

## KARTA KURSU (realizowanego w specjalności)

### Data Science

(nazwa specjalności)

|                 |                                    |
|-----------------|------------------------------------|
| Nazwa           | <b>Metody zbierania informacji</b> |
| Nazwa w j. ang. | <i>Methods of Data Collection</i>  |

|                 |                                             |                                             |
|-----------------|---------------------------------------------|---------------------------------------------|
| Koordinator     | dr Roman Czapla                             | Zespół dydaktyczny                          |
|                 |                                             | mgr Katarzyna Marczak<br>mgr Patryk Mazurek |
| Punktacja ECTS* | st. stacjonarne: 3<br>st. niestacjonarne: 3 |                                             |

#### Opis kursu (cele kształcenia)

Celem kursu jest zapoznanie studentów z metodami pozyskiwania danych z różnych źródeł – zarówno automatycznych, jak i opartych na interakcji z użytkownikami. Studenci uczą się wykorzystywać narzędzia i biblioteki języka Python do web scrapingu i pobierania danych z API, a także poznają zasady projektowania formularzy i ankiet. Kurs porusza także kwestie jakości danych, ich organizacji oraz etycznych i prawnych aspektów ich pozyskiwania.

Zajęcia rozwijają praktyczne umiejętności w zakresie automatyzacji pobierania informacji, przetwarzania ich do formatu analitycznego oraz odpowiedzialnego zarządzania danymi.

#### Warunki wstępne

|              |                                                                                                                                                                                                        |
|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Wiedza       | Student zna podstawy języka Python oraz orientuje się w podstawowych pojęciach związanych z analizą danych i typami danych.                                                                            |
| Umiejętności | Student potrafi uruchamiać skrypty w Pythonie, pracować z prostymi strukturami danych (np. listy, słowniki) oraz korzystać z podstawowych narzędzi do pracy z plikami i środowiskiem programistycznym. |
| Kursy        | Wymagane zaliczenie kursu: Podstawy Data Science                                                                                                                                                       |

#### Efekty uczenia się

|        | Efekt uczenia się dla kursu                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Odniesienie do efektów kierunkowych |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Wiedza | Po zakończeniu kursu student:<br><b>W01:</b> Student zna podstawowe źródła danych i metody ich pozyskiwania, w tym web scraping, API oraz formularze.<br><b>W02:</b> Student zna strukturę stron internetowych oraz zasady działania narzędzi do automatycznego pobierania danych.<br><b>W03:</b> Student zna podstawy prawne i etyczne związane z pozyskiwaniem i wykorzystywaniem danych.<br><b>W04:</b> Student zna zasady oceny jakości danych oraz metody ich przechowywania i organizacji.<br><b>W05:</b> Student zna możliwości techniczne zapisu i automatyzacji procesu zbierania danych. |                                     |

|              | Efekt uczenia się dla kursu                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Odniesienie do efektów kierunkowych |
|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Umiejętności | <p>Po zakończeniu kursu:</p> <p><b>U01:</b> Student potrafi pozyskiwać dane z internetu za pomocą bibliotek Python (requests, BeautifulSoup, Selenium).</p> <p><b>U02:</b> Student potrafi korzystać z API do pobierania danych i przetwarzać je w formatach JSON/CSV.</p> <p><b>U03:</b> Student potrafi zaprojektować prosty formularz lub ankietę i przygotować zebrane dane do analizy.</p> <p><b>U04:</b> Student potrafi ocenić kompletność i spójność danych oraz stosować podstawowe metody walidacji.</p> <p><b>U05:</b> Student potrafi organizować i przechowywać dane w uporządkowanej formie oraz zadbać o ich automatyczne pobieranie i backup.</p> |                                     |

|                       | Efekt uczenia się dla kursu                                                                                                                                                                                                                                                                   | Odniesienie do efektów kierunkowych |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Kompetencje społeczne | <p>Po zakończeniu kursu:</p> <p><b>K01:</b> Student potrafi odpowiedzialnie podejmować decyzje dotyczące pozyskiwania danych, z uwzględnieniem aspektów prawnych i etycznych.</p> <p><b>K02:</b> Student wykazuje ostrożność i krytycyzm wobec jakości źródeł danych i ich wiarygodności.</p> |                                     |

### Studia stacjonarne

| Organizacja   |            |                     |   |    |   |   |   |
|---------------|------------|---------------------|---|----|---|---|---|
| Forma zajęć   | Wykład (W) | Ćwiczenia w grupach |   |    |   |   |   |
|               |            | A                   | K | L  | S | P | E |
| Liczba godzin | 20         |                     |   | 20 |   |   |   |

### Studia niestacjonarne

| Organizacja   |            |                     |   |    |   |   |   |
|---------------|------------|---------------------|---|----|---|---|---|
| Forma zajęć   | Wykład (W) | Ćwiczenia w grupach |   |    |   |   |   |
|               |            | A                   | K | L  | S | P | E |
| Liczba godzin | 10         |                     |   | 15 |   |   |   |

### Opis metod prowadzenia zajęć

Zajęcia prowadzone są w formie wykładów oraz ćwiczeń praktycznych. W trakcie wykładów omawiane są metody pozyskiwania danych, ich klasyfikacja, aspekty prawne i etyczne, a także zagadnienia związane z jakością i organizacją danych. Ćwiczenia mają charakter warsztatowy i obejmują m.in. pozyskiwanie danych z Internetu za pomocą narzędzi Python (web scraping, API), projektowanie formularzy oraz automatyzację pobierania i zapisu danych.

Studenci realizują projekt zaliczeniowy, w którym praktycznie stosują wybrane metody zbierania informacji, dokumentując proces i efekty swojej pracy.

## Formy sprawdzania efektów uczenia się

|     | E – learning | Gry dydaktyczne | Ćwiczenia w szkole | Zajęcia terenowe | Praca laboratoryjna | Projekt indywidualny | Projekt grupowy | Udział w dyskusji | Referat | Praca pisemna (esej) | Egzamin ustny | Egzamin pisemny | Zadania problemowe |
|-----|--------------|-----------------|--------------------|------------------|---------------------|----------------------|-----------------|-------------------|---------|----------------------|---------------|-----------------|--------------------|
| W01 |              |                 |                    |                  |                     |                      |                 |                   |         |                      |               |                 |                    |
| W02 |              |                 |                    |                  |                     |                      |                 |                   |         |                      |               |                 |                    |
| U01 |              |                 |                    |                  |                     |                      |                 |                   |         |                      |               |                 |                    |
| U02 |              |                 |                    |                  |                     |                      |                 |                   |         |                      |               |                 |                    |
| K01 |              |                 |                    |                  |                     |                      |                 |                   |         |                      |               |                 |                    |

|                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kryteria oceny | <p>Osiągnięcie efektów kształcenia podanych powyżej uprawnia studentów do uzyskania oceny nie wyższej niż dostateczna.</p> <p>Warunkiem zaliczenia przedmiotu jest:</p> <ul style="list-style-type: none"> <li>• zaliczenie kolokwium sprawdzającego znajomość metod i narzędzi pozyskiwania danych,</li> <li>• przygotowanie i oddanie projektu zaliczeniowego polegającego na samodzielnym pozyskaniu danych z wybranego źródła (np. scraping, API, formularz), ich przetworzeniu i dokumentacji,</li> <li>• aktywny udział w zajęciach oraz wykonywanie bieżących zadań.</li> </ul> <p>Zaliczenie na ocenę dobrą lub bardzo dobrą otrzymuje student, który spełnia warunki oceny dostatecznej, a ponadto:</p> <ul style="list-style-type: none"> <li>• samodzielnie i poprawnie dobiera metody pozyskiwania danych do kontekstu problemu,</li> <li>• wykazuje inicjatywę i dbałość o jakość, czytelność i rzetelność przygotowanych danych,</li> <li>• stosuje dobre praktyki organizacji danych i potrafi uzasadnić wybory techniczne i etyczne,</li> <li>• przygotowuje projekt kompleksowy, zgodny z wymaganiami i dokumentujący cały proces pozyskiwania danych.</li> </ul> |
|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|       |  |
|-------|--|
| Uwagi |  |
|-------|--|

## Treści merytoryczne (wykaz tematów)

1. Źródła i klasyfikacja danych: przegląd rodzajów danych (pierwotne, wtórne, ustrukturyzowane i nieustrukturyzowane), źródła danych publicznych, komercyjnych i otwartych, podstawowe techniki pozyskiwania informacji.
2. Web scraping w Pythonie: zasady działania stron internetowych, narzędzia do automatycznego pobierania danych (requests, BeautifulSoup, lxml, Selenium), techniki ekstrakcji danych oraz ich zapis i organizacja, ograniczenia prawne i techniczne.
3. Zbieranie danych z API: podstawy komunikacji z interfejsami API, metody dostępu do danych z wykorzystaniem języka Python, uwierzytelnianie i praca z danymi w formacie JSON, przykładowe źródła API.
4. Zbieranie danych od użytkowników: projektowanie formularzy i ankiet, wybór narzędzi, typy pytań, przygotowanie i eksport danych do dalszej analizy.
5. Ocena jakości danych: kryteria jakości i wiarygodności danych z różnych źródeł, identyfikacja błędów i niespójności, wstępna walidacja danych.
6. Aspekty etyczne i prawne: ochrona danych osobowych, legalność pozyskiwania danych, licencje i zasady odpowiedzialnego wykorzystania danych.
7. Organizacja i przechowywanie danych: struktury plików, formaty danych, dobre praktyki organizacyjne i dokumentacyjne, automatyzacja zadań, backup i wersjonowanie.

Wykaz literatury podstawowej

|    |
|----|
| 1. |
|----|

Wykaz literatury uzupełniającej

|    |
|----|
| 1. |
|----|

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) – **studia stacjonarne**

|                                                             |                                                                                              |    |
|-------------------------------------------------------------|----------------------------------------------------------------------------------------------|----|
| Liczba godzin w kontakcie z prowadzącymi                    | Wykład                                                                                       | 20 |
|                                                             | Konwersatorium (ćwiczenia, laboratorium itd.)                                                | 20 |
|                                                             | Pozostałe godziny kontaktu studenta z prowadzącym                                            |    |
| Liczba godzin pracy studenta bez kontaktu z prowadzącymi    | Lektura w ramach przygotowania do zajęć                                                      |    |
|                                                             | Realizacja zadań domowych (problemowych) po zapoznaniu się z niezbędną literaturą przedmiotu |    |
|                                                             | Przygotowanie projektu lub prezentacji na podany temat (praca indywidualna lub w grupie)     |    |
|                                                             | Przygotowanie do egzaminu/zaliczenia                                                         |    |
| Ogółem bilans czasu pracy                                   |                                                                                              |    |
| Liczba punktów ECTS w zależności od przyjętego przelicznika |                                                                                              | 3  |

Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) – **studia niestacjonarne**

|                                                             |                                                                                              |    |
|-------------------------------------------------------------|----------------------------------------------------------------------------------------------|----|
| Liczba godzin w kontakcie z prowadzącymi                    | Wykład                                                                                       | 10 |
|                                                             | Konwersatorium (ćwiczenia, laboratorium itd.)                                                | 15 |
|                                                             | Pozostałe godziny kontaktu studenta z prowadzącym                                            |    |
| Liczba godzin pracy studenta bez kontaktu z prowadzącymi    | Lektura w ramach przygotowania do zajęć                                                      |    |
|                                                             | Realizacja zadań domowych (problemowych) po zapoznaniu się z niezbędną literaturą przedmiotu |    |
|                                                             | Przygotowanie projektu lub prezentacji na podany temat (praca indywidualna lub w grupie)     |    |
|                                                             | Przygotowanie do egzaminu/zaliczenia                                                         |    |
| Ogółem bilans czasu pracy                                   |                                                                                              |    |
| Liczba punktów ECTS w zależności od przyjętego przelicznika |                                                                                              | 3  |