

ŁUKASZ SZOSZKIEWICZ

DOSTĘPNOŚĆ DANYCH
W CZASACH SZTUCZNEJ
INTELIGENCJI
A PRAWA CZŁOWIEKA
W DZIEDZINIE NAUKI



DOSTĘPNOŚĆ DANYCH
W CZASACH SZTUCZNEJ
INTELIGENCJI
A PRAWA CZŁOWIEKA
W DZIEDZINIE NAUKI

ŁUKASZ SZOSZKIEWICZ

DOSTĘPNOŚĆ DANYCH
W CZASACH SZTUCZNEJ
INTELIGENCJI
A PRAWA CZŁOWIEKA
W DZIEDZINIE NAUKI



dr Łukasz Szoszkiewicz,
Poznańskie Centrum Praw Człowieka INP PAN,
Wydział Prawa i Administracji Uniwersytetu im. Adama Mickiewicza w Poznaniu
ORCID: 0000-0001-6671-2893

Recenzenci:

prof. dr hab. Grażyna Szpor, Uniwersytet Kardynała Stefana Wyszyńskiego
dr hab. Adam Bodnar, prof. USWPS, Uniwersytet SWPS

Wydanie monografii zostało sfinansowane przez Instytut Nauk Prawnych PAN.

Redakcja i korekta językowa: Zofia Jaworowska, Wydawnictwo INP PAN

Projekt layoutu, skład i łamanie tekstu: Grzegorz Gromulski, Wydawnictwo INP PAN

Rada Naukowa Wydawnictwa INP PAN:

dr hab. Celina Nowak, prof. INP PAN
prof. dr hab. Władysław Czapliński
prof. dr hab. Irena Rzeplińska
dr hab. Anna Mlynarska-Sobaczewska, prof. INP PAN
dr hab. Mateusz Błachucki, prof. INP PAN
dr Wojciech Drobny

Seria Rozprawy Prawnicze INP PAN

Redaktor prowadzący serii: dr hab. Mateusz Błachucki, prof. INP PAN

eISBN: 978-83-66300-57-6

DOI: 10.5281/zenodo.5808312

Wydawnictwo INP PAN
ul. Nowy Świat 72, 00-330 Warszawa
e-mail: wydawnictwo@inp.pan.pl
tel.: (22) 65-72-738

Wydanie I, Warszawa 2021 © Copyright INP PAN



SPIS TREŚCI

Wykaz skrótów i skrótowców	4
Podziękowania.....	13
1. WSTĘP	15
1.1. Znaczenie podejmowanego tematu dla międzynarodowej ochrony praw człowieka.....	15
1.2. Paradygmat <i>big data</i>	19
1.3. Zakres przedmiotowy monografii	23
1.4. Cel pracy i hipotezy badawcze.....	25
2. ROSNĄCA ROLA DANYCH W NAUCE – KSZTAŁTOWANIE SIĘ PARADYGMATU <i>BIG DATA</i>.....	27
2.1. Uwagi wstępne.....	27
2.2. <i>Big data</i> – próby definicji	33
2.2.1. <i>Big data</i> jako zjawisko technologiczne.....	33
2.2.2. <i>Big data</i> jako zjawisko społeczne	35
2.2.3. <i>Big data</i> jako rewolucja w nauce	36
2.2.4. <i>Big data</i> a rewolucja w języku.....	41
2.3. Rozwój systemów opartych na algorytmach oraz SI	47
2.3.1. Algorytmy.....	49
2.3.2. Zarządzanie przez algorytmy.....	53
2.3.3. Sztuczna inteligencja	56
2.4. Dostępność danych jako kluczowy element w rozwoju paradygmatu <i>big data</i>	58

3. WYZWANIA DLA DOSTĘPNOŚCI I EKSPLORACJI DANYCH W ŚWIETLE ART. 17 I ART. 19 MPPOIP ORAZ ART. 15 UST. 1 LIT. C MPPGSIK	63
3.1. Uwagi wstępne.....	63
3.2. Dane a prawo do prywatności (art. 17 MPPOiP).....	65
3.2.1. Zasada minimalizacji danych i ograniczenia celu.....	68
3.2.2. Anonimizacja i pseudonimizacja danych osobowych.....	71
3.2.3. Prawo do przenoszenia danych.....	77
3.3. Dane a prawo do poszukiwania, pozyskiwania i upowszechniania informacji (art. 19 MPPOiP)	80
3.3.1. Zakres i forma udostępniania danych sektora publicznego	81
3.3.2. Rola państwa w zarządzaniu danymi sektora publicznego.....	83
3.3.3. Wykonalność obowiązków dotyczących zarządzania danymi	85
3.3.4. Zakres ustanawiania ograniczeń w odniesieniu do zarządzania danymi.....	87
3.4. Dane a prawa twórców (art. 15 ust. 1 lit. c MPPGSiK).....	89
3.4.1. Ochrona informacji nieujawnionej (tajemnica przedsiębiorstwa).....	91
3.4.2. Prawo autorskie	94
3.4.3. Prawo <i>sui generis</i> do ochrony bazy danych.....	100
4. BIG DATA W KONTEKCIE PRAW W DZIEDZINIE NAUKI	105
4.1. Uwagi wstępne.....	105
4.2. Zakres podmiotowy.....	111
4.3. Zakres przedmiotowy.....	116
4.3.1. Dane jako środek opisu rzeczywistości i przedmiot badań	118
4.3.2. Dane jako zasób ludzkości	119
4.3.3. Dane jako zasób ludności terytorialnej	121
4.3.4. Dane a rola państwa w obszarze ochrony praw człowieka	124
4.4. Udostępnianie korzyści postępu naukowego (art. 15 ust. 1 lit. b MPPGSiK)	127
4.5. Partycipacja w postępie naukowym (art. 15 ust. 1 lit. b MPPGSiK).....	136
4.6. Ochrona przed nadużyciami nauki (art. 15 ust. 3 MPPGSiK)	141
4.6.1. Integralność osoby jako stan podlegający ochronie	142
4.6.2. Badania behawioralne i ich zastosowania a nadużycie nauki ...	145
4.6.3. Wolność badań naukowych i działalności twórczej.....	151

4.6.4. Klauzula limitacyjna (art. 4 MPPGSiK)	153
4.6.5. Zakaz nadużywania praw (art. 5 ust. 1 MPPGSiK)	160
4.7. Tworzenie przyjaznego otoczenia dla zachowania, rozwoju i upowszechniania nauki (art. 15 ust. 2 oraz ust. 4 MPPGSiK).....	161
4.7.1. Eksterytorialny wymiar praw w dziedzinie nauki	163
4.7.2. Geopolityczny wymiar rozwoju sztucznej inteligencji	167
5. KRYTERIA REALIZACJI OBOWIĄZKÓW PAŃSTW WYNIKAJĄCYCH Z PRAW W DZIEDZINIE NAUKI NA PRZYKŁADZIE ICH WYPEŁNIENIA PRZEZ RZECZPOSPOLITĄ POLSKĄ.....	171
5.1. Uwagi wstępne.....	171
5.1.1. Dostępność fizyczna	173
5.1.2. Infrastruktura do wytwarzania i udostępniania danych sektora publicznego.....	174
5.1.3. Badania naukowe (statystyczne)	181
5.1.4. Edukacja cyfrowa	188
5.1.5. Infrastruktura w szkołach	190
5.1.6. Programy kształcenia	193
5.1.7. Budowanie kompetencji cyfrowych personelu nauczającego...194	
5.2. Jakość danych.....	196
5.2.1. Dane sektora publicznego	198
5.2.2. Dane osobowe	204
5.3. Osiągalność danych.....	207
5.3.1. Dane sektora publicznego	209
5.3.2. Dane sektora prywatnego	212
5.3.3. Prawo autorskie	214
5.3.4. Patent	217
5.3.5. Informacja nieujawniona (tajemnica przedsiębiorstwa)	223
5.3.6. Ochrona baz danych.....	225
5.4. Akceptowalność danych.....	227
5.4.1. Ochrona autonomii informacyjnej jednostki	227
5.4.2. Komunikacja ze społeczeństwem	232
5.4.3. Zakaz nadużywania nauki.....	234
5.4.4. Zakaz ingerencji władz publicznych w naukę	234
5.4.5. Zakaz nadużywania nauki przez badaczy	236
5.5. Podsumowanie.....	244

6. WNIOSKI.....	249
Wykaz aktów prawnych	257
Orzecznictwo	262
Publikacje urzędowe.....	266
Bibliografia	287
Netografia.....	308

WYKAZ SKRÓTÓW, SKRÓTOWCÓW I AKRONIMÓW

AKTY PRAWNE

Karta NZ – Karta Narodów Zjednoczonych

KEL – Kodeks Etyki Lekarskiej

k.p. – Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy

KPP – Karta praw podstawowych UE

MPPGSiK – Międzynarodowy Pakt Praw Gospodarczych, Socjalnych i Kulturalnych ONZ

MPPOiP – Międzynarodowy Pakt Praw Obywatelskich i Politycznych ONZ

PDPCz – Powszechna Deklaracja Praw Człowieka

pr.aut. – Ustawa z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych

p.w.p. – Ustawa z dnia 30 czerwca 2000 r. – Prawo własności przemysłowej

RODO – Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE

u.d.i.p. – Ustawa z dnia 6 września 2001 r. o dostępie do informacji publicznej

u.p.w.i.s.p. – Ustawa z dnia 25 lutego 2016 r. o ponownym wykorzystywaniu informacji sektora publicznego

u.s.p. – Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej

u.z.n.k. – Ustawa z dnia 16 kwietnia 1993 r. o zwalczaniu nieuczciwej konkurencji

PODMIOTY POSIADAJĄCE OSOBOWOŚĆ PRAWNĄ

BATX – akronim największych chińskich korporacji technologicznych: Baidu, Alibaba, Tencent, Xiaomi

BGH – Federalny Trybunał Sprawiedliwości (Bundesgerichtshof)

CLEPA – European Association of Automotive Suppliers

- ENISA – Agencja Unii Europejskiej ds. Cyberbezpieczeństwa (European Union Agency for Network and Information Security)
- ETPC – Europejski Trybunał Praw Człowieka
- FRA – Agencja Praw Podstawowych UE (EU Agency for Fundamental Rights)
- GAFAM – akronim największych amerykańskich korporacji technologicznych: Google, Amazon, Facebook, Apple, Microsoft
- GUS – Główny Urząd Statystyczny
- IBC – Międzynarodowy Komitet Bioetyki UNESCO (International Bioethics Committee)
- ITU – Międzynarodowy Związek Telekomunikacyjny (International Telecommunication Union)
- Komitet PGSK – Komitet Praw Gospodarczych, Socjalnych i Kulturalnych ONZ
- MC – Ministerstwo Cyfryzacji
- MKB – Międzynarodowy Komitet Bioetyki
- MOP – Międzynarodowa Organizacja Pracy
- MTPC – Międzyamerykański Trybunał Praw Człowieka
- NSA – Naczelny Sąd Administracyjny
- OECD – Organizacja Współpracy Gospodarczej i Rozwoju
- ONZ – Organizacja Narodów Zjednoczonych
- OSE – Ogólnopolska Sieć Edukacyjna
- RP – Rzeczpospolita Polska
- RPO – Rzecznik Praw Obywatelskich
- RE – Rada Europy
- SN – Sąd Najwyższy
- TK – Trybunał Konstytucyjny
- TRIPS – Porozumienie w sprawie Handlowych Aspektów Praw Właściwości Intelektualnej
- TSUE – Trybunał Sprawiedliwości Unii Europejskiej
- UOKiK – Urząd Ochrony Konkurencji i Konsumenta
- UE – Unia Europejska
- UNESCO – Organizacja Narodów Zjednoczonych dla Wychowania, Nauki i Kultury
- UNICEF – Fundusz Narodów Zjednoczonych na rzecz Dzieci
- USA – Stany Zjednoczone Ameryki
- WIPO – Światowa Organizacja Właściwości Intelektualnej
- WSA – Wojewódzki Sąd Administracyjny
- WTO – Światowa Organizacja Handlu

PUBLIKATORY

CETS – Council of Europe Treaty Series

Dz.U. – Dziennik Ustaw

Dz. Urz. UE – Dziennik Urzędowy Unii Europejskiej

ECLI – europejski identyfikator orzecznictwa (ang. *European Case Law Identifier*)

M.P. – Monitor Polski

OSNKW – „Orzecznictwo Sądu Najwyższego. Izba Karna i Wojskowa”

OTK – „Orzecznictwo Trybunału Konstytucyjnego. Zbiór Urzędowy”

OTK-A – „Orzecznictwo Trybunału Konstytucyjnego. Zbiór Urzędowy. Seria A”

UNTS – United Nations Treaty Series

ZD – „Zeitschrift für Datenschutz”

INNE

b.m.w. – brak informacji o miejscu wydania publikacji

CC – Creative Commons

COCA – Korpus Historycznego Języka Angielskiego

COHA – Korpus Współczesnego Amerykańskiego Języka Angielskiego

FRAND – licencje przyznawane na sprawiedliwych, rozsądnych i niedyskryminacyjnych warunkach (ang. *fair, reasonable, non-discrimination*)

KRI – Krajowe Ramy Interoperacyjności

PPP – partnerstwo publiczno-prywatne

prawa GSK – prawa gospodarcze, socjalne i kulturalne

PZIP – Program Zintegrowanej Informatyzacji Państwa

SI – sztuczna inteligencja

PODZIĘKOWANIA

Pomysł podjęcia tematyki dostępności danych narodził się w 2015 r., kiedy rewolucja cyfrowa wciąż znajdowała się w punkcie, w którym trudno było określić jej wpływ na nasze życie. Było to przed historycznym zwycięstwem programu AlphaGo nad człowiekiem w starochińskiej grze go, a także przed referendum ws. Brexitu oraz wyborami prezydenckimi w USA (2016 r.), w których przedsiębiorstwo Cambridge Analytica wykorzystało analitykę danych do wpływu na decyzje podejmowane przez obywateli. Z tego względu chciałbym podziękować Panu Profesorowi Zdzisławowi Kędzim, mojemu mentorowi naukowemu, za wiarę i wsparcie w realizacji celu badawczego, który wówczas uznawany był za zbyt abstrakcyjny. Dzięki otwartości Pana Profesora na nowe tematy oraz metody badawcze, a także umiejętność tworzenia atmosfery życiowej współpracy okres ostatnich sześciu lat miał decydujący wpływ na rozwój mojego warsztatu naukowego.

Podziękowania składam recenzentom rozprawy doktorskiej – Panu Profesorowi Adamowi Bodnarowi oraz Pani Profesor Grażynie Szpor. Wnikliwe recenzje były nieocenione w przygotowaniu niniejszego tekstu do publikacji.

Wyrazy wdzięczności kieruję do Pani Profesor Hanny Suchockiej oraz uczestników seminariów doktoranckich organizowanych w Zakładzie Prawa Konstytucyjnego WPiA UAM, w szczególności Moniki Bonieckiej, Artura Pietruszki, Patryka Rejsa oraz Hanny Wiczanowskiej. Prezentując wnioski ze swoich badań podczas seminariów, zawsze mogłem liczyć na życiowe i szczegółowe uwagi od Pani Profesor, a także koleżanek i kolegów.

Dziękuję także zespołowi Poznańskiego Centrum Praw Człowieka INP PAN, w szczególności dr Katarzynie Sękowskiej-Kozłowskiej, dr Aleksandrze Gliszczyskiej-Grabias, dr Grażynie Baranowskiej, dr Annie Hernandez-Połczyńskiej oraz Joannie Grygiel za stworzenie komfortowych warunków do ukończenia rozprawy doktorskiej, a także konsultacje podczas przygotowania do obrony doktorskiej oraz oddawania niniejszej monografii do druku.

Za wsparcie w podejmowaniu trudnych wyzwań oraz niegasnącą wiarę we mnie jestem wdzięczny Rodzicom. Za inspirację do podjęcia badań metodą jazykoznawstwa korpusowego, niezliczone rozmowy o nowych technologiąch i ich wpływie na nasze życie, a także cierpliwość podczas długich wieczorów spędżonych nad biurkiem jestem

wdzięczny Żonie, a na ostatnim etapie – również córce, której narodziny zbiegły się z obroną rozprawy doktorskiej.

Niniejsza książka została oparta na rozprawie doktorskiej obronionej na Wydziale Prawa i Administracji Uniwersytetu im. Adama Mickiewicza w Poznaniu w marcu 2021 r.

1. WSTĘP¹

1.1. ZNACZENIE PODEJMOWANEGO TEMATU DLA MIĘDZYNARODOWEJ OCHRONY PRAW CZŁOWIEKA

Postęp naukowy, jakiego obecnie doświadczamy, często jest określany mianem „czwartej rewolucji przemysłowej”², której motorem napędowym jest *big data*. Termin ten związany jest z wykładniczym wzrostem danych generowanych przez urządzenia elektroniczne. Szacuje się, że 90% istniejących na świecie danych zostało wygenerowanych w przeciągu ostatnich dwóch lat i z każdym rokiem ilość danych będzie się podwajać³. Połączenie dużej ilości danych ze znymi technikami ich przetwarzania, w szczególności przez algorytmy uczenia maszynowego, pozwala na tworzenie systemów komputerowych będących zdolnych z coraz większą precyzją wykonywać zadania, dla których zostało zaprojektowane m.in. rozpoznawać obiekty na fotografiach i nagraniach filmowych czy określić prawdopodobieństwo wystąpienia określonych zdarzeń w otoczeniu bądź cech u osoby. Wykładniczy wzrost mocy obliczeniowej komputerów pozwala na łączenie i przetwarzanie zbiorów danych, których analiza nie była możliwa jeszcze kilka lat temu. Najbardziej zaawansowane z technik takiej analizy (tzw. sztuczne sieci neuronowe) są często określane mianem „czarnych skrzynek” (ang. *black boxes*) lub sztuczną inteligencją ze względu na niemożność wytłumaczenia, dlaczego rezultat ich działania jest taki, a nie inny.

¹ Fragmenty rozprawy (rozdz. 4) są rezultatem badań przeprowadzonych w ramach projektu badawczego pt. „Rola nauki i danych w czasach Sztucznej Inteligencji - obowiązki państw w zakresie zarządzania danymi w świetle art. 15 ust. 1 lit. b) Międzynarodowego Paktu Praw Gospodarczych, Socjalnych i Kulturalnych” finansowanego przez Narodowe Centrum Nauki (nr rej. 2018/31/N/HS5/01592).

² Za pierwszą rewolucję przemysłową uważa się okres następujący po wynalezieniu maszyny parowej, za „drugą rewolucję przemysłową” (XIX/XX w.) i upowszechnienie się elektryczności oraz takich wynalazków jak telefon czy radio, z kolei termin „trzeciej rewolucji przemysłowej” odnosi się do upowszechnienia komputerów i Internetu. Zob. K. Schwab, *The Fourth Industrial Revolution*, Currency, Genewa 2016, s. 11-12.

³ J.S. Perry, *What is Big Data?*, <https://developer.ibm.com/dwblog/2017/what-is-big-data-insight/> [dostęp: 25.05.2020].

Pomimo wątpliwości, rosnąca precyzja tych systemów sprawia, że jesteśmy coraz bardziej skłonni do powierzania im podejmowania decyzji. Systemy wykorzystujące analitykę danych są powszechnie wykorzystywane w sądach – do oszacowania prawdopodobieństwa powrotu sprawcy do przestępstwa⁴, przez banki – do określenia wiarygodności kredytowej klientów⁵, przez media – do personalizacji treści dostarczanych użytkownikom⁶, przez administrację publiczną – do przyznawania świadczeń socjalnych i wykrywania nadużyć⁷. Oparty o analitykę danych system Prometea pomaga dzisiaj w procesie wstępnej kontroli skarg oraz przygotowania wyroków i decyzji m.in. w Miedzymerykańskim Trybunale Praw Człowieka, Trybunale Konstytucyjnym Kolumbii czy biurze Prokuratora Generalnego Buenos Aires, w szczególności w sprawach z zakresu praw gospodarczych, socjalnych i kulturalnych⁸. Jednym z najbardziej kontrowersyjnych zastosowań *big data* jest wprowadzany w Chińskiej Republice Ludowej system oceniania obywateli, tzw. System Zaufania Społecznego (ang. *Social Credit System*), w którym wysoki wynik uprawnia do uzyskania takich korzyści jak niższe oprocentowanie kredytu w banku czy przyspieszenie procedur wizowych⁹. System ten, podobnie jak i inne wymienione zastosowania wykorzystujące techniki analizy danych, były przedmiotem raportu Specjalnego Sprawozdawcy ONZ ds. prawa do prywatności, który kończy następującą konkluzją: „Dane są i pozostaną kluczowym zasobem ekonomicznym, tak jak kapitał i praca. Prywatność i innowacja

⁴ Np. algorytm COMPAS stosowany w stanie Wisconsin. Zob. opinia Sądu Najwyższego Wisconsin z 13 lipca 2016 r. w sprawie *Stan Wisconsin przeciwko Ericowi L. Loomisowi*, nr 2015AP157–CR, <https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690> [dostęp: 4.12.2021].

⁵ Np. algorytm SCHUFA, którego ujawnienie było przedmiotem postępowania przed niemieckimi sądami. Zob. wyrok BGH z 28 stycznia 2014 r., VI ZR 156/13, <https://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&nr=66910&pos=0&anz=1> [dostęp: 4.12.2021]; zob. także wyrok fińskiego Narodowego Trybunału Antydyskryminacyjnego i Równościowego z 21 marca 2018 r., 216/2017, <https://www.yvltlk.fi/en/index/opinionsanddecisions/decisions.html> [dostęp: 4.12.2021].

⁶ M. Ali, P. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, A. Rieke, *Discrimination Through Optimization: How Facebook's Ad Delivery Can Lead to Skewed Outcomes*, „Proceedings of the ACM on Human-Computer Interaction” 2019, t. 3, s. 1-30, <https://www.doi.org/10.1145/3359301>.

⁷ Np. algorytm SyRI wykorzystywany w Królestwie Niderlandów do automatycznego wykrywania nadużyć w zakresie pobierania świadczeń socjalnych. Zob. wyrok Rechtbank Den Haag z 5 maja 2020 r. w sprawie *Federation of the Dutch Trade Movement przeciwko Królestwu Niderlandów*, C-09-550982-HA ZA 18-388, <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878> [dostęp: 4.12.2021].

⁸ J.G. Corvalán, *Keynote: PROMETEA – Artificial Intelligence to Transform Public Organizations*, [w:] L. Terán, A. Meier, J. Pincay (red.), *2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG)*, IEEE Computer Society, Quito 2019, s. 15, <https://www.doi.org/10.1109/ICEDEG.2019.8734429>.

⁹ J.A. Cannataci, *Report of the Special Rapporteur on the right to privacy*, 19.10.2017, A/72/43103, pkt 104.

mogą iść i idą ramię w ramię. Zrozumienie jak skutecznie wykorzystywać *big data* i czerpać z tego korzyści, nie osłabiając przy tym ochrony praw człowieka, będzie trudne, ale ostatecznie warte wysiłku”¹⁰.

W zagranicznej literaturze przedmiotu z obszaru ochrony praw człowieka *big data* jest często podejmowanym tematem. Przeważająca część opracowań ogranicza się jednak do opisywania tego zjawiska jako zagrożenia dla prawa do prywatności jednostki¹¹. Wiele opracowań dotyczy *stricto* systemu ochrony danych osobowych funkcjonującego w Unii Europejskiej¹². Taka perspektywa znajduje swoje uzasadnienie zarówno w dynamice zmian dokonywanych w unijnym prawie, jak i w zaawansowaniu systemu ochrony. Drugim najczęściej podejmowanym tematem jest wpływ omawianego zagadnienia na prawo do ochrony zdrowia¹³. Znacznie mniej pozycji koncentruje się natomiast na szansach, jakie *big data* tworzy dla ochrony praw człowieka¹⁴. Z tego też powodu, niniejsza praca traktuje *big data* jako zjawisko będące wyzwaniem, które odpowiednio uregulowane może pozytywnie wpływać na korzystanie z praw człowieka, a uregulowane nieodpowiednio – stanowić zagrożenie dla tych praw. W monografii przyjąłem perspektywę obowiązków wynikających dla państw z traktatów w dziedzinie praw człowieka. Spojrzenie takie jest stosunkowo rzadko spotykane w literaturze przedmiotu, natomiast o tyle właściwe, że to

¹⁰ Tł. własna autora za: UN Report of the Special Rapporteur on the right to privacy, 28.02.2018, A/HRC/37/62, Appendix 4: Interim Report and Preliminary Recommendations of Big Data Open Data Thematic Action Stream Taskforce, pkt 124, http://www.ohchr.org/Documents/Issues/Privacy/SR_Privacy/2018AnnualReportAppendix4.docx [dostęp: 4.12.2021].

¹¹ V. Torra, *Data Privacy: Foundations, New Developments and the Big Data Challenge*, Springer, Cham 2017; zob. też: L. Taylor, L. Floridi, B. van der Sloot (red.), *Group Privacy: New Challenges of Data Technologies*, Springer, Cham 2017. T.M. Payton, T. Claypoole, R.H. Schmidt, *Privacy in the Age of Big Data: Recognizing Threats, Defending Your Rights, and Protecting Your Family*, Routledge, Abingdon 2017; G. Hull, *Successful Failure: What Foucault Can Teach Us About Privacy Self-Management in a World of Facebook and Big Data*, „Ethics and Information Technology” 2015, nr 17(2), s. 89-101; F.H. Cate, V. Mayer-Schönberger, *Notice and Consent in a World of Big Data*, „International Data Privacy Law” 2013, nr 3(2), s. 67-73.

¹² V. Mayer-Schönberger, Y. Padova, *Regime Change: Enabling Big Data through Europe’s New Data Protection Regulation*, „Columbia Science and Technology Law Review” 2016, nr 2(17), s. 188-228; zob. też: M. Leese, *The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union*, „Security Dialogue” 2014, nr 45(5), s. 494-511.

¹³ M. Mostert, A.L. Bredenoord, B. van der Sloot, J.M. van Delden, *From Privacy to Data Protection in the EU: Implications for Big Data Health Research*, „European Journal of Health Law” 2017, nr 24, s. 1-13; zob. też: M. Mostert, A.A. Bredenoord, M. Biesaart, J.M. van Delden, *Big Data in Medical Research and EU Data Protection Law: Challenges to the Consent or Anonymise Approach*, „European Journal of Human Genetics” 2016, nr 24, s. 956-960; E. Vayena, M. Salathé, L. C. Madoff, J. S. Brownstein, *Ethical Challenges of Big Data in Public Health*, „PLoS Computational Biology” 2015, nr 11(2), <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003904> [dostęp: 30.11.2021].

¹⁴ D. Bollier, *The Promise and Peril of Big Data*, The Aspen Institute, Washington 2010.

państwa ostatecznie ustanawiają regulacje prawne wiążące podmioty znajdujące się pod ich jurysdykcją. Tym samym zapewnienie realizacji obowiązków przez państwa powinno wzmacnić ochronę jednostek i grup, wobec których naruszeń coraz częściej dopuszczają się nie tylko podmioty publiczne, a podmioty prywatne. Obowiązki te powinny być zazwyczaj uzupełnione, biorąc pod uwagę specyfikę wyzwań, jakie stawia przed nami *big data*.

W polskiej doktrynie prawa, zagadnienia związane z analityką danych oraz rozwojem sztucznej inteligencji zaczynają być coraz częściej podejmowanym tematem, w szczególności w zakresie ponownego przetwarzania danych sektora publicznego¹⁵, prawa cywilnego i autorskiego¹⁶, a także ochrony danych osobowych i cyberbezpieczeństwa¹⁷. W 2019 r. na polskim rynku wydawniczym ukazała się monografia pt. *Internet. Analityka danych* pod redakcją G. Szpor oraz K. Czaplickiego, która uwzględnia szerokie spektrum zagadnień (m.in. analitykę danych w bankowości i ubezpieczeniach, ochronie zdrowia, administracji publicznej i cyberbezpieczeństwstwie)¹⁸. Zainteresowanie tą tematyką nasiliło się w związku z pandemią COVID-19 i cyfryzacją różnych obszarów życia społecznego¹⁹. Pojawiają się również opracowania poruszające wybrane aspekty prawne z zakresu rozwoju szeroko pojętej SI²⁰. Do refleksji nad koniecznością rewizji istniejących regulacji prawnych skłaniają również badania z zakresu nauk psychologicznych²¹. Niewiele jest natomiast opracowań

¹⁵ W ramach realizowanego w latach 2013-2016 projektu badawczo-rozwojowego pt. „Model regulacji jawności i jej ograniczeń w demokratycznym państwie prawnym” pod kierownictwem prof. dr hab. Grażyny Szpor ukazało się 12 tomów serii *Jawność i jej ograniczenia* (wyd. C.H. Beck).

¹⁶ I. Bałos, *Trenowanie wybranych modeli sztucznej inteligencji a uprawnienia twórców*, „*Studia Prawnicze. Rozprawy i Materiały*” 2019, nr 2(25), s. 15-28; R. Markiewicz, *Sztuczna inteligencja i własność intelektualna*, [w:] A. Adamczak, *100 lat ochrony własności przemysłowej w Polsce. Księga jubileuszowa Urzędu Patentowego Rzeczypospolitej Polskiej*, Wolters Kluwer, Warszawa 2018, s. 1434-1458; M. Barański, M. Jankowska, *Utwór pracowniczy powstał z wykorzystaniem sztucznej inteligencji oraz informacji sektora publicznego*, „*Zeszyty Naukowe KUL*” 2018, nr 4(244), s. 195-208; K. Gienas, *Data mining jako problem prawa autorskiego*, „*Zeszyty Naukowe Uniwersytetu Jagiellońskiego*” 2017, nr 1, s. 24-43.

¹⁷ K. Czaplicki, G. Szpor (red.), *Internet. Przetwarzanie danych osobowych*, C.H. Beck, Warszawa 2019; K. Flaga-Gieruszyńska, J. Gołaczyński, D. Szostek (red.), *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, C.H. Beck, Warszawa 2019, s. 193-240; A. Gryszczyńska, G. Szpor (red.), *Internet. Strategie bezpieczeństwa*, C.H. Beck, Warszawa 2017; A. Wójtowicz, W. Cellary, *New Challenges for User Privacy in Cyberspace*, [w:] A. Moallem, (red.), *HumanComputer Interaction and Cybersecurity. Handbook*, CRC Press, Boca Raton 2019, s. 77-98.

¹⁸ K. Czaplicki, G. Szpor (red.), *Internet. Analityka danych*, C.H. Beck, Warszawa 2019.

¹⁹ A. Gryszczyńska, G. Szpor (red.), *Internet. Cyberpandemia. Cyberpandemic*, C.H. Beck, Warszawa 2021.

²⁰ A. Chłopecki, *Sztuczna inteligencja szkice prawnicze i futurologiczne*, C.H. Beck, Warszawa 2018; zob. także: L. Lai, M. Świerczyński (red.), *Prawo Sztucznej Inteligencji*, C.H. Beck, Warszawa 2020.

²¹ A. Przegalińska, L. Ciechanowski, A. Stróż, P. Gloor, G. Mazurek, *In Bot We Trust: A New Methodology of Chatbot Performance Measures*, „*Business Horizons*” 2019, nr 62(6), s. 785-797; L. Ciechanowski, A. Przegalińska, M. Magnuski, P. Gloor, *In the Shades of the Uncanny Valley: An Experimental Study of*

opisujących zjawisko *big data* z perspektywy praw człowieka, a te istniejące najczęściej ograniczają się do analizy prowadzonej z perspektywy prawa do prywatności, prawa do swobody wypowiedzi oraz prawa dostępu do informacji publicznej²². Niewiele jest również opracowań dotyczących praw człowieka w dziedzinie nauki jako takich. Najbardziej kompleksową analizę wolności badań naukowych w prawie międzynarodowym przeprowadziła J. Rezmer, zawężając jednak swoją analizę do obszarów bioetyki, prawa morza, przestrzeni kosmicznej oraz badań w Antarktyce²³. Zagadnienie wolności badań naukowych i etyki było przedmiotem analizy w ramach unijnego projektu SATORI (2014–2017), w realizacji którego brali udział polscy badacze z Helsińskiej Fundacji Praw Człowieka²⁴.

1.2. PARADYGMAT BIG DATA

Kiedy termin *big data* pojawił się w dyskursie medialnym i akademickim, wielu przepowidało, że zniknie równie szybko jak się pojawił²⁵. Dotychczasowe obserwacje wskazują jednak, że postęp naukowy związany z wykorzystaniem danych powoduje głębokie przemiany nie tylko w kolejnych dziedzinach życia, ale rewolucję w sposobie, w jaki postrzegamy i interpretujemy otaczającą nas rzeczywistość. Coraz częściej badacze wskazują na narodziny nowego paradymatu naukowego, który opiera się na przeświadczeniu, że wiedza

Human–Chatbot Interaction, „Future Generation Computer Systems” 2019, nr 92, s. 539-548; M. Kosiński, D. Stillwell, T. Graepel, *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior*, „Proceedings of the National Academy of Sciences” 2013, nr 110(15), s. 5802-5805; M. Kosiński, *Facial Recognition Technology Can Expose Political Orientation from Naturalistic Facial Images*, „Scientific Reports” 2021, nr 11(100), <https://doi.org/10.1038/s41598-020-79310-1>.

²² G. Szpor (red.), *Internet. Cloud computing. Przetwarzanie w chmurach*, C.H. Beck, Warszawa 2013, s. 49-64, 145-154, 201-232. Rozdział drugi pracy zbiorowej pt. *Konstytucyjne wartości w chmurach* uwzględnia przede wszystkim perspektywę prywatności oraz wolności informacyjnej. Zob. też M. Grzymałowska-Cybulska, M. Cybulski, *Prawo do decydowania o zakresie i zasięgu informacji udostępnianych innym osobom na temat swojego życia w kontekście Big Data*, „*Studia Prawa Publicznego*” 2014, nr 8, s. 83-101; W.R. Wiewiórowski poruszał zagadnienia prywatności oraz dostępu do informacji publicznej. Zob. W.R. Wiewiórowski, *Założenia wstępne dla zrównoważonego przetwarzania informacji ze źródeł publicznych w czasach big data*, [w:] T. Bąkowski (red.), *Jawność i jej ograniczenia*. T. XII. *Model regulacji*, C.H. Beck, Warszawa 2016, s. 1-70. Należy zarazem zaznaczyć, że pojedyncze prace dotyczą innych obszarów praw człowieka. Por. m.in. K. Łakomiec, *Biobanki w dobie Big Data z perspektywy prawa konstytucyjnego*, „*Studia Iuridica*” 2018, nr 73, s. 105-118; K. Stępnik, *Prawo do Internetu jako środka zapewniającego partycypację w państwie demokratycznym*, „*Studia Prawnicze i Administracyjne*” 2017, nr 21(3), s. 65-70.

²³ J. Rezmer, *Wolność badań naukowych w świetle prawa międzynarodowego*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2016.

²⁴ SATORI, <https://satoriproject.eu/> [dostęp: 1.09.2021].

²⁵ J. De Goes, ‘*Big Data’ Is Dead. What’s Next?*’, <https://venturebeat.com/2013/02/22/big-data-is-dead-whats-next/> [dostęp: 24.09.2021]; L. Gavin, *Big Data Will Be Dead in 5 Years*, <https://towardsdatascience.com/big-data-will-be-dead-in-5-years-ef4344269aef> [dostęp: 24.09.2021].

„zrodzona z danych” (ang. *knowledge born from data*) pozwoli na przekroczenie dotychczasowych barier w poznaniu świata²⁶.

Pojęcie paradygmatu, wprowadzone do filozofii nauki przez amerykańskiego badacza T. Kuhna w latach 60. ubiegłego wieku, jest często kwestionowane, m.in. ze względu na brak jednej definicji²⁷. Jednocześnie jego koncepcja, którą opisuję w początkowej części pracy²⁸ wydaje mi się dobrze charakteryzować moment, w którym nauka znalazła się dzisiaj. Zgodnie z teorią przedstawioną przez T. Kuhna paradygmatem są „powszechnie uznawane osiągnięcia naukowe, które w pewnym czasie dostarczają społeczności uczonych modelowych problemów i rozwiązań”²⁹. Pozostając świadomym trudności, ale i elastyczności terminu paradygmat *big data*, zdecydowałem się użyć go w niniejszej monografii do opisu zjawisk i związanych z nimi metod badawczych, wpływających na głębokie i trwałe zmiany zachodzące w uprawianiu nauki, w tym również w uprawianiu nauk prawnych. Nowe podejście zapewne nie zastąpi badań prowadzonych tradycyjnymi, dotychczasowymi metodami, ale umożliwi przyjęcie całkiem nowej perspektywy dla poznawania istniejących zjawisk. Zmiany te można scharakteryzować przy pomocy trzech elementów:

- 1) Nowe rodzaje danych (np. teksty, obrazy). Rozpoznawanie obrazów czy przetwarzanie języka naturalnego samo w sobie nie jest czymś nowym, ale dopiero przy przekroczeniu pewnej masy krytycznej dostępnych danych i mocy obliczeniowej komputerów staje się efektywne i powszechnie, i to właśnie upowszechnienie stanowi o nazywaniu danego zespołu metod czy teorii mianem paradygmatu. Moc obliczeniowa komputerów pozwala nam obecnie traktować pojedyncze słowa w tekście jak zmienne i w oparciu o metody statystyczne wskazywać na istotne różnice pomiędzy dokumentami, współwystępowanie pewnych słów, a także umożliwia wykrywanie wzorców w danych, np. zmian w posługiwaniu się terminologią na przestrzeni czasu. Takie metody są już dzisiaj wykorzystywane do automatycznej kategoryzacji dokumentów (np. system Prometea w Trybunale Konstytucyjnym Kolumbii) czy sporządzania projektów umów (np. system LegalMation rozwijany przez informatyczne przedsiębiorstwo IBM)³⁰. Metody te mogą również pełnić istotne narzędzie w interpretacji tekstów prawnych, np. poprzez wskazywanie na podobieństwa lub różnice w terminologii w bardzo dużych zbiorach

²⁶ R. Kitchin, *Big Data, New Epistemologies and Paradigm Shifts*, „Big Data & Society” 2014, nr 1(1), s. 2.

²⁷ D. Shapere, *The Paradigm Concept*, „Science” 1971, nr 172(3984), s. 706-709.

²⁸ Zob. podrozdział 2.1.3. *Big Data jako rewolucja w nauce*.

²⁹ „[Paradigm refers to – przyp. aut.] universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of practitioners.” T. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago-Londyn 1970, s. 10.

³⁰ *Tackling a New Legal Landscape with AI*, <https://www.ibm.com/case-studies/legalmation/> [dostęp: 1.09.2021].

danych czy rekonstrukcje siatki cytowań dla identyfikacji najbardziej znaczących prac badawczych i orzeczeń sądowych³¹. Podobnie jest z obrazami, których duże zbiorów wykorzystywane są do rozwoju systemów rozpoznawania obiektów czy twarzy. Możliwość wykorzystywania tekstów i obrazów jako danych stanowi wyzwanie dla prawa własności intelektualnej, które chroni wytwory ludzkiego umysłu.

- 2) Badania napędzane przez dane (ang. *data-driven research*). Jedną z konsekwencji zjawiska *big data* w nauce jest odwrócenie tradycyjnego procesu badawczego polegającego na postawieniu hipotezy, a następnie jej weryfikacji w oparciu o starannie wybrane dane. Nowy paradymat sprawia, że najpierw aplikujemy algorytmy do dużych zbiorów danych tekstowych, a wychwycone przez nie zależności (np. podobieństwo pewnej kategorii dokumentów czy współwystępowania określonych słów) pozwala nam na postawienie czy ekstrapolację hipotezy badawczej. W takiej rzeczywistości zmianie ulega również rola badacza. Jego podstawowym zadaniem staje się bowiem selekcja i standaryzacja danych – inaczej niż w tradycyjnej nauce, w której istotnym wyzwaniem było ich zebranie³². Taki model działania można było zauważać np. w podejmowanych przez badaczy wysiłkach opracowania szczepionki na SARS-CoV-2. W jednym z badań wykorzystano moc obliczeniową najpotężniejszego komputera na świecie o nazwie Summit, a jego celem było sprawdzenie na modelach statystycznych działania kilku tysięcy różnych substancji i zidentyfikowania tych, które mają największe szanse przyznać się do opracowania szczepionki³³. Dziedzina zajmująca się problematyką analizy danych w celu znalezienia wzorców i wydobycia z nich informacji, w szczególności przy użyciu maszynowych technik przetwarzania, nazywana jest analityką danych (ang. *data science*)³⁴. Kluczowym elementem w badaniach napędzanych przez dane jest zapewnienie naukowcom dostępu do dużych zbiorów danych.
- 3) Ocena doskonałości badań naukowych przez pryzmat skuteczności algorytmu. Wartość wiedzy wydobytej z danych mierzy się precyją przewidywania wystąpienia zjawisk

³¹ Jedną z najbardziej kompleksowych monografii ukazujących różnorodne możliwości wykorzystania metod statystycznych do badania prawa przygotowali amerykańscy badacze M.A. Livermore, D.N. Rockmore (red.), *Law as Data: Computation, Text, and the Future of Legal Analysis*, SFI Press, Santa Fe 2019.

³² M. Szreder, *Nowe źródła informacji i ich wykorzystywanie w podejmowaniu decyzji*, „Wiadomości Statystyczne” 2017, nr 7(674), s. 5-17.

³³ M. Smith, J.C. Smith, *Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the Sars-Cov-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface*, „ChemRxiv” 2020, <https://chemrxiv.org/engage/chemrxiv/article-details/60c74980f96a00352b28727c> [dostęp: 30.11.2021].

³⁴ Por. definicje proponowane w: M. Niezgódka, *E-infrastruktury analityki danych: wybrane problemy strukturalne*, [w:] K. Czaplicki, G. Szpor (red.), *Internet. Analityka danych...*, op.cit., s. 11; I. Pawełoszek, *Rola kreatywności w analityce biznesowej w kontekście analizy danych marketingowych*, „*Studia i Prace Kolegium Zarządzania i Finansów*” 2018, nr 4(162), s. 91.

w oparciu o korelację różnych, często abstrakcyjnych, czynników³⁵. Jak wskazuje E. Siegel, odnosząc się do rezultatów analiz przeprowadzanych na dużych zbiorach danych: „Zazwyczaj nie wiemy nic o związku przyczynowym, i nie przejmujemy się tym [...], celem jest raczej przewidywanie niż zrozumienie świata”³⁶. Przewidywanie wystąpienia określonych zjawisk opiera się na ich powtarzalności – z tego powodu celem badaczy podejmujących analizę dużych zbiorów danych jest identyfikacja nieznanych wcześniej wzorców występujących w danych (ang. *patterns*). Proces ten, nazywany eksploracją danych (ang. *data exploration*), polega na zastosowaniu różnych metod statystycznych oraz algorytmów uczenia maszynowego, najczęściej przy niewielkiej kontroli ze strony badacza. W tym kontekście warto przywołać np. eksperymenty prowadzone przez M. Kosińskiego. W jednym z badań wykazał on, że algorytm może być znacznie bardziej skuteczny niż człowiek w przewidywaniu preferencji politycznych danej osoby (liberalnych albo konserwatywnych) opierając się wyłącznie na analizie jednego zdjęcia profilowego na Facebooku. W przeprowadzonym eksperymencie prognozy dokonane przez człowieka były nieznacznie lepsze niż przypadkowa odpowiedź (55%), z kolei algorytm oparty o sztuczne sieci neuronowe osiągnął trafność na poziomie 72%³⁷. Wiele zmiennych, jakie bierze pod uwagę algorytm jest trudnych lub wręcz niemożliwych do wyjaśnienia. W takich przypadkach o jakości badania i odkrycia decyduje nie tyle możliwość wyjaśnienia zjawiska, ale właśnie skuteczność algorytmu.

Upowszechnianie paradygmatu *big data* sprawia, że dostępność danych decyduje zarówno o kierunkach prowadzonych badań naukowych, jak i o funkcjonowaniu algorytmów wykorzystywanych przez systemy, urządzenia i aplikacje w różnych obszarach życia społecznego. Należy przypuszczać, że w niedalekiej przyszłości tendencja zastępowania ludzkiego umysłu przez komputerową analizę danych wzmacni się w związku z postępującą robotyzacją poszczególnych procesów zarządzania, produkcji, a także usług. Nadreprezentacja lub niedostateczna ilość danych o poszczególnych grupach społecznych, a także brak lub utrudniony dostęp do wybranych kategorii danych może sprawić, że niektóre grupy zamiast korzystać z nowych osiągnięć postępu naukowego, doświadczają dyskryminacji opartej na braku dostępności danych o odpowiedniej jakości.

³⁵ R. Kitchin, *Big Data...*, op.cit., s. 2.

³⁶ E. Siegel, *Predictive Analytics*, Wiley, Hoboken 2013, s. 90. Por. także W.R. Wiewiórowski, *Założenia wstępne...*, Legalis, nb. §9. Należy zarazem zauważyć, że analizy dużych zbiorów danych mogą służyć również opisywaniu świata oraz weryfikacji hipotez.

³⁷ M. Kosiński, *Facial Recognition Technology Can Expose Political Orientation from Naturalistic Facial Images*, „Scientific Reports” 2021, nr 11(100), <https://www.nature.com/articles/s41598-020-79310-1> [dostęp: 30.11.2021].

1.3. ZAKRES PRZEDMIOTOWY MONOGRAFII

Praca została przygotowana przy wykorzystaniu metody dogmatyczнопrawnej, prawnoporównawczej, historycznej, a w ograniczonym zakresie także metod językoznawstwa korpusowego. Ostatnia z wymienionych metod została użyta w celu dokonania wykładni językowej terminów „dane” (ang. *data*) oraz „informacja” (ang. *information*), a jej za stosowanie pozwoliło na wykazanie zmian w znaczeniu słów począwszy od 1810 r. aż do współczesności. Fragment pracy wykorzystującej tę metodę jest zarazem ilustracją zastosowania paradygmatu *big data* w badaniach naukowych.

Ramy prawne pracy wyznaczają przede wszystkim międzynarodowe paktyle praw człowieka, przyjęte przez ONZ w 1966 r., stanowiące podstawowe źródła międzynarodowego prawa traktatowego dotyczącego praw człowieka³⁸. Postanowienia traktatowe ustanawiają dla państw wiążące obowiązki w zakresie realizacji praw człowieka. W kontekście tematu monografii szczegółowe znaczenie mają prawo do prywatności oraz prawo do poszukiwania, pozyskiwania i upowszechniania informacji (odpowiednio art. 17 i 19 MPPPiP³⁹), a także prawa w dziedzinie kultury i nauki (art. 15 MPPGSiK)⁴⁰. Analizę przeprowadziłem na trzech płaszczyznach: 1) regulacji wypracowanych na forach organizacji międzynarodowych (przede wszystkim ONZ); 2) regulacji wypracowanych na forach organizacji regionalnych (przede wszystkim Unii Europejskiej); 3) regulacji obowiązujących w Rzeczypospolitej Polskiej.

W ramach pierwszej z tych płaszczyzn, która jest dominującą, analizie poddałem wskazane powyżej międzynarodowe paktyle, a także dorobek organów traktatowych ONZ (w szczególności Komitetu Praw Człowieka oraz Komitetu Praw Gospodarczych, Socjalnych i Kulturalnych) – zarówno pod kątem interpretacji zawartych w komentarzach ogólnych, jak i rekomendacji formułowanych w ramach rozpatrywania sprawozdań

³⁸ Traktaty te usankcjonowały prawa człowieka proklamowane w Powszechniej Deklaracji Praw Człowieka w 1948 r. W pierwotnym zamiarze deklaracja nie była dokumentem kreującym wiążące obowiązki państw. Z biegiem czasu dokument ten (lub jego wybrane postanowienia) zyskał status źródła zwyczajowego prawa międzynarodowego, które jest powszechnie wiążące. Zob. Z. Kędzia, *70 lat Powszechniej Deklaracji Praw Człowieka – pomnik czy żywy dokument?*, „Ruch Prawniczy, Ekonomiczny i Socjologiczny” 2018, nr 4, s. 5-23.

³⁹ Międzynarodowy Pakt Praw Obywatelskich i Politycznych otwarty do podpisu w Nowym Jorku dnia 19 grudnia 1966 r., Dz.U. 1977 Nr 38 poz. 167.

⁴⁰ Oficjalne polskie tłumaczenie drugiego z traktatów brzmi: Międzynarodowy Pakt Praw Gospodarczych, Społecznych i Kulturalnych, Dz.U. 1977 Nr 38, poz. 169, jednak za Z. Kędzią zdecydowałem się posłużyć w pracy bardziej precyzyjnym tłumaczeniem nazwy traktatu: Międzynarodowy Pakt Praw Gospodarczych, Socjalnych i Kulturalnych. Por. Z. Kędzia, *Wprowadzenie*, [w:] Z. Kędzia, A. Hernandez-Połczyńska (red.), *Międzynarodowy Pakt Praw Gospodarczych, Socjalnych i Kulturalnych. Komentarz*, C.H. Beck, Warszawa 2018, s. 2.

państw. Analiza objęła również wybrane dokumenty wypracowane na forum UNESCO (w szczególności Rekomendację dotyczącą statusu badaczy naukowych z 2017 r.)⁴¹ oraz Światowej Organizacji Handlu (w szczególności Porozumienie w sprawie Handlowych Aspektów Praw Własności Intelektualnej)⁴². Pomocniczo wykorzystałem dorobek Specjalnych Sprawozdawców ONZ ds. prawa do prywatności oraz ds. praw kulturalnych.

Analizie poddałem także regulacje wypracowane na forum Unii Europejskiej, które w omawianym obszarze są szczególnie rozwinięte i stanowią ramę dla rozwiązań przyjmowanych w państwach członkowskich UE, w tym również w Polsce. Szczególną uwagę poświęciłem rozporządzeniu w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych⁴³, dyrektywie w sprawie ochrony prawnej baz danych⁴⁴, dyrektywie w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego⁴⁵, a także dyrektywie w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym⁴⁶. Analizę prawodawstwa Rady Europy (RE) ograniczyłem do zagadnień z zakresu prawa do prywatności i objęła ona Protokół zmieniający Konwencję Rady Europy nr 108 o ochronie osób w związku z automatycznym przetwarzaniem danych osobowych (konwencja 108)⁴⁷ oraz Rekomendację Komitetu Ministrów RE dotyczącą wpływu systemów algorytmicznych na prawa człowieka⁴⁸.

⁴¹ UNESCO, Revision of the Recommendation on the Status of Scientific Researchers (1974), 17.05.2017, SHS/BIO/SSR/2017/1, <https://unesdoc.unesco.org/ark:/48223/pf0000248910?1=null&queryId=c4a3bbc1-3879-4e35-98df-9bf71cd07e78> [dostęp: 4.12.2021].

⁴² Agreement on Trade-Related Aspects of Intellectual Property Rights, 15.04.1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 UNTS 299.

⁴³ Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE, Dz. Urz. UE L 119 z 4.05.2016 r., s. 1-88.

⁴⁴ Dyrektywa 96/9/WE Parlamentu Europejskiego i Rady z dnia 11 marca 1996 r. w sprawie ochrony prawnej baz danych, Dz. Urz. UE L 77 z 27.03.1996 r., s. 20-28.

⁴⁵ Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego, Dz. Urz. UE L 172 z 26.06.2019 r., s. 56-83.

⁴⁶ Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/790 z dnia 17 kwietnia 2019 r. w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym oraz zmiany dyrektyw 96/9/WE i 2001/29/WE, Dz. Urz. UE L 130 z 17.05.2019 r., s. 92-125.

⁴⁷ Protokół zmieniający Konwencję Rady Europy nr 108 o ochronie osób w związku z automatycznym przetwarzaniem danych osobowych, 18.05.2018 r., CM(2018)2-final, https://www.europarl.europa.eu/doceo/document/TA-8-2019-0142_PL.pdf [dostęp: 9.12.2021]. Polska ratyfikowała go 10 czerwca 2020 r.

⁴⁸ Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, 8.04.2020, https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154 [dostęp: 4.12.2021].

Na płaszczyźnie polskiej badałem przede wszystkim przepisy Konstytucji RP oraz właściwe ustawodawstwo. Analizę przeprowadziłem przede wszystkim w ostatnim rozdziale pracy, a jej celem było wykazanie, w jakim stopniu polski ustawodawca realizuje obowiązki w zakresie zapewnienia dostępności danych, jakie można wyprowadzić z art. 15 MPPGSiK. Analizą objąłem przede wszystkim prawa w obszarze kultury i nauki (art. 73 Konstytucji RP), prawo do prywatności oraz ochrony danych osobowych (art. 47 oraz art. 51), prawo dostępu do informacji publicznej (art. 61), a także prawo własności (art. 64). Na poziomie ustawodawstwa badałem wybrane przepisy z zakresu dostępu do informacji publicznej⁴⁹, ponownego wykorzystywania informacji sektora publicznego⁵⁰, funkcjonowania statystyki publicznej⁵¹, a także wybrane aspekty prawa własności intelektualnej z obszaru prawa autorskiego⁵², prawa własności przemysłowej⁵³ oraz prawa do ochrony baz danych⁵⁴.

1.4. CEL PRACY I HIPOTEZY BADAWCZE

Celem niniejszej pracy jest analiza wpływu paradygmatu *big data* na rolę i obowiązki państwa wynikające z praw człowieka chronionych przez międzynarodowe paktu praw człowieka ONZ. Będzie ona zmierzała do określenia obowiązków państwa w zarządzaniu danymi (ang. *data governance*), w szczególności tych związanych z zapewnieniem dostępności danych oraz uregulowaniu dopuszczalności eksploracji danych przy użyciu technik uczenia maszynowego. W tym kontekście niniejsza praca zmierza do weryfikacji następujących hipotez:

- 1) Obowiązki państw wynikające z prawa do prywatności oraz prawa do poszukiwania, pozyskiwania i upowszechniania informacji (odpowiednio art. 17 i art. 19 MPPOiP) oraz prawa do ochrony interesów twórców (art. 15 ust. 1 lit. c MPPGSiK) nie zapewniają, w świetle ich obecnej interpretacji, dostatecznej dostępności danych oraz równego korzystania z osiągnięć postępu naukowego opartego na eksploracji danych przez różne grupy społeczne.
- 2) Obowiązki państw, które można wyprowadzić z praw w dziedzinie nauki ustanowionych w art. 15 MPPGSiK, mogą wzmacnić dostępność danych oraz zapewnić równe korzystanie z osiągnięć postępu naukowego opartego na eksploracji danych.

⁴⁹ Ustawa z dnia 6 września 2001 r. o dostępie do informacji publicznej, t.j. Dz.U. 2020, poz. 2176 z późn. zm.

⁵⁰ Ustawa z dnia 25 lutego 2016 r. o ponownym wykorzystywaniu informacji sektora publicznego, t.j. Dz.U. 2019, poz. 1446.

⁵¹ Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej, t.j. Dz.U. 2021, poz. 955.

⁵² Ustawa z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych, t.j. Dz.U. 2021, poz. 1062.

⁵³ Ustawa z dnia 30 czerwca 2000 r. – Prawo własności przemysłowej, t.j. Dz.U. 2021, poz. 324.

⁵⁴ Ustawa z dnia 27 lipca 2001 r. o ochronie baz danych, t.j. Dz.U. 2021, poz. 386.

W celu weryfikacji pierwszej hipotezy wskazałem luki w systemie ochrony praw człowieka powstające przy próbie uregulowania wyzwań związanych z paradygmatem *big data* tylko w oparciu o prawo do prywatności (art. 17 MPOPiP), prawo do poszukiwania, pozyskiwania i upowszechniania informacji (art. 19 MPOPiP) oraz prawa do ochrony interesów twórców (art. 15 ust. 1 lit. c MPPGSiK).

Dla weryfikacji drugiej hipotezy wskazałem, w jaki sposób zidentyfikowane wyzwania związane z *big data* mogą zostać uregulowane w oparciu o prawa w dziedzinie nauki (art. 15 MPPGSiK). Szczególną uwagę poświęciłem zagadnieniu zwiększenia dostępności (otwartości) danych. W ostatniej części monografii zaproponowałem szczegółowe kryteria oceny realizacji obowiązków w zakresie zapewnienia dostępności danych, które są rozwinięciem metodologii przedstawionej przez Komitet Praw Gospodarczych, Socjalnych i Kulturalnych ONZ (Komitet PGSK) w Komentarzu Ogólnym poświęconym prawom w dziedzinie nauki⁵⁵. Kryteria zaprezentowałem na przykładzie analizy polskiego prawodawstwa, co pozwoliło zarazem na ocenę wywiązywania się z obowiązków władz publicznych RP.

W związku z tym, że tematyka pracy dynamicznie rozwijającej się dziedziny nauki oraz techniki, w których dominuje język angielski, wiele nowych terminów nie posiada jeszcze ugruntowanych polskich odpowiedników. Niekiedy angielskie terminy mogą być tłumaczone na różne sposoby. Sam termin *big data*, mimo podejmowanych prób jego tłumaczeń, w polskiej literaturze prawniczej oraz technicznej występuje najczęściej w języku angielskim i tak też będzie przywoływany w niniejszej monografii⁵⁶. Jednocześnie w uzasadnionych sytuacjach, obok polskich tłumaczeń różnych terminów technicznych będą przywoływane ich angielskie odpowiedniki, aby czytelnik miał pewność, jakich zjawisk one dotyczą.

⁵⁵ UN Economic and Social Council, General comment No. 25 (2020) on science and economic, social and cultural rights (article 15 (1) (b), (2), (3) and (4) of the International Covenant on Economic, Social and Cultural Rights), 7.04.2020, E/C.12/GC/25, <https://undocs.org/E/C.12/GC/25> [dostęp: 4.12.2021].

⁵⁶ M. Tabakow, J. Korczak, B. Franczyk, *Big Data – Definicje, wyzwania i technologie informatyczne*, „Informatyka Ekonomiczna” 2014, nr 1(31), s. 138-153.