

Annales Universitatis Paedagogicae Cracoviensis

Studia ad Bibliothecarum Scientiam Pertinentia 19 (2021)

ISSN 2081-1861

DOI 10.24917/20811861.19.37

Katarzyna Jarczewska-Walendziak

Uniwersytet Mikołaja Kopernika w Toruniu

ORCID 0000-0002-9192-5097

Małgorzata Kowalska-Chrzanowska

Uniwersytet Mikołaja Kopernika w Toruniu

ORCID 0000-0002-2839-5732

Przemysław Krysiński

Uniwersytet Mikołaja Kopernika w Toruniu

ORCID 0000-0001-7946-3515

Kształcenie specjalistów z zakresu *data science* na poziomie studiów wyższych w Europie

Wprowadzenie

Nauka o danych (ang. *data science*) to interdyscyplinarna dziedzina badań, która wykorzystuje metody naukowe oraz procesy i algorytmy do wyodrębniania wartości z danych. Ujawnia trendy i generuje wnioski, na podstawie których firmy i inne organizacje mogą podejmować bardziej trafne decyzje oraz tworzyć bardziej innowacyjne produkty i usługi¹. Mimo że wywodzi się z analizy statystycznej i eksploracji danych², za jej początki przyjmuje się rok 2002, w którym Komitet ds. Danych dla Nauki i Technologii Międzynarodowej Rady Nauki (ang. International Council for Science: Committee on Data for Science and Technology) rozpoczął wydawanie periodyku pt. „Data Science Journal”, poświęconego takim zagadnieniom, jak opis systemów danych, ich wdrożeń i publikacji, aplikacji, infrastruktury, oprogramowania, kwestii prawnych, odtwarzalności i przejrzystości, dostępności i użyteczności zbiorów danych³.

Termin *data scientist* (w języku polskim *data scientist*, *analityk data science*, *specjalista z zakresu data science*, *mistrz danych*, *operator danych*) jako określenie

1 A. Adhikari, J. DeNero, D. Wagner, *Computational and Inferential Thinking: The Foundations of Data Science*, 2018, [on-line:] <https://www.inferentialthinking.com/> – 30.03.2021.

2 P. Naur, *The Science of Datalogy*, „Communications of the ACM” 1966, vol. 9, no. 7, s. 485.

3 K. D. Foote, *A Brief History of Data Science*, 2016, [on-line:] <https://www.dataversity.net/brief-history-data-science/#> – 30.03.2021; R. A. Irizarry, *The Role of Academia in Data Science Education*, „Harvard Data Science Review” 2020, vol. 2, no. 1.

charakteryzujące osobę, która potrafi zarządzać różnorodnymi zbiorami danych, pracować na nich i je analizować, zaistniał w 2008 roku za sprawą DJ Patila i Jeffa Hammerbachera – analityków danych w serwisach LinkedIn i Facebook. Następnie, w 2012 roku, został spopularyzowany przez „Harvard Business Review”⁴. Na przestrzeni kolejnych lat określenie weszło do oficjalnego języka, a w wielu krajach uruchomiono projekty mające na celu przyspieszenie działań w zakresie tworzenia profesji *data scientist*⁵. Z czasem fach ten został oficjalnie usankcjonowany w rozmaitych klasyfikacjach zawodów i specjalności jako zawód związany z branżą informatyczną⁶.

Chociaż w wielu firmach specjalistę w zakresie *data science* utożsamia się z istniejącym od wielu lat zawodem analityka danych, nie jest to do końca uprawnione. O ile bowiem praca analityka danych sprowadza się przede wszystkim do stosowania metod analizy statystycznej do danych już ustrukturyzowanych, rolą specjalisty z zakresu *data science* jest poruszanie się w niespójnych i różnorodnych zbiorach danych. Obowiązki analityka *data science* mogą obejmować opracowywanie strategii analizy danych, przygotowywanie danych do analizy, badanie, analizę i tworzenie wizualizacji danych, budowę modeli z danymi za pomocą języków programowania oraz wdrażanie modeli do aplikacji. By sprostać tym zadaniom, konieczne jest łączenie szeregu umiejętności z zakresu statystyki, informatyki i biznesu. W rzeczywistości analizę dużych zbiorów danych najskuteczniej przeprowadza się w zespołach. W takim zespole oprócz analityka *data science* mogą pracować analityk biznesowy, który definiuje problem, inżynier danych, który przygotowuje dane oraz zapewnia do nich dostęp, architekt IT, który nadzoruje bazowe procesy i infrastrukturę, oraz programista aplikacji, który wdraża modele lub wyniki analizy w aplikacjach i produktach⁷.

4 T. Davenport, DJ Patil, *Data Scientist: The Sexiest Job of the 21st Century*, „Harvard Business Review” 2012, October, [on-line:] <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century> – 30.03.2021.

5 Np. W 2012 r. powstał program stażowego stworzony przez firmy technologiczne z Doliny Krzemowej z myślą o doktorantach i pracownikach naukowych z tytułem doktora. Propozycja zakładała wykorzystanie potencjału środowiska akademickiego w eksploracji danych na zasadach komercyjnych. Z kolei w latach 2015–2017 z inicjatywy Uniwersytetu w Amsterdamie realizowany był projekt EDISON. Z myślą o przyszłych specjalistach opracowano zbiór dokumentów stanowiących wskazówki dla nauczycieli i trenerów, pracodawców i menedżerów oraz samych naukowców zajmujących się danymi. Szerzej zob. H. Tomoyuki, *Data Scientist: A Key Factor in Innovation Driven by Big Data*, „Journal of Information Processing & Management” 2013, vol. 56, no. 1, s. 2–11; Y. Demchenko, *The Emerging Role of the Data Scientist and the Experience of Data Science Education at the University of Amsterdam*, [w:] *LEARN Toolkit of Best Practice for Research Data Management*, London 2017, s. 105–115.

6 Np. *2010 Standard Occupational Classification System*, 2018, [on-line:] https://www.bls.gov/soc/2010/2010_major_groups.htm – 30.03.2021; *Klassifikation der Berufe 2010: Systematik und Verzeichnisse der KldB 2010*, 2013, [on-line:] <https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010/Systematik-Verzeichnisse/Systematik-Verzeichnisse-Nav.html> – 30.03.2021.

7 R. A. Irizarry, dz. cyt.

Jak wskazują rozmaite doniesienia, zapotrzebowanie na specjalistów z zakresu *data science* stale rośnie. Dane z raportu McKinsey & Company dowodzą, że między styczniem 2015 roku a styczniem 2018 roku liczba ofert pracy dla *data scientists* w samym serwisie Indeed.com wzrosła o 75%⁸. Według serwisu Glassdoor z kolei, popyt na analityków danych od 2016 roku do końca 2020 roku nieprzerwanie rósł. Jeszcze w listopadzie 2020 roku był to jeden z trzech najbardziej poszukiwanych zawodów na świecie (obok programisty oraz twórcy witryn i aplikacji internetowych)⁹.

Szacuje się, że w dobie rosnących zbiorów danych i szybkiego rozwoju usług do ich analizy segment *data science* będzie nadal intensywnie się rozwijał. Wraz z nim będzie rosło także zapotrzebowanie na specjalistów z tego zakresu oraz na różne formy kształcenia przygotowujące do wykonywania tego zawodu. Szacuje się, że w roku 2020 istniało ponad 830 odrębnych programów kształcenia z zakresu nauki o danych oferowanych przez około 500 uniwersytetów na całym świecie¹⁰. Niniejszy artykuł stanowi próbę odpowiedzi na pytanie, jak w krajach europejskich w roku akademickim 2020/2021 wyglądała oferta kształcenia specjalistów w zakresie *data science*. Autorzy skupiają się w nim na analizie 100 losowo wybranych programów studiów pod względem proponowanych treści dydaktycznych, oferowanych kompetencji oraz potencjalnych korzyści płynących z ukończenia tych studiów. Poza zainteresowaniem autorów pozostają natomiast takie kwestie, jak liczba uczelni uruchamiających studia w poszczególnych krajach, ich status, geograficzna lokalizacja czy zasięg oddziaływania.

Kształcenie *data scientists* w świetle literatury przedmiotu

Zagadnienie kształcenia specjalistów *data science* było już przedmiotem opracowań naukowych. Temat ten był podejmowany przede wszystkim w kontekście ogólnej potrzeby kształcenia tego rodzaju specjalistów w związku ze wzrostem znaczenia danych w świecie oraz koniecznością modyfikacji istniejących programów kształcenia specjalistów informacji. Dokonywano również analiz ofert kształcenia, prace te były jednak ograniczone do konkretnego obszaru – najczęściej Stanów Zjednoczonych łącznie z Kanadą.

Jedną z osób, które zajęły się problemem kształcenia specjalistów *data science*, była Barbara Brynko. Dokonując analizy istniejących w 2013 roku programów edukacyjnych Arizona State University, Drexel University i San Jose State University, badaczka wykazała, że wiele programów kształcenia specjalistów *data science* ma wymiar interdyscyplinarny, obejmuje bowiem zagadnienia wpisujące się w obszar

⁸ *The Age of Analytics: Competing in a Data-Driven World*, 2016, [on-line:] <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world> – 30.03.2021.

⁹ A. Woodie, *Why Data Science Is Still a Top Job*, 2020, [on-line:] <https://www.datanami.com/2020/11/16/why-data-science-is-still-a-top-job/> – 30.03.2021.

¹⁰ *Compare All Types of Data Science Degrees*, 2021, [on-line:] <https://www.datascienceprograms.org/> – 30.03.2021.

zainteresowania wielu dyscyplin – nie tylko tych związanych wyłącznie z technologią, inżynierią czy matematyką. Taki kształt oferty dydaktycznej Brynko oceniła pozytywnie, zauważając, że jest on o tyle uzasadniony, że wykorzystywanie danych często wiąże się nie tylko z kwestiami, co gromadzić, jak selekcjonować i przetwarzać dane, lecz także jak je mierzyć, oceniać, interpretować oraz jak motywować do ich zbierania¹¹.

Do analogicznych wniosków doszli również Rong Tang i Watinee Sae-Lim, którzy w 2016 roku przeprowadzili badania 30 losowo wybranych programów kształcenia w Stanach Zjednoczonych. W opisach poszczególnych programów użyto szeregu specjalistycznych terminów, przy czym terminy są wspólne dla różnych dyscyplin. Część programów kładła nacisk na rozwój umiejętności informacyjnych i informatycznych (w tym z zakresu matematyki i statystyki), inne – na umiejętności komunikacyjne. Poza tym programy różniły się pod względem liczby godzin dydaktycznych, zajęć praktycznych i form ukończenia studiów¹².

Najbardziej kompleksowo kształcenie analityków danych zostało omówione w raporcie *Data Science for Undergraduates: Opportunities and Options* przygotowanym w 2018 roku przez amerykańską Narodową Akademię Nauki, Inżynierii i Medycyny (ang. National Academies of Sciences, Engineering, and Medicine). Analiza ofert na poziomie akademickim dowiodła, że kształcenie z zakresu analizy danych w Stanach Zjednoczonych prowadzone było w różnych formach, jako:

- kursy wprowadzające,
- studia pierwszego oraz drugiego stopnia,
- dwuletnie kursy dokształcające,
- szkolenia kończące się otrzymaniem certyfikatu z zakresu nauki o danych,
- masowe kursy online,
- tzw. szkoły letnie i obozy szkoleniowe¹³.

Jak wykazano w raporcie, poza programami *stricte* dedykowanymi przetwarzaniu i analizie danych, w ofercie dydaktycznej wielu uczelni figurowały także studia, które w sposób pośredni nawiązywały do tej tematyki. Najwięcej tego typu treści można było znaleźć na studiach z zakresu matematyki, statystyki oraz informatyki. W programach dominowały takie zagadnienia, jak: zaawansowana analityka danych, duże zbiory danych, eksploracja danych, modelowanie symulacji i myślenie komputacyjne. Wiele zagadnień poświęconych nauce o danych podejmowano także na studiach z zakresu ekonomii, biznesu, psychologii, biologii, geografii oraz nauk o Ziemi¹⁴.

11 B. Brynko, *Data Scientists: You Sexy Thing (Cover Story)*, „Information Today” 2013, vol. 30, no. 10, s. 1–36.

12 R. Tang, W. Sae-Lim, *Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus*, „Education for Information” 2016, vol. 32, no. 3, s. 269–290.

13 *Data Science for Undergraduates: Opportunities and Options*, 2018, [on-line:] <https://www.nap.edu/catalog/25104/data-science-for-undergraduates-opportunities-and-options> – 30.03.2021.

14 Tamże.

W styczniu 2020 roku Rafael A. Irizarry na łamach „Harvard Data Science Review” opublikował artykuł pt. *The Role of Academia in Data Science Education* – wskazał w nim kryteria, które powinny być spełniane przez programy kształcenia analityków danych. Irizarry wyszedł z założenia, że analizą danych mogą zajmować się zarówno analitycy danych, specjaliści uczenia maszynowego, jak i programiści tworzący oprogramowanie na potrzeby nauki o danych. Badacz sformułował pogląd, że program kształcenia analityków danych powinien uwzględniać naukę algorytmów oraz zagadnień związanych z optymalizacją i zarządzaniem strukturami danych; w mniejszym stopniu powinien skupiać się na teorii, a w większym stawiać na kształcenie praktycznych umiejętności (np. Python, Spark, TensorFlow, Keras i inne języki programowania). Ze względu na konieczność posiadania odpowiedniego doświadczenia, studia z zakresu analizy danych powinny być uruchamiane na poziomie studiów magisterskich i doktoranckich oraz być prowadzone przez silne jednostki naukowe¹⁵.

Analogiczne wytyczne znalazły się także w opublikowanym w styczniu 2021 roku raporcie pt. *Computing Competencies for Undergraduate Data Science Curricula*, wydanym przez grupę zadaniową ACM Data Science Task Force, skupiającą naukowców ze Stanów Zjednoczonych, Wielkiej Brytanii oraz Chin. Jak podkreślono w dokumencie, pojawienie się nauki o danych stwarza duże możliwości, ale stanowi też swego rodzaju wyzwanie dla przedstawicieli wszystkich dyscyplin naukowych, dlatego programy kształcenia powinny realizować konkretny cel: pomóc studentom w zdobywaniu umiejętności rozpoznawania potencjału danych, gromadzenia danych wysokiej jakości oraz poszukiwania nowej wiedzy na podstawie tych danych. Tworzone w takim duchu programy na pierwszym etapie kształcenia powinny obejmować podstawy informatyki, statystyki, matematyki i uczenia maszynowego, a w szczególności m.in. teorię gromadzenia danych, rachunek prawdopodobieństwa i analizę statystyczną, badanie struktur danych, analizę kontekstową, programowanie i kodowanie, wizualizację, bezpieczeństwo gromadzenia danych i etykę ich wykorzystywania. Nade wszystko powinny one rozwijać umiejętność rzeczywistego wykorzystywania danych. Na dalszych etapach kształcenia nauczanie powinno uwzględniać szersze konteksty nauki o danych, a więc m.in. projektowanie struktur danych i aplikacji, testowanie algorytmów, sztuczną inteligencję, systemy Big Data, *data mining*, projektowanie zorientowane na użytkownika, badania jakości, kompletności i spójności danych, prototypowanie, automatyczne wnioskowanie. Poza treściami obowiązkowymi, w zależności od osadzenia danego kierunku kształcenia w danej dyscyplinie nauki (biologia, chemia, ekonomia), w programach studiów powinny znaleźć się także przedmioty specyficzne dla nich¹⁶.

15 R. A. Irizarry, dz. cyt.; zob. też S. C. Hicks, R. A. Irizarry, *A Guide to Teaching Data Science*, „The American Statistician” 2018, vol. 72, no. 4, s. 382–391.

16 A. Danyluk, P. Leidig, *Computing Competencies for Undergraduate Data Science Curricula*, 2021, [on-line:] http://dstf.acm.org/DSTF_Final_Report.pdf – 30.03.2021.

Metodologia

Ze względu na istnienie różnych systemów kształcenia na świecie, przedmiot analiz autorów stanowiła wyłącznie oferta dydaktyczna uczelni z krajów europejskich, objętych bolońskim systemem edukacji.

Poddane analizie programy studiów zidentyfikowane zostały pod koniec 2020 roku w drodze poszukiwań w zasobach kilku portali edukacyjnych: Academiccourses.com¹⁷, MastersPortal.com¹⁸, Postgrad.com¹⁹ i Studies in Europe²⁰. Do wyszukiwania kierunków studiów zastosowano następujące słowa kluczowe: *data analyst, data analytics, analytics, data science, data scientist*. W dalszej kolejności, na podstawie informacji zamieszczonych na stronach WWW poszczególnych uczelni, zweryfikowano istnienie i aktualność danej oferty. W etapie końcowym dokonano analizy poszczególnych programów kształcenia, biorąc pod uwagę takie kryteria, jak nazwa kierunku, jednostka prowadząca, opis programu kształcenia (wymagania wstępne, oferowana wiedza i umiejętności, potencjalne miejsca zatrudnienia), czas trwania nauki, tryb studiów, przedmioty nauczania. Do prezentacji wybrano po trzy przykłady najciekawszych ofert kształcenia. Programy przedstawione zostały w podziale na studia pierwszego i drugiego stopnia oraz studia podyplomowe.

Kształcenie specjalistów z zakresu *data science* na studiach pierwszego stopnia

Studia z zakresu *data science* na poziomie studiów pierwszego stopnia oferuje 30 ośrodków europejskich, m.in. w takich krajach jak Wielka Brytania, Włochy, Czechy, Hiszpania, Francja czy Polska (por. tabela 1).

Najczęściej w nazwie kierunku pojawiają się określenia *data science* lub *data analytics*. Większość studiów ma charakter studiów odpłatnych (średnio ok. 1300 euro za semestr). W większości placówek kształcenie na kierunku analityk danych realizowane jest w trybie stacjonarnym (sporadycznie wieczorowym), trwa średnio trzy lub cztery lata (ok. 120 godzin w semestrze) i kończy się uzyskaniem tytułu licencjata.

W programach studiów dominują przedmioty informatyczne, takie jak programowanie, kombinatoryka, bazy danych, eksploracja tekstu, wizualizacja, uczenie maszynowe itp. Moduł przedmiotów specjalistycznych wzbogacają przedmioty z ekonometrii (jak statystyka czy rachunek prawdopodobieństwa) oraz przedmioty ogólne (np. metodologia prowadzenia badań naukowych, teoria grup społecznych, komunikacja społeczna, seminarium i języki obce).

17 Academiccourses.com, 2020, [on-line:] <https://www.academiccourses.com/> – 30.03.2021.

18 MastersPortal.com, 2020, [on-line:] <https://www.mastersportal.com/> – 30.03.2021.

19 Postgrad.com, 2020, [on-line:] <https://www.postgrad.com/> – 30.03.2021.

20 Studies in Europe, 2020, [on-line:] <https://www.studies-in-europe.eu/> – 30.03.2021.

Twórcy programów podkreślają, że w stale zmieniającym się globalnym środowisku, gdzie przetwarzanie danych odbywa się za pomocą coraz bardziej złożonych i wyrafinowanych modeli algorytmicznych oraz z wykorzystaniem zaawansowanych technologii, istnieje ogromne zapotrzebowanie na specjalistów, którzy mogą przetwarzać, rozumieć i prezentować te dane w odpowiedni sposób.

Przyszłych studentów do podejmowania nauki mają zachęcać potencjalne miejsca pracy, którymi według twórców programów kształcenia są przedsiębiorstwa produkcyjne i usługowe, biura rachunkowe, banki i zakłady ubezpieczeń, instytucje administracji publicznej, gdzie absolwenci analityki będą mogli zajmować takie stanowiska, jak analityk finansowy, analityk biznesowy, specjalista ds. analiz i sprawozdawczości, specjalista ds. analiz rynkowych, analityk danych, analityk informacji.

Kształcenie specjalistów z zakresu *data science* na studiach drugiego stopnia

Oferta kształcenia analityków *data science* na poziomie studiów drugiego stopnia jest o wiele szersza niż w przypadku studiów pierwszego stopnia. Kształcenie w tym zakresie w Europie oferuje około 170 ośrodków, m.in. z Wielkiej Brytanii, Francji, Hiszpanii, Belgii, Włoch, Finlandii, Cypru i Polski (por. tabela 2).

Podobnie jak w przypadku studiów licencjackich, w nazwach kierunków studiów występuje określenie *data science* lub *data analytics*, z tą jednak różnicą, że w nazwach tych bardzo często pojawiają się wyrażenia sugerujące sprofilowanie kształcenia na ściśle określony typ informacji (dominuje ujęcie biznesowe). Większość studiów ma charakter studiów odpłatnych (średnio ok. 1500 euro za semestr) i prowadzona jest przez uczelnie prywatne. Kształcenie na kierunku analityka danych realizowane jest z reguły w trybie stacjonarnym, trwa średnio rok lub dwa lata (ok. 200 godzin w semestrze) i kończy się uzyskaniem tytułu magistra.

W programach studiów dominują przedmioty technologiczne, które w zależności od specjalizacji wzbogacane są modułami przedmiotów ogólnych, związanych z danym profilem kształcenia (np. z zakresu nauk ekonomicznych, matematycznych czy geograficznych).

W opisie studiów, podobnie jak w przypadku studiów licencjackich, bardzo często podkreśla się, że dane są podstawowym surowcem XXI wieku, a w nich ukryte są przydatne informacje, które mogą wspierać podejmowanie decyzji w wielu sektorach gospodarki. Wskazuje się także, że postępy w technologii komputerowej i eksplozja informacyjna danych wymuszają potrzebę kształcenia specjalistów w zakresie eksploracji i przetwarzania danych.

Według twórców programów kształcenia, proponowane kierunki studiów gwarantują nabycie umiejętności analitycznych i technicznych pozwalających na przekształcanie danych w cenne informacje, które będą mogły być wykorzystywane w procesach wspomagania decyzji i rozumienia działań organizacji w różnych dziedzinach, w tym w obszarach zdrowia publicznego, rozwoju produktów, badań naukowych, marketingu i neuronauki.

Tabela 1. Przykładowe programy kształcenia specjalistów z zakresu *data science* w roku akademickim 2020/2021 na studiach pierwszego stopnia

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Skrócony opis oferty kształcenia	Czas trwania	Tryb	Przykładowe przedmioty / bloki tematyczne
Czechy	Czech University of Life Sciences (Praga)	Analityka danych GIS (ang. <i>Environmental Data Science</i>)	Wydział Inżynierii	Studia dostarczają wiedzy na temat różnych źródeł pozyskiwania informacji z zakresu nauk o ziemi (czujników, satelitów i symulacji komputerowych, aplikacji, wirtualnych laboratoriów teledetekcji i systemów wczesnego ostrzegania przed zagrożeniami naturalnymi). Pozwalają nabyć umiejętności analizowania dużych ilości danych i odkrywania wzorców, które pomogą w przyszłości służyć przewidywaniu wyzwań środowiskowych, rozwiązywaniu problemów, projektowaniu systemów, zarządzaniu projektami i kreatywnemu myśleniu ^[4] .	3 lata	stacjonarny	<ul style="list-style-type: none"> eksploracyjna analiza danych, hydrologia, matematyka z aplikacjami, meteorologia i klimatologia, nauki o ziemi, podstawy geografii fizycznej programowanie w języku R, programowanie w Pyhtonie, rozwój algorytmów^[4]
Wielka Brytania	Bournemouth University (Poole)	Nauka o danych i analityka (ang. <i>Data Science & Analytics</i>)	Wydział Nauki i Technologii	Studia przygotowują do pracy z zasobami Big Data. Zapoznają z najnowocześniejszymi technikami uczenia maszynowego, w tym implementacją algorytmów w sztucznych sieciach neuronowych. W zakresie analizy wychodzą poza analizę danych biznesowych i obejmują analizę obrazu, wideo, mowy i tekstu. Przygotowują do tego, jak pozyskiwać dane z różnych aplikacji, jak wykrywać fałszywe wiadomości i jak dokonywać analizy danych medycznych oraz analizy wizualnej ^[3] .	3 lata z absolutorium lub 4 lata z dyplomem	stacjonarny	<ul style="list-style-type: none"> podstawy programowania, dane i bazy danych, sieci i bezpieczeństwo cybernetyczne, analiza i projektowanie systemów biznesowych, uczenie maszynowe, narzędzia i technologie analizy danych, zarządzanie danymi, zarządzanie projektem i praca zespołowa, tworzenie stron internetowych, projektowanie aplikacji, projektowanie systemów, wizualizacja danych, głębokie uczenie, inteligencja maszyn, eksploracja danych, kierunki rozwoju przetwarzania danych^[4]

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Strócony opis oferty kształcenia	Czas trwania	Tryb	Przykładowe przedmioty / bloki tematyczne
Włochy	University of Campania „Luigi Vanvitelli” (Santa Maria Capua Vetere)	Analityka danych (ang. <i>Data Analytics</i>)	Wydział Matematyki i Fizyki	Program studiów koncentruje się na metodach i technikach nauczania w dziedzinie matematyki, statystyki wnioskowania i analizy danych eksploracyjnych. Obejmuje metody i narzędzia informatyczne do projektowania i zarządzania bazami danych, podstawy programowania i ekonometrię. Dostarcza umiejętności statystycznych w zakresie: przetwarzania i analizy wielowymiarowych lub złożonych danych pochodzących z różnych źródeł; eksploracji danych i technik optymalizacji; przewidywania i monitorowania zjawisk ewolucyjnych; zarządzania bazami danych i systemami rozproszonych przetwarzania danych; przekazywania wyników analizy poprzez prezentację i raporty ^[5] .	3 lata	stacjonarny	<ul style="list-style-type: none"> • algebra, • analiza, • bazy danych i systemy informacyjne, • bioinformatyka, • ekonometria w Business Intelligence, • eksploracja danych i Big Data, • metodologia badań społecznych, • metody numeryczne w analizie danych, • modelowanie systemów komputerowych i sieci semantyczne, • podstawy programowania, • statystyka, • teoria prawdopodobieństwa^[6]

^[1] *Czech University of Life Sciences: Environmental Data Science*, 2020, [on-line:] <https://www.fzp.czu.cz/en/r-9408-study/r-9495-study-programmes/r-9744-bachelor-s-study-programmes/r-14239-environmental-data-science-30.03.2021>]; ^[2] *Czech University of Life Sciences: Study plan BEDS*, 2020, [on-line:] <https://www.fzp.czu.cz/dl/71175?lang=en-30.03.2021>; ^[3] *Bournemouth University: Data Science & Analytics: Course Details*, 2020, [on-line:] <https://www.bournemouth.ac.uk/study/courses/bsc-hons-data-science-analytics-30.03.2021>; ^[4] *Bournemouth University: Data Science and Analytics: Programme Specification – Purpose and Guidance for Completion*, 2020, [on-line:] <https://intranet.bournemouth.ac.uk/progspecs/bsc-data-science-and-analytics.pdf-30.03.2021>; ^[5] *University of Campania „Luigi Vanvitelli”*: Bachelor’s Degree in Data Analytics, 2020, [on-line:] <https://www.matfis.unicampania.it/didattica/corsi-di-studio/data-analytics/2-non-categorizzato/117-bachelor-s-degree-in-data-analytics-30.03.2021>; ^[6] *University of Campania „Luigi Vanvitelli”*: Bachelor’s Degree in Data Analytics: Study Plan, 2020, [on-line:] https://www.matfis.unicampania.it/images/didattica/triennale_analytics/piano_studi/piano_studi_Analytics_2019_2020.pdf-30.03.2021.

Źródło: opracowanie własne.

Tabela 2. Przykładowe programy kształcenia specjalistów z zakresu *data science* w roku akademickim 2020/2021 na studiach drugiego stopnia

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Skrócony opis oferty kształcenia	Czas trwania	Tryb	Przykładowe przedmioty / bloki tematyczne
Hiszpania	Universidad Pontificia Comillas (Madryt)	Technologie Big Data i zaawansowana analityka (ang. <i>Big Data Technologies and Advanced Analytics</i>)	Wydział Inżynierii	<p>Studia obejmują zagadnienia z zakresu inżynierii, technologii i biznesu. Przeznaczone są dla absolwentów inżynierii przemysłowej, inżynierii telekomunikacyjnej, inżynierii komputerowej, nauk matematycznych i fizyki. Pozwalają nabyć umiejętności z zakresu wykorzystywania technologii Big Data i zaawansowanych technik eksploracji danych i uczenia maszynowego. Przygotowują do wykonywania takich zadań, jak inżynier danych, analityk danych i architekt danych^[1].</p>	rok	stacjonarny	<ul style="list-style-type: none"> • analiza danych nieustrukturyzowanych, • architektura Big Data, • Big Data i zarządzanie danymi, • matematyczne podstawy analizy danych, • przetwarzanie strumieniowe, • systemy pamięci masowej NoSQL, • technologie przetwarzania dużych zbiorów danych, • uczenie maszynowe^[2]
Polska	Uniwersytet Warszawski (Warszawa)	Data Science and Business Analytics	Wydział Nauk Ekonomicznych	<p>Studia stanowią naturalną i pożądaną kontynuację studiów pierwszego stopnia na kierunku informatyka i ekonometria oraz na kierunkach pokrewnych. Ponadto studia te są atrakcyjną propozycją dla osób posiadających dobre wykształcenie w zakresie matematyki i nauk ścisłych, a chcących poznać ich praktyczne zastosowanie w ekonomii, zarządzaniu, finansach i ubezpieczeniach. Absolwent nabeździe wiedzę z zakresu metod analizy danych oraz praktyczne umiejętności gromadzenia, przetwarzania i dogłębnej analizy danych. Wykształci zdolność twórczego i kreatywnego myślenia, która wzbogacona specjalistyczną wiedzą z analizy danych ułatwi mu podejmowanie decyzji o charakterze strategicznym dla instytucji i przedsiębiorstw oraz umiejętność swobodnego łączenia wiedzy teoretycznej z kompleksowym podejściem do danych (od pozyskania, przetwarzania, analizowania aż po eksplorację danych). Zatrudnieniem absolwenta mogą być zainteresowane: instytucje finansowe, firmy telekomunikacyjne, marketingowe, farmaceutyczne, doradcze, domy mediowe i wszystkie inne, które potrzebują kompleksowej i dokładnej analizy danych^[3].</p>	2 lata	stacjonarny; wieczorowy	<ul style="list-style-type: none"> • algorytmy, • auto prezentacja, • Big Data, • eksploracja danych, • komunikacja, • negocjacje, • programowanie w języku R, • pozyskiwanie danych z Internetu i mediów społecznościowych, • statystyka / analiza danych eksploracyjnych, • uczenie maszynowe, • wizualizacja danych, • wprowadzenie do Python i SQL, • zaawansowana ekonometria^[4]

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Skrócony opis oferty kształcenia	Czas trwania	Tryb	Przykładowe przedmioty / bloki tematyczne
Wielka Brytania	Birmingham City University (Birmingham)	Analityka Big Data (ang. <i>Big Data Analytics</i>)	Wydział Informatyki i Technologii Cyfrowej	Studia obejmują zagadnienia eksploracji danych, metody i narzędzia zarządzania Big Data i zaawansowaną statystykę. Przeznaczone są dla absolwentów informatyki, biznesu i inżynierii lub kierunków pokrewnych, w których istotny element stanowiła informatyka. Pozwalają nabyć umiejętności analityczne i rozwiązywania problemów, a także w zakresie przeprowadzania badań. Studia przygotowują do wykonywania takich zawodów, jak specjalista ds. hurtowni danych, analityk danych, konsultant danych, architekt rozwiązań technologicznych ^[5] .	rok	stacjonarny, niestacjonarny	<ul style="list-style-type: none"> • analityka i wizualizacja danych z mediów społecznościowych, • bazy danych dla przedsiębiorstw, • eksploracja danych, • indywidualny projekt magisterski, • nowoczesna optymalizacja, • statystyka stosowana, • zarządzanie Big Data^[6]

^[1] Universidad Pontificia Comillas: Master in Big Data Technologies and Advanced Analytics, 2020, [on-line:] <https://www.comillas.edu/postgrado/master-en-big-data-tecnologia-y-analitica-avanzada> – 30.03.2021; ^[2] Tamże; ^[3] Uniwersytet Warszawski: *Data Science and Business Analytics*, [on-line:] <https://www.wne.uw.edu.pl/en/candidates/data-science/> – 15.05.2021; ^[4] Tamże; ^[5] Birmingham City University: Big Data Analytics, 2020, [on-line:] <https://www.bcu.ac.uk/courses/big-data-analytics-msc-2020-21> – 30.03.2021; ^[6] Tamże.

Źródło: opracowanie własne.

Tabela 3. Przykładowe programy kształcenia specjalistów z zakresu *data science* w roku akademickim 2020/2021 na studiach podyplomowych

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Skrócony opis oferty kształcenia	Czas trwania	Przykładowe przedmioty / bloki tematyczne
Belgia	KU Leuven (Katholieke Universiteit Leuven)	Big Data i analityka w biznesie i zarządzaniu (ang. <i>Big Data & Analytics in Business and Management</i>)	Wydział Ekonomii i Biznesu	<p>Program studiów zapewnia wiedzę z zakresu zastosowań technologii Big Data i technologii analitycznych. Koncentruje się na aplikacjach biznesowych i zarządzanych, pokazując, w jaki sposób duże zbiory danych i techniki analityczne mogą tworzyć wartość biznesową i być wykorzystywane w zarządzaniu projektami.</p> <p>Studia skierowane są do osób posiadających minimum tytuł licencjata oraz minimum roczne doświadczenie w zakresie analizy biznesowej, analizy danych, statystyki, zarządzania danymi, zarządzania projektami itp. Podjęciem studiów powinny być zainteresowane przede wszystkim osoby wykonujące takie zawody, jak główny inspektor danych, główny analityk, menedżer ds. ryzyka, architekt danych, menedżer projektów badawczych, analityk biznesowy, menedżer produktu, kierownik projektu^[1].</p>	poł roku	<ul style="list-style-type: none"> • analityka marketingowa, • analityka predykcyjna, • analityka ubezpieczeniowa, • analiza i integracja IT, • analizy dużych zbiorów danych za pomocą Apache Spark i H2O, • analizy opisowe, • analizy sieci społecznościowych, • bazy danych NoSQL, • eksploracja sieci, • eksploracja tekstu, • etyka i prywatność, • graficzne bazy danych, • Hadoop i jego ekosystem, • jakość i bezpieczeństwo danych, • model procesu analitycznego, • paradygmat MapReduce, • prywatności i zgodność danych, • systemy rekomendujące, • techniki uczenia się • waloryzacja analityki, • wykrywanie oszustw^[2]

Kraj	Uczelnia	Nazwa kierunku	Jednostka prowadząca	Skrócony opis oferty kształcenia	Czas trwania	Przykładowe przedmioty / bloki tematyczne
Pol-ska	Politechnika Warszawska (Warszawa)	Data Science – algorytmy, narzędzia i aplikacje klasy Big Data	Wydział Elektroniki i Techniki Informatycznych	Oferta edukacyjna skierowana do osób, które chcą wykorzystywać wiedzę zawartą w dużych wolumenach danych w celu wspierania podejmowania decyzji, w szczególności dla analityków i decydentów z obszaru finansów, bankowości, ubezpieczeń, produkcji, marketingu, handlu, usług, opieki zdrowotnej, branży energetycznej, nauki i innych obszarów działalności. Celem studiów jest przekazanie słuchaczom praktycznych umiejętności z zakresu przetwarzania i analizy dużych danych i zapoznanie ich z rolą danych w procesach podejmowania decyzji oraz wykorzystaniem danych jako cennych aktywów strategicznych. Studia przewidziane są dla osób posiadających doświadczenie w pracy z technologiami (np. podstawowa umiejętność programowania w dowolnym języku, podstawowa znajomość zagadnień związanych z bazami danych i językiem SQL), jak i nieposiadających takiego doświadczenia ^[4] .	rok	<ul style="list-style-type: none"> • analiza danych – podstawy statystyczne, • Big Data – koncepcje, technologie, • <i>data mining</i> – metody eksploracji danych, • metody sztucznej inteligencji, • programowanie w R i Python, • studium przypadku, • <i>text mining</i> – odkrywanie wiedzy z tekstowych zbiorów danych, • wstęp do wizualizacji danych, • zaawansowana analityka z SAS Vija^[4]
Wielka Brytania (Londyn)	Birkbeck, University of London	Stosowana nauka o danych (ang. <i>Applied Data Science</i>)	Wydział Informatyki i Systemów Informatycznych	Oprócz zdobycia szerokiej wiedzy na temat komputerów, studia umożliwiają nabycie umiejętności programowania i analizy danych przy użyciu języka programowania Python. Rozwijają umiejętność rozwiązywania ogólnych problemów obliczeniowych, ze szczególnym uwzględnieniem rozwiązań w zakresie czyszczenia danych, analiz, uczenia maszynowego i wizualizacji danych. Studia przewidziane są dla absolwentów, którzy nie posiadają zaawansowanej wiedzy i umiejętności z informatyki ^[5] .	rok	<ul style="list-style-type: none"> • myślenie komputacyjne i algorytmy, • programowanie obiektowe, • statystyki i prawdopodobieństwa w języku Python, • systemy zarządzania bazami danych i SQL, • wprowadzenie do analizy Big Data, • wprowadzenie do przetwarzania w chmurze, • wprowadzenie do sieci neuronowych i głębokiego uczenia^[6]

[4] KU Leuven: Postgraduate Studies: Big Data & Analytics in Business and Management, 2020, [on-line:] <https://feb.kuleuven.be/permanente-vorming/bigdata-analytics> – 30.03.2021; [2] KU Leuven: Postgraduate Studies: Big Data & Analytics in Business and Management: Brochure, 2020, [on-line:] <https://feb.kuleuven.be/permanente-vorming/postgraduatebigdatabrochure/> – 30.03.2021; [3] Politechnika Warszawska: Data Science, 2020, [on-line:] <http://datascience.ii.pw.edu.pl/datascience.html> – 30.03.2021; [4] Tamże; [5] Birkbeck, University of London: Applied Data Science, 2020, [on-line:] http://www.bbk.ac.uk/study/2020/postgraduateprogrammes/TPCCOMIP_C – 30.03.2021; [6] Tamże.

Źródło: opracowanie własne.

Wśród stanowisk, które będą zajmować absolwenci analityki danych, obok – pojawiających się już w ofercie studiów pierwszego stopnia – inżynierów danych, *data scientists*, analityków danych i architektów danych wymieniani są jeszcze: kierownicy projektów telekomunikacyjnych, szefowie usług telekomunikacyjnych, analitycy danych i zarządzania inteligencją biznesową, specjaliści ds. zarządzania bezpieczeństwem informacji, konsultanci ds. technologii, inżynierowie systemów Big Data, architekci rozwiązań Big Data, menedżerowie bezpieczeństwa w projektach Big Data, konsultanci danych, menedżerowie ds. prywatności w rozwiązaniach Big Data i audytorzy systemów Big Data.

Kształcenie specjalistów z zakresu *data science* na studiach podyplomowych

Studia podyplomowe z zakresu szeroko rozumianej *data science* oferuje w Europie ponad 350 różnego typu wyższych szkół publicznych i prywatnych, m.in. z takich krajów jak Anglia, Irlandia, Szkocja, Czechy, Belgia, Hiszpania, Francja, Węgry, Szwajcaria, Grecja i Polska (por. tabela 3).

W porównaniu do programów kształcenia na studiach pierwszego i drugiego stopnia, w ramach studiów podyplomowych dostrzegalna jest większa specjalizacja. W wielu przypadkach nazwy kierunków studiów nie ograniczają się już do określenia *data science* czy *data analytics*, lecz wskazują na sprofilowanie ścieżki kształcenia (np. analityka biznesowa, analityka gospodarcza, analityka finansowa, analityka danych strategicznych, analityka danych statystycznych, analityka w bezpieczeństwie). Taka specjalizacja przekłada się zarówno na program kształcenia (wśród przedmiotów pojawiają się przedmioty ogólne – związane z określoną dziedziną, oraz technologiczne – wskazujące na zastosowanie technologii w określonej branży), jak i koszt studiów (średni koszt studiów wynosi około 4 tys. euro za rok nauki). Studia podyplomowe z analityki danych trwają z reguły rok lub dwa lata i obejmują do 250 godzin dydaktycznych. Wieńczy je uzyskanie certyfikatu z zakresu analizy danych.

Podobnie jak w przypadku studiów licencjackich i magisterskich, zachętą do podjęcia kształcenia z zakresu analityki na studiach podyplomowych ma być nabycie niezbędnych współcześnie umiejętności pozyskiwania i przetwarzania informacji z dużych źródeł danych. Takie wykształcenie ma zapewnić zatrudnienie na intratnych stanowiskach analityka biznesowego, analityka finansowego, specjalisty Business Intelligence, analityka danych, analityka raportów, specjalisty ds. produkcji, analityka rynku pracy, analityka rynku pieniężnego, walutowego, kapitałowego, specjalisty ds. zarządzania aktywami finansowymi, specjalisty ds. zarządzania ryzykiem, doradcy finansowego (gospodarczego). W przypadku wielu uczelni o przyjęcie na studia podyplomowe z analityki mogą ubiegać się osoby, które legitymują się wykształceniem ekonomicznym lub informatycznym, a także posiadają kilkuletnie doświadczenie na stanowisku związanym z analizą danych.

Podsumowanie

Przegląd istniejących ofert kształcenia dla osób aspirujących do zawodu *data scientist* wskazuje na ogromną różnorodność zarówno pod względem liczby godzin dydaktycznych, jak i oferowanych przedmiotów nauczania. W poszczególnych krajach europejskich nie dostrzega się jednak zasadniczych różnic w realizowanych treściach dydaktycznych. Większość programów koncentruje się na metodach i technikach wykorzystywanych w matematyce (analiza i rachunek różniczkowy), wnioskowaniu statystycznym i eksploracyjnej analizie danych. Dużo miejsca zajmują tu też zagadnienia dotyczące metod i narzędzi informatycznych do projektowania i zarządzania bazami danych, podstaw programowania, epistemologii badań, ekonometrii oraz eksperymentalnych i dynamicznych modeli danych.

Z przeglądu ofert wynika, że studia pierwszego stopnia, zgodnie z zaleceniami badaczy czy organizacji międzynarodowych przywoływanymi już na łamach niniejszego tekstu, obejmują przede wszystkim matematykę, statystykę, uczenie maszynowe i kursy programowania. Studia drugiego stopnia z kolei poszerzają zdobytą już wiedzę i umiejętności o zagadnienia z zakresu myślenia komputacyjnego, programowania obiektowego, sztucznej inteligencji, funkcjonowania sieci neuronowych i głębokiego uczenia.

Największe zróżnicowanie ofert kształcenia obserwuje się na studiach podyplomowych. Tym, co zwraca uwagę, jest sprofilowanie wszystkich istniejących programów studiów pod względem określonych dyscyplin nauki – przede wszystkim ekonomii i statystyki oraz informatyki i matematyki. Dlatego w przypadku tych studiów od kandydatów bardzo często wymaga się już posiadania zaawansowanych umiejętności z zakresu ICT i wykształcenia ekonomicznego lub informatycznego (ewentualnie ekwiwalentnego), a nierzadko również doświadczenia zawodowego w zakresie analizy danych.

Studia z zakresu *data science* uruchamiane są w większości przez wydziały prowadzące kształcenie w obszarze nauk społecznych (ekonomia i biznes) oraz technicznych (matematyka, elektronika, informatyka). Ich beneficjentami są osoby, które chciałyby pogłębić wiedzę i umiejętności z zakresu nauk informatycznych i ścisłych, poznać sposoby gromadzenia, przechowywania oraz przetwarzania danych, nauczyć się zasad poprawnego i skutecznego programowania, wykonywania analiz statystycznych oraz interpretacji ich wyników.

Przegląd wybranych ofert kształcenia pozwala sformułować ogólny wniosek, że istniejące w Europie programy studiów z zakresu *data science* pozwalają:

- osiągnąć wiedzę na temat systemów zarządzania bazami danych i przetwarzania rozproszonego, także w środowisku chmury obliczeniowej,
- zapoznać się z metodologią eksploracji danych i technikami rozwiązywania przy ich pomocy złożonych problemów,
- nabyć umiejętności przetwarzania i analizy wysokowymiarowych lub złożonych danych pochodzących z różnych i nieustrukturyzowanych źródeł,

- zdobyć doświadczenie w analizie dużych zasobów danych pochodzących z sieci i na tej podstawie wykształcić umiejętność przewidywania ewolucji różnych zjawisk i procesów,
- nauczyć się wykorzystywania oprogramowania statystycznego i określonych języków programowania, zwłaszcza R i Python,
- nabyć umiejętności wizualizacji i przekazywania wyników analiz.

Oprócz wiedzy kierunkowej studenci *data science* mają szansę pogłębienia swoich kompetencji miękkich, ze szczególnym uwzględnieniem zdolności komunikacyjnych i negocjacyjnych.

Tym, co zaskakuje wśród analizowanych ofert kształcenia, jest niedostatek propozycji dla absolwentów szeroko rozumianych studiów informacyjnych (zarządzania informacją, infobrokerstwa, bibliotekoznawstwa, architektury informacji, informatologii). Na razie, zgodnie z profilami programów kształcenia, analiza danych jest domeną osób z wykształceniem informatycznym lub ekonomicznym. A przecież wbrew obiegowym opiniom na stanowisku *data scientist* mogą pracować również humaniści. Skierowanie oferty kształcenia do tych osób wydaje się więc jak najbardziej potrzebne, choćby z tego powodu, że nierzadko osoby te biegle poruszają się po źródłach informacji, mają znajomość metod i technik wyszukiwania, oceny i wizualizacji informacji, a także wiedzę z zakresu podstaw programowania i tworzenia aplikacji. Potrzeba im w zasadzie jedynie możliwości uzupełnienia wiedzy z zakresu metod analitycznych czy planowania strategicznego.

Bibliografia

- 2010 Standard Occupational Classification System*, 2018, [on-line:] https://www.bls.gov/soc/2010/2010_major_groups.htm – 30.03.2021.
- Academiccourses.com, 2020, [on-line:] <https://www.academiccourses.com/> – 30.03.2021.
- Adhikari A., DeNero J., Wagner D., *Computational and Inferential Thinking: The Foundations of Data Science*, 2021, [on-line:] <https://www.inferentialthinking.com/> – 30.03.2021.
- Birkbeck, University of London: Applied Data Science*, 2020, [on-line:] http://www.bbk.ac.uk/study/2020/postgraduate/programmes/TPCCOMIP_C – 30.03.2021.
- Birmingham City University: Big Data Analytics*, 2020, [on-line:] <https://www.bcu.ac.uk/courses/big-data-analytics-msc-2020-21> – 30.03.2021.
- Bournemouth University: Data Science & Analytics: Course Details*, 2020, [on-line:] <https://www.bournemouth.ac.uk/study/courses/bsc-hons-data-science-analytics> – 30.03.2021.
- Bournemouth University: Data Science and Analytics: Programme Specification – Purpose and Guidance for Completion*, 2020, [on-line:] <https://intranetsp.bournemouth.ac.uk/progspecs/bsc-data-science-and-analytics.pdf> – 30.03.2021.
- Brynko B., *Data Scientists: You Sexy Thing (Cover Story)*, „Information Today” 2013, vol. 30, no. 10, s. 1–36.
- Compare All Types of Data Science Degrees*, 2021, [on-line:] <https://www.datascience-programs.org/> – 30.03.2021.

- Czech University of Life Sciences: Environmental Data Science*, 2020, [on-line:] <https://www.fzp.czu.cz/en/r-9408-study/r-9495-study-programmes/r-9744-bachelor-s-study-programmes/r-14239-environmental-data-science> – 30.03.2021.
- Czech University of Life Sciences: Study plan BEDS*, 2020, [on-line:] <https://www.fzp.czu.cz/dl/71175?lang=en> – 30.03.2021.
- Danyluk A., Leidig P. i in., *Computing Competencies for Undergraduate Data Science Programs: An ACM Task Force Final Report*, [w:] *SIGCSE 2021 – Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, New York 2021, s. 1119–1120.
- Data Science for Undergraduates: Opportunities and Options*, 2018, [on-line:] <https://www.nap.edu/catalog/25104/data-science-for-undergraduates-opportunities-and-options> – 30.03.201.
- Davenport T. H, DJ Patil, *Data Scientist: The Sexiest Job of the 21st Century*, „Harvard Business Review” 2012, October, [on-line:] <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century> – 30.03.2021.
- Demchenko Y., *The Emerging Role of the Data Scientist and the Experience of Data Science Education at the University of Amsterdam*, [w:] *LEARN Toolkit of Best Practice for Research Data Management*, London 2017, s. 105–115.
- Foote K. D., *A Brief History of Data Science*, 2016, [on-line:] <https://www.dataversity.net/brief-history-data-science/> – 30.03.2021.
- Hicks S. C., Irizarry R. A., *A Guide to Teaching Data Science*, „The American Statistician” 2018, vol. 72, no. 4, s. 382–391.
- Irizarry R. A., *The Role of Academia in Data Science Education*, „Harvard Data Science Review” 2020, vol. 2, no. 1.
- Klassifikation der Berufe 2010: Systematik und Verzeichnisse der KldB 2010*, 2013, [on-line:] <https://statistik.arbeitsagentur.de/DE/Navigation/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010/Systematik-Verzeichnisse/Systematik-Verzeichnisse-Nav.html> – 30.03.2021.
- KU Leuven: Postgraduate Studies: Big Data & Analytics in Business and Management*, 2020, [on-line:] <https://feb.kuleuven.be/permanente-vorming/bigdataanalytics> – 30.03.2021.
- KU Leuven: Postgraduate Studies: Big Data & Analytics in Business and Management: Brochure*, 2020, [on-line:] <https://feb.kuleuven.be/permanente-vorming/postgraduatebigdatabrochure/> – 30.03.2021.
- MastersPortal.com, 2020, [on-line:] <https://www.mastersportal.com/> – 30.03.2021.
- Meng X.-L., *Data Science: An Artificial Ecosystem*, „Harvard Data Science Review” 2019, vol. 1, no. 1.
- Naur P., *The Science of Datalogy*, „Communications of the ACM” 1966, vol. 9, no. 7, s. 485.
- Politechnika Warszawska: Data Science*, 2020, [on-line:] <http://datascience.ii.pw.edu.pl/datascience.html> – 30.03.2021.
- Postgrad.com, 2020, [on-line:] <https://www.postgrad.com/> – 30.03.2021.
- Studies in Europe, 2020, [on-line:] <https://www.studies-in-europe.eu/> – 30.03.2021.
- Tang R., Sae-Lim W., *Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus*, „Education for Information” 2016, vol. 32, no. 3, s. 269–290.

- The Age of Analytics: Competing in a Data-Driven World*, 2016, [on-line:] <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world> – 30.03.2021.
- Tomoyuki H., *Data Scientist: A Key Factor in Innovation Driven by Big Data*, „Journal of Information Processing & Management” 2013, vol. 56, no. 1, s. 2–11.
- Universidad Pontificia Comillas: Master in Big Data Technologies and Advanced Analytics*, 2020, [on-line:] <https://www.comillas.edu/postgrado/master-en-big-data-tecnologia-y-analitica-avanzada> – 30.03.2021.
- University of Campania „Luigi Vanvitelli”: Bachelor’s Degree in Data Analytics*, 2020, [on-line:] <https://www.matfis.unicampania.it/didattica/corsi-di-studio/data-analytics/2-non-categorizzato/117-bachelor-s-degree-in-data-analytics> – 30.03.2021.
- University of Campania „Luigi Vanvitelli”: Bachelor’s Degree in Data Analytics: Study Plan*, 2020, [on-line:] https://www.matfis.unicampania.it/images/didattica/triennale_analytics/piano_studi/piano_studi_Data_Analytics_2019_2020.pdf – 30.03.2021.
- Uniwersytet Warszawski: Data Science and Business Analytics*, [on-line:] <https://www.wne.uw.edu.pl/en/candidates/data-science/> – 12.05.2020.
- Woodie A., *Why Data Science Is Still a Top Job*, 2020, [on-line:] <https://www.datanami.com/2020/11/16/why-data-science-is-still-a-top-job/> – 30.03.2021.

Data Scientists’ Higher Education in Europe

Abstract

The aim of the paper was to analyse data science educational programs at institutions of higher learning in Europe for the 2020–2021 academic year. The authors attempted to answer the following questions: What are the education options for individuals who would like to follow the professional path of data science analytics? Where and how may one obtain the required competencies? What competencies may be earned during higher studies? What professions are predestined by data science studies?

The authors analysed 100 randomly selected bachelor’s, major and post-graduate curricula found in the prospectus of various institutions of higher learning. The basic information about the offered courses were obtained through individual searches on several Internet portals (e.g., Studies in Europe, Academiccourses.com, Studia.gov.pl, Postgrad.com and MastersPortal.com). Based on the information available on the institutions’ websites, the authors grouped the data according to the following criteria: branch name, academic unit, description of curriculum (preliminary requirements, offered knowledge and skills, potential employment), duration of learning, course type, and subject taught.

The research proves the diversity in the presentation of data scientists’ education. In Europe, this curriculum is offered at bachelor level by ca. 30 institutions of higher learning, at major level – by ca. 170 institutions and at post-graduate level – by ca. 350 institutions. The most varied educational contents are offered in post-graduate programmes. All study programmes identified by the authors are profiled for the needs of certain disciplines – especially IT and Mathematics, as well as Statistics and Economics. The candidates for higher learning courses are often expected to possess advanced skills in ICT, and often also professional experience in the field of data analysis. Most courses are paid, and depending on the level of education, the courses last from one to four years. Furthermore, study programmes are dominated by IT subjects and enriched with general curriculum subject modules. Most branches make

it possible to earn knowledge and skills in the field of IT and exact sciences, especially the methods of aggregating, storing and processing of Big Data, the rules of correct and effective programming, and performing statistical analyses and interpreting their results. Among the identified branches of studies, surprisingly, there is a small number of opportunities directly addressed to graduates of the humanities, including the widely known information studies (information management, information broking, library science, information architecture, information science).

Keywords: data science, data scientist, higher education, curricula, Europe.