

**Bezpieczeństwo teleinformatyczne** - świadomość zagrożeń czekających w codziennej pracy z narzędziami informatycznymi. Studenci dowiedzą się z zajęć w jaki sposób cyberprzestępcy mogą pozyskać wrażliwe dane firmy lub przejąć kontrolę nad komputerem (np. żądając później okupu) oraz jak się przed nimi chronić. Dokumentacja związana z bezpieczeństwem

**CSR – raportowanie niefinansowe** - celem zajęć jest omówienie pojęcia Społecznej Odpowiedzialności Biznesu, analiza działań firm społecznie odpowiedzialnych ze wskazaniem dobrze zaplanowanej strategii CSR

**Zintegrowane systemy biblioteczne i informacji archiwalnej** - celem przedmiotu jest zaznajomienie studentów z narzędziami umożliwiającymi organizację i udostępnianie informacji o dokumentach gromadzonych w bibliotekach oraz archiwach. Szczególna uwaga zostanie skierowana w kierunku zintegrowanych systemów bibliotecznych, systemów informacji archiwalnej, systemów typu discovery i delivery, link solverów, pełnotekstowych baz bibliograficznych i innych baz danych, bibliotek i repozytoriów cyfrowych itp.

**Praktyki** - praktyki mają za zadanie poszerzenie wiedzy zdobytej na studiach i rozwijanie umiejętności jej wykorzystania, zapoznanie studenta ze specyfiką środowiska zawodowego, kształtowanie umiejętności zawodowych związanych z miejscem odbywania praktyki, kształtowanie umiejętności skutecznego komunikowania się w organizacji i poznanie zasad jej funkcjonowania, a także doskonalenie umiejętności organizacji pracy własnej i zespołowej, efektywnego zarządzania czasem, sumienności, odpowiedzialności za powierzone zadania, doskonalenie umiejętności posługiwania się językiem obcym w sytuacjach zawodowych.