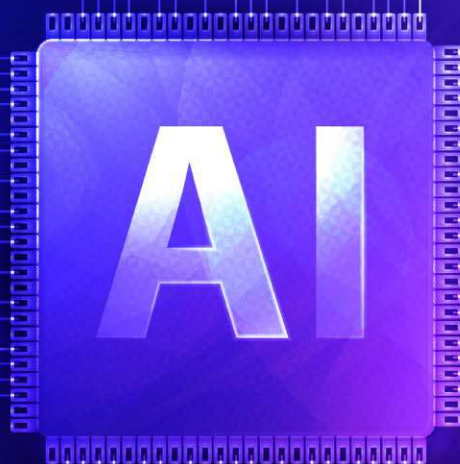


RAPORT Z BADANIA

REKOMENDACJE, ZALECENIA I DOBRE PRAKTYKI DOTYCZĄCE WYMOGÓW ETYCZNYCH WPROWADZANYCH ROZWIĄZAŃ OPARTYCH O AI W SEKTORZE BANKOWYM – ANALIZA ISTNIEJĄCYCH ROZWIĄZAŃ ZAGRANICZNYCH I MIĘDZYNARODOWYCH

Sygn. PAB/WIB/20/2024



ISBN 978-83-973493-1-5

PROF. UW DR HAB. KRZYSZTOF KOŹMIŃSKI (red.)
DR MICHAŁ JABŁOŃSKI | JAN CZARNOCKI | SEWERYN SASIN

Raport opracowany na zlecenie Programu Analityczno-Badawczego
Fundacji Warszawski Instytut Bankowości

Warszawa, 20.10.2024



PROGRAM
ANALITYCZNO
BADAWCZY

O raporcie

Raport „**Rekomendacje, zalecenia i dobre praktyki dotyczące wymogów etycznych wobec wprowadzanych rozwiązań opartych o AI w sektorze bankowym – analiza istniejących rozwiązań zagranicznych i międzynarodowych**” został opracowany na zlecenie Programu Analityczno-Badawczego Fundacji Warszawski Instytut Bankowości.

Autorzy

Prof. ucz. dr hab. Krzysztof Koźmiński

Profesor Uniwersytetu Warszawskiego, kierownik Zakładu Ekonomicznej Analizy Prawa, a także kierownik Centrum Oceny Skutków Regulacji oraz kierownik Podyplomowego Studium Zagadnień Legislacyjnych na Wydziale Prawa i Administracji Uniwersytetu Warszawskiego. Radca prawny, partner zarządzający w kancelarii Jabłoński Koźmiński.

Dr Michał Jabłoński

Doktor nauk prawnych, adiunkt w Katedrze Prawa i Postępowania Administracyjnego Uniwersytetu Warszawskiego. Adwokat, partner zarządzający w kancelarii Jabłoński Koźmiński.

Jan Czarnocki

Doktorant i stypendysta Marii Skłodowskiej-Curie w KU Leuven Centre for IT & IP Law. Zajmuje się badaniem prywatności i ochrony danych, koncentrując się na danych biometrycznych i zdrowotnych w kontekście IoT-AI, w ramach projektu Horizon 2020 Privacy Matters ITN.

Seweryn Sasin

Prawnik, doktorant w Szkole Doktorskiej Nauk Społecznych na UMK w Toruniu, przygotowuje rozprawę doktorską poświęconą wykorzystaniu sztucznej inteligencji w wydawaniu automatycznych aktów administracyjnych.

O programie

Program Analityczno-Badawczy przy Fundacji Warszawski Instytut Bankowości powstał w 2019 r. jako odpowiedź na potrzeby sektora bankowego w zakresie wysokiej jakości analiz i badań z obszaru cyberbezpieczeństwa i nowych technologii, a także szeroko rozumianego otoczenia sektora bankowego, kształtującego warunki działania banków w Polsce. Prace analityczno-badawcze w ramach programu realizowane są pod kątem możliwości praktycznego wykorzystania wyników w celu rozwoju sektora bankowego, podnoszenia poziomu bezpieczeństwa usług oraz – w ostateczności – kreowania wartości dla klientów – końcowych użytkowników bankowości.

W ramach programu realizowane są analizy i badania w następujących obszarach:

-  Ekonomiczna analiza rozwoju sektora bankowego, stabilności banków, otoczenia makroekonomicznego;
-  Nowe technologie w sektorze bankowym, cyberbezpieczeństwo banków i systemy płatnicze;
-  Prawno-regulacyjne i legislacyjne uwarunkowania działalności sektora bankowego oraz wyzwania wynikające z orzecznictwa sądów krajowych i europejskich;
-  ESG i sustainable finance;
-  Prace analityczno-badawcze w zakresie poszczególnych obszarów funkcjonowania banków oraz wskaźników referencyjnych;

Więcej na temat działalności programu na stronie www.pabwib.pl

Kontakt

dr Tomasz Pawlonka
Dyrektor Programu
m: 505 917 778
e: tpawlonka@wib.org.pl

Warszawski Instytut Bankowości
00-380 Warszawa, ul. Kruczkowskiego 8
t: (22) 182 31 70
e: pab@wib.org.pl

Agnieszka Nierodka
m: 607 484 391
e: anierodka@wib.org.pl

dr Agnieszka Wicha
m: 661 646 474
e: agnieszka.wicha@zbp.pl

Bartosz Przyborowski
m: 795 933 104
e: bartosz.przyborowski@zbp.pl

Spis treści

	Streszczenie kierownicze	5
	1. Sztuczna inteligencja w bankowości, a kwestie etyczne. Uwagi wstępne.	9
	2. Doświadczenia zagraniczne: odpowiedzi Respondentów.	16
	3. Regulacje na poziomie ponadnarodowym	25
	3.1. Zalecenia UNESCO w sprawie etyki sztucznej inteligencji	26
	3.2. Dokumenty Unii Europejskiej	28
	3.3. Stanowisko regulatorów i organów nadzoru	31
	3.4. Stanowisko przedstawicieli sektora bankowego	32
	3.5. Stanowisko partnerów społecznych	32
	3.6. Podsumowanie	33
	4. Republika Federalna Niemiec	34
	4.1. Ogólna strategia wobec sztucznej inteligencji	35
	4.2. Podejście organów regulacji i nadzoru	35
	4.3. Problematyka ochrony danych	36
	4.4. Problematyka udziału systemów AI w procesach decyzyjnych	37
	4.5. Stanowisko organizacji konsumentów	39
	4.6. Podsumowanie	39
	5. Wielka Brytania	40
	5.1. Rząd	41
	5.2. Bank Centralny Anglii i Komisja Nadzoru Finansowego	44
	5.3. Działania podejmowane przez inne instytucje w Wielkiej Brytanii	50
	5.4. Wielka Brytania: podsumowanie	50
	6. Francja	52
	6.1. Strategia i polityka państwa w zakresie rozwoju AI	53
	6.2. Działalność i rekomendacje francuskich instytucji	55
	6.3. Podsumowanie	58
	7. Hiszpania	59
	7.1. Uwagi wprowadzające	60
	7.2. Hiszpańska Karta Praw Cyfrowych w kontekście AI	60
	7.3. Hiszpańska Agencja Nadzoru nad Sztuczną Inteligencją	61
	7.4. Raport Hiszpańskiego Stowarzyszenia Bankowego	62
	7.5. Kontekst stosowania AI w bankowości w hiszpańskiej literaturze	63
	7.6. Rządowa Strategia na rzecz sztucznej inteligencji 2024	64
	7.7. Współpraca Hiszpanii z KE (pierwsze kontrolowane środowisko testowe dot. zastosowania sztucznej inteligencji wysokiego ryzyka)	65
	7.8. Działania hiszpańskiego banku centralnego Banco de España	65
	7.9. Podsumowanie	66
	8. Dania	67
	9. Włochy	74
	10. Wnioski i rekomendacje	77

Słowa kluczowe

sztuczna inteligencja, AI, SI, bankowość, sektor bankowy, etyka, compliance, soft law, rekomendacje, kodeks etyczny, legislacja, regulacja, nadzór, AI ACT, chatbot, robodoradztwo, cyberbezpieczeństwo, voicebot, biometria, automatyzacja, ryzyka, konsumenci, RODO, dane osobowe, proporcjonalność, wyjaśnialność, dyskryminacja, rynek finansowy, personalizacja

Streszczenie kierownicze





Cel przeprowadzonego badania

Celem, który postawili sobie autorzy niniejszego raportu, było udzielenie odpowiedzi na pytania o aktualny stan regulacji – w tym szczególnie kształt tzw. miękkiego prawa (*soft law*), czyli niemających charakteru wiążących norm *ius cogens*, nakładających precyzyjne prawa, obowiązki oraz wprowadzających sankcje, a zaleceń, rekomendacji, wytycznych, kodeksów etycznych, dobrych praktyk – w obszarze etyki AI w sektorze bankowym, w innych niż Polska państwach europejskich.

Jednym ze szczegółowych celów było też ustalenie, jak na tym tle wypadają rozwiązania przyjęte w polskim Kodeksie Etyki Bankowej oraz czy wnioski wynikające z analizy komparatystycznej (porównawczej) skłaniają do postulatu nowelizacji Kodeksu, zaprojektowania nowej oddzielnej regulacji dedykowanej etyce AI, czy raczej sytuacja w Polsce jest analogiczna do stanu w innych krajach tej części świata.

Odbiorcy raportu

Raport powstał z myślą przede wszystkim o: podmiotach rynku finansowego (zwłaszcza bankach, ale i towarzystwach ubezpieczeń i emerytalnych), działających w obszarze tego rynku organach władzy publicznej, a także innych organów stosujących oraz stanowiących prawo, w tym Komisji Etyki Bankowej Związku Banków Polskich.

Przedmiot analizy

Analizie poddano aktualnie stosowane rozwiązania, kwalifikowane zwłaszcza jako szeroko rozumiane „miękkie prawo” (*soft law*) odnoszące się do etyki AI w sektorze bankowym, w wybranych krajach Zachodniej Europy – państwach członkowskich Unii Europejskiej (m.in. Republika Federalna Niemiec, Francja, Dania, Hiszpania), a także Wielkiej Brytanii. Ponadto, uwzględniono analogiczne regulacje na szczeblu ponadnarodowym, w tym akty Unii Europejskiej lub Organizacji Narodów Zjednoczonych, a także literaturę naukową i poglądy formułowane przez zagranicznych ekspertów oraz reprezentantów sektora finansowego.

Zastosowano nie tylko metodę formalno-dogmatyczną (polegającą na analizie tekstów prawnych), lecz również komparatystyczną (prawno-porów-

nawczą), a nawet elementy metody empirycznej (ankieta z pytaniami przekazana wybranym Respondentom, co zaowocowało zebraniem kilkunastu wywiadów, składających się z odpowiedzi na pytania z ankiety).

Wnioski wynikające z badania

Przeprowadzone badania wykazały, że:

- Respondenci zagraniczni preferują raczej twarde i precyzyjne regulacje w prawie powszechnie obowiązującym, zamiast *soft law* lub samoregulacji;
- w większości krajów Unii Europejskiej regulacja sztucznej inteligencji na poziomie krajowym nie jest zaawansowana, ponieważ lokalni prawodawcy – w oczekiwaniu na AI Act – powstrzymywali się od wprowadzania regulacji mogących wchodzić w sprzeczność z regulacją unijną, natomiast po przyjęciu AI Act jego przepisy są raczej aktualnie analizowane w kontekście kierunków dalszego przełożenia na prawo krajowe niż już dostosowane na szczeblu krajowym;
- legislacja krajowa w obszarze AI jest obecnie w relatywnie wczesnym stadium rozwoju – Respondenci podkreślali, że wobec niedawnego przyjęcia AI Act oraz braku stosownego orzecznictwa TSUE lub sądów krajowych aktualny stan prawny określić można mianem „stanu przejściowego” lub „regulacji eksperymentalnej”, natomiast docelowo regulacje te muszą być przyjmowane w drodze ponadnarodowego dialogu (z uwzględnieniem prawa i orzecznictwa unijnego, doświadczeń innych krajów członkowskich oraz być może krajów niebędących członkami Unii Europejskiej), a także w ramach dialogu z przedsiębiorcami, ich organizacjami oraz innymi interesariuszami (eksperci, firmy technologiczne, uczelnie wyższe i instytuty naukowe, konsumenci i ich organizacje);
- wyrażany jest delikatny sceptycyzm względem idei regulacji AI w sektorze bankowym w drodze aktów prawa miękkiego (*soft law*), poparty argumentem, iż – z uwagi na delikatność materii dla indywidualnych praw i wolności (prywatność, dane osobowe itp.) – materia AI wymaga przede wszystkim jednoznacznego stanowiska ustawodawcy, który w drodze stanowczych i egzekwo-



wanych za pomocą sankcji przepisów doprecyzuje „co jest dozwolone, a co zakazane”, nie pozostawiając wątpliwości, niedomówień, ani swobody adresatom regulacji;

- aktualny stan regulacji AI w sektorze bankowym (lub szerzej finansowym) jest często przedmiotem krytyki ze względu na wielość przepisów oraz defragmentację regulacji, która rozproszona jest w kilku równolegle obowiązujących aktach prawnych na poziomie zarówno Unii Europejskiej, jak i prawa krajowego (o różnym statusie), co utrudnia uzyskanie poprawnej i spójnej ich wykładni oraz ustalenia w sposób pewny jakie są prawa i obowiązki banków stosujących rozwiązania oparte o sztuczną inteligencję;
- żaden z uwzględnionych w badaniu krajów nie wprowadził jednego uniwersalnego aktu prawnego, który regulowałby etykę AI dla sektora finansowego/bankowego;
- brak jest obecnie systemowego, zamkniętego i koherentnego zestawu takich norm zarówno w „prawie twardym”, jak i w „prawie miękkim”, odnoszącego się do etyki AI w bankowości;
- mało prawdopodobne wydaje się, aby w najbliższym czasie (szacowanym na rok-trzy lata) udało się wypracować jeden europejski (ogólnounijny) kodeks etyki AI w sektorze bankowym, choć nie oznacza to, że nie ma takiej potrzeby, a tym bardziej, że organy unijne, krajowe oraz banki powinny ze sobą współpracować w tym zakresie, prowadzić dyskusję, utworzyć forum (płaszczyzny) współpracy, dialog w celu wymiany doświadczeń itd.;
- aktualnie regulacja AI, w tym również dotycząca etyki AI, jest rozdrobniona: pojawia się w wielu różnych odrębnych aktach (dotyczących danych osobowych, praw konsumentów, kredytów bankowych, płatności *online* itp.);
- poszczególne organy i instytucje zajmujące się rynkiem finansowym (krajowe banki centralne, organy nadzoru finansowego, organizacje bankowe) przyjmują różne akty *quasi*-prawne lub prawo miękkie (a także wydają publikacje/stanowiska/komentarze eksperckie itp.) dotyczące etyki AI, ale mają one charakter ogólny, wskazujący zasady i wartości (a nie konkretne reguły);
- rozwiązania AI w bankowości zmieniają się i rozwijają tak szybko, że trudno wyobrazić sobie jeden spójny i precyzyjny kodeks etyki – raczej powinien on bazować i wyrażać pewne ogólne zasady i wartości, jak np.: odpowiedzialności i rozliczalności (ang. *responsibility and accountability*), uczciwości oraz wyjaśnialności (*bias, fairness and explainability*) itp.;
- dla etyki AI kluczowe są – jako kryteria oceny etyczności AI – pewne uniwersalne ogólne zasady i wartości, takie jak: wyjaśnialność, poszanowanie godności i praw człowieka (w tym prywatności i danych osobowych), równość, niedyskryminacja, nadzór ludzki itp.;
- doświadczenia zagraniczne państw zachodniej Europy oraz standardy na szczeblu ponadnarodowym pozwalają sformułować przypuszczenie, że nie należy spodziewać w najbliższej przyszłości wydania koherentnego, precyzyjnego i uniwersalnego kodeksu etycznego dla AI, który stanowić mógłby wzorzec dla innych dokumentów lub analogicznych aktów poszczególnych państw europejskich, w tym Polski.

Rekomendacje

Przed wszystkim sugeruje się, by Związek Banków Polskich, Komisja Etyki Bankowej, banki oraz organy władzy publicznej monitorowały prawodawstwo unijne oraz zagraniczne, obserwując tendencję legislacyjną oraz praktykę korzystania z AI.

Ponadto zaleca się promocję, wskazanych w Raporcie, ogólnych standardów, zasad i wartości z obszaru etyki AI, a także upowszechnianie wiedzy o etyce AI, w tym poprzez szkolenia, konferencje, konkursy, publikacje naukowe, raporty.

Rekomenduje się aktywny udział polskiego sektora bankowego w inicjatywach na poziomie unijnym, prowadzących do wypracowania standardów w zakresie etyki AI na szczeblu ponadnarodowym.

Zaleca się szczególnie respektowanie zasad, wartości i standardów wskazanych w Raporcie, w tym m.in.: zaangażowania człowieka w decyzje z użyciem systemów AI (ang. *human in the loop*) w krytycznych dla konsumentów obszarach, takich jak ocena wiarygodności kredytowej i decyzja o udziele-



niu kredytu, odpowiedzialność, rozliczalność, przejrzystość, wyjaśnialność, ochrona prywatności i równe traktowanie.

Istotne jest informowanie klientów banków na temat zastosowania systemów SI w ważnych decyzjach podejmowanych przez bank, w tym zwłaszcza uwzględnianych w nim parametrów, kryteriów i okoliczności. W tym celu szczególnie ważne jest proaktywne edukowanie konsumentów na temat roli systemów SI w sektorze bankowym;

Rekomenduje się ponadto jasne i precyzyjne określanie odpowiedzialności za systemy SI wewnątrz banków oraz ich transparentne komunikowanie, a także inicjowanie regularnych audytów sprawdzających – zarówno od strony technicznej, prawnej, jak i etycznej – systemy AI wykorzystywane przez podmioty sektora bankowego.

Sugeruje się również konieczność powtórnego przeprowadzenia analogicznego badania (w celu weryfikacji aktualności ustaleń i tez przedstawionych w raporcie) w przyszłości, za dwa-trzy lata.

1. Sztuczna inteligencja w bankowości, a kwestie etyczne. Uwagi wstępne.



AI
ETHICS



Sztuczna inteligencja (dalej również jako: „AI” lub „SI”) i etyka jej zastosowania jest bardzo popularnym i wzbudzającym emocje tematem. W dobie dynamicznego rozwoju technologicznego AI staje się nieodłącznym elementem wielu dziedzin naszego życia, w tym sektora bankowego. Dlatego koniecznym jest zadanie sobie pytania o etykę wykorzystania AI w tym krytycznym dla społeczeństwa sektorze gospodarki. Jednak zanim przejdziemy do sedna tego zagadnienia, ustalić najpierw należy czym jest etyka i co ona oznacza w kontekście systemów AI i dlaczego jest ważna.

Etyka, według Arystotelesa i jego słynnej „*Etyki Nikomachejskiej*”¹, jest nauką o szczęściu i dobrym życiu. Niezależnie od bogactwa definicji, koncepcji filozoficznych oraz myślicieli próbujących przeformować własne ujęcie etyki na przestrzeni dziejów² – w potocznym znaczeniu o etyce mówimy, kiedy określamy co „jest dobre a co złe”, a także więc jakie zachowania są dobre i złe, czyli moralne, dopuszczalne, akceptowalne społecznie, aprobowane albo negowane itd. Nie wchodząc w skomplikowane rozważania filozoficzne, można owo dobro i zło określać metafizycznie albo bardziej potocznie; w praktycznym rozumieniu jako zachowania dobre i złe określamy takie jakie prowadzą, lub oddalają podmiot od osiągnięcia jakichś pożądaných celów, takich jak np. rozwój cnoty, albo maksymalizacja określonych stanów, które są powszechnie aprobowane.

Tymczasem etykę bankową definiować można jako zespół „wartości, którymi powinni kierować się zarówno właściciele (zarządzający) bankiem, jak i jego pracownicy, podejmując codzienne czynności w ramach prowadzonej przez bank działalności”³ lub „standardy zachowań, które zapobiegają naruszaniu ram prawnych rynku finansowego lub przynajmniej dyktują godne wyjście z sytuacji, w których już je naruszono”⁴.

Czym natomiast jest etyka sztucznej inteligencji? Na potrzeby tego raportu będziemy rozumieli przez to relację owej technologii z życiem indywidualnego człowieka oraz społeczności ludzi, a także zachodzące zmiany, szanse i zagrożenia, które ta technologia wprowadza dla ich dobrego życia, w tym respektowania godności, wolności oraz praw. Analogicznie przez sztuczną inteligencję rozumiemy szeroką gamę rozwiązań technologicznych, które są tak klasyfikowane przez prawodawcę unijnego za pomocą ogólnej i pojemnej definicji legalnej⁵ w Rozporządzeniu Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji (dalej jako: „AI Act”)⁶, a także w drodze generalnych przepisów wprowadzonych przez prawodawców krajowych lub doktrynę prawa, tak aby objąć jak największy zbiór technologii związanych z rozwojem społeczeństwa informacyjnego, które w bardzo istotny sposób zmieniają nie tylko człowieka, ale jego relacje ze światem, i społecznym i nieożywionym.

Dlatego przez sztuczną inteligencję rozumiemy, za czołowym badaczem etyki tej dziedziny Luciano Floridim⁷, „*sprawczość bez inteligencji*”, w której podmioty (ang. *agents*) niebędące ludźmi-osobami fizycznymi mogą wykonywać zadania lub działać w określonych ramach, nie posiadając przy tym prawdziwego zrozumienia ani świadomości. Floridi wskazuje ponadto, że ci „agenci” są zdolni do generowania wyników, które mogą mieć moralne znaczenie – dobre lub złe – mimo że nie są inteligentni w świadomy sposób, w jaki są ludzie. Floridi postrzega sztuczną inteligencję jako zjawisko, które oddziela sprawczość od inteligencji.

5 Zgodnie z art. 3 pkt 1 AI Act:

„Do celów niniejszego rozporządzenia stosuje się następujące definicje:
1) „system AI” oznacza system maszynowy, który został zaprojektowany do działania z różnym poziomem autonomii po jego wdrożeniu oraz który może wykazywać zdolność adaptacji po jego wdrożeniu, a także który – na potrzeby wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie otrzymanych danych wejściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne; (...)”.

6 Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Tekst mający znaczenie dla EOG), Dz.U. L, 2024/1689, 12.7.2024, ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>, [dostęp: 2 października 2024 r.].

7 L. Floridi, *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*, Oxford 2023, s. 3, <https://doi.org/10.1093/oso/9780198883098.001.0001>, [dostęp: 15 października 2024 r.].

1 Arystoteles, *Etyka Nikomachejska*, Warszawa 2024.

2 Na ten temat przykładowo: T. Angier (red.), *Ethics: The Key Thinkers*, Londyn 2022; B. Duignan (red.), *Thinkers and Theories in Ethics The Britannica Guide to Ethics*, Nowy Jork 2011; J. Skorupski, *The Routledge Companion to Ethics*, Londyn 2012; R. Crisp (red.), *The Oxford Handbook of the History of Ethics*, Oxford 2013; J. Deigh, *An Introduction to Ethics*, Cambridge 2010; J. Suikkanen, *This is Ethics. An Introduction*, Oxford 2015.

3 A. Mierzejewska, *Etyka w bankowości*, „Język. Religia. Tożsamość”, 2/2021, s. 168.

4 V. Fetiniuc, I. Luchian, „*Banking Ethics: Main Conceptions and Problems*”, „Annals of the University of Petroșani, Economics”, 14/2014, s. 92.



Prace Floridiego zagłębiają się w koncepcję „sztucznych agentów”, które uważa za podmioty zdolne do działań w ramach moralnych, mimo że nie posiadają stanów umysłowych ani uczuć. Można ich rozumieć jako „moralnych agentów”, ponieważ ich działania wpływają na świat w sposób mający etyczne konsekwencje, chociaż nie mają oni odpowiedzialności w sensie ludzkim. Takie podejście pozwala analizować rolę SI w społeczeństwie bez potrzeby przypisywania tym systemom świadomości czy emocji podobnych do ludzkich.

To co z etycznego punktu widzenia jest kluczowe dla Floridiego to fakt, że SI wyposaża ludzi (lub instytucje, takie jak banki) w to co on nazywa „*agency on tap*”. Oznacza to nie mniej ni więcej, że po prostu SI pozwala zrobić znacznie więcej rzeczy automatycznie i na absolutnie masową skalę, bez wykorzystania pracy ludzkiej. Taka aktywność jest głównym źródłem wyzwań etycznych związanych z SI.

Nie sam tylko fakt istnienia i dynamicznego rozwoju SI jest źródłem etycznych wyzwań. Bezpośrednio powiązanym zjawiskiem, które przyczynia się do wyzwań etycznych związanych z SI jest tzw. dataizacja (ang. *datafication*) rzeczywistości. Oznacza to, że w ciągu ostatnich paru dekad, dzięki urządzeniom bazującym na technologiach informacyjnych, powstało i powstaje niewyobrażalna ilość danych.

Z tych danych SI może rozpoznawać regularności, czy prawidłowości (ang. *patterns*) i działać na ich podstawie. Oznacza to, że ludzie, czy instytucje potrafiące posługiwać się systemami SI i z wielkimi zbiorami danych (ang. *big data*) mogą mieć, po pierwsze, ogromną wiedzę, niedostępną dla tych, którzy danych i SI nie mają, a po drugie, dzięki, SI zdolność do działania na podstawie tej wiedzy.

Dochodzimy finalnie do sedna problemu i określenia jakie etyczne problemy i ryzyka tworzy SI (szanse i możliwości są ogromne, lecz nie ma przestrzeni na ich omówienie tutaj). Cofając się jednak trochę najpierw, stwierdzić należy, że generalnie bardzo dużą część problemów czy wyzwań życia społecznego można sprowadzić do faktu, że z różnych przyczyn jedni ludzie, czy instytucje, mają więcej władzy niż inni i jeżeli posiadacze tej władzy postępują nieetycznie, czyli źle i bez poszanowania dobra i godności innych osób, nadużywają w ten sposób władzy i czynią krzywdę, czy po prostu zło.

To stwierdzenie jest ważne, bo AI wyposaża posiadaczy ogromnych zbiorów danych, systemów SI

i zdolności do ich wykorzystania w nieproporcjonalną władzę względem innych ludzi, ponieważ takie podmioty mogą po prostu więcej. SI amplifikuje agencję, czyli zdolności działania, poza to co normalnie byłoby możliwe, bez pomocy SI. Mówiąc kolokwialnie: posiadacze AI, umiejący z niej korzystać i mający dużo danych, wiedzą więcej i praktycznie mogą więcej. Rodzi to po ich stronie wiele obowiązków etycznych.

Pierwszym z nich jest obowiązek samoświadomości jeśli chodzi o posiadaną władzę. Drugim takim obowiązkiem jest wstrzymanie się od nadużywania tej władzy. Dlatego dużą część uwagi etycznego wykorzystania SI należy poświęcić właśnie kwestii władzy jaką daje SI i tego jak tej władzy nie nadużyć, starając utrzymać się realistycznie równe, proporcjonalne i symetryczne relacje między posiadaczami systemów AI, a tymi na których te systemy mają wpływ.

Kolejnym po kwestii władzy i jej nadużycia czołowym problemem etyki SI jest kwestia transformacji relacji, a także środowiska w którym człowiek żyje, ze świata zdominowanego przez relacje pomiędzy człowiekiem, a człowiekiem, albo pomiędzy człowiekiem, a instytucją, do relacji pomiędzy człowiekiem i maszyną, lub pomiędzy człowiekiem poprzez maszynę do drugiego człowieka, czy pomiędzy człowiekiem przez maszynę do instytucji. Zbiorowo tę grę problemów nazywa się w etyce SI „podejmowaniem decyzji w sposób automatyczny” (ang. *automated-decision making*).

Jest to fundamentalna transformacja, ponieważ istnienie relacji między ludźmi jest jedną z najważniejszych kwestii szczęśliwego życia jako takiego. Dlatego transformacja tych relacji w relacje często z, lub przez maszynę, rodzi konkretne etyczne wyzwania. W etyce AI dlatego postuluje się o przejrzystość, rozliczalność i odpowiedzialność, jeśli chodzi o sposób wykorzystania systemów SI i decyzji jakie one podejmują, jeżeli takie decyzje mają wpływ na ludzi.

Z tych dwóch czołowych wyzwań, czyli ogromnej wiedzy i władzy podmiotów posiadających systemy SI, a także masowej transformacji relacji międzyludzkich w relacje mediowane przez maszynę, wynikają też inne problemy. Przykładowo są to ryzyka dla bezpieczeństwa prywatności czy autonomii, lub mówiąc bardziej bezpośrednio ryzyko inwigilacji i manipulowania ludźmi. Efektem może być właśnie nadużycie władzy i alienacja człowieka, poprzez eliminację relacji międzyludzkich na rzecz relacji z maszynami, lub przez maszyny.



Co to praktycznie oznacza dla sektora bankowego i etycznego wykorzystania AI? Sektor bankowy jest jedną z gałęzi gospodarki, przez którą prawdopodobnie przepływa największa ilość różnego rodzaju wartościowych danych na temat ludzi. Jednocześnie jest to sektor, który ma dostęp do kapitału i dlatego jest w stanie zbierać te dane, korzystać z SI i w ten sposób działać w stosunku do klienta, konsumenta, czy człowieka, obywatela, a także członka rodziny. Specjalnie wymieniamy te różne role społeczne, ponieważ sektor bankowy jest w stanie wpłynąć, i z, i bez wykorzystania systemów SI, na to jakie życie człowiek będzie prowadził w każdej z tych ról. Te możliwości banków SI bezprecedensowo amplifikuje.

Wyzwaniem dla sektora bankowego jest dlatego takie wykorzystanie danych i SI, które bierze pod uwagę te wszystkie zmiany jakie AI wprowadza, wykorzystuje je do zwiększenia rentowności i w interesie klienta, a jednocześnie zapobiega wszystkim ryzykom, które mogą być powiązane z nadużyciem wiedzy i władzy, a także alienacją człowieka, poprzez odseparowanie go od drugiego człowieka, nawet jeżeli tym drugim człowiekiem jest pracownik banku, lub bank jako zbiorowość ludzi i instytucja. Jest to problem stary, ze względu na jego naturę, ale nowy, jeżeli wziąć pod uwagę skalę wiedzy, władzy i możliwości podejmowania automatycznych decyzji z AI.

Kolejnym palącym problemem jest kwestia uprzedzeń i dyskryminacji (ang. *bias and discrimination*) w systemach SI. Algorytmy SI uczą się na podstawie danych historycznych, które mogą odzwierciedlać istniejące nierówności społeczne. W rezultacie mogą one nieświadomie utrzymywać te nierówności, np. odmawiając kredytów pewnym grupom społecznym na podstawie czynników, które są niesprawiedliwe lub nieistotne. Oznacza to, że człowiek nie jest traktowany jako jednostka odpowiedzialna za swoje życie i zdolna do odpowiedzialności i z indywidualną historią, lecz tylko i wyłącznie jako część bliżej niezdefiniowanej (a „uśrednionej”) zbiorowości, do której system SI, często arbitralnie, przypisał człowieka i na tej podstawie rozstrzygnął decyzje o np. wiarygodności kredytowej.

Powiązaniem z powyższym, lecz innym dylematem jest odpowiedzialność za decyzje podjęte przez AI. Jest to kwestia odpowiedzialności i rozliczalności (ang. *responsibility and accountability*). Kiedy decyzje są podejmowane przez systemy SI, pojawia się pytanie, kto ponosi za nie odpowiedzialność, zwłaszcza gdy decyzje te mają istotny wpływ na ży-

cie ludzi. W tradycyjnym modelu bankowości odpowiedzialność za decyzje spoczywa na konkretnych pracownikach czy zarządzie banku. Wprowadzenie AI komplikuje tę kwestię, ponieważ decyzje są podejmowane automatycznie przez algorytmy, które nie posiadają świadomości ani zdolności do ponoszenia odpowiedzialności w sensie moralnym. Dlatego w przypadku automatyzacji procesów decyzyjnych pojawia się pytanie, kto jest odpowiedzialny za błędy lub nieetyczne decyzje algorytmu.

Ta kwestia musi być jasno określona. Banki biorą na siebie pełną odpowiedzialność za działania swoich systemów AI. Oznacza to, że banki powinny zapewnić odpowiednie mechanizmy nadzoru nad działaniem AI, a także być gotowe do rozliczenia się z konsekwencji decyzji podejmowanych przez te systemy. Odpowiedzialność ta obejmuje zarówno aspekt prawny, jak i etyczny, wymagając od banków transparentności i uczciwości w relacjach z klientami.

Kontynuując tę refleksję, warto zwrócić uwagę na konkretne zasady etycznej sztucznej inteligencji, takie jak przejrzystość (ang. *transparency*) i wyjaśnialność (ang. *explainability*), które mają fundamentalne znaczenie w kontekście sektora bankowego. Te zasady nie tylko odpowiadają na wyzwania związane z władzą i relacjami międzyludzkimi, ale także stanowią praktyczne narzędzia do budowania zaufania między bankami a ich klientami. Przejrzystość odnosi się do otwartości i jasności w działaniu systemów SI. W kontekście bankowości oznacza to, że klienci powinni mieć dostęp do informacji o tym, jak ich dane są wykorzystywane i w jaki sposób algorytmy wpływają na decyzje podejmowane przez bank. Przykładowo, jeżeli decyzja o przyznaniu kredytu jest oparta na algorytmie AI, klient powinien wiedzieć, jakie czynniki zostały wzięte pod uwagę i dlaczego podjęto taką a nie inną decyzję.

Wyjaśnialność idzie o krok dalej, umożliwiając zrozumienie wewnętrznych mechanizmów działania algorytmów. W praktyce oznacza to, że banki powinny być w stanie przedstawić jasne i zrozumiałe wyjaśnienia dotyczące tego, jak ich systemy SI dochodzą do konkretnych wniosków. To nie tylko zwiększa zaufanie klientów, ale również pozwala na identyfikację i korektę potencjalnych błędów czy uprzedzeń w algorytmach.

Kwestia prywatności danych również stanowi istotne wyzwanie. Banki gromadzą i przetwarzają ogromne ilości danych osobowych. Wykorzystanie SI do



analizy tych danych może prowadzić do naruszenia prywatności klientów, zwłaszcza jeśli dane są wykorzystywane w sposób, o którym klienci nie są informowani lub na który nie wyrazili zgody.

Innym ważnym wyzwaniem jest ryzyko systemowe (ang. *systemic risk*), które może pojawić się w wyniku powszechnego zastosowania SI w sektorze bankowym. Gdy wiele instytucji korzysta z podobnych algorytmów i modeli SI, istnieje ryzyko, że błędy lub uprzedzenia w tych systemach będą miały szerokie, negatywne konsekwencje dla całego systemu finansowego. W skrajnym przypadku może to prowadzić do destabilizacji rynków finansowych lub wywołania kryzysu gospodarczego.

Banki muszą zatem nie tylko dbać o etyczność własnych systemów AI, ale także być świadome wpływu, jaki ich działania mogą mieć na cały sektor. Współpraca z innymi instytucjami finansowymi oraz organami regulacyjnymi jest kluczowa dla zminimalizowania ryzyka systemowego i zapewnienia stabilności gospodarczej. Dlatego zalecamy wypracowanie międzybankowych standardów związanych z parametrami systemów SI i sposobem ich wykorzystania.

Konflikt interesów (ang. *conflict of interest*) stanowi kolejne istotne wyzwanie etyczne. Systemy SI mogą być wykorzystywane do automatycznego doradztwa finansowego czy rekomendowania produktów. Istnieje ryzyko, że algorytmy będą faworyzować produkty bardziej opłacalne dla banku, a niekoniecznie najlepsze dla klienta. Może to prowadzić do sytuacji, w której dobro klienta jest podporządkowane interesom instytucji finansowej.

Aby temu zapobiec, banki powinny zapewnić, że ich systemy SI działają w sposób obiektywny i transparentny, a wszelkie rekomendacje są oparte na rzeczywistych potrzebach i korzyściach dla klienta. Etyka wymaga tutaj uczciwości i priorytetowego traktowania interesu klienta nad zyskiem własnym.

Tym bardziej, że misją i celem banku w społecznej gospodarce rynkowej jest przechowywanie kapitału oszczędzających konsumentów i mnożenie go poprzez udostępnianie go tam, gdzie może być on produktywny, czyli innym konsumentom. Dzieje się to w formie kredytu, które banki udzielają.

Funkcje te systemy SI mogą wspomóc, zwiększając produktywność, skuteczność, efektywność i zmniejszając ryzyko dla banków, lecz także

udzielając kredytów większej ilości konsumentów, poprzez lepsze określenie ryzyka i zdolności kredytowej. Jednak w zależności od użytych parametrów, determinujących systemy SI, efekt może być też odwrotny.

Dlatego to w tej właśnie centralnej funkcji banków, jaką jest udzielanie kredytów i pożyczek, najlepiej objawiają się wszystkie wyżej wymienione wyzwania etyczne wykorzystania SI w bankach i potrzeba wprowadzenia i poszanowania dla odpowiednich zasad etycznych. Zastosowanie tych zasad jest kluczowe dla banku i w jego relacjach z indywidualnymi klientami i w roli banku w ogólnym systemie finansowym.

Przykładowo systemy SI niosą etyczne ryzyko podejmowania decyzji w sposób automatyczny na masową skalę. Tworzy to problem odpowiedzialności i rozliczalności za takie decyzje, zwłaszcza w przypadku, kiedy w długiej perspektywie te decyzje mogą okazać się błędne. Dlatego banki powinny w sposób optymistycznie ostrożny podejść do wykorzystania tej technologii w tym obszarze i jasno określić i monitorować parametry na podstawie jakich decyzja jest podejmowana, z odpowiednim audytem danych na podstawie, których to się dzieje. Oznacza też to ostrożność w wykorzystaniu systemów AI, których wyjaśnialność jest ograniczona.

Kolejnym istotnym aspektem są wymogi przejrzystości (*transparency*) oraz wyjaśnialności (*the principle of explicability*) systemów AI, co – jak wynika ze stanowiska Niezależnej Grupy Ekspertów Wysockiego Szczepła ds. Sztucznej Inteligencji (Grupa Ekspertów ds. AI)⁸ – oznacza, że „procesy muszą być przejrzyste, możliwości i cele systemów SI otwarcie komunikowane, a decyzje w jak największym stopniu możliwe do wyjaśnienia osobom, na które mają one bezpośredni i pośredni wpływ”⁹.

Prywatność i jej ochrona to kolejne z wyzwań na polu etycznych relacji banków ze swoimi klientami. Również i w tym zakresie odwołać się można do

8 Grupa Ekspertów Wysockiego Szczepła ds. Sztucznej Inteligencji. Wytoczne w zakresie etyki dotyczącej godnej zaufania sztucznej inteligencji, https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI-PL.pdf [dostęp: 2 września 2024 r.].

9 *Ibidem*, s. 16.
Na ten temat zob. też: A. Auleytner, M.J. Stępień, Wyjaśnialność sztucznej inteligencji a prawo autorskie. Rozważania na tle projektowanych uregulowań prawnych możliwości wyjaśnienia systemów sztucznej inteligencji, „Prawo Nowych Technologii”, 1/2021, s. 46 i n.



rekomendacji Grupy Ekspertów ds. AI, postulujących, by wykorzystywane systemy SI gwarantowały ochronę prywatności i danych przez cały swój „cykl życia”. Obejmuje to zarówno informacje podane początkowo przez użytkownika, lecz także dane na temat użytkownika w trakcie jego interakcji z systemem (np. wyniki wygenerowane przez system AI dla konkretnych użytkowników lub ich preferencji). W związku z tym, że dzięki cyfrowym rejestrom zachowań ludzkich systemy SI będą mogły ustalić nie tylko indywidualne preferencje jednostek, ale prawdopodobnie również ich cechy osobiste, orientację seksualną, wiek, płeć, poglądy religijne lub polityczne – należy zapewnić, aby dane gromadzone na ich temat nie były wykorzystywane do dyskryminowania ich w sposób niesprawiedliwy, arbitralny lub bezprawny¹⁰.

Tak samo, banki muszą samoograniczać się, jeśli chodzi o ilość gromadzonej wiedzy o klientach, tak aby nie móc manipulować nimi, oferując niepotrzebne i niezgodne z interesami klientów produkty i usługi, wykorzystując podatność klientów spowodowaną asymetrią wiedzy i możliwości technologicznych między klientem, a bankiem, umożliwiającymi różne formy manipulacji decyzjami klientów.

Aby sprostać tym wyzwaniom, konieczne jest opracowanie i wdrożenie etycznych ram dla wykorzystania SI w sektorze bankowym. Obejmuje to nie tylko ustanowienie zasad przejrzystości i wyjaśnialności, ale także stworzenie mechanizmów nadzoru i kontroli nad działaniem algorytmów.

Banki powinny inwestować w rozwój tzw. „wyjaśnialnej sztucznej inteligencji” (ang. *explainable AI*), która łączy skuteczność z możliwością interpretacji decyzji. Ponadto, konieczne jest przeprowadzanie regularnych audytów etycznych i technicznych algorytmów, aby wykrywać i eliminować potencjalne uprzedzenia.

Ważne jest również angażowanie różnych interesariuszy, w tym klientów, regulatorów i ekspertów etycznych, w proces tworzenia i wdrażania systemów SI. Taka wielostronna współpraca może pomóc w identyfikacji potencjalnych problemów i wypracowaniu rozwiązań, które będą akceptowalne społecznie.

W ciągu najbliższych lat sektor bankowy staje również przed wyzwaniem wdrożenia AI Act, który reguluje sposób wykorzystania takich systemów, które stanowią ryzyko dla praw podstawowych. Akt ten można interpretować jako etyczne minimum wymagane od banków. Jednak wiele norm tego aktu pozostawia szeroką sferę wolności co do dokładnego sposobu wdrożenia i ograniczania ryzyka, a także ochrony praw podstawowych.

Sztuczna inteligencja niesie ze sobą ogromny potencjał dla sektora bankowego, ale jednocześnie stawia przed nami poważne wyzwania etyczne. Przejrzystość i wyjaśnialność to kluczowe zasady, które mogą pomóc w zbudowaniu zaufania i zapewnieniu, że SI będzie służyć dobru człowieka. Jednak wdrożenie tych zasad wymaga świadomego działania i zaangażowania ze strony banków, regulatorów, reprezentantów środowiska naukowego oraz społeczeństwa.

Niniejszy raport stanowi kontynuację oraz rozszerzenie wcześniejszych badań, których wyniki prezentował raport pt. „Aktualne wyzwania prawne związane z zastosowaniem sztucznej inteligencji w polskim sektorze bankowym”, również opracowany na zlecenie Programu Analityczno-Badawczego Fundacji Warszawski Instytut Bankowości oraz opublikowany w kwietniu 2024 r.¹¹.

Celem wspomnianego opracowania było ustalenie odpowiedzi na m.in. następujące pytania: na ile rozwiązania oparte o sztuczną inteligencję są w praktyce aktualnie wykorzystywane w polskim sektorze bankowym, jak banki krajowe wypadają na tle doświadczeń zagranicznych, a także czy przygotowane są na poważne zmiany legislacyjne związane z wejściem w życie kompleksowej regulacji unijnej w postaci tzw. AI Act. Nowatorskim elementem ówczesnych badań było podjęcie bezpośredniego dialogu z przedstawicielami polskiego sektora bankowego, tj. przeprowadzenie wywiadów pogłębionych z respondentami – reprezentantami środowiska finansowego, którzy dzieląc się własnymi doświadczeniami, przemyśleniami i przewidywaniami, prezentują oceny na temat zalet i wad nowoczesnej

10 Grupa Ekspertów Wysokiego Szczebla ds. Sztucznej Inteligencji. Wytyczne w zakresie etyki..., s. 21.

11 K. Koźmiński (red.), *Aktualne wyzwania prawne związane z zastosowaniem sztucznej inteligencji w polskim sektorze bankowym*, sygn. PAB/WIB 8/2024, Warszawa 2024, https://pabwib.pl/wp-content/uploads/2024/01/PAB_WIB_Wyzwania_prawne_AI_Kozminski.pdf [dostęp: 2 września 2024 r.].



technologii (w tym zwłaszcza sztucznej inteligencji), formułowali rekomendacje oraz postulaty *de lege ferenda* skierowane do prawodawcy oraz instytucji nadzorczych.

Inaczej niż poprzednio, niniejszy Raport koncentruje się nie tyle na ramach prawnych wytyczonych przez prawodawcę, tj. organy władzy publicznej, które w drodze konkretnych i na ogół precyzyjnych przepisów *ius cogens* określają prawa, obowiązki oraz sankcje – ale na standardach, wytycznych, zaleceniach, rekomendacjach, kodeksach etycznych i innych podobnych normach, niemających charakteru wiążącego „twardego” prawa (czyli ustaw, rozporządzeń itd.), ale „miękkiego”, czyli tzw. *soft law*. W polskich warunkach aktem takim jest np. Kodeks Etyki Bankowej (Zasady Dobrej Praktyki Bankowej)¹², a zatem dokument który nie wyznacza wzorca zachowania nakazanego i wymaganego pod groźbą sankcji-kary przez państwo, ale jest raczej głosem sektora, wyrazem nie tego co „trzeba”, „zakazane” lub „obligatoryjne”, ale „sugerowane”, „powinno się”, „zaleca” itd. Warto dostrzec, że również i w owym Kodeksie – choć zdawkowo i ogólnie – znalazły się dyrektywy odnoszące do innowacji i nowych technologii (fragment VI Kodeksu Etyki Bankowej).

Już w tym miejscu podkreślić trzeba, że jedną z poszukiwanych odpowiedzi w trakcie prowadzonych badań była kwestia ewentualnej konieczności rozbudowy polskiego Kodeksu Etyki Bankowej o bardziej szczegółowe rekomendacje odnoszące się właśnie do etyki AI.

Biorąc pod uwagę fakt, że niniejszy Raport powstał nie tylko na bazie analizy aktów prawnych, dostępnej literatury, opracowań eksperckich oraz innych

dokumentów pozostających w domenie publicznej (np. raporty instytucji publicznych, stanowiska organów władzy publicznej, oficjalne wypowiedzi decydentów lub ich reprezentantów), lecz również korespondencji mailowej z instytucjami kompetentnymi w obszarze bankowości, sektora finansowego oraz AI z zagranicznych krajów Europy Zachodniej – **Autorzy Raportu czują się w obowiązku złożyć podziękowania wszystkim Instytucjom oraz Osobom, które udzieliły odpowiedzi na pytania przekazane w ankiecie, a także podzieliły się innymi informacjami, wiedzą, opiniami, w tym własnymi (a bardzo cennymi i inspirującymi) przemyśleniami na temat etyki AI.**

Podziękowania są tym bardziej uzasadnione, że sformułowane pytania nie były przedstawiane w ramach oficjalnych trybów, nakładających na Respondentów obowiązek odpowiedzi (analogicznie do np. procedury uregulowanej w polskich przepisach ustawy z dnia 6 września 2001 r. o dostępie do informacji publicznej¹³), a odpowiedzi Respondentów z zagranicy były merytorycznie wartościowe, precyzyjne, w tym niektóre z nich rozbudowane i przekrojowe, uwzględniające zarówno lokalny kontekst prawny, jak i technologiczny oraz społeczno-gospodarczy. Ich udzielenie wymagało czasu oraz zaangażowania, co wymaga docenienia jako samej w sobie – niezależnie od dobrych praktyk i rekomendacji niniejszego Raportu – dobrej praktyki, zasługującej na promocję w sektorze bankowym, jako otwartość, jawność, życzliwość oraz woła pomocy.

Wnioski z owej analizy, w tym wybrane przetłumaczone wypowiedzi, ze wskazaniem Respondentów, zostaną zaprezentowane w Rozdziale 2 Raportu pt. „Doświadczenia zagraniczne: odpowiedzi Respondentów”.

12 https://zbp.pl/getmedia/c2d2f502-0cfd-4ee3-9bab-6ab8a4db4ce3/Kodeks-Etyki-Bankowej-2023?_gl=1*1cwsq70*_up*MQ..*_ga*MTgzMTYwMzA1NC4xNzI5NDMzMjMw*_ga_VLYTTWXB1S*MTcyOTQzMzIzMC4xLjAuMTcyOTQzMzIzMC4wLjAuMA.. [dostęp: 2 września 2024 r.].

13 Dz.U. z 2022 r., poz. 902.

2. Doświadczenia zagraniczne: odpowiedzi Respondentów





Niezależnie od przeprowadzonej analizy zagranicznych źródeł oraz literatury w wybranych państwach Unii Europejskiej, zaprezentowanej w dalszych częściach Raportu – zespół Autorów Raportu zwrócił się z pytaniami skierowanymi do kilkudziesięciu instytucji (odpowiedników polskich: Związku Banków Polskich i Komisji Etyki Bankowej przy Związku Banków Polskich, Narodowego Banku Polski, Urzędu Ochrony Konkurencji i Konsumentów, Komisji Nadzoru Finansowego, Urzędu Ochrony Danych Osobowych, Rzecznika Finansowego i innych, a także wybranych organizacji pozarządowych, do których zadań statutowych należy problematyka sektora finansowego lub zagadnienia AI) oraz kilku uznanych badaczy akademickich z różnych państw europejskich (w tym Anglii, Austrii, Portugalii, Belgii, Hiszpanii, Włoch), w celu uzyskania od nich bezpośrednio informacji oraz ocen dotyczących krajowej regulacji AI w sektorze bankowym, ewentualnych *soft law*, a także formułowanych przez krajowych ekspertów ocen i postulatów.

Prośba o udzielenie informacji, wraz z krótką ankietą otwartych pytań (pozostawiających respondentowi swobodę wypowiedzi) oraz syntetyczną informacją o realizowanym projekcie badawczym, została wysłana w drodze wiadomości mailowych do grupy liczącej łącznie siedemdziesięciu dwóch (72) adresatów, a następnie poszerzona – z powodu zidentyfikowania dodatkowych adresów mailowych potencjalnych odbiorców, przekierowania przez pierwotnych adresatów lub wskazania innych osób kompetentnych – o dwudziestu (20) dodatkowych adresatów.

Choć zdecydowana większość z nich nie odpowiedziała na pytania, udało się uzyskać kilkanaście odpowiedzi Respondentów z następujących krajów: Bułgarii (Evdokiya Slavcheva - *International Coordination Directorate-Financial Supervision Commission* z siedzibą w Sofii), Chorwacji (anonimowy respondent z ogólnego maila informacyjnego-*Hrvatska Narodna Banka* z siedzibą w Zagrzebiu), Irlandii (Fiona Haughey - *Communications Associate* z *Central Bank of Ireland*), Holandii (Frans van Bruggen z *De Nederlandsche Bank*); Węgier (dr Vig Gyula-Magyar Nemzeti Bank); Hiszpanii (Lorena Mullor z *Asociación Española de Banca*); Rumunii (Cati Ursu - *Executive Director* z *Romanian Association of Banks*); Finlandii (Tatu Räsänen - *Digitalisation Economist in the Financial Stability and Statistics Department* oraz Heli Snellman - *Head of Cash Division in the Payment Systems Department* z Banku Finlandii), Danii (Cammilla Neuenschwander - *Chief Special Adviser* z Da-

nish Financial Supervisory Authority), a także odpowiedzi merytoryczne osób niepodpisanych z imienia i nazwiska, przekazane z oficjalnych skrzynek mailowych. Respondentami byli w przeważającej części osoby reprezentujące oficjalnie instytucje publiczne.

Odpowiedzi sformułowane były w różny sposób – zarówno pod względem ich objętości, złożoności oraz formy (niektóre bardziej rozbudowane, inne lakoniczne, niektóre udzielone na oficjalnych formularzach, zawierające załączniki wraz z aktami normatywnymi lub raportami, inne w formie krótkiej odpowiedzi mailowej sporządzonej w nieformalny sposób).

Podkreślić trzeba jednak, że:

- kilkakrotnie **akcentowano wagę zagadnienia, wskazując równoległą działalność kilku centralnych krajowych organów władzy publicznej, które są w obszarze AI w sektorze finansowym właściwe** (m.in. właściwych w obszarze nadzoru nad rynkiem finansowym, ochrony konsumenta, ochrony danych osobowych);
- wśród odpowiedzi dominowało stanowisko, zgodnie z którym regulacja sztucznej inteligencji na poziomie krajowym nie jest w danym państwie zaawansowana, ponieważ **lokalni prawodawcy – w oczekiwaniu na AI Act – powstrzymywali się od wprowadzania regulacji mogących wchodzić w sprzeczność z regulacją unijną, natomiast po przyjęciu AI Act jego przepisy są raczej aktualnie analizowane w kontekście kierunków dalszego przełożenia na prawo krajowe niż już dostosowane na szczeblu krajowym**;
- niemal wszyscy Respondenci podkreślali **wagę AI Act, a także doniosłość zmian technologicznych związanych z rozwojem sztucznej inteligencji dla branży finansowej**, w tym zwłaszcza sektora bankowego, wskazując zarówno związane z nim szanse, jak i zagrożenia (w tym etyczne);
- niektórzy Respondenci wskazywali na **trudność w jednoznacznej ocenie (również z punktu widzenia etyki) poszczególnych rozwiązań stosujących AI**, podkreślając, że **dopiero w przyszłości, dzięki zdobytemu doświadczeniu na konkretnych przypadkach praktycznych, możliwe będzie sformułowanie ram prawnych i etycznych dla sztucznej inteligencji w sektorze bankowym**;

- niektórzy Respondenci w swoich odpowiedziach koncentrowali się na obowiązującym „twardym prawie”, tj. zarówno przyjętym AI Act lub przepisach krajowych o charakterze *ius cogens*, które wytyczają w sposób wiążący prawa i obowiązki związane ze stosowaniem sztucznej inteligencji, pomijając lub zdawkowo odnosząc się do „prawa miękkiego” (*soft law*);
- na podstawie większości odpowiedzi sformułować można generalny pogląd, iż legislacja krajowa w obszarze AI jest obecnie w relatywnie wczesnym stadium rozwoju – Respondenci podkreślali, że wobec niedawnego przyjęcia AI Act oraz braku stosownego orzecznictwa TSUE lub sądów krajowych **aktualny stan prawny określić można mianem „stanu przejściowego” lub „regulacji eksperymentalnej”, natomiast docelowo regulacje te muszą być przyjmowane w drodze ponadnarodowego dialogu** (z uwzględnieniem prawa i orzecznictwa unijnego, doświadczeń innych krajów członkowskich oraz być może krajów niebędących członkami Unii Europejskiej), **a także w ramach dialogu z przedsiębiorcami, ich organizacjami oraz innymi interesariuszami (eksperci, firmy technologiczne, uczelnie wyższe i instytuty naukowe, konsumenci i ich organizacje)**;
- w odpowiedziach kilku Respondentów **wyrażony został delikatny sceptycyzm względem idei regulacji AI w sektorze bankowym w drodze aktów prawa miękkiego (*soft law*)**, poparty argumentem, iż – z uwagi na delikatność materii dla indywidualnych praw i wolności (prywatność, dane osobowe itp.) – **materia AI wymaga przede wszystkim jednoznacznego stanowiska ustawodawcy, który w drodze stanowczych i egzekwowanych za pomocą sankcji przepisów doprecyzuje „co jest dozwolone, a co zakazane”, nie pozostawiając wątpliwości, niedomówień, ani swobody adresatom regulacji**;
- część Respondentów, relacjonując aktualny stan regulacji AI w sektorze bankowym (lub szerszej finansowym), akcentowała **wielość przepisów oraz defragmentację regulacji, która rozproszona jest w kilku równolegle obowiązujących aktach prawnych na poziomie zarówno Unii Europejskiej, jak i prawa krajowego (o różnym statusie), co utrudnia uzyskanie poprawnej i spójnej ich wykładni oraz ustalenia w sposób pewny jakie są prawa i obowiązki banków stosujących rozwiązania oparte o sztuczną inteligencję**.

Przykładowe przetłumaczone (poniżej oraz w dalszej części Raportu tłumaczenia przeprowadzone przez Autorów Raportu) wypowiedzi:

„Z uwagi na fakt, że chorwacka ustawa o sztucznej inteligencji nie została jeszcze wdrożona, instytucja krajowa nie została jeszcze zdefiniowana, lecz – ponieważ chorwacki bank narodowy jest właściwym organem, który zgodnie z aktualnym stanem prawnym weryfikuje modele wewnętrzne w bankach i sprawuje nadzór nad bankami – prawdopodobnie to właśnie nasz krajowy bank centralny będzie on odpowiedzialny za wdrożenie ustawy o sztucznej inteligencji w bankach. (...) Podkreślam jednak, że nie jest to informacja wiążąca, ponieważ proces legislacyjny w tym obszarze jeszcze się nawet nie rozpoczął”.

Wiadomość przekazana przez Departament Komunikacji Chorwackiego Banku Narodowego (Hrvatska Narodna Bank)

„Wytyczne etyczne dotyczące wykorzystania AI w bankach/szerszym kontekście społecznym są wprowadzane na szczeblu Unii Europejskiej przez Komisję Europejską w sposób jasny i ogólnodostępny. AI Act odnosi się bezpośrednio do prac Komisji w motywach: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> Aktualnie nie ma konkretnych szczegółowych krajowych przepisów, które stanowiłyby ramy prawne dla etyki mechanizmów AI stosowanych w sektorze bankowym w Irlandii.

Jednak nasze regulacje bazują na stanowisku „neutralnym technologicznie”, co oznacza, że – niezależnie od używanej technologii – firmy muszą przestrzegać obiektywnie obowiązujących przepisów prawa. W dziedzinie etyki Centralny Bank Irlandii wydał kilka publikacji, które mogą być interesujące:

*Consumer Protection Outlook Report:
<https://www.centralbank.ie/regulation/consumer-protection/consumer-protection-outlook-report>*

*Consumer Protection Code Review:
<https://www.centralbank.ie/regulation/consumer-protection/consumer-protection-codes-regulations/consumer-protection-code-review>*

*Individual Accountability Framework:
<https://www.centralbank.ie/regulation/how-we-regulate/individual-accountability-framework> (...).*



Zagadnienie wykorzystania AI w sektorze bankowym nie było do tej pory przedmiotem naszych analiz, choć w zeszłym roku Centralny Bank Irlandii przeprowadził badania dotyczące związków etyki oraz digitalizacji w sektorze ubezpieczeń – w razie zainteresowania, prosimy o zapoznanie się z raportem tutaj:

Data Ethics Within Insurance:

https://www.centralbank.ie/docs/default-source/regulation/industry-market-sectors/insurance-re-insurance/solvency-ii/communications/data-ethics-within-insurance.pdf?sfvrsn=54219f1d_4

Dostosowanie prawa krajowego do AI Act jest przedmiotem prac irlandzkiego rządu, niedawno odpowiedzialny za to departament prowadził konsultacje publiczne w tym obszarze, tu znajdziesz więcej informacji na ten temat: <https://enterprise.gov.ie/en/consultations/public-consultation-on-eu-ai-act.html>

W konsultacjach wziął udział również Bank Centralny – w którego stanowisku zostały zaakcentowane kwestie ryzyka stronnictwa, uczciwości oraz wyjaśnialności (bias, fairness and explainability). Tutaj znajdziesz pełną informację o stanowisku Banku Centralnego: https://edit.centralbank.ie/docs/default-source/regulation/innovation-hub/central-bank-of-ireland-response-to-the-public-consultation-on-the-ai-act.pdf?sfvrsn=2aaa611a_3.

Wiadomość od Fiona Haughey,
Central Bank of Ireland

„Obecnie w Holandii brak jest prawnej regulacji etyki AI stosowanej w sektorze bankowym. Nie przyjęły takich ani Holenderski Bank Centralny, ani Urząd Nadzoru Rynków Finansowych, ani też organ odpowiedzialny w obszarze prywatności danych, ani instytucje kompetentne w sferze gwarancji praw człowieka.

Obecnie EBA pracuje nad wytycznymi AI-Act dla sektora bankowego, w tym nad uczciwością algorytmów”.

Wiadomość od Frans van Bruggen
z De Nederlandsche Bank

„Centralny Bank Węgier podziela pogląd, że ważne jest, aby systemy sztucznej inteligencji wykorzystywane w sektorze bankowym były

zgodne ze standardami etycznymi, dlatego nieustannie poszukujemy sposobów promowania etycznej sztucznej inteligencji w naszej pracy nadzorczej. (...)

Nie mamy wiedzy o żadnych ramach regulacyjnych specjalnie promujących wykorzystanie etycznej sztucznej inteligencji. Zgodnie z punktami 5 i 6 badania „State of AI Hungary 2023: How do companies harness the power of artificial intelligence?” (<https://www2.deloitte.com/hu/en/pages/risk/articles/ai-study.html>) większość zaangażowanych firm nie ma żadnej strategii w tym zakresie i przeprowadziła swoje prace rozwojowe, zarządzając ryzykiem specyficznym dla sztucznej inteligencji na niskim poziomie. Te stwierdzenia również potwierdzają aktualność Waszych (tj. Autorów Raportu) tez.

Jednak w 2020 r. rząd przyjął „Hungary’s Artificial Intelligence Strategy 2020-2030” (Strategia; <https://ai-hungary.com/files/api/v1/companies/15/files/146074/download.pdf>).

Punkt 4.1.6 Strategii określa cel strategiczny ustanowienia krajowych ram regulacyjnych i etycznych, tj. „stworzenie skutecznego, wydajnego, zorientowanego na człowieka i wspierającego krajowego środowiska regulacyjnego dla działania sztucznej inteligencji, wraz z niezbędnymi ramami etycznymi; uwzględniając odpowiednie przepisy UE”. Wraz z przyjęciem Strategii rząd utworzył stowarzyszenie instytucji rządowych i uczestników rynku (Koalicja AI), które pełni również funkcję centrum kompetencji. Koalicja AI powołała specjalną grupę roboczą ds. regulacji i etyki (Workgroup), która jest gotowa publikować zalecenia dotyczące omawianej kwestii.

Centralny Bank Węgier jest bardzo skoncentrowany na kwestii etycznej sztucznej inteligencji. Niedawno przeprowadziliśmy ankiety i konsultowaliśmy się z instytucjami kredytowymi na temat rozwoju sztucznej inteligencji. Na podstawie tych opinii i międzynarodowych najlepszych praktyk, Centralny Bank Węgier zamierza wesprzeć instytucje kredytowe, wydając sektorowy, niewiążący pakiet rekomendacji i wytycznych, aby zapewnić elastyczne ramy promujące przejrzyste, niezawodne i zorientowane na klienta wykorzystanie sztucznej inteligencji w sektorze bankowym. Obecnie przygotowujemy nowelizację Rekomendacji: Recommendation No 4/2021. (III. 30.) of the Magyar Nemzeti Bank on the digital transfor-

mation of credit institutions (<https://www.mnb.hu/letoltes/mnb-recommendation-4-2021-en.pdf>) w tym zakresie, która ma zostać wydana w tym roku po szerokich konsultacjach branżowych.

Chcielibyśmy również wspomnieć, że Węgierski Urząd Ochrony Danych (NAIH) zajmuje się kwestiami zgodności AI w indywidualnych przypadkach ochrony danych, a Węgierski Urząd ds. Konkurencji (GVH) rozpoczął analizę rynku na początku tego roku na temat wpływu AI na konkurencję rynkową i zachowania konsumentów. Ta ostatnia analiza może mieć szczególne znaczenie w sektorze bankowym, ponieważ badane ryzyka dotyczą sektora bankowego i całego systemu finansowego. Ponadto AI Coalition działa aktywnie w obszarze podnoszenia świadomości na temat AI i wdrażania Strategii.

Na koniec, rząd węgierski właśnie przyjął dekret rządowy 1301/2024 (IX.30) w sprawie środków wymaganych do wdrożenia ustawy o sztucznej inteligencji. Zgodnie z tym dekretem rządowym, krajowa implementacja ustawy o sztucznej inteligencji będzie odpowiedzialnością odrębnej organizacji utworzonej na mocy odrębnego rozporządzenia pod nadzorem Ministerstwa Gospodarki Narodowej. Ponadto utworzono węgierską Radę ds. Sztucznej Inteligencji, do której Bank Centralny Węgier również deleguje członków.

Nie mamy wiedzy o żadnych działaniach mających na celu regulację AI w sektorze bankowym w związku z przyjęciem ustawy o AI. Nie można jednak wykluczyć, że zalecenia wydane przez grupę roboczą mogą zawierać elementy tej ustawy.

Chcielibyśmy również zwrócić uwagę na pakt AI Komisji, którego celem jest zachęcanie i wspieranie organizacji w planowaniu z wyprzedzeniem wdrożenia środków ustawy o AI”.

Wiadomość od dra Vig Gyula
z Magyar Nemzeti Bank

„Ponieważ AI Act ma na celu uniknięcie negatywnych skutków przyjęcia sztucznej inteligencji na zdrowie, bezpieczeństwo i podstawowe prawa obywateli europejskich, można ją rozumieć jako zawierającą pewien komponent etyczny. Ponadto, przed uchwaleniem AI Act, grupa ekspertów wysokiego szczebla Komisji Europejskiej wydała niewiążące wytyczne dotyczące rozwoju godnej

zaufania sztucznej inteligencji, które obejmowały również pewne kwestie etyczne. AI Act bazuje na tych ustaleniach.

Poza tym, pomimo że rząd oświadczył, że należy zająć się etycznymi implikacjami sztucznej inteligencji, Hiszpania nie uchwaliła żadnego dodatkowego aktu prawnego ani miękkiego prawa w tej sprawie. Z drugiej strony, w sektorze finansowym obowiązują już regulacje. Np. wytyczne EBA dotyczące udzielania i monitorowania pożyczek, które między innymi mają na celu zapewnienie, że praktyki instytucji są zgodne z zasadami ochrony konsumentów i respektują uczciwe traktowanie konsumentów. Zgodnie z ust. 39: »w ramach swoich polityk i procedur dotyczących ryzyka kredytowego oraz opierając się na strategii ryzyka kredytowego, instytucje powinny również brać pod uwagę zasady odpowiedzialnego pożyczania. W szczególności: a. powinny brać pod uwagę konkretną sytuację pożyczkobiorcy, taką jak uczciwe traktowanie pożyczkobiorców, którzy mają trudności ekonomiczne; b. powinny projektować produkty kredytowe oferowane konsumentom w sposób odpowiedzialny.«

Naszym zdaniem ważne jest, aby uniknąć opracowywania norm krajowych, które ostatecznie rozdrobnią ramy dla wykorzystania sztucznej inteligencji w Europie. To właśnie jest jeden z celów ustawy o AI, aby zapewnić zharmonizowane ramy na poziomie europejskim. (...)

W Hiszpanii mamy nowego międzysektorowego nadzorcę skupionego na wykorzystaniu AI w kraju, Agencia Española de Supervisión de la IA - AESIA (hiszpańska agencja nadzoru nad AI), a postępowanie banków jest już nadzorowane przez organy finansowe. Niemniej jednak w tym momencie implikacje etyczne dotyczące wykorzystania AI przez banki nie wydają się dla nich priorytetem. (...)

DG FISMA przeprowadziła konsultacje, próbując ustalić, czy konieczna jest konkretna interwencja organów finansowych w zakresie stosowania ustawy o sztucznej inteligencji w sektorze finansowym i jej wzajemnego oddziaływania z regulacjami finansowymi lub w obszarach nieobjętych obecnymi ramami regulacyjnymi dotyczącymi finansów. Niemniej jednak w takich konsultacjach nie było żadnych wskazań dotyczących zainteresowania DG FISMA rozwiązywaniem kwestii etycznych związanych ze sztuczną inteligencją



w sektorze finansowym. Naszym zdaniem etyczne implikacje stosowania sztucznej inteligencji zostaną objęte ustawą o sztucznej inteligencji we wszystkich sektorach. Nie widzimy potrzeby dodatkowych szczegółowych regulacji dla sektora finansowego. Wyzwaniem byłoby posiadanie różnych oczekiwań co do stronniczości, możliwości wyjaśnienia, nadzoru... w różnych sektorach. W nadchodzących miesiącach jest wiele do zrobienia w związku z ustawą o sztucznej inteligencji, opracowaniem regulacji drugiego poziomu, wytycznych, szablonów, nowych standardów (...) które będą miały kluczowe znaczenie dla wyniku tej regulacji. Regulatorzy finansowi i organy nadzoru muszą być zaangażowani w tę pracę, aby zapewnić, że praca ta zostanie wdrożona również w sektorze finansowym”.

Wiadomość od Lorena Mullor
z Asociación Española de Banca

„Ze względu na zasadę neutralności technologicznej w regulacjach, różne rozporządzenia UE, prawa krajowe, rozporządzenia niższego szczebla i wytyczne mają zastosowanie do wykorzystania sztucznej inteligencji w sektorze bankowym. Przepisy te obejmują, ale nie ograniczają się do, Dyrektywy 2014/65 Parlamentu Europejskiego i Rady w sprawie rynków instrumentów finansowych (MiFiD II; patrz np. przepisy dotyczące handlu algorytmicznego), Ogólnego rozporządzenia o ochronie danych UE 2016/679 (RODO) i Rozporządzenia (UE) 2022/2554 Parlamentu Europejskiego i Rady w sprawie regulacji cyfrowej odporności operacyjnej sektora finansowego (DORA), które zaczęły obowiązywać w styczniu 2025 r. Ponadto przepisy dotyczące dobrych praktyk pożyczkowych i wykorzystania automatycznego podejmowania decyzji w Dyrektywie (UE) 2023/2225 Parlamentu Europejskiego i Rady z dnia 18 października 2023 r. w sprawie umów o kredyt konsumencki (dyrektywa w sprawie kredytu konsumenckiego) obejmują wykorzystanie sztucznej inteligencji. Ponadto raport Europejskiego Urzędu Nadzoru Bankowego (EUNB) na temat stosowania modeli uczenia maszynowego w wewnętrznych modelach ratingu kredytowego banków (IRB) stanowi konkretny przykład udoskonaleń i ogólnych zaleceń wydanych przez organ nadzoru w odniesieniu do stosowania sztucznej inteligencji https://www.eba.europa.eu/sites/default/files/document_library/Publications/Reports/2023/1061483/

Follow-up%20report%20on%20machine%20learning%20for%20IRB%20models.pdf (...)

Niezależnie od generalnej zasady neutralności technologicznej regulacji, istnieje kilka krajowych instytucji publicznych, które zajmują się etyką mechanizmów sztucznej inteligencji wśród swoich innych obowiązków nadzorczych. Te krajowe instytucje publiczne obejmują np. Fiński Urząd Nadzoru Finansowego (FIN-FSA), Fiński Urząd ds. Konkurencji i Konsumentów (FCCA) i Biuro Rzecznika Ochrony Danych. (...)

AI Act został opublikowany w Dzienniku Urzędowym Unii Europejskiej w dniu 12 lipca 2024 i weszła w życie z dniem 1 sierpnia 2024 r. Szczegółowy harmonogram procesu legislacyjnego wdrażania AI Act do ustawodawstwa krajowego zostanie ustalony przez parlament Finlandii. Jednak szczegóły tego procesu legislacyjnego (np. konsultacje i obieg wniosku w celu uzyskania komentarzy) nie zostały na razie określone. (...)

Gdybyś był zainteresowany, polecamy zapoznanie się z następującymi materiałami w języku angielskim:

<https://vm.fi/en/ethical-guidelines-for-ai-in-public-administration>

https://www.oecd.org/en/publications/regulatory-approaches-to-artificial-intelligence-in-finance_f1498c02-en.html

oraz fińskim:

<https://www.eurojatalous.fi/fi/2024/artikkelit/generatiivinen-tekoaly-voi-mullistaa-rahoitus-sektoria/>”.

Wiadomość od Tatu Räsänen
oraz Heli Snellman Banku Finlandii

„Aby odpowiedzieć na Twoje pytania, poniżej znajdziesz kilka przydatnych informacji (niestety linki są dostępne tylko w języku rumuńskim).

Na posiedzeniu w dniu 11 lipca 2024 r. rząd Rumunii zatwierdził Narodową Strategię w dziedzinie Sztucznej Inteligencji (AI).

Rumuński Komitet ds. Sztucznej Inteligencji jest organem konsultacyjnym pod wysokim patronatem Premiera Rumunii, a sekretariat techniczny jest zapewniany przez Ministerstwo Badań Naukowych, Innowacji i Cyfryzacji (MRID). (...)

Polecam te strony: <https://lege5.ro/gratuit/gezt-mmrugm2tg/ordinul-nr-20484-2023-privind-in-fiintarea-organizarea-si-functionarea-comitetului-roman-pentru-inteligenta-artificiala>

Misją Komitetu jest wspieranie MRID w wykonywaniu jego obowiązków - w ramach jego roli organu państwowego w dziedzinie badań i rozwoju, innowacji i cyfryzacji - poprzez propozycje dotyczące regulacji, koordynacji, monitorowania i oceny działań związanych ze sztuczną inteligencją (AI), a także jej rozwojem w Rumunii, zgodnie z Memorandum Rządowym nr 20D/31.781/MN z dnia 8.11.2022 r. w sprawie utworzenia Rumuńskiego Komitetu ds. Sztucznej Inteligencji i niektórych inicjatyw związanych z obszarem sztucznej inteligencji w Rumunii.

Głównym celem Komitetu jest stworzenie ekosystemu sztucznej inteligencji opartego na doskonałości, zaufaniu i przestrzeganiu zasad etycznych, ekosystemu, który jest spójny, dobrze działający i zrównoważony, generujący wartość dodaną społecznie i ekonomicznie, poprzez wykorzystanie technologii sztucznej inteligencji, promowanie i wspieranie badań, rozwoju i innowacji oraz edukacji/szkoleń w tej dziedzinie, zwiększanie liczby lokalnych ekspertów i uwzględnianie wkładu sektora prywatnego i obywateli Rumunii mieszkających za granicą, mając na celu zachęcenie do szerokiej adopcji dziedziny sztucznej inteligencji w Rumunii.

Narodowe ramy strategiczne w dziedzinie sztucznej inteligencji na lata 2023–2027:

<https://www.adr.gov.ro/wp-content/uploads/2024/06/Strategia-Nationala-pentru-Inteligenta-Artificiala.pdf>

Narodowe Ramy Strategiczne Sztucznej Inteligencji (NSF-AI) zostały opracowane w szerszym kontekście przyznanym przez projekt o nazwie „Strategiczne ramy przyjmowania i wykorzystywania innowacyjnych technologii w administracji publicznej 2021-2027 – rozwiązania w celu usprawnienia działań”, kod SIPOCA 704, finansowany w ramach Programu Operacyjnego „Możliwości Administracyjne 2014-2020”, wdrażanego przez Urząd ds. Cyfryzacji Rumunii we współpracy z Politechniką w Klużu-Napoce.

Z okresem wdrażania zaprojektowanym na lata 2023-2027, NSF-AI jest potrzebnym i aktualnym kamieniem milowym w przygotowaniu rumuń-

skiego społeczeństwa do zrozumienia, zaakceptowania i wykorzystania transformacyjnych procesów generowanych przez sztuczną inteligencję.

Urząd ds. Digitalizacji Rumunii, we współpracy z Politechniką w Klużu-Napoce, wdrożył w dniach 14 lipca – 30 września 2021 r. pierwszy etap konsultacji online, aby uzyskać opinie środowiska akademickiego, osób pracujących w obszarze badań, rozwoju i innowacji oraz administracji publicznej na szczeblu centralnym i lokalnym, a także środowiska biznesowego, na temat obecnego etapu i potencjału sztucznej inteligencji (AI) w Rumunii.

Rząd Rumunii musi zdecydować, czy powołać podmiot odpowiedzialny za nadzór nad stosowaniem przepisów rozporządzenia UE, czy też przyznać te nowe obowiązki organowi, który już istnieje.

ROZDZIAŁ 12 Implikacje dla ram prawnych

Utworzenie „Urzędu Regulacyjnego ds. AI” jest bezpośrednią konsekwencją wdrożenia ustawy o AI. Utworzenie tego organu w 2024 r. jest przydatne – początkowo w zmniejszonym rozmiarze, zdolnym do aktywowania specyfikacji ustawy o AI (monitorowanie rynku, udział w przestrzeniach testowych dotyczących regulacji, akredytowanie jednostek notyfikowanych itp.) – i które można następnie rozszerzyć, funkcja krajowego ekosystemu AI.

Przygotowanie przyszłej współpracy między Urzędem Regulacyjnym ds. AI a Urzędem Ochrony Konsumentów byłoby przydatne w celu monitorowania i zapewnienia nowych praw osobom w kontekście AI: prawa do otrzymywania wyjaśnień po decyzjach podjętych za pomocą lub przez system AI, prawa do bycia poinformowanym, że ma nastąpić interakcja z systemem opartym na AI itp.

Rumuński HUB na rzecz AU – HRIA: <https://upb.ro/hub-roman-de-inteligenta-artificiala-hria/>

Instytucje będące członkami Międzyresortowej Komisji ds. Koordynacji Wdrażania NS-AI:

1. Sekretariat Generalny Rządu
2. Ministerstwo Spraw Wewnętrznych
3. Ministerstwo Finansów
4. Ministerstwo Transportu i Infrastruktury

5. Ministerstwo Sprawiedliwości
6. Ministerstwo Obrony Narodowej
7. Ministerstwo Edukacji
8. Ministerstwo Spraw Zagranicznych
9. Ministerstwo Gospodarki, Przedsiębiorczości i Turystyki
10. Ministerstwo Inwestycji i Projektów Europejskich
11. Ministerstwo Rozwoju, Robót Publicznych i Administracji
12. Ministerstwo Ochrony Zdrowia
13. Ministerstwo Rolnictwa i Rozwoju Wsi
14. Ministerstwo Energii
15. Ministerstwo Badań Naukowych, Innowacji i Cyfryzacji
16. Ministerstwo Środowiska, Wód i Lasów
17. Ministerstwo Pracy i Solidarności Społecznej
18. Urząd ds. Cyfryzacji Rumunii
19. Krajowa Agencja Zatrudnienia
20. Krajowa Agencja Zamówień Publicznych
21. Krajowy Instytut Administracji
22. Trybunał Obrachunkowy Rumunii
23. Krajowy Dom Ubezpieczenia zdrowotne
24. Krajowy Urząd Regulacji Energetyki
25. Narodowy Instytut Statystyki
26. Rada ds. Konkurencji
27. Rumuńskie Stowarzyszenie Normalizacyjne
28. Krajowa Rada Audiowizualna
29. Rumuńska Agencja Wywiadowcza
30. Zewnętrzna Agencja Wywiadowcza
31. Specjalna Służba Telekomunikacyjna
32. Krajowy Urząd Zarządzania i Regulacji w Komunikacji
33. Krajowa Dyrekcja ds. Cyberbezpieczeństwa
34. Rumuńska Agencja Kosmiczna.”

Wiadomość od Cati Ursu
z Romanian Association of Banks

„Duński Urząd Nadzoru Finansowego (Finanstilsynet) nadzoruje banki i inne instytucje finansowe. Nadzorujemy również korzystanie ze sztucznej inteligencji przez instytucje finansowe w obszarach objętych ustawodawstwem finansowym. Przewidujemy, że zostaniemy mianowani organem nadzoru rynku w obszarach finansowych, kredytowych i ubezpieczeniowych, zgodnie z regulacjami zawartymi w załączniku III do ustawy o sztucznej inteligencji. (...)”

Agencja Rządu Cyfrowego (Digitaliseringsstyrelsen) zarządza wieloma aspektami cyfryzacji. Będzie ona głównym organem nadzoru rynku dla AI Act i prawdopodobnie będzie również jedynym

punktem kontaktowym dla regulacji (komunikat prasowy w języku duńskim: <https://digst.dk/nyheder/nyhedsarkiv/2024/april/rollen-som-national-tilsynsmyndighed-med-eu-s-ai-forordning-skal-varetages-af-digitaliseringsstyrelsen/>) (...)”

Duński Urząd ds. Biznesu (Erhvervsstyrelsen) odpowiada za ogólne duńskie prawodawstwo biznesowe, sprawozdania roczne, Centralny Rejestr Biznesowy itp. Opublikował wytyczne dotyczące etyki danych.

W 2019 r. duński rząd powołał Radę ds. Etyki Danych (Dataetisk Råd), aby „monitorować rozwój kwestii etyki danych w celu rozwoju i wspierania, w sposób etycznie odpowiedzialny, wielu korzyści oferowanych przez wykorzystanie danych”.

W tym samym roku rząd duński opublikował Narodową Strategię dla AI obejmującą sześć zasad etycznych dla AI. Celem tych zasad było ustanowienie wspólnych ram dla rozwoju i wykorzystania sztucznej inteligencji. Zasady są wymienione i wyjaśnione na stronie 27 i następnych. Strategia została odnowiona w 2023 r.

O ile mi wiadomo, nie ma żadnych konkretnych krajowych ustaw ani rozporządzeń wykonawczych regulujących konkretnie wykorzystanie AI poza duńskimi regulacjami dotyczącymi dostosowania RODO. MiFID zawiera zasady dotyczące zasad raportowania dotyczących handlu algorytmicznego.

Jednakże duńskie przepisy dotyczące sprawozdań rocznych, zarówno dla firm w ogóle, jak i dla sektora finansowego w szczególności, obejmują „zasadę przestrzegania lub wyjaśnienia” dotyczącą etyki danych. Duński Urząd ds. Biznesu monitoruje ogólne przepisy (artykuł 99d „årsregnskabsloven”), a duński FSA monitoruje rozporządzenia wykonawcze regulujące konkretnie sektor finansowy. W przypadku banków odnosi się do artykułu 154 rozporządzenia wykonawczego w sprawie sprawozdań finansowych instytucji kredytowych (...)”

Banki o określonej wielkości muszą dołączyć do swoich sprawozdań rocznych oświadczenie dotyczące swojej pracy i polityki dotyczącej etyki danych. W przypadku, gdy nie wykonały takiej pracy, są zobowiązane wyjaśnić powód tego. (...)”

W ciągu ostatnich lat my w Duńskim Urzędzie Nadzoru Finansowego opublikowaliśmy kilka



dokumentów z zaleceniami i raportami na temat wykorzystania AI w sektorze finansowym, w tym na temat etyki danych:

2019: Dokument (w języku duńskim) z zaleceniami dotyczącymi korzystania z nadzorowanego uczenia maszynowego: „God praksis ved brug af superviseret machine” („Dobre praktyki przy korzystaniu z nadzorowanego uczenia maszynowego”).

2023: Dokument (w języku duńskim) z pomysłami na pracę z etyką danych: „Rapport om dataetik ved brug af AI i den finansielle sektor”. We wrześniu tego roku opublikowaliśmy angielskie tłumaczenie Raportu na temat etyki danych przy korzystaniu z AI w sektorze finansowym.

2024: Zaktualizowana wersja (w języku duńskim) dokumentu z 2019 r. z zaleceniami dotyczącymi

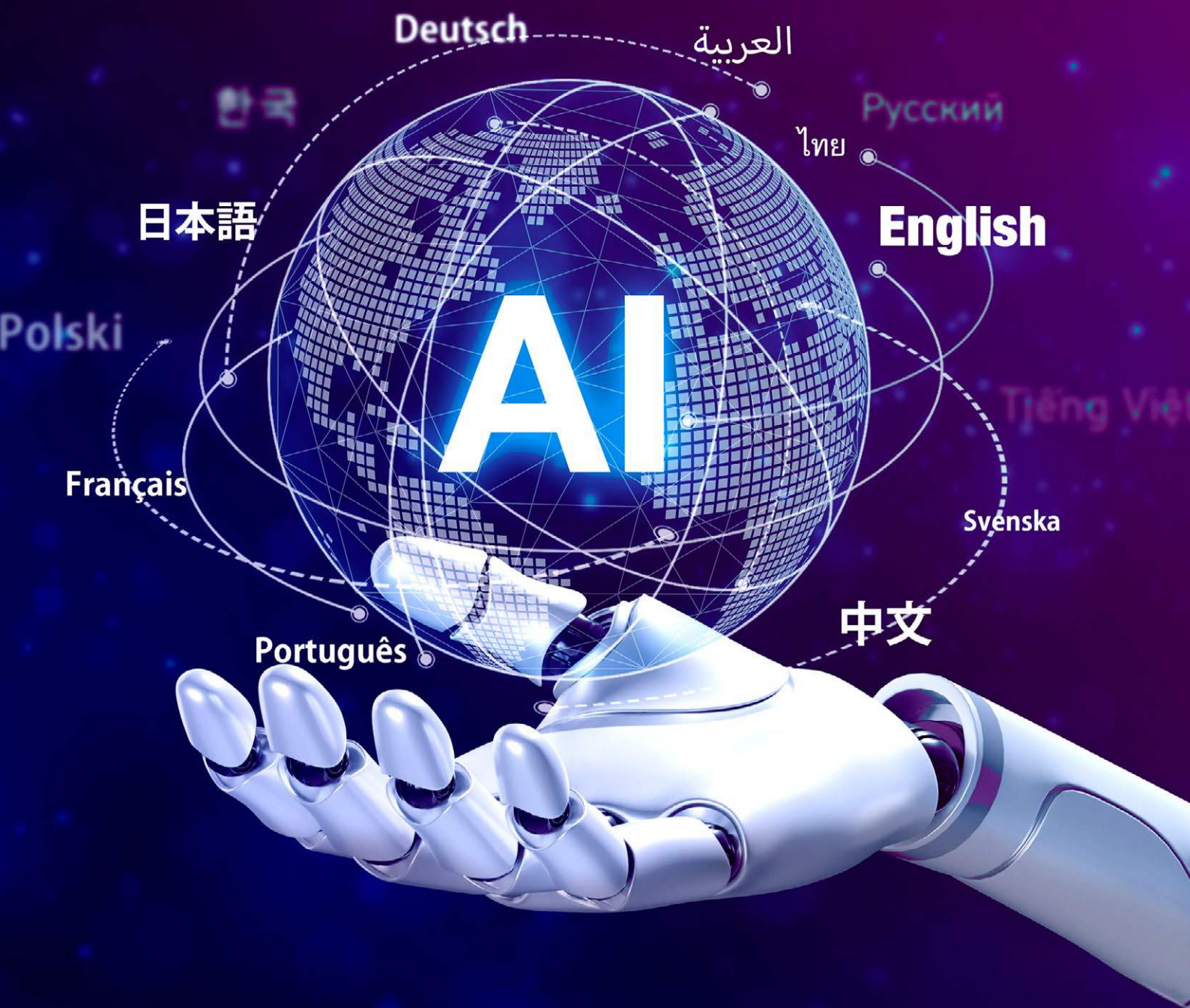
korzystania z AI w ogólności: „God praksis ved brug af kunstig intelligens” („Dobre praktyki przy korzystaniu ze sztucznej inteligencji”).

2024: Raport (w języku duńskim) z naszego badania dotyczącego wykorzystania sztucznej inteligencji w sektorze finansowym: „Brug af kunstig intelligens i den finansielle sektor i 2024” (...).

Z pewnością będą krajowe regulacje związane z przyjęciem ustawy o AI. Agencja Rządu Cyfrowego będzie wiodącą rolę w tych regulacjach, prawdopodobnie przy współudziale różnych organów, takich jak Duński Urząd Nadzoru Finansowego. Niestety, w tym momencie nie jesteśmy w stanie dostarczyć Ci więcej informacji na temat tych przepisów.”

Wiadomość od Camilla Neuenschwander z Danish Financial Supervisory Authority

3. Regulacje na poziomie ponadnarodowym





3.1. Zalecenia UNESCO w sprawie etyki sztucznej inteligencji

Porównanie różnych podejść regulacyjnych do problemu zastosowania sztucznej inteligencji w sektorze finansowym (ze szczególnym uwzględnieniem bankowości) warto zacząć od analizy mających charakter ogólny i uniwersalny „Zaleceń UNESCO w sprawie etyki sztucznej inteligencji”¹⁴. Dokument ten nie odnosi się konkretnie do sektora bankowego, ale dąży do ustalenia ogólnych standardów mających zastosowanie w wielu różnych obszarach, w których wprowadzane są mechanizmy oparte na sztucznej inteligencji, a zwłaszcza na polach edukacji, nauki, kultury oraz komunikacji i informacji będących domenami działalności UNESCO (Organizacja Narodów Zjednoczonych do spraw Edukacji, Nauki i Kultury, ang. *United Nations Educational, Scientific and Cultural Organization*, por. pkt 3 zaleceń).

Zalecenia UNESCO są przyjmowane przez Konferencję Generalną organizacji i nie podlegają ratyfikacji przez państwa członkowskie, **nie stają się więc elementem prawa powszechnie obowiązującego w Polsce, czerpią jednak swój autorytet z uniwersalnego charakteru gremium (193 państwa), autorytetu Organizacji Narodów Zjednoczonych jako organizacji międzynarodowej i mają na celu „wpływanie na rozwój prawa i praktyki krajowej”**¹⁵.

Państwa członkowskie są także zobowiązane do przedstawiania co 4 lata raportów z postępów w implementacji zaleceń. Zgodnie ze zobowiązaniem zawartym w zaleceniach, UNESCO przygotowało również **metodologię oceny oddziaływania AI pod względem etycznym (ethical impact assessment – EIA)**¹⁶.

Zalecenia ze względu na dynamicznie zmieniającą się rzeczywistość nie podejmują próby przedstawienia precyzyjnej definicji sztucznej inteligencji, zamiast tego wskazują jedynie na zasadnicze cechy technologii będących w zakresie zainteresowania dokumentu¹⁷. Są to systemy „*posiadające zdolność do przetwarzania danych i informacji w sposób, który przypomina inteligentne zachowanie i zazwyczaj obejmuje aspekty takie jak rozumowanie, uczenie się, postrzeganie, przewidywanie, planowanie czy kontrola*” (pkt 2). Zalecenia dotyczą przy tym **wszystkich etapów cyklu życia systemów AI** – od badań i projektowania po wyłączenie z obrotu oraz utylizowanie (pkt 2 ppkt a).

Zasadniczo dokument opiera się na podejściu do etyki sztucznej inteligencji stawiającym „**w centrum uwagi ludzką godność i prawa człowieka**” (Preambuła). Oprócz tego wskazuje się też inne podstawowe zasady mające kierować wykorzystaniem systemów AI – **dobro środowiska i ekosystemów, zapewnienie różnorodności i włączenia społecznego oraz życie w społeczeństwach pokojowych, sprawiedliwych i spójnych wewnętrznie**.

Użycie technologii opartych na sztucznej inteligencji powinno być także zgodne z **zasadą proporcjonalności** (technologia musi być więc właściwa do osiągnięcia danego uprawnionego celu, nie może naruszać podstawowych wartości, w tym zwłaszcza praw człowieka, a jej podstawą mają być rzetelne badania naukowe – pkt 26). Twórcy dokumentu zdają sobie sprawę z potencjalnego napięcia mogącego powstawać między regulacją a rozwojem technologicznym, ale, jak stwierdza się w zaleceniach, „*uwzględnianie zagrożeń i problemów natury etycznej nie powinno hamować innowacyjności i rozwoju, lecz raczej stwarzać nowe możliwości*” (Preambuła, zob. też np. pkt 56).

Zobowiązanie do „przestrzegania, chronienia i promowania praw człowieka i podstawowych wolności” odnosi się również do sektora prywatnego. Stąd zachęca się podmioty gospodarcze do przeprowadzania oceny oddziaływania (także pod względem etycznym) wdrażanych rozwiązań dotyczących AI, powoływania niezależnych **inspektorów do spraw etyki AI** oraz opracowywania **wytycznych w sprawie**

14 Wersja w języku angielskim: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> [dostęp: 6 października 2024 r.]; Tłumaczenie na język polski opracowane przez Stały Sekretariat Polskiego Komitetu do spraw UNESCO: https://www.unesco.pl/fileadmin/user_upload/pdf/Sztuczna_inteligencja-4.pdf [dostęp: 6 października 2024 r.]. Jeżeli nie zaznaczono inaczej, cytaty z zaleceń w dalszej części tekstu pochodzą z polskiego tłumaczenia.

15 Informacje ogólne na temat instrumentu prawnego, jakim są zalecenia - na stronie UNESCO - <https://www.unesco.org/en/legal-affairs/standard-setting/recommendations> [dostęp: 6 października 2024 r.], zob. także Konwencja dotycząca utworzenia Organizacji Narodów Zjednoczonych dla Wychowania, Nauki i Kultury, art. IV B.4.

16 Dostępne na stronie internetowej UNESCO - https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000386276&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_353afb2-65da-4cee-a99f-5347218f574d%3F_%3D386276eng.pdf&locale=en&multi=true&ark=/ark:/48223/pf0000386276/PDF/386276eng.pdf#917_23%20E%20Ethical%20Impact%20Assessment.indd%3A.58841%3A192 [dostęp: 6 października 2024 r.].

17 Por. L. Floridi, *The Ethics of Artificial Intelligence Principles, Challenges, and Opportunities*, Oxford 2023, s. 14-16, gdzie autor wskazuje na problem ze sformulowaniem powszechnie akceptowalnej definicji sztucznej inteligencji.

etyki AI. Pożądane jest również włączenie różnych interesariuszy do procesu zarządzania AI (pkt 58).

Wydaje się, że spośród ogólnych zaleceń dotyczących etyki sztucznej inteligencji można wskazać kilka wątków szczególnie istotnych z punktu widzenia zastosowań technologii opartych na AI w sektorze finansowym.

Po pierwsze, zalecenia wielokrotnie podkreślają konieczność **przeciwdziałania dyskryminacji**, która może być efektem wypaczonych zastosowań lub nieodpowiednich mechanizmów działania AI, utrwalających uprzedzenia i schematy nierównego traktowania występujące w rzeczywistości analogowej. Genezą takiego efektu może być również nieprawidłowy dobór zestawów danych szkoleniowych dla systemów AI¹⁸. Nierówności, które należy minimalizować, powstają także na tle braku dostępu do technologii oraz wiedzy dotyczącej sztucznej inteligencji w określonych grupach społecznych (por. pkt 28-30, pkt 66-67).

Innym ważnym aspektem etyki AI jest kwestia **ludzkiego nadzoru i możliwości przypisania odpowiedzialności człowiekowi**. Chodzi więc m.in. o przypisanie odpowiedzialności za szkody wywołane przez technologie oparte na AI, co jest jednym z najintensywniej dyskutowanych zagadnień w obszarze prawa nowych technologii. Zalecenia UNESCO stają tu wyrażnie na stanowisku **wykluczającym możliwość przyznania systemom AI osobowości prawnej**. Jak podkreślają „ostatecznie odpowiedzialność i obowiązek rozliczalności musi zawsze spoczywać na osobach fizycznych lub prawnych” (pkt 68), a regulacje i praktyka państw członkowskich powinny umożliwiać przypisanie osobom fizycznym lub prawnym odpowiedzialności etycznej i prawnej za każdy etap cyklu życia systemów AI. Ze względów efektywności w określonych sytuacjach ludzie mogą zdać się na technologie AI, ale zrzeczenie się kontroli nie oznacza w żadnym wypadku zrzeczenia się odpowiedzialności (pkt 35-36). Istnieją też obszary, w których ostateczną decyzję musi podejmować człowiek – przypadki nieodwracalnych lub trudnych do odwrócenia skutków oraz sprawy życia i śmierci (pkt 26). Dodatkowo w sytuacji wpływu na prawa i wolności osoby fizycznej powinna istnieć możliwość „przedstawienia swoich uwag wyznaczonemu pracownikowi firmy prywatnej lub instytucji publicznej, który będzie w stanie dokonać weryfikacji i korekty decyzji” (pkt 38).

Z problemem odpowiedzialności silnie związana jest **transparentność i wyjaśnialność** (*explainability* – w polskim tłumaczeniu zaleceń także jako „**wytłumaczalność**”)¹⁹. Zasady te obejmują szeroką gamę wymagań dotyczących zastosowania sztucznej inteligencji. Po pierwsze, podmioty wykorzystujące AI powinny mieć obowiązek **poinformowania** użytkowników, gdy produkt lub usługa są dostarczane z wykorzystaniem tej technologii (pkt 38). Kluczowe znaczenie ma natomiast **możliwość wytłumaczenia**, w jaki sposób system AI generuje określone wyniki i **prześledzenia poszczególnych etapów** dochodzenia do rezultatów końcowych. Wymóg ten nie ma charakteru absolutnego i powinien być **dostosowany do kontekstu** (proporcjonalność) – w wymaganiach dostrzeżono bowiem, że w określonych sytuacjach może wchodzić w konflikt z innymi zasadami, jak przykładowo ochrona prywatności (pkt 38). Zwłaszcza jednak w przypadku szczególnie istotnych (nieodwracalnych, dotyczących praw człowieka) decyzji adresat powinien otrzymać uzasadnienie decyzji i móc złożyć swoje uwagi. Jest to bowiem warunek możliwości podważania konkretnych rozstrzygnięć, a ostatecznie także korzystania z **prawa do rzetelnego procesu sądowego i zadośćuczynienia** (pkt 37). Kiedy poważnie zagrożone są prawa człowieka, zasady transparentności i wyjaśnialności mogą także oznaczać konieczność udostępnienia kodu i zbiorów danych (pkt 39). Zastosowanie powyższych wskazań ma, oprócz zapewniania gwarancji praw jednostek, przyczyniać się do **wzmacniania publicznej kontroli nad zastosowaniem AI** oraz do **budowania w społeczeństwie zaufania** do nowych narzędzi technologicznych (pkt 39).

Warto także zwrócić uwagę na kwestię **ochrony danych i prywatności** – w zaleceniach wielokrotnie podkreślono konieczność **zapewnienia zgodności wprowadzanych rozwiązań z obowiązującym prawem w tym zakresie**. Szczególny nacisk należy położyć na ochronę danych, gdy są one wykorzystywane do celów komercyjnych, takich jak reklama oparta na mikrotargetowaniu (pkt 73). Podmioty działające w obszarze AI powinno się też zachęcać do przeprowadzania **oceny oddziaływania wdrażanych rozwiązań na prywatność** (pkt 72).

Kwestią, w której zalecono państwom członkowskim opracowanie odpowiednich wytycznych, są również **interakcje człowiek-robot (maszyna)**. W tym zakresie mieści się m.in. zastosowanie chat-

18 Por. *Ethical Impact Assessment*, sekcja 6, s. 21.

19 Por. L. Floridi, *The Ethics of ...*, s. 63 oraz pkt 41 zaleceń.

botów, które coraz częściej wykorzystuje się do obsługi klientów²⁰. Interakcje te powinny opierać się na **zgodności z prawami człowieka i podstawowymi wolnościami**. Natomiast ryzykiem, jakie dostrzeżono w tym obszarze, jest zwłaszcza posługiwanie się sztuczną inteligencją do manipulowania ludźmi i wykorzystywanie ludzkiej niewiedzy (pkt 125-126).

Choć zalecenia UNESCO są adresowane przede wszystkim do państw członkowskich (odpowiedzialnych za tworzenie ram prawnych i prowadzenie polityki) – to mogą stanowić również źródło inspiracji dla podmiotów prywatnych do etycznej refleksji przy wdrażaniu systemów AI w sposób służący „*ludzkości, jednostkom, społeczeństwom, środowisku i ekosystemom*” (pkt 5).

3.2. Dokumenty Unii Europejskiej

Regulacja prawna AI na poziomie Unii Europejskiej podlega w ostatnim czasie istotnym zmianom. Kluczowe znaczenie ma tu zwłaszcza tzw. AI Act, którego nie sposób wyczerpująco scharakteryzować w ramach przeglądowego omówienia dokumentów typu *soft law*. Fakt przeobrażeń, jakim są poddane unijne ramy prawne funkcjonowania sztucznej inteligencji, należy jednak uwzględnić jako wpływający na częściową dezaktualizację wielu opracowań, deklaracji i wytycznych przedstawionych nawet w nieodległej przeszłości. Niektóre z prezentowanych w nich wskazań etycznych nabrały już bowiem prawnego charakteru, pewne postulowane działania doczekały się zaś realizacji.

Przed przejściem do analizy zagadnień dotyczących konkretnie zastosowań sztucznej inteligencji w sektorze finansowym (zwłaszcza w bankowości) warto przyjrzeć się pewnym ogólnym założeniom znajdującym zastosowanie na wszystkich polach wykorzystania AI.

Generalne kierunki polityki wobec rozwoju sztucznej inteligencji i podejście do wyzwań z nim związanych zostały zaprezentowane w dokumentach takich jak: „**Biała księga w sprawie sztucznej inteligencji. Eu-**

ropejskie podejście do doskonałości i zaufania”²¹, „**Skoordynowany plan w sprawie sztucznej inteligencji**”²² oraz strategia „**Sztuczna inteligencja dla Europy**”²³.

Wynika z nich, że Unia Europejska dąży do osiągnięcia pozycji światowego lidera w stosowaniu nowych technologii i do zdobycia **przewagi konkurencyjnej w obszarze AI** przy jednoczesnym **respektowaniu swoich wartości**. Z jednej strony zakłada się więc promowanie stosowania sztucznej inteligencji, z drugiej zaś przeciwdziałanie zagrożeniom związanym z jej wykorzystaniem²⁴. AI ma być „**zorientowana na człowieka, zrównoważona, bezpieczna, sprzyjająca włączeniu społecznemu i wiarygodna**”²⁵. Unia chce stać się „*orędownikiem podejścia do AI, które przynosi korzyści zarówno indywidualnym obywatelom, jak i całemu społeczeństwu*”²⁶ oraz zamierza promować „*ambitne globalne zasady i normy*”²⁷. Tego typu strategia ma zapewnić nie tylko zgodność z wartościami, ale też przekładać się na przewagi konkurencyjne, ponieważ warunkiem wdrożenia sztucznej inteligencji na szeroką skalę jest wiarygodność tej technologii. Należy więc tworzyć odpowiedni „**ekosystem zaufania**”²⁸. W „Białej księdze” podkreślono również konieczność skoordynowanych działań, które umożliwią wykorzystanie efektu skali i zapobiegą niepewności prawa na jednolitym rynku²⁹.

Charakter bardziej szczegółowych reguł horyzontalnych (a więc nie sektorowych) mają wypracowane przez Niezależną Grupę Ekspertów Wysokiego Szczebla ds. Sztucznej Inteligencji (powołaną przez Komisję Europejską) „**Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji**”³⁰. W związku z tym, że, jak zauważono w tym dokumencie, prawo nie zawsze nadąża za zmianami techno-

21 https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_pl?filename=commission-white-paper-artificial-intelligence-feb2020_pl.pdf [dostęp: 6 października 2024 r.].

22 <https://ec.europa.eu/newsroom/dae/redirection/document/95793> [dostęp: 6 października 2024 r.].

23 <https://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:52018DC0237> [dostęp: 6 października 2024 r.].

24 Zob. *Biała księga* ..., s. 1.

25 *Skoordynowany plan* ..., s. 2.

26 *Sztuczna inteligencja dla Europy*, s. 3.

27 *Skoordynowany plan* ..., s. 43.

28 *Biała księga* ..., s. 3,10.

29 *Biała księga* ..., s. 3.

30 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60436 [dostęp: 6 października 2024 r.].

20 Zob. Raport z badania *Aktualne wyzwania prawne związane z zastosowaniem sztucznej inteligencji w polskim sektorze bankowym*, zwłaszcza rozdział 2.3, z którego wynika, że zautomatyzowana obsługa klienta jest jednym z najczęstszych sposobów wykorzystania AI przez sektor bankowy.

logicznymi i nie w każdej sytuacji jest adekwatnym środkiem, zastosowania sztucznej inteligencji powinny również spełniać **kryteria etyczne**. Sformułowanie tego typu wymagań jest właśnie celem autorów wytycznych. Zaprezentowane wnioski mogą być m.in. użyte przez podmioty wykorzystujące AI do aktualizacji wewnętrznych dokumentów programowych i kodeksów. Nawet jednak sformułowanie zbioru reguł nie zastąpi, co zasygnalizowano w wytycznych, **etycznego rozumowania uwzględniającego okoliczności konkretnej sytuacji**, niezbędne jest więc utrzymanie „etycznej kultury i sposobu myślenia poprzez debatę publiczną, edukację i praktyczne uczenie się” (pkt 36). Ważną rolę odgrywa tu **zasada proporcjonalności**, której zastosowanie można odnaleźć w wielu fragmentach wytycznych. Dokument stanowi też próbę odpowiedzi na pytanie o koncepcję i wizję normatywną AI przyjmowaną w UE. W tym kontekście pojawia się pojęcie „**godnej zaufania sztucznej inteligencji**”.

Ogólnym założeniem wytycznych jest podejście **ukierunkowane na człowieka**, z czego wynika, że podstawą etycznego stosunku do AI są **prawa podstawowe „zakorzenione w nieodłącznym statusie moralnym człowieka”** (pkt 39). Wymaga to m.in. poszanowania **godności i wolności jednostki**, jak również zapewnienia **równości i niedyskryminacji**.

Na tym oparte zostały cztery zasady etyczne³¹:

- 1) **Poszanowanie autonomii człowieka** (pkt 50), które wyklucza manipulowanie ludźmi i ich oszukiwanie (wprowadzanie w błąd) przy użyciu AI, każe pozostawić człowiekowi realną możliwość wyboru i wskazuje na konieczność zapewnienia **ludzkiego nadzoru** nad działaniami sztucznej inteligencji.
- 2) **Zapobieganie szkodom** (pkt 51) wymagające, by systemy były solidne technicznie oraz odporne na wykorzystanie w złym zamiarze. Szczególną uwagę należy tu zwrócić na przeciwdziałanie powodowaniu lub pogłębianiu niekorzystnego wpływu związanego z **asymetrią władzy lub in-**

formacji (np. w relacji pracodawca – pracownik, przedsiębiorca – konsument).

- 3) **Sprawiedliwość** (pkt 52), na którą składa się sprawiedliwość **materialna** (a więc m.in. brak dyskryminacji, równy i sprawiedliwy podział korzyści i kosztów) oraz **proceduralna** (możliwość kwestionowania wyników działania systemów AI, a także dochodzenia roszczeń związanych ze spowodowanymi przez nie szkodami).
- 4) **Możliwość wyjaśnienia** (pkt 53), która zakłada, że działania systemów AI powinny być w jak największym stopniu (proporcjonalnie do kontekstu i wagi działania) możliwe do wyjaśnienia osobom, na które mają bezpośredni i pośredni wpływ (co jest warunkiem możliwości ich kwestionowania).

Przedstawione zasady mogą w pewnych przypadkach wchodzić ze sobą w konflikt. W takiej sytuacji konieczna jest oparta na dowodach refleksja, przy założeniu, że w wypracowanych rozwiązaniach **korzyści powinny znacznie przeważać nad indywidualnymi przewidywanymi zagrożeniami**.

Wskazano również siedem bardziej szczegółowych wymogów, z których większość ma zastosowanie do wszystkich systemów AI, **szczególną uwagę należy jednak poświęcić tym systemom, które wywierają bezpośredni lub pośredni wpływ na jednostki**.

Są to:

- 1) **Przewodnia i nadzorcza rola człowieka** (pkt 62-65), która zaleca m.in. przeprowadzanie oceny skutków w zakresie praw podstawowych (z uwzględnieniem proporcjonalności). Zakłada też **niepodleganie decyzjom opartym wyłącznie na automatycznym przetwarzaniu danych**, w przypadku gdy wywołują one wobec użytkowników skutki prawne lub w podobny sposób znacząco na nich wpływają (zasada ta znalazła swój prawny wyraz w art. 22 RODO)³², należy też zapewnić odpowiedni **ludzki nadzór nad systemem** (stosownie do okoliczności, przy założeniu, że im mniejsza możliwość nadzoru, tym

31 Warto wskazać, że zasady te w znacznej mierze pokrywają się z czterema zasadami wyróżnionymi przez Luciano Floridiego. Po przeanalizowaniu sześciu różnych zestawów zasad etycznych dotyczących sztucznej inteligencji stwierdził on, że dają się one generalnie sprowadzić do pięciu zasadniczych wytycznych: przynoszenia korzyści (beneficence), nieszkodzenia, autonomii, sprawiedliwości oraz wytłumaczalności (explicability) – L. Floridi, *The Ethics of Artificial Intelligence Principles, Challenges, and Opportunities*, Oxford 2023, s. 59-60.

32 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 (ogólne rozporządzenie o ochronie danych), Dz.Urz. UE L 119 z 4.5.2016, s. 1-88.

bardziej rygorystyczne zarządzanie systemem i szerszy zakres jego testów).

- 2) **Techniczna solidność i bezpieczeństwo** (pkt 66-70), które każą minimalizować niezamierzone szkody i zapobiegać szkodom niedopuszczalnym (w przypadku sytuacji obejmujących szczególne ryzyko wskazane jest proaktywne podejście do środków bezpieczeństwa). Kluczowe jest w tym aspekcie także zapewnienie **odtworzalności** (określa ona, czy system AI, w doświadczeniu powtórzonym w tych samych warunkach, zachowuje się w identyczny sposób).
- 3) **Ochrona prywatności i zarządzanie danymi** (pkt 71-74) nakazują dbać o **jakość i integralność wykorzystywanych danych**, a także ich **istotność** w stosunku do dziedziny, w której AI będzie używane. W związku z tym wykluczona jest dyskryminacja jednostek przy pomocy gromadzonych danych. Oprócz tego dane poszczególnych osób powinny być dostępne jedynie dla wykwalifikowanego personelu i **tylko wtedy, gdy jest to konieczne**.
- 4) **Przejrzystość** (pkt 75-78) obejmuje wymóg **dokumentowania** zbiorów danych i procesów działania systemu, ale również konieczność przedstawienia **odpowiedniego wytłumaczenia** procesu decyzyjnego sztucznej inteligencji, które jest **dostosowane do wiedzy fachowej zainteresowanych osób**, ilekroć system AI ma znaczący wpływ na życie ludzi. Wytłumaczenie może dotyczyć także tego, w jakim stopniu AI wpływa na proces decyzyjny w organizacji. Użytkownicy mają również prawo do **informacji, że wchodzą w interakcję z systemem AI**, a w niektórych sytuacjach wymogi praw podstawowych wymuszają zapewnienie **możliwości kontaktu z człowiekiem**.
- 5) **Różnorodność, niedyskryminacja i sprawiedliwość** (pkt 79-82) oznaczają unikanie nierównego traktowania wynikającego między innymi z niekompletności lub historycznie uwarunkowanej stronniczości danych. W celu ograniczenia zagrożeń wskazane jest zapewnienie udziału **osób reprezentujących różne kultury, środowiska i dyscypliny**, a także projektowanie systemów w sposób pozwalający na korzystanie z opartych na nich usług i produktów **bez względu na wiek, płeć czy umiejętności**.

6) **Dobrostan społeczny i środowiskowy** (pkt 83-86) w odniesieniu od AI oznacza uznanie za zainteresowane strony w całym cyklu życia tej technologii również ogółu społeczeństwa, innych istot zdolnych do odczuwania oraz środowiska. Wymaga to więc m.in. krytycznej **kontroli zużycia zasobów** oraz uwzględniania wpływu sztucznej inteligencji na **umiejętności społeczne ludzi, ich zdrowie fizyczne i psychiczne, demokrację oraz instytucje**.

7) Wymóg **odpowiedzialności** (pkt 87-91) oznacza konieczność zapewnienia **niezależnej kontroli** systemów AI w przypadku zastosowań mających wpływ na prawa podstawowe, gwarancję odpowiedniej ochrony dla osób zgłaszających uzasadnione obawy związane z systemem AI, a także **procedurę dochodzenia roszczeń** w przypadku szkód związanych ze sztuczną inteligencją. Należy także podkreślić, że w przypadku braku możliwości wypracowania etycznie dopuszczalnego kompromisu, nie powinno się wdrażać danej technologii. Decydenci ponoszą przy tym **odpowiedzialność** za realizowane kompromisy.

W kwestii podejścia do AI we wszystkich dziedzinach zastosowania można również wskazać raport Biura Analiz Parlamentu Europejskiego „**Etyka sztucznej inteligencji. Zagadnienia i inicjatywy**”³³. Wskazano w nim na wiele aspektów etycznego wpływu AI (oddziaływanie na społeczeństwo, psychikę, gospodarkę, środowisko i zaufanie) oraz przeanalizowano podejmowane w różnych państwach inicjatywy będące odpowiedzią na to oddziaływanie.

Wśród wyzwań wymieniono m.in. potencjalne **zwolnienia, degradację środowiska, powiększanie nierówności oraz problemy związane z prywatnością i relacjami międzyludzkimi**. Jako obszary niedostatecznie dostrzeżone przez wcześniej powstałe międzynarodowe wytyczne wyróżniono np. kwestie zwiększonego zużycia energii związanego z funkcjonowaniem systemów AI, nierównej dystrybucji korzyści płynących z zastosowania sztucznej inteligencji oraz możliwego wykorzystywania pracowników. Ostatecznie podkreślono, że **namysł etyczny** powinien stanowić kluczowy element polityki wobec AI, a nieuwzględnianie tego wymiaru nie tylko stano-

33 [http://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf) [dostęp: 6 października 2024 r.].

wi naruszenie zobowiązań moralnych, ale też wpływa negatywnie na potencjał konkurencyjny danego państwa na globalnym rynku.

3.3. Stanowisko regulatorów i organów nadzoru

Oprócz wskazanych wyżej ogólnych wytycznych, rozwój i zastosowania AI stały się też obiektem zainteresowania bardziej wyspecjalizowanych organów regulacji i nadzoru, które zaangażowały się m.in. w prace badawcze czy opracowywanie raportów i rekomendacji. Wskazać tu trzeba przede wszystkim na Europejski Urząd Nadzoru Bankowego³⁴ (ang. *European Banking Authority*, dalej także jako: „EBA”) oraz Europejski Bank Centralny³⁵.

EBA w swoich raportach dostrzega wpływ, jaki na sektor bankowy wywiera rozwój sztucznej inteligencji. Wiąże się to z potencjałem poprawy standardów świadczonych usług w wielu obszarach (np. skrócenie czasu oczekiwania na decyzję kredytową), ale także z konkretnymi zagrożeniami. W związku z tym formułowane są różne wskazania i zalecenia, a także zwraca się uwagę na kluczowe wartości. W raporcie dotyczącym *big data* oraz zaawansowanych narzędzi analitycznych w sektorze bankowym (w którego tematyce mieści się także AI) wyróżniono siedem elementów odpowiadających za zaufanie do tego typu rozwiązań. Są to:

1) **Etyka** – odwołano się tu przede wszystkim do umówionych wyżej „Wytycznych w zakresie etyki dla godnej zaufania sztucznej inteligencji”, wskazano, że „moralne zobowiązania mogą wykraczać ponad i poza wypełnianie znajdujących zastosowanie wymogów prawnych”³⁶ oraz zarekomendowano uwzględnianie perspektywy etycznej od samego początku prac nad projek-

tami związanymi z AI (**ethical by design**). Jako mogące ułatwić zapewnienie zgodności z wymaganiami etyki zasugerowano tworzenie **komisji ds. etyki**, które mogłyby odpowiadać m.in. za zatwierdzanie decyzji o wdrożeniu AI.

2) **Możliwość wyjaśnienia (explainability) i interpretacji**, która musi być dostosowana do kontekstu i poziomu automatyzacji (np. wyższy poziom powinien być zastosowany, kiedy decyzja ma bezpośredni wpływ na człowieka; wyjaśnialność istotna jest również m.in. gdy człowiek ma podjąć decyzję, opierając się na wynikach przedstawionych przez AI). W tym miejscu warto wskazać, że w rekomendacjach zawartych w raporcie EBA dotyczącym zastosowania uczenia maszynowego w modelach ratingów wewnętrznych (rekomendacje te również oparto na omawianych siedmiu elementach zaufania) zalecono **unikanie niepotrzebnej złożoności** i zaznaczono, że wzrost złożoności powinien być usprawiedliwiony znaczącym (*significant*) wzrostem możliwości przewidywania³⁷. Trzeba też doprecyzować, że możliwość wyjaśnienia jest elementem szerszego pojęcia **transparentności**, które odnosi się nie tylko do modelu, ale też do całego procesu jego wykorzystania.

3) **Bezstronność i unikanie uprzedzeń**, które mogą powstać m.in. na skutek użycia danych niereprezentatywnych dla analizowanej populacji. W tym kontekście wskazuje się na konieczność **kontroli przygotowywania danych**, bieżące **monitorowanie** działania systemu oraz przeprowadzanie analiz statystycznych służących wykryciu potencjalnych uprzedzeń (*bias*).

4) **Możliwość śledzenia (traceability) i weryfikacji (auditability)**, która powinna zapewnić zdolność do „prześledzenia wszystkich kroków, kryteriów i wyborów w trakcie procesu”³⁸ będącą warunkiem przeprowadzenia kontroli. Istotne znaczenie ma w związku z tym m.in. **odpowiednia dokumentacja** (np. zmian, którym podlega model).

5) **Ochrona i jakość danych** – w tym zakresie odniesiono się przede wszystkim do **standardów wynikających z RODO**.

34 Zob. np. raporty: *Raport na temat Big Data i zaawansowanych rozwiązań analitycznych* (https://eba.europa.eu/sites/default/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf) [dostęp: 6 października 2024 r.]; *Uczenie maszynowe w modelach ratingów wewnętrznych. Zaktualizowany raport po konsultacjach wstępnego raportu na temat uczenia maszynowego w modelach ratingów wewnętrznych* (https://eba.europa.eu/sites/default/files/document_library/Publications/Reports/2023/1061483/Follow-up%20report%20on%20machine%20learning%20for%20IRB%20models.pdf) [dostęp: 6 października 2024 r.].

35 Zob. np. opublikowane przez EBC artykuł C. Papathanassiou, *Digital innovation and banking regulation*, „ECB Occasional Paper”, 351/2024, (<https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op351~c46b57f061.en.pdf>) [dostęp: 6 października 2024 r.].

36 *Raport na temat Big Data* ..., s. 45.

37 *Uczenie maszynowe w modelach* ..., s. 21.

38 *Raport na temat Big Data* ..., s. 39.

- 6) **Bezpieczeństwo**, którego zapewnienie wymaga uwzględnienia nowych rodzajów cyberataków specyficznych dla technologii opartych na AI (np. *data poisoning*).
- 7) **Ochrona konsumenta** – problemy w tym obszarze mogą pojawić się m.in. w związku z **wykluczeniem konsumentów**, którzy nie będą chcieli podzielić się danymi, albo ich nie mają (*non-digital consumers*).

3.4. Stanowisko przedstawicieli sektora bankowego

Swój punkt widzenia na problematykę sztucznej inteligencji przedstawili również przedstawiciele sektora bankowego – warto tu przywołać „**Stanowisko Europejskiej Federacji Bankowej w sprawie AI w sektorze bankowym**”³⁹. Zauważa się w nim, że sztuczna inteligencja ma wpływ na przemiany sektora bankowego. Wiąże się z nią istotne możliwości, takie jak: zapewnienie klientom bardziej spersonalizowanych produktów, precyzyjniejsza analiza zdolności kredytowych, lepsze doradztwo finansowe, redukcja kosztów czy skuteczniejsze wykrywanie oszustw. Dostrzegane są jednak również zagrożenia, a także specyficzne problemy, przed którymi stoi sektor bankowy. W tym kontekście odnotowano, że nowi uczestnicy rynku często nie są poddani rygorystycznym wymaganiom stosowanym wobec banków (nawet jeśli oferują takie same usługi). To sprawia, że także w obszarze wykorzystania AI banki mają znacznie ograniczone możliwości, z czego wynikać może nierównowaga konkurencyjna. Z tego względu podkreślona została konieczność realizacji założenia „**takie same usługi, takie same zagrożenia, takie same reguły, taki sam nadzór**” (*same services, same risks, same rules and same supervision*).

W omawianym stanowisku Europejskiej Federacji Bankowej (*European Banking Federation*, dalej także jako EBF) odniesiono się również obszernie do etycznego wymiaru sztucznej inteligencji. Proponowane podejście zakłada oparcie się przede wszystkim na **ogólnych wartościach** (*high-level principles*) zamiast przedstawiania szczegółowych wytycznych. Ma to zapewnić odpowiednią elastyczność i zapobiegać hamowaniu innowacji. Podkreślono również postulat **stosunku neutralnego technologicznie**. W związku bowiem z trudnościami z jednoznacznym zdefi-

niowaniem pojęcia sztucznej inteligencji standardy powinny być takie same dla wszystkich technologii (***technology-agnostic ethics standards***). W procesie zastosowania wskazań etycznych istotna jest też oczywiście reguła proporcjonalności.

W wytycznych poruszono również kilka bardziej szczegółowych zagadnień. Wskazano na problem **zapewnienia bezstronności**, który związany jest m.in. z uprzedzeniami zawartymi w zbiorach danych. Zaznaczono jednak, że trudno oczekiwać wypracowania rozwiązań całkowicie wolnych od uprzedzeń. Akceptowalne powinny więc być technologie wykazujące poziom uprzedzeń nie większy niż dostępne alternatywy.

Jako istotną zidentyfikowano również kwestię **transparentności i możliwości wyjaśnienia**. Dostrzeżono przy tym jednak konflikt pomiędzy poziomem wyjaśnialności a dokładnością systemu. W związku z tym możliwość wyjaśnienia powinna być zapewniona proporcjonalnie do istotności zagadnienia i wpływu wyników działania systemu na jednostki. Zaznaczono również, że samo zastosowanie AI nie powinno automatycznie zwiększać wymagań w zakresie wyjaśnialności (podejście neutralnie technologicznie).

W wielu aspektach dotyczących regulacji wykorzystania sztucznej inteligencji wskazano na potrzebę zapewnienia odpowiedniej **równowagi między pozostającymi w konflikcie wartościami** – np. między transparentnością a ochroną własności intelektualnej i tajemnicy handlowej czy między personalizacją usług a zapobieganiem dyskryminacji.

3.5. Stanowisko partnerów społecznych

Jeszcze nieco inną perspektywę ukazuje „**Wspólna deklaracja w sprawie aspektów zatrudnienia związanych ze sztuczną inteligencją opracowana przez europejskich partnerów społecznych w sektorze bankowym**”⁴⁰. Opiera się ona na stanowisku, że **ludzie pozostaną w centrum organizacji pracy** (Preambuła) i na związanej z tym zasadzie **kontroli sprawowanej przez człowieka**. Aplikacje oparte o AI powinny być przejrzyste, uczciwe i bezstronne, pracownicy zaś muszą dopilnować by były odpowiednio wykorzystywane. W dokumencie pojawia się pojęcie „**sprawiedliwej sztucznej inteligencji**” (tak w polskim tłumaczeniu, w oryginale *fair AI*), w którym mieści się **unikanie stronniczości** oraz szanowanie

39 https://www.ebf.eu/wp-content/uploads/2020/03/EBF_037419-Artificial-Intelligence-in-the-banking-sector-EBF.pdf [dostęp: 6 października 2024 r.].

40 <https://www.ebf.eu/wp-content/uploads/2024/07/Signed-Joint-Declaration-Polish.pdf> [dostęp: 6 października 2024 r.].

i **traktowanie priorytetowo praw człowieka**. Termin ten obejmuje więc sposoby projektowania, wdrażania i obsługi systemów AI mające na celu **eliminowanie dyskryminacji**, zapewnianie decyzji wolnych od uprzedzeń, **poszanowanie prawa i etyki** przy gromadzeniu, przechowywaniu i wykorzystywaniu danych osobowych, zagwarantowanie **przejrzystości** i zrozumienia procesów decyzyjnych, utrzymanie przez ludzi kontroli i **odpowiedzialności** za działanie AI i jego wyniki, umożliwienie pracownikom zdobycia umiejętności korzystania z AI, **poszanowanie ludzkiej godności i autonomii** (w tym nienaruszanie praw i wolności), zwiększanie ludzkich możliwości zamiast ich zastępowania oraz utrzymanie **bezpieczeństwa w interakcji człowiek-maszyna**.

Bardziej szczegółowe zalecenia obejmują m.in. przeprowadzanie **ocen ryzyka w zakresie bezpieczeństwa i higieny pracy** dotyczących również skutków zarządzania algorytmicznego oraz zapewnianie pracownikom niezbędnych szkoleń. Zwrócono także uwagę na zagrożenia związane z możliwością wykorzystywania sztucznej inteligencji do **monitorowania pracowników** – tego typu rozwiązania powinny być ograniczone, przejrzyste i proporcjonalne, a decyzje co do ich zastosowania muszą być podejmowane przez człowieka i być zgodne z zasadą kontroli sprawowanej przez człowieka. Wymieniono również prawo pracowników do niepodlegania decyzjom, które mają na nich istotny i prawny wpływ, a które oparte są wyłącznie na zautomatyzowanych zmiennych. Podporządkowane zasadzie kontroli przez człowieka powinny być zwłaszcza rozstrzygnięcia dotyczące takich spraw jak awans wewnętrzny, zmiany funkcji i poziomu zaszerogowania zawodowego, system sankcji i oceny wyników.

3.6. Podsumowanie

Soft law tworzone przez organy Unii Europejskiej oraz przez działające na poziomie europejskim organizacje branżowe, a także praktyka ponadnarodowych organów i innych organizacji międzynarodowych, ma istotne znaczenie z punktu widzenia Polski, w tym krajowego sektora bankowego.

Nawet bowiem jeśli nie stanowi norm prawnie wiążących, to przez udział przedstawicieli krajowych instytucji w jego tworzeniu i wdrażaniu, a także przez inspirujący wpływ, jaki wywiera na narodowe strategie i wewnętrzne korporacyjne wytyczne, wyznacza trendy i kierunki w podejściu do przemian technologicznych. Z pewnością aktywność instytucji unijnych związana ze sztuczną inteligencją jest bardzo intensywna, czego dowodzić może przyjęcie AI Act. Nawet jeśli te dynamiczne zmiany regulacyjne spowodowały częściową dezaktualizację istotnych unijnych dokumentów programowych, to jednak znaczna część zawartych w nich zaleceń etycznych zachowuje swoje znaczenie.

Niezależnie od pewnych różnic między perspektywami prezentowanymi przez unijną administrację, przedstawicieli sektora bankowego i partnerów społecznych, wszystkie strony dostrzegają istotną rolę etycznego wymiaru wdrażania systemów AI. Podkreślane są w związku z tym takie wyzwania jak **zapobieganie dyskryminacji**, zapewnienie **nałego poziomu ochrony danych**, dbałość o **transparentność i wyjaśnialność** proporcjonalną do istotności działań wykonywanych przez system, a także uwzględnienie odpowiedniego **udziału człowieka** i właściwej dystrybucji **odpowiedzialności**.

4. Republika Federalna Niemiec



Rzeczywistość prawna Republiki Federalnej Niemiec jako państwa członkowskiego UE jest w dużej mierze kształtowana przez unijną legislację. Zwłaszcza w obszarze regulacji systemów AI zasadnicze znaczenia miało współuczestniczenie w kształtowaniu, a później wdrażanie tzw. AI Act (istotne są oczywiście również inne akty prawa europejskiego – przykładowo RODO). Unijne zaangażowanie w tym obszarze sprawiło, że działania legislacyjne na poziomie krajowym miały ograniczony zakres, choć można wskazać sfery, w których przyjęto nowe regulacje (o pionierskim, można powiedzieć, charakterze) – np. ramy prawne dla autonomicznych pojazdów⁴¹.

4.1. Ogólna strategia wobec sztucznej inteligencji

Podstawowe założenia niemieckiej polityki w kwestii podejścia do rozwoju AI przedstawia „Strategia rządu federalnego dotycząca sztucznej inteligencji” (niem. *Strategie Künstliche Intelligenz der Bundesregierung* pierwsza wersja: 2018 r., wersja zaktualizowana, w tym wzbogacona o pierwsze doświadczenia związane z pandemią COVID-19: 2020 r.)⁴². Jest to dokument o charakterze ogólnym, który wskazuje przede wszystkim na kierunki przyszłych działań. Motywem przewodnim strategii jest **podejście do AI skoncentrowane na człowieku** (*human-centric approach*). Rozwój technologii wykorzystujących AI ma więc wzmacniać konkurencyjność niemieckiej gospodarki, ale przy uwzględnieniu **korzyści dla całego społeczeństwa oraz przestrzegania praw człowieka**. Pod uwagę brany jest również **zrównoważony rozwój**, w tym aspekt ochrony środowiska.

Cały dokument podzielono na pięć segmentów, które dotyczą odpowiednio: wiedzy technicznej (*minds*), badań (*research*), zastosowania (*transfer and application*), ram regulacyjnych (*regulatory framework*) oraz społeczeństwa (*society*). Szczególnie interesująca jest tu część dotycząca regulacji, która zawiera ogólne założenia tworzenia ram prawnych dla AI. Celem regulacji tego sektora jest oczywiście uniknię-

cie związanego z jego rozwojem ryzyka. W strategii zaznaczono jednak, że w dziedzinach cechujących się znacznym potencjałem innowacyjności związanym z AI **regulacje powinny stymulować innowacje, a nie je uniemożliwiać**. Chodzi więc o stworzenie stabilnych ram prawnych dla podmiotów zaangażowanych w rozwój i wykorzystanie AI (w tym także dla społeczeństwa i administracji publicznej). Należy przy tym zwrócić uwagę na zapewnienie odpowiedniego poziomu **transparentności i możliwości śledzenia** (*traceability*) działania tego typu technologii, jak również na **przeciwdziałanie dyskryminacji, manipulacji** i innym potencjalnym nadużyciom oraz na **ochronę praw człowieka**. Niezbędny jest więc monitoring obowiązującego prawa pod kątem zmian pożądanых ze względu na rozwój technologiczny (ze szczególnym uwzględnieniem dziedzin takich jak np. prawo pracy). Istotną rolę w odpowiedzialnym rozwoju systemów AI ma też odgrywać standaryzacja. W związku z tym opracowano „Mapę drogową standaryzacji w obszarze sztucznej inteligencji” (niem. *Deutschen Normungsroadmap Künstliche Intelligenz*, w 2022 roku ukazała się jej druga edycja)⁴³, której wdrażanie zaleca strategia.

Analizując zarówno prawo powszechnie obowiązujące, jak i *soft law* w Niemczech, trzeba wziąć pod uwagę federalną strukturę kraju. Rozdział kompetencji między władze centralne i landy sprawia bowiem, że realizacja konkretnych założeń politycznych może być bardziej złożona lub utrudniona⁴⁴, a dzielona odpowiedzialność dotyczy również sfer istotnych z punktu widzenia AI, jak np. ochrona danych osobowych. W tym kontekście warto także wspomnieć, że 5 spośród 16 niemieckich landów przyjęło własne strategie dotyczące sztucznej inteligencji⁴⁵.

4.1. Podejście organów regulacji i nadzoru

Oprócz tych ogólnych strategii na poziomie federalnym i landowym, zainteresowanie konsekwencjami rozwoju systemów AI jest także wyraźnie dostrzegalne pośród organów regulacji i nadzoru. Wyraża się ono przede wszystkim w przygotowy-

41 Raport OECD dotyczący sztucznej inteligencji w Niemczech *OECD Artificial Intelligence Review of Germany* (https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/06/oecd-artificial-intelligence-review-of-germany_c1c35ccf/609808d6-en.pdf) [dostęp: 6 października 2024 r.], s. 98.

42 https://www.ki-strategie-deutschland.de/?file=files/downloads/Fortschreibung_KI-Strategie_engl.pdf&cid=955 [dostęp: 6 października 2024 r.].

43 <https://www.din.de/resource/blob/916798/ed09ae58b60f0d3a498fa90fa5085b7c/nrm-ki-engl-2023-final-web-250-neu-data.pdf> [dostęp: 6 października 2024 r.].

44 Raport OECD, s. 97.

45 Tamże (we wskazanym miejscu dostępne krótkie omówienie działań landów i gmin w tym zakresie).

waniu raportów, organizacji spotkań, publikacji materiałów prasowych, czasem przybiera także formę bardziej skonkretyzowanych zaleceń. Aktywność w tym obszarze podejmują m.in. Federalne Biuro ds. Bezpieczeństwa Informacji (*Bundesamt für Sicherheit in der Informationstechnik*)⁴⁶, Federalny Komisarz ds. Ochrony Danych i Wolności Informacji (*Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit*)⁴⁷ we współpracy z landowymi organami działającymi w tym obszarze⁴⁸, Federalny Urząd ds. Konkurencji (*Bundeskartellamt*)⁴⁹, Niemiecki Bank Federalny (*Deutsche Bundesbank*)⁵⁰ oraz Federalny Urząd Nadzoru Finansowego (*Bundesanstalt für Finanzdienstleistungsaufsicht*, dalej także jako *BaFin*)⁵¹.

W raportach i opracowaniach powyższych instytucji dostrzega się, że rozwój technologii opartych na AI wywiera istotny wpływ na rynek finansowy, w tym bankowy⁵². Zmieniają się oczekiwania odbiorców

usług bankowych zwłaszcza w obszarze obsługi klienta (dotyczy to m.in. możliwości korzystania z produktów *online* i 24/7, intuicyjności stron internetowych i aplikacji, a także personalizacji i indywidualizacji) oraz co do szybkości procesów podejmowania decyzji⁵³.

To sprawia, że ważną rolę na rynku zaczynają odgrywać **nowe podmioty**, zmienia się dystrybucja zysków, w pewnym stopniu następuje oderwanie procesów *middle* i *back office* od kontaktów z klientem, coraz bardziej znaczące stają także **wielkie koncerny technologiczne (*bigtech*)**⁵⁴. Z drugiej strony wzrasta **rola danych** gromadzonych przez banki (np. dotyczących płatności), które mogą być wykorzystywane przez systemy AI m.in. do przewidywania zachowań konsumentów. Otwiera to drogę dla nowych modeli biznesowych (np. dostarczanie usług finansowych bez opłat przez *bigtechy* w celu zbierania danych możliwych do wykorzystania w ich podstawowej działalności), w tym również dla możliwości czerpania zysków przez tradycyjne banki poprzez udostępnianie za opłatą gromadzonych danych podmiotom trzecim (oczywiście wymagałoby to zgody podmiotów, których dane byłyby udostępniane)⁵⁵. Dostrzegana jest również możliwość wykorzystania AI w procedurach **compliance**. Z wymienionymi zmianami wiążą się określone wyzwania etyczne i prawne.

Wydaje się, że można wyróżnić dwa obszary, które są najczęściej dotykane w tym kontekście:

- a) **Ochrona danych osobowych i bezpieczeństwo informacji** – generalnie nie są to kwestie charakterystyczne wyłącznie dla sektora finansowego, ale ze względu na szczególną wrażliwość gromadzonych przez banki danych (np. dotyczących sytuacji finansowej jednostek) poziom ochrony jest w tym segmencie szczególnie istotny.
- b) **Wykorzystanie mechanizmów opartych na AI w procesach decyzyjnych.**

4.3. Problematyka ochrony danych

Związek AI z problematyką **ochrony danych** jest bardzo silny ze względu na ilości danych, jakie są

46 Zob. przykładowo raporty: *Generative AI Models Opportunities and Risks for Industry and Authorities* (https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/Generative_AI_Models.pdf?__blob=publicationFile&v=4) [dostęp: 6 października 2024 r.], *Security of AI-Systems: Fundamentals Security Considerations for Symbolic and Hybrid AI* (https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/Security-of-AI-systems_fundamentals_considerations_symbolic_hybrid.pdf?__blob=publicationFile&v=2) [dostęp: 6 października 2024 r.].

47 Zob. sekcja 4.2 (s. 33-35) dotyczącą AI w rocznym raporcie z aktywności za rok 2021 (https://www.bfdi.bund.de/SharedDocs/Downloads/EN/Taetigkeitsberichte/30TB_21.pdf?__blob=publicationFile&v=3) [dostęp: 6 października 2024 r.], wystąpienie Federalnego Komisarza dotyczące sztucznej inteligencji na I Międzynarodowym Kongresie Ochrony Danych - https://www.bfdi.bund.de/SharedDocs/Downloads/EN/DokumenteBfDI/Reden_Gastbeitraege/2021/Artificial-Intelligence.pdf?__blob=publicationFile&v=3 [dostęp: 6 października 2024 r.].

48 Zob. Deklaracja z Hambach w sprawie sztucznej inteligencji (https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/publikationen/DSK/2019/2019-DSK-Hambach_Declaration_AI-en.pdf) [dostęp: 6 października 2024 r.].

49 Zob. przykładowo raport roboczy *Algorytmy i konkurencja* (https://www.bundeskartellamt.de/SharedDocs/Publikation/EN/Berichte/Algorithms_and_Competition_Working-Paper.pdf?__blob=publicationFile&v=3) [dostęp: 6 października 2024 r.].

50 Zorganizowano np. konferencję *Uczenie maszynowe, sztuczna inteligencja i big data – perspektywa banków centralnych* (<https://www.bundesbank.de/en/bundesbank/international-central-bank-dialogue/international-central-banking-courses-/machine-learning-artificial-intelligence-and-big-data-the-perspective-of-central-banks-917800>) [dostęp: 6 października 2024 r.].

51 W tym przypadku wskazać można wiele przykładów aktywności, które ze względu na rolę wskazanego organu dla sektora finansowego zostaną bardziej szczegółowo omówione w dalszej części tekstu.

52 W szczególności zob. raport BaFin *Big data meets artificial intelligence. Challenges and implications for the supervision and regulation of financial services* (https://www.bafin.de/SharedDocs/Downloads/EN/dl_bdai_studie_en.pdf?sessionId=006321B7DFFD98CA2CAB324969C14611.internet981?__blob=publicationFile&v=3) [dostęp: 6 października 2024 r.] – zwłaszcza rozdział 5.2 stanowiący analizę wpływu AI i big data na rynek bankowy, a także rozdział 4.8 wspominanej wyżej *Mapy drogowej standaryzacji ...* poświęcony analizie obszaru usług finansowych.

53 *Big data meets ...*, s. 63-64.

54 Tamże, s. 65-70.

55 Tamże, s. 71, 76-77.

wykorzystywane przez tego typu systemy. Dane są bowiem używane nie tylko w toku pracy już dopuszczonego do działania systemu, ale są niezbędne na etapach jego tworzenia, kalibracji i walidacji. Rola regulacji związanych z ochroną danych osobowych jest więc w tym obszarze szczególnie. Jako dotyczącą tego zagadnienia wprost wskazać można **Deklarację z Hambach** przyjętą na 97. Konferencji Federalnych i Landowych Niezależnych Organów Ochrony Danych (odesłanie w przypisie 35.). Wyodróżniono w niej siedem wymagań ochrony danych w kontekście sztucznej inteligencji (stanowią one przede wszystkim zastosowanie wymogów RODO w obszarze AI):

- AI nie może traktować ludzi przedmiotowo,
- AI może być wykorzystywane jedynie do dopuszczalnych konstytucyjnie celów i nie może naruszać wymagań ograniczenia celu,
- AI musi być transparentne, zrozumiałe i wyjaśnialne,
- AI musi unikać dyskryminacji,
- zasada minimalizacji ilości danych odnosi się także do AI,
- AI wymaga odpowiedzialności,
- AI wymaga standardów technicznych i organizacyjnych.

Warto tu zwrócić uwagę na kilka aspektów. Po pierwsze, podkreśla się **wymóg zrozumiałości** w kontekście danych, podmiot danych musi mieć więc dostęp do zrozumiałych informacji na temat procesu, w jakim jego dane są przetwarzane. Oprócz tego zaakcentowana została konieczność **minimalizacji ilości wykorzystywanych danych**. Przetwarzanie danych osobowych musi być **ograniczone do tego, co niezbędne**. Dotyczy to także AI, w tym danych szkoleniowych. Wreszcie wykorzystanie danych przez AI wymaga jasnego wskazania kontrolera, który jest odpowiedzialny za ten proces.

Poza tymi założeniami, warto również wskazać na wyzwanie, jakie w swoim raporcie zidentyfikował BaFin, tj. rosnącą presję wywieraną na konsumenta, by ujawniał swoje dane. W tym kontekście pojawia się pytanie, czy organy nadzoru powinny podejmować działania w celu zagwarantowania odbiorcom możliwości wyboru konwencjonalnych lub wykorzystujących

cyh mniej danych osobowych usług (*services that are economical with personal data*)⁵⁶.

W raportach i analizach dostrzegane są również zagrożenia dotyczące **cyberbezpieczeństwa** (w tym bezpieczeństwa wykorzystywanych przez AI danych). Można tu przykładowo wskazać na nowe rodzaje zagrożeń charakterystyczne dla AI. Jednym z nich jest atak typu *data poisoning*, którego celem jest manipulacja danymi szkoleniowymi tak, by wypaczyć późniejsze działanie systemu⁵⁷.

4.4. Problematyka udziału systemów AI w procesach decyzyjnych

Drugi wskazany obszar wyzwań, a więc **udział systemów AI w procesach podejmowania decyzji**, stanowi punkt, w którym spotykają się kluczowe problemy etyczne dotyczące sztucznej inteligencji, takie jak np. **ochrona ludzkiej godności i autonomii oraz praw człowieka, ryzyko dyskryminacji, a także odpowiedzialność za działania AI**⁵⁸.

Mamy tu więc do czynienia z charakterystycznymi wyzwaniami – złożonością systemów, krótkimi okresami rekalkulacji, problemem wyjaśnialności działania systemów oraz możliwością przenoszenia obecnych w społeczeństwie uprzedzeń (*bias*) i stereotypów do mechanizmów działania AI⁵⁹. Obszar zaś bankowości jest pod tym względem newralgiczny ze względu na charakter decyzji, jakie są w nim podejmowane, które oddziałują na „ważne warunki wstępne partycypacji społecznej”⁶⁰ (np. dostęp do kredytu). W związku z tym zagadnienie to stało się obiektem zainteresowania organów regulacji i nadzoru.

Wskazać tu należy zwłaszcza na wytyczne BaFin dotyczące wykorzystania algorytmów w procesie podejmowania decyzji mające charakter *soft law* mogącego „służyć jako przewodnik dla podmiotów

56 *Big data meets ...*, s. 179.

57 *Mapa drogowa ...*, s. 243.

58 Por. stanowisko Federacji Niemieckich Organizacji Konsumentów (Verbraucherzentrale Bundesverband) *Algorytmiczne podejmowanie decyzji dla dobra konsumentów* (https://www.vzbv.de/sites/default/files/downloads/2019/07/19-06-25_vzbv_positions_adm_control_summary_en.pdf) [dostęp: 6 października 2024 r.].

59 Por. *Mapa drogowa ...*, s. 232.

60 Tamże, s. 233.

kontrolowanych przez BaFin⁶¹. Dokument ten jako zbieżny z tematyką niniejszego opracowania wymaga bardziej szczegółowego omówienia.

Definiując zakres zastosowania zaleceń, trzeba stwierdzić, że dokument ten, jak wiele podobnych tekstów, napotyka na problem z precyzyjnym zdefiniowaniem terminu *sztuczna inteligencja*⁶². Po przez odwołanie do wcześniejszego raportu BaFin⁶³ pojęcie to zostaje wyjaśnione jako „połączenie *big data*, zasobów obliczeniowych oraz uczenia maszynowego”. Kluczowymi zaś cechami nowoczesnej sztucznej inteligencji, których rola jest istotna z perspektywy monitorowania ryzyka, są wysoka złożoność algorytmu, krótki cykl rekalkulacji oraz wysoki poziom automatyzacji. Przyznano jednak, że **nie jest możliwe przeprowadzenie jasnego rozróżnienia między AI a tradycyjnymi procesami statystycznymi**. W związku z tym zalecenia znajdują zastosowanie ogólnie do algorytmów, a w szczególności do tych wykazujących wymienione wyżej cechy.

Dokument opiera się na założeniu **proporcjonalnego, zorientowanego na ryzyko i neutralnego technologicznie** (wymagania stawiane podmiotom kontrolowanym są więc takie same niezależnie od wykorzystywanych technologii) podejścia do kontroli. Dodatkowo proces podejmowania decyzji powinien być rozpatrywany jako całość, co oznacza uwzględnienie **kontekstu** działania algorytmu.

Wymienione zostały także kluczowe wartości. Proces podejmowania decyzji wymaga więc, przede wszystkim, **jasnej odpowiedzialności**. Personel zarządzający jest odpowiedzialny za decyzje biznesowe nawet, jeżeli są one oparte na algorytmach, członkowie kierownictwa powinni w związku z tym posiadać **niezbędną techniczną wiedzę**. Zauważono w tym kontekście, że zmiany wprowadzone w 2017 roku przez BaFin ułatwiły specjalistom z dziedziny IT zajmowanie stanowisk w kierownictwie kontrolowanych przez ten organ podmiotów. Również pośród odpowiedzialnych za wewnętrzną kontrolę powinny znaleźć się osoby o odpowiedniej znajomości technologii. Na wszystkich etapach

funkcjonowania algorytmu powinno się dążyć do **zapobiegania decyzjom opartym na uprzedzeniach** (*bias*), skutkować one mogą bowiem opieraniem rozstrzygnięć na fałszywych przesłankach, ale także dyskryminacją (ta ostatnia powinna być wyeliminowana poprzez wprowadzenie odpowiednich procedur weryfikacji).

Szczegółowe zalecenia przedstawiono z podziałem na dwie fazy: projektowania i wdrożenia (*development phase*) oraz zastosowania (*application phase*).

W pierwszej fazie wskazano na konieczność **odpowiedniej dokumentacji** obejmującej etapy wyboru modelu, jego treningu i kalibracji oraz walidacji. Wybierając algorytm, należy odpowiedzieć na pytanie, czy dany algorytm jest **odpowiedni w stosunku do zamierzonego celu**, przeanalizować jego **jakość** oraz **porównać zwiększenie dokładności wyników ze wzrostem złożoności modelu**. Wybór modelu złożonego powinien być zawsze wyjaśniony oraz uzasadniony. Jest to więc procedura w pewien sposób analogiczna do znanego w prawie testu proporcjonalności.

Wraz z procesem walidacji modelu powinno się również wyznaczyć częstotliwość następnych walidacji, a także wskazać okoliczności, które będą uzasadniały walidację *ad hoc* (a w razie potrzeby również rekalkulację lub nawet zmianę modelu) – przykładem takiej przyczyny może być zmiana ram prawnych, w jakich funkcjonuje dany algorytm. Trzeba również pamiętać o potrzebie **doboru odpowiednich danych** na potrzeby kalibracji i walidacji – muszą one być **reprezentatywne** (uwzględniać wszystkie istotne podgrupy), biorąc pod uwagę planowane zastosowanie. Na wszystkich etapach należy również zapewnić zgodność z **wymogami ochrony danych**.

W fazie zastosowania, natomiast, należy zapewnić **odpowiedni poziom zaangażowania pracowników w podejmowanie decyzji**, proporcjonalny do istotności danego problemu i ryzyka z nim związanego. Udział czynnika ludzkiego powinien przy tym przynosić realne korzyści, a nie być sprowadzony jedynie do mechanicznej akceptacji każdej decyzji algorytmu. Istotne jest również zdefiniowanie okoliczności, w których decyzje wymagają dodatkowej, bardziej pogłębionej analizy.

W przypadku procesów o kluczowym znaczeniu należy także przewidzieć **rozwiązania zastępcze** na wypadek pojawienia się problemów z procedurą opartą na AI. Funkcjonujący algorytm powinien być

61 *Big data and artificial intelligence: Principles for the use of algorithms in decision-making processes* (https://www.bafin.de/SharedDocs/Downloads/EN/Aufsichtsrecht/dl_Prinzipienpapier_BDAI_en.pdf?jsessionid=A0EEFF780BB35F090431B5CAD7F90E61.internet982?__blob=publicationFile&v=3) [dostęp: 6 października 2024 r.].

62 Podobny problem został dostrzeżony przy okazji omawiania zaleceń UNESCO.

63 *Big data meets ...*, s. 25 i n.

również poddawany bieżącej ewaluacji i odpowiednio dostosowywany.

Z powyższymi założeniami korespondują wskazania przedstawione przez BaFin w „Minimalnych wymaganiach w obszarze zarządzania ryzykiem”⁶⁴. W ramach punktu AT 4.3.5 scharakteryzowano zasady używania modeli w sektorze bankowym (zaznaczono, że odnoszą się one także do sztucznej inteligencji). Wskazano więc na **odpowiedzialność instytucji za wybór modelu**, konieczność zapewnienia **danych odpowiedniej jakości** oraz potrzebę **regularnej ewaluacji**. Mocniej niż we wcześniej omawianych zaleceniach zaakcentowano natomiast konieczność **zapewnienia wyjaśnialności** (*explainability*) modelu, którą zdefiniowano jako możliwość wskazania wzajemnych zależności między parametrami wejść (*input*) i wyjść (*output*)⁶⁵.

4.5. Stanowisko organizacji konsumentów

Podobieństwo do przedstawionych wyżej zaleceń wykazują postulaty wyrażone w stanowisku Federacji Niemieckich Organizacji Konsumentckich dotyczącym analogicznych kwestii.

Można tu wymienić: konieczność odpowiedniej **oceny ryzyka**, przeprowadzania **procesów kontrolnych** oraz przypisania **odpowiedzialności** za działanie systemów AI. Duży nacisk położono jednak, oprócz tego, na aspekt, który nie został zaakcentowany w dokumentach BaFin, a mianowicie na **transpa-**

rentność użycia algorytmów zarówno w stosunku do jednostek będących w ich strefie oddziaływania, jak i w odniesieniu do opinii publicznej.

4.6. Podsumowanie

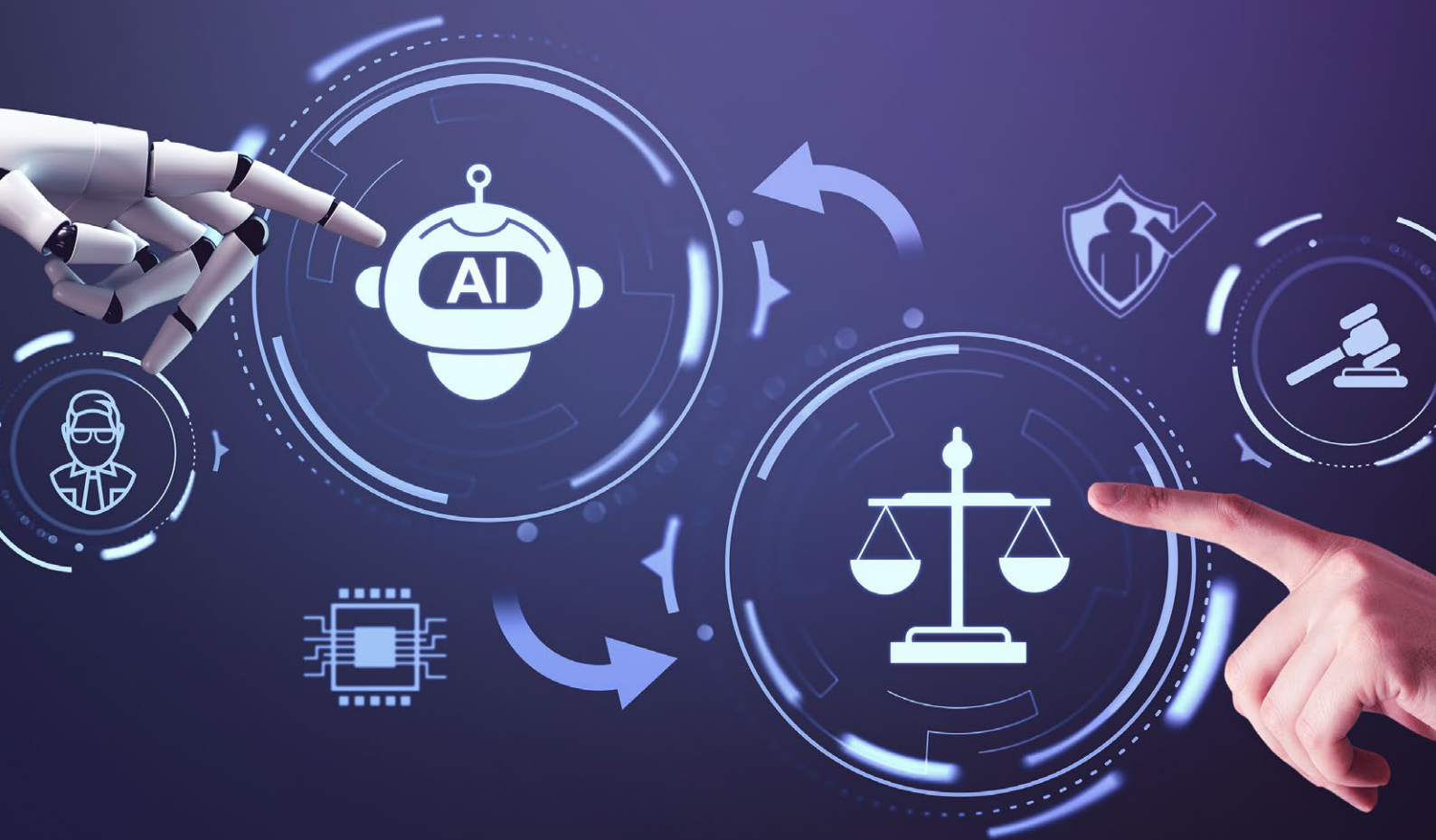
Podsumowując, przedstawiciele zarówno rządu federalnego, jak i poszczególnych organów regulacji i nadzoru bez wątpienia dostrzegają dynamiczny rozwój technologii opartych o sztuczną inteligencję oraz ich wpływ – aktualny i potencjalny – na różne aspekty życia. Zaangażowanie na polu prowadzenia badań oraz tworzenia strategii i wytycznych należy uznać za znaczne, także wśród instytucji odpowiedzialnych za nadzór rynku finansowego. Dostrzeżone przez nich problemy oraz wypracowane rozwiązania mogą stanowić inspirację dla refleksji i działań w Polsce.

W tworzonych raportach i strategiach został wzięty pod uwagę także etyczny wymiar wdrażania sztucznej inteligencji, w tym możliwe napięcia pomiędzy zastosowaniem AI a ochroną godności człowieka i jego praw. W poszczególnych publikacjach i wytycznych mamy do czynienia z różnym rozłożeniem akcentów, ogólnie jednak zwraca się uwagę zwłaszcza na zapewnienie odpowiedniej **transparentności i wyjaśnialności** działania systemów sztucznej inteligencji, **przeciwdziałanie dyskryminacji** (w tym przez używanie danych odpowiedniej jakości), prowadzenie **bieżącego monitoringu, kontroli i ewaluacji**, a także zapewnienie przypisania **odpowiedzialności** za poszczególne etapy funkcjonowania technologii.

64 Minimum Requirements for Risk Management (*Mindestanforderungen an das Risikomanagement – MaRisk*) https://www.bafin.de/SharedDocs/Downloads/EN/Rundschreiben/dl_rs_0523_marisk_ba_en.pdf?__blob=publicationFile&v=3 (wersja z 29.06.2023) [dostęp: 6 października 2024 r.].

65 Zagadnienie znaczenia terminu *explainability* nie jest jednak bezsporne – por. *Mapa drogowa ...*, s. 237, gdzie zakłada się, że należy unikać używania tego pojęcia i nie powinno się ono pojawiać w standaryzacji (preferowane jest zamiast tego pojęcie *transparentności* *transparency*).

5. Wielka Brytania





Choć Wielka Brytania nie jest już państwem członkowskim Unii Europejskiej – z uwagi na jej geograficzną bliskość, niedawne doświadczenia obecności w Unii, a przede wszystkim ze względu na potencjał gospodarczy oraz rolę jaką londyńskie City odgrywa na globalnym rynku finansowym – obowiązujące w niej rozwiązania w zakresie regulacji sztucznej inteligencji zostaną w niniejszym Raporcie uwzględnione.

Zanim przejdziemy do szczegółowej analizy poszczególnych rozwiązań w zakresie korzystania z AI i etyki z tym związanej w sektorze bankowym w Zjednoczonym Królestwie Wielkiej Brytanii i Irlandii Północnej (dalej jako: „UK” lub „Wielka Brytania”), należy poczynić kilka następujących uwag wprowadzających.

Po pierwsze, Wielka Brytania jest państwem z kręgu kultury prawnej *common law*, co niesie ze sobą – tradycyjne, ale wciąż dostrzegalne do dnia dzisiejszego – różnego rodzaju odmienności w sposobie regulacji prawnej, szczególnie z uwagi na kluczową rolę precedensów.

Po drugie, relacje systemu prawnego UK w stosunku do systemu prawnego UE, a także stosunki ekonomiczne, gospodarcze czy społeczne uległy istotnym zmianom w wyniku wyjścia Wielkiej Brytanii z UE (tzw. *Brexit*) w 2020 r. Okoliczność ta ma również fundamentalne znaczenie z punktu widzenia wpływu, jakie na prawo obowiązujące w UK wywiera i będzie wywierać w przyszłości prawo unijne, w tym AI Act.

Po trzecie, Wielka Brytania składa się z czterech części składowych, tj. Anglii, Szkocji, Walii oraz Irlandii Północnej, w których zachodzą pewnego rodzaju odrębności na poziomie administracyjnym, ale także i prawnym. Niemniej, na potrzeby niniejszego Raportu analiza wprowadzanych szeroko pojętych rozwiązań przez podmioty publiczne oraz prywatne w Wielkiej Brytanii, może przynieść cenne oraz inspirowane wnioski nie tylko w kontekście wprowadzenia ewentualnych zmian na poziomie prawa powszechnie obowiązującego, lecz również *soft law*, w tym polskiego Kodeksu Etyki Bankowej⁶⁶.

Poniższa analiza została podzielona na rozwiązania podejmowane przez poszczególne instytucje,

na różnym poziomie, zarówno te publiczne, jak i prywatne, poczynając od rozwiązań wprowadzanych na szczeblu państwowym.

5.1. Rząd

Na wstępie należy zaznaczyć, że rząd Wielkiej Brytanii podszedł do problematyki związanej z AI systemowo, tj. poprzez **podejmowanie działań kompleksowych, integrujących poszczególne zagadnienia oraz mając na uwadze różne gałęzie prawa i obszary ryzyka**. W tym celu wyodrębniono specjalne, dedykowane poszczególnym problemom instytucje lub jednostki (departamenty) w ministerstwach z wyraźnie określonym zakresem zadań i kompetencji, tj. biuro ds. AI (*Office for Artificial Intelligence*) oraz centrum ds. etyki danych i innowacji (*Centre for Data Ethics and Innovation*) – aktualnie oba są częściami departamentu nauki, innowacji i technologii (*Department for Science, Innovation and Technology*), a także utworzono lub rozwinięto wiele podmiotów skupiających swoje działania na badaniach nad AI, np. AI Safety Institute⁶⁷, The Alan Turing Institute⁶⁸, Government Data Quality Hub⁶⁹, AI Standards Hub⁷⁰. Aktywność tych podmiotów skupia się zarówno na działaniach doraźnych, np. organizacji konferencji, seminariów lub innych forum dialogu i wymiany myśli, a także innym wsparciu pracowników akademickich, w tym finansowaniu projektów badawczych lub wdrożeniowych, jak również na działaniach długoterminowych, np. **ustalaniu przyszłych polityk, długofalowych strategii, przygotowywaniu raportów zawierających rekomendacje oraz postulaty *de lege ferenda*** (tj. postulatów regulacji danej materii w przyszłości).

Punktem wyjścia dla rządu Wielkiej Brytanii w obszarze regulacji oraz rozwoju AI było ustanowienie narodowej strategii dot. AI (*National AI Strategy*)⁷¹ we wrześniu 2021 r. Już sama grafika tytułowa strategii obrazuje holistyczne, systemowe podejście do problemu w UK:

67 Zob. www.aisi.gov.uk/ [dostęp: 17 października 2024 r.].

68 Zob. www.turing.ac.uk/ [dostęp: 17 października 2024 r.].

69 Zob. www.gov.uk/government/organisations/government-data-quality-hub [dostęp: 17 października 2024 r.].

70 Zob. ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/ [dostęp: 17 października 2024 r.].

71 Zob. *National AI Strategy*, 2021, assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf [dostęp: 17 października 2024 r.].

66 Zob. zbp.pl/Dla-Bankow/Prawo-i-legislacja/Komisja-Etyki-Bankowej [dostęp: 17 października 2024 r.].

National AI Strategy



W jej ramach wskazano m. in. że w okresie 10 lat UK ma zamiar stać się światowym liderem, jeśli chodzi o rozwój AI. Oprócz jednak samych deklaracji dotyczących tego, jak rozwój AI może przysłużyć się Wielkiej Brytanii, w dokumencie zwrócono także uwagę na kluczowe kwestie związane z zagrożeniami z tym idącymi. Podkreślono, że: (i) jednym z drogowskazów dla podejmowanych działań regulacyjnych oraz rozwojowych ma być „moralny imperatyw” do zapewnienia ludziom ze wszystkich środowisk i części UK **możliwości uczestnictwa i rozwoju** w nowej gospodarce opartej na AI, (ii) w celu uniknięcia etycznych naruszeń, np. w postaci dyskryminacji, **tym bardziej kluczowe jest zapewnienie szerokiego dostępu różnych środowisk społecznych UK**, mając na uwadze uniknięcia selektywności i arbitralności danych, na

których mają się opierać poszczególne modele AI. Zwrócono również uwagę na niebezpieczeństwa na poziomie państwowym związane z *deepfake*, dezinformacją i atakami skierowanymi w konkretne grupy społeczne czy infrastrukturę krytyczną. W tym kontekście warto zauważyć, że szczególnie narażonym na takie niebezpieczeństwa narażony jest rynek finansowy i bankowy. **Dlatego też tak istotne jest wykorzystywanie AI w sposób etyczny, np. poprzez odpowiednią filtrację danych, aby uniknąć dostania się ich w posiadanie organizacji terrorystycznych czy innych niepożądanych podmiotów.**

Kolejnym krokiem, a zarazem głównym i najważniejszym przedsięwzięciem rządu Wielkiej Brytanii, było wydanie tzw. Białej Księgi (*White Paper*), pt. „Proin-



nowacyjne podejście do regulacji AI" (*A pro-innovation approach to AI regulation*)⁷². W przedmowie byłej Sekretarz Stanu Michelle Donelan zidentyfikowano jak ważną rolę będzie odgrywać AI m.in. w sektorze publicznym. Biała księga ma stanowić punkt wyjścia dla Wielkiej Brytanii do stania się „miejszem najbar-dziej przyjaznym dla badań i rozwoju” AI. Cel ten ma zostać osiągnięty w 2030 r. Co więcej, wskazano również, że Wielka Brytania od 2014 r. zainwestowała ponad 2,5 miliarda funtów w rozwój AI. **Zgodnie z zamierzeniami twórców Białej Księgi powinna ona być postrzegana jako dokument ramowy, w którym zdefiniowano zagrożenia płynące z AI.** Jednocześnie zadbane o to, żeby rozwiązania tam przewidziane były proporcjonalne oraz proinnowacyjne, tak żeby nie były hamulcem dla rozwoju AI. Skupiono się więc nie na konkretnych rozwiązaniach technologicznych, a na umieszczeniu AI w odpowiednim kontekście, w którym jest wykorzystywana. Podkreślono także, że bardzo istotna jest w tym zakresie współpraca pomiędzy rządem, organami regulacyjnymi oraz biznesem.

W kontekście etyki zastosowania AI, kluczowe jest zidentyfikowanie przez Białą Księgę tego, że pomimo wielu zalet i korzyści jakie niesie ze sobą AI, to równocześnie stwarza ona wiele zagrożeń dla praw człowieka, szczególnie w kontekście zdrowia fizycznego i psychicznego oraz prywatności. Dlatego też, w opinii brytyjskiego rządu, ważna jest budowa zaufania do AI wśród społeczeństwa, co będzie możliwe wyłącznie wtedy, gdy AI zostanie uregulowane w kompleksowy i dostosowany do potrzeb społecznych sposób uwzględniając jej dynamikę rozwoju.

Głównymi zasadami, na których opiera się Biała Księga, a także na których opierać się będzie regulacja AI w przyszłości są:

- **bezpieczeństwo, ochrona i solidność,**
- **odpowiednia transparentność i wytłumaczalność,**
- **uczciwość,**
- **rozliczalność (odpowiedzialność) i dobre zarządzanie,**

- **zaskarżalność i możliwość dochodzenia roszczeń**⁷³.

Stanowią one wytyczne, dzięki którym wykorzystanie AI będzie „etyczne”. Powyższe zasady początkowo mają zostać wydane jako zalecenia i zostać wdrożone przez istniejące organy regulacyjne.

W tym kontekście warto cofnąć się do 2022 roku, kiedy to wydano dokument określający politykę regulacji AI (*AI Regulation Policy Paper*). Jego celem było zaproponowanie modelu, który koordynowałby organy regulacyjne w sprawach dotyczących AI. Końcowym wynikiem było to, że same organy regulacyjne zwracały uwagę, że efektywniejsza byłaby większa centralizacja w tym zakresie. Dlatego też w Białej Księdze wskazano szereg rozwiązań, które mają zapewnić, że regulacje dotyczące AI oferować będą proporcjonalne, a co najważniejsze elastyczne odpowiedzi na ryzyka wynikające ze stosowania AI z uwzględnieniem odpowiedniego poziomu koordynacji i nadzoru nad organami regulacyjnymi. Szczególnie istotne dla prawodawcy brytyjskiego jest stworzenie takich ram, które będą aktualne nie tylko w czasie uchwalania przepisów, ale będzie można je zastosować również w przyszłości, również w stosunku do nowych wyzwań.

Wśród takich rozwiązań wyróżniono:

- monitorowanie i ocena ogólnej skuteczności ram regulacyjnych oraz implementacji zasad,
- ocena i monitorowanie ryzyk dla gospodarki płynących z wykorzystywania AI,
- przeprowadzanie badań horyzontalnych i analiza luk, w tym poprzez organizację konferencji branżowych, w celu przekazania kompleksowych odpowiedzi na nowe trendy w AI,
- wsparcie platform testowych i inicjatyw typu *sandbox*,
- edukowanie i podnoszenie świadomości w tematyce AI,
- promowanie zgodności z międzynarodowymi ramami regulacyjnymi.

⁷² Zob. *A pro-innovation approach to AI regulation*, 2023, assets.publishing.service.gov.uk/media/64cb71a547915a00142a91c4/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf [dostęp: 17 października 2024 r.].

⁷³ W wersji oryginalnej: „Safety, security and robustness, Appropriate transparency and explainability, Fairness, Accountability and governance, Contestability and redress”.

Oprócz wyżej wskazanych ogólnych sfer, na które negatywny wpływ może mieć wykorzystanie AI, wskazuje się także ryzyka związane m.in. z tzw. zasadą uczciwości. Za przykład podano sytuację, w której **ze względu na niepoprawne dane, na których AI było trenowane, instytucja finansowa oferowałaby pożyczki klientom na różnych warunkach z uwagi na kolor skóry lub płeć.**

Podsumowując, w całym dokumencie konsekwentnie podkreśla się, że rozwój AI jest absolutnym priorytetem, a także to, że nie jesteśmy w stanie dzisiaj wskazać możliwych ścieżek rozwoju w zakresie technologii, wszystkich sposobów wykorzystania, itp. Jednocześnie zwraca się uwagę na zagrożenia z tym idące w zakresie **prawa do prywatności, przetwarzania danych osobowych, ochrony konkurencji i konsumentów, dyskryminacji** i wielu innych. W kontekście więc etycznego stosowania AI należy każdorazowo dokonać testu korzyści jakie niesie ze sobą wprowadzenie danego rozwiązania, przy czym nie chodzi tutaj o prosty rachunek zysków i strat, ale zapewnienie tego, że nie zostaną naruszone jakiegokolwiek prawa człowieka, a w sytuacji, gdyby jednak miało to miejsce, to osoby poszkodowane miałyby odpowiednie środki dochodzenia ochrony swoich praw.

Na marginesie, należy dodać, że pomimo tego, iż Biała Księga nie stanowi źródła prawa, co raczej rodzaj „wytocznych na przyszłość” dla prawodawcy, innych organów władzy publicznej oraz podmiotów stosujących AI, a zarazem rodzaj aktu „miękkiego prawa” (*soft law*) – to Wielka Brytania do tej pory nie była pozbawiona regulacji mających lub mogących mieć wpływ na rozwój AI. Wśród takich aktów prawnych należy wymienić ustawę o ochronie danych (*Data Protection Act 2018*)⁷⁴ oraz ustawę o usługach i rynkach finansowych (*Financial Services and Markets Act 2000*)⁷⁵.

W Wielkiej Brytanii obowiązuje już zatem zrębowa **regulacja prawna w postaci kilku aktów normatywnych, zawierających przepisy *ius cogens* (a zatem tzw. „twarde prawo”)**, które prawdopodobnie będą rozwijane, tj. nowelizowane, uszczegóławiane i rozbudowywane. Przypuszczać należy zatem, że dalsze prace legislacyjne dotyczące materii AI będą polegać raczej na udoskonalaniu dotychczas obowiązujących rozwiązań (w ramach aktów prawnych nor-

mujących szczegółowe zagadnienia takie jak dane osobowe czy usługi finansowe, które „przy okazji” mogą dotyczyć materii AI), zamiast na tworzeniu całkowicie nowych aktów prawnych dedykowanych wyłącznie regulacji sztucznej inteligencji.

5.2. Bank Centralny Anglii i Komisja Nadzoru Finansowego

Dwie centralne instytucje państwowe na rynku finansowym, tj. Bank Centralny Anglii (*Bank of England*, dalej jako: „**Bank Anglii**”) wraz Komisją Nadzoru Finansowego (*Financial Conduct Authority*, dalej jako: „**FCA**”) aktywnie uczestniczą w procesie dostosowywania aktualnych regulacji w Wielkiej Brytanii do wyzwań związanych z wykorzystywaniem AI, jak również są inicjatorami oraz aktywnymi uczestnikami dyskusji nad przyszłą regulacją AI. Oprócz indywidualnych inicjatyw, oba te organy podejmują również wspólne działania.

Bank Anglii, analogicznie jak banki centralne w innych państwach europejskich (np. w Polsce Narodowy Bank Polski), odgrywa kluczową rolę jeśli chodzi o rynek finansowy oraz bankowy. Dlatego też warto przyjrzeć się z bliska proponowanym tam rozwiązaniom w podejściu do AI, a szczególnie w świetle etyki jej stosowania.

Na wstępie należy zacząć od Kodeksu Etyki Banku Anglii⁷⁶, który stanowi punkt wyjścia działań podejmowanych przez Bank Anglii. Został on oparty na czterech filarach, tj.:

- **działania w sposób uczciwy i bezstronny,**
- **otwartości i odpowiedzialności,**
- **bezpieczeństwa i niezawodności (*safe and secure*),**
- **tworzenie inkluzywnej i upodmiotawiającej/ wspierającej kultury.**

Szczególnie istotna w kontekście etycznego stosowania AI jest kwestia „bezpieczeństwa i niezawodności” (*safe and secure*). Za przykład niepożądanego sytuacji wpływającej na tę kwestię Kodeks Etyki Banku Anglii wskazuje sytuację, w której informacje bankowe zostałyby utracone lub ujawnione na skutek dodania ich do modelu AI, który nie został poddany wcześniej odpowiedniej weryfikacji.

⁷⁴ Zob. www.legislation.gov.uk/ukpga/2018/12/contents [dostęp: 17 października 2024 r.].

⁷⁵ Zob. www.legislation.gov.uk/ukpga/2000/8/contents [dostęp: 17 października 2024 r.].

⁷⁶ Zob. www.bankofengland.co.uk/about/governance-and-funding/staff-codes-and-policies/our-code [dostęp: 17 października 2024 r.].

The loss or compromise of Bank information can occur in many forms – physically (where we lose papers or a notebook), verbally (conversations overheard in public or within Bank premises), electronically (emails sent to the wrong recipient or information uploaded to an unapproved Artificial Intelligence solution) or via a third party (where a supplier of ours has suffered a data breach).

Źródło: www.bankofengland.co.uk/about/governance-and-funding/staff-codes-and-policies/our-code, rozdział *Being safe and secure*.

Ponadto Bank Anglii w otwarty sposób komunikuje swoją perspektywę na rozwój AI, określając także jaką on sam powinien pełnić w tym rolę, w szczególności poprzez udział przedstawicieli Banku Anglii na różnego rodzaju wydarzeniach, począwszy od wykładów uniwersyteckich po międzynarodowe konferencje. Przykładowo:

Home / Monsters in the deep? – speech by Jonathan Hall

Monsters in the deep? – speech by Jonathan Hall

Given at University of Exeter Business School

Źródło: www.bankofengland.co.uk/speech/2024/may/jon-hall-speech-at-the-university-of-exeter⁷⁷.

Home / A weathervane for a changing world: refreshing our data and analytics strategy – speech by James Benford

A weathervane for a changing world: refreshing our data and analytics strategy – speech by James Benford

Given at Big Data and AI World, part of Tech Show London, at the ExCeL centre, Royal Victoria Dock, London

Źródło: www.bankofengland.co.uk/speech/2024/march/james-benford-speech-at-the-big-data-ai-world-conference⁷⁸.

FCA prowadzi aktywną działalność w zakresie publikacji materiałów informacyjnych dotyczących rozwoju AI skierowanych do podmiotów działających na rynku finansowym, a także organizuje konferencje czy spotkania branżowe poświęcone tej tematyce.

77 Jonathan Hall (*External member of the Financial Policy Committee*) w swoim wystąpieniu skupił się na tym, jak rozwój AI może wpłynąć na stabilność finansową. Wykład jest szczególnie wartościowy z uwagi na szczegółowy opis zarówno części dotyczącej AI, w szczególności tzw. *deep neural networks*, a także części dotyczącej tego, w jaki sposób rynki finansowe wpływają na ekonomię. Wśród zagrożeń związanych z wykorzystaniem AI, Jonathan wskazał w szczególności na AI działającą jako makler giełdowy (*deep trading agents*) – czyli przy wykorzystaniu uczenia głębokiego, jako podkategorii ML (*deep learning*).

78 James Benford (*Executive Director for Data and Analytics Transformation and Chief Data Officer*) oprócz nakreślenia historii Banku Anglii w zakresie zarządzania danymi, skupił się na opisie siedmiu kroków, które były podstawą dla specjalnie utworzonego oddziału Banku Anglii (*Independent Evaluation Office*) w zakresie zmiany podejścia do przetwarzania danych, również przy wykorzystaniu narzędzi AI.

Speech by Jessica Rusu, FCA Chief Data, Information and Intelligence Officer, at The Alan Turing Institute's Framework for Responsible Adoption of Artificial Intelligence in the Financial Services Industry (FAIR) event.



Źródło: www.fca.org.uk/news/speeches/building-better-foundations-ai⁷⁹.

Speech by Jessica Rusu, FCA Chief Data, Information and Intelligence Officer, at the City and Financial Global AI Regulation Summit 2023.



Źródło: www.fca.org.uk/news/speeches/ai-flipping-coin-financial-services⁸⁰.

Annual public meeting Annual Public Meeting 2022 October 12, 2022 10:00 AM (Europe/London)	FCA / PSR APP Fraud TechSprint Demo Showcase 29 September 2022 Authorised Push Payment Fraud TechSprint Demo... September 29, 2022 01:30 PM (Europe/London)	Restructuring and Insolvency of FCA regulated... September 22, 2022 10:00 AM (Europe/London)	Innovation Open Day 13 July 2022 Innovation Open Day July 13, 2022 10:00 AM (Europe/London)
RegTech Forum RegTech Forum – FCA Approach to Innovation &... February 10, 2022 02:00 PM (Europe/London)	Operational Resilience January 27, 2022 10:30 AM (Europe/London)	IFPR – Practical implementation November 30, 2021 02:00 PM (Europe/London)	IFPR – Technical session November 30, 2021 10:00 AM (Europe/London)

Źródło: webinars.fca.org.uk/webinars.

- 79 Jessica Rusu (FCA Chief Data, Information and Intelligence Officer) podkreślała w swoim wystąpieniu, że (i) regulacja AI na rynku usług finansowych powinna być tworzona poprzez **współpracę podmiotów**, które na tym rynku funkcjonują, w szczególności organów regulacyjnych, (ii) obszarami, na które trzeba zwrócić szczególną uwagę dalej rozwijając AI są **zasady inkluzywności i różnorodności**, (iii) coraz to lepsze wyniki organów regulacyjnych w zakresie zastosowania AI są uzależnione od wdrażania i inwestowania w kolejne innowacje.
- 80 Jessica Rusu (FCA Chief Data, Information and Intelligence Officer) wskazywała w 2023 r., że był to czas przełomowy jeśli chodzi o obranie kierunku, w którym powinno podążać podejście do regulacji nie tylko AI, ale również innych nowych technologii. Podkreśliła również, że obowiązkiem organów regulacyjnych jest **zapewnienie bezpiecznego i odpowiedzialnego stosowania AI na rynku usług finansowych**, mając na uwadze w szczególności **zapewnienie odpowiedniej infrastruktury cyfrowej, odporności systemów AI i bezpieczeństwa konsumentów**.

Warta uwagi jest również polityka FCA polegająca na ciągłym monitoringu **podmiotów, które działają na rynku finansowym, ale nie są w żaden sposób przez nią autoryzowane**⁸¹. Podmioty te próbują wykorzystać popularność tematyki AI i oferują różnego rodzaju instrumenty finansowe, np. udzielanie pożyczek lub inwestowanie w kryptowaluty. Liczba tego rodzaju podmiotów wskazuje na powagę zagrożenia, jaki niesie ze sobą dynamiczny rozwój AI. Dlatego też podmioty rynku finansowego, w szczególności bankowego, powinny wziąć pod uwagę nie tylko zagrożenia „wewnętrzne”, takie jak niewłaściwe zastosowanie AI w procesie udzielania klientowi kredytu lub innej usługi, ale też **zagrożenia „zewnętrzne” w postaci wyłaniania się podmiotów upodabniających się do banków lub oferujących podobne produkty wykorzystując przy tym AI**, np. poprzez trenowanie modelu AI na podstawie danych, które wyciekły lub w inny sposób zostały udostępnione nieupoważnionym osobom trzecim.

Warto również zwrócić uwagę w tym kontekście na **transparentność w zakresie reklamowania swoich usług przy wykorzystaniu AI lub promowanie instytucji poprzez zawarcie w reklamie, że w jej działalności wykorzystywana jest AI**. Podobne bowiem reklamy zamieszczają również ww. nieautoryzowane przez FCA podmioty na różnego rodzaju platformach internetowych, w tym *social mediach* (m. in. Tik Tok). W interesie więc instytucji finansowych (w tym banków) powinno być takie skonstruowanie przekazu do klienta, aby mógł on już na pierwszy rzut oka odróżnić autoryzowaną instytucję od tzw. *scamu* lub podmiotów prowadzących działalność parabankową.

W kwietniu 2024 r. FCA opublikowała również dokument uaktualniający podejście FCA do AI (*FCA AI Update*)⁸². Wskazano w nim, że dotychczas FCA skupiała swoją uwagę na promocji bezpiecznego i odpowiedzialnego użycia AI na rynku finansowym, dzięki któremu AI będzie napędzało korzystne innowacje. Owe korzystne innowacje mają bowiem pozytywnie wpłynąć na konkurencyjność rynku finansowego. Do tej pory FCA mogła się pochwalić

wspieraniem takich innowacji, m. in. poprzez tzw. sandbox regulacyjny, sandbox cyfrowy, TechSprints. Określono również cele na najbliższe 12 miesięcy, w których podkreślono następujące kwestie:

- kontynuację prac związanych z wdrażaniem AI na rynku finansowym w UK, m. in. poprzez otwarcie trzeciej edycji ankiety nt. uczenia maszynowego (dalej jako: „ML”) wspólnie z Bankiem Centralnym oraz organem regulacyjnym do spraw płatności (*Payment System Regulator*)⁸³,
- kontynuację prac na aktualnie dostępnych technologiach, np. na dużych modelach językowych (tzw. LLMs),
- zacieśnienie współpracy z innymi organami regulacyjnymi, instytucjami naukowymi oraz podmiotami rynku finansowego,
- rozszerzenie międzynarodowej współpracy.

Wspólne działania Banku Centralnego i FCA w zakresie rozwoju i regulacji AI rozpoczęły się już w 2020 r. kiedy to ustanowiono Publiczno-Prywatne Forum AI (*AI Public-Private Forum*, dalej jako: „AIPPF”). W jego ramach grupa ekspertów z sektora finansowego, technologicznego oraz pracowników naukowych, opracowała raport opublikowany w lutym 2022 r., w którym to zbadano różnego rodzaju przeszkody oraz zidentyfikowano wyzwania i zagrożenia płynące z zastosowania AI w sektorze finansowym. Raport AIPPF powstał na podstawie czterech kwartalnych spotkań ekspertów oraz licznych warsztatów. Obszary badań i dyskusji zostały podzielone na trzy sfery, tj. (i) dane (*Data*), (ii) model (*Model risk*), (iii) zarządzanie (*Governance*). W ramach każdego z nich opracowano zalecenia dla praktyki, które obrazuje poniższa tabela oraz zobrazowane odpowiednie części raportu.

Co ważne, jednym z wniosków AIPPF było to, że **sektor prywatny oczekuje pomocy ze strony organów regulacyjnych w bezpiecznej implementacji regulacji AI w sektorze finansowym**.

81 Zob. np. www.fca.org.uk/news/warnings/ai-daily-robot-aidailyrobotcom; www.fca.org.uk/news/warnings/trade101-trade101ai; www.fca.org.uk/news/warnings/quopi-ltd-quopiai; www.fca.org.uk/news/warnings/ai-forex-limited; www.fca.org.uk/news/warnings/crypto-fx-ai-trades; www.fca.org.uk/news/warnings/crypto-comai-cryptocomai [dostęp: 17 października 2024 r.].

82 Zob. *FCA AI Update*, www.fca.org.uk/publication/corporate/ai-update.pdf [dostęp: 17 października 2024 r.].

83 Ankieta nt. uczenia maszynowego jest przeprowadzana przez FCA i Bank Anglii od 2019 r. (I edycja). Każda edycja ma na celu lepsze zrozumienie zmian jakie nastąpiły w zakresie wykorzystywania ML-u na brytyjskim rynku usług finansowym, w tym aktualny stan co do ilości i jakości przypadków użycia ML-u, a także korzyści i zagrożenia jakie z tego płyną. Respondentami tej ankiety są instytucje finansowe oraz inne podmioty działające na rynku finansowym.

	ZALECENIA
DANE	- Wdrożenie procesów śledzenia i pomiaru przepływu danych zarówno wewnątrz, jak i na zewnątrz organizacji. - Przeprowadzanie regularnych audytów danych i oceny ich wykorzystania, w tym zgodności z celami biznesowymi. - Stworzenie systemu oceny wartości posiadanych i wykorzystywanych danych, zarówno dla konkretnych projektów, jak i dla całej instytucji. - Zagwarantowanie odpowiedniego poziomu zrozumienia i dokumentacji pochodzenia danych wykorzystywanych przez modele AI, zwłaszcza w przypadku, gdy pochodzą one od podmiotów trzecich.
MODEL	- Udokumentowany i uzgodniony proces weryfikacji i zatwierdzania AI dla wszystkich nowych zastosowań. - Kompletny wykaz wszystkich używanych i opracowywanych modeli AI. - Przejrzyste udokumentowane metody i procesy identyfikacji i zarządzania stronniczością (błędami poznawczymi spowodowanymi subiektywizmem lub arbitralnością) w danych wejściowych i danych wyjściowych. - Regularne przeprowadzanie ocen wydajności modeli AI, w tym potencjalnego wpływu zewnętrznego. - Klarowne przedstawienie i wyjaśnienie ryzyka związanego z modelami AI oraz zastosowaniem instrumentów minimalizujących ryzyko. Dokumentacja i ocena danych wejściowych modeli AI, w tym jakości i przydatności danych.
ZARZĄDZANIE	- Wzmocnienie kontaktu pomiędzy zespołami zajmującymi się analizą danych a zespołami ds. ryzyka od wczesnych etapów rozwoju modelu AI przy jednoczesnym zachowaniu niezależności. - Ustanowienie centralnego komitetu nadzorującego rozwój i wykorzystanie AI. Instytucje powinny posiadać specyficzne dla AI elementy ram ryzyka, a także ramy prywatności, zasady operacyjne, i co byłoby idealnym rozwiązaniem, zestaw zasad etycznych, które mogą pomóc w podejmowaniu decyzji. Zapewnienie szkoleń i odpowiedniego zrozumienia AI (w tym odpowiedzialności i zasad rozliczalności) w celu zapewnienia odpowiedniego poziomu umiejętności w całej organizacji. Dzielenie się dobrymi praktykami wewnątrz organizacji, np. w zakresie zarządzania danymi, rozwoju oprogramowania i MLOps.

Firms should:

- Consider adoption and use of AI holistically, harmonising, where possible, data and model processes. In particular, they should aim to coordinate data management and strategy with AI management and strategy.
- Have processes in place for tracking and measuring data flows within, as well as into and out of, the organisation.
- Carry out regular data audits and assessments of data usage, including alignment with business objectives.
- Have a clear assessment of the value of the data it holds and uses, both for specific projects, and to the firm as a whole. This should inform the cost/benefit analysis of the AI project.
- Have clear understanding and documentation of the provenance of data used by AI models, especially in the case of third-party data.
- Have a clear understanding of the limitations and challenges of using alternative and/or synthetic data.

- Firms should have:
 - A documented and agreed AI review and sign-off process for all new applications.
 - A complete inventory of all AI applications in use and in development.
 - Clearly documented methods and processes for identifying and managing bias in inputs and outputs.
 - Regular assessments of AI application performance, including potential external impact.
 - A clear explanation of AI application risks and mitigation (to be updated as appropriate).
 - Documentation and assessment of AI application inputs, including data quality and suitability.

Firms should:

- Strengthen contact between data science teams and risk teams from the early stages of the model development cycle while maintaining independent challenge.
- Establish a central committee to oversee firm-wide development and use of AI. Firms should have AI-specific elements of the risk framework; as well as a privacy framework; operational principles; and, ideally, a set of ethical principles which can help guide decision-making.
- Provide training and understanding of AI (including responsibility and accountability) with the aim of embedding a sufficient level of skills throughout the organisation.
- Should share good practice across the organisation, for example, from data management, through to software development and MLOps.

Źródło: AIPPF Raport, www.bankofengland.co.uk/-/media/boe/files/fintech/ai-public-private-forum-final-report.pdf, s. 20, 28, 38.

Odpowiedzią na AIPPF było opublikowanie w 2022 r. kolejnego wspólnego dokumentu, tzw. DP 5/22 (DP5/22 – *Artificial Intelligence and Machine Learning*)⁸⁴ (dalej jako: „**DP 5/22**”) przez Bank Anglii wraz z FCA, którego zadaniem było rozpoczęcie dyskusji nad regulacją AI oraz ML. Określono w nim pięć kluczowych obszarów, dla których zarówno ryzyka, jak i potencjalne korzyści związane z wykorzystywaniem AI są szczególnie widoczne.

Pierwszym z nich jest **ochrona konsumentów**, dotyczy to przede wszystkim zagrożeń związanych z niespotykanymi do tej pory możliwościami w zakresie identyfikacji cech osobistych konsumentów i ich preferencji oraz wydawania automatycznych decyzji, które mogą mieć charakter dyskryminacyjny.

Drugim obszarem jest **konkurencja**, dla której szczególnym wyzwaniem może się okazać m.in. wykrywa-

nie cen przez konkurentów czy podwyższenie progu wejścia na rynek.

Kolejnym jest szeroko pojęte **bezpieczeństwo i stabilność sektora finansowego**, który narażony jest np. poprzez problemy z precyzyjnością systemu w zakresie procesu weryfikacji zdolności kredytowej klienta.

Czwartym obszarem zasługującym na uwzględnienie jest **rynek ubezpieczeniowy**, w którym ponownie zidentyfikowano zagrożenia związane z automatycznymi decyzjami niosącymi ryzyko błędów, w szczególności wydawania decyzji na podstawie fałszywych lub niewłaściwych danych.

Ostatni z nich wskazano jako **stabilność finansowa i integralność rynku**. W tej sferze, zasadniczo aktualne pozostają wszystkie zagrożenia z innych obszarów, ponadto wskazano także na zagrożenie płynące z unifikacji rynku AI, przez co ewentualne awarie operacyjne lub cyberataki mogą skutkować zakłóceniami w świadczeniu usług.

Kolejnym krokiem było opublikowanie dokumentu FS 2/23 (FS 2/23 – *Artificial Intelligence and Ma-*

⁸⁴ Zob. DP 5/22 – *Artificial Intelligence and Machine Learning*, 2022, www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence [dostęp: 17 października 2024 r.].

chine Learning)⁸⁵, w którym zebrano odpowiedzi respondentów na pytania zawarte w DP 5/22. Warto wskazać, że według respondentów: (i) **niepotrzebna byłaby regulacyjna definicja AI, lepszym rozwiązaniem byłoby skoncentrowanie się na konkretnych cechach AI**, (ii) **organy regulacyjne, ze względu na dynamiczne zmiany AI, mogłyby reagować poprzez opracowywanie i utrzymywanie „żywych” wytycznych**, (iii) **istniejące struktury zarządzania firmami (i ramy regulacyjne, takie jak system wyższej kadry kierowniczej i certyfikacji (SM&CR)) są wystarczające do przeciwdziałania ryzyku związanemu z AI**.

5.3. Działania podejmowane przez inne instytucje w Wielkiej Brytanii

Oprócz wyżej wskazanych podmiotów, warto również zwrócić uwagę na działalność brytyjskiego Urzędu Ochrony Danych Osobowych (*Information Commissioner's Office*, dalej jako: „brytyjskie UODO”), brytyjskiego związku banków (*UK Finance*, dalej jako: „UK Finance”)⁸⁶ oraz Instytutu Alana Turinga (*The Alan Turing Institute*).

Główną aktywnością brytyjskiego UODO oraz UK Finance dotyczącą AI jest publikowanie materiałów informacyjnych oraz raportów, które są ogólnodostępne dla każdego zainteresowanego na stronach internetowych poszczególnych instytucji⁸⁷.

Więcej uwagi warto poświęcić działaniom **Instytutu Alana Turinga, który jest państwową instytucją zajmującą się nauką o danych oraz AI**. W ramach swojej działalności Instytut Alana Turinga współpracuje z innymi państwowymi jednostkami, a także podmiotami prywatnymi, tworząc raporty, podsumowania, itp. Oprócz tego prowadzi szeroko zakrojoną politykę informacyjną, zwłaszcza dotyczącą AI. Jednym z wymienionych pionów, w ramach których funkcjonuje Instytut Alana Turinga jest *Artificial Intelligence (Safe and ethical)*⁸⁸.

Podstawami dla prac prowadzonych w tym pionie jest interdyscyplinarność oraz współpraca pomiędzy ekspertami, poszczególnymi branżami, rządem, organami regulacyjnymi oraz społeczeństwem. Głównym celem z tym związanym jest utworzenie centrum doskonałości dla studiów nad technicznymi aspektami **bezpiecznej i etycznej AI**.

W celu przeprowadzenia efektywnych badań wyznaczono cztery kluczowe obszary, tj.:

- **zwiększanie przejrzystości i zrozumiałości AI;**
- **poprawa sprawiedliwości systemów algorytmicznych, w tym sposoby pomiaru i łagodzenia stronniczości (*bias*);**
- **rozwijanie niezawodnych systemów, które dobrze dostosowują się do nowego środowiska, zabezpieczają się przed atakami i szanują prywatność;**
- **rozwijanie systemów, które skutecznie współpracują z ludźmi, utrzymują odpowiedni poziom kontroli nad nimi przez człowieka i zapobiegają niepożądanemu wpływowi z zewnątrz.**

W kontekście powyższego, eksperci Instytutu Alana Turinga przeprowadzili (i nadal przeprowadzają) szereg warsztatów dotyczących m.in. kwestii etycznych korzystania z AI, które są kluczowe dla sektora finansowego, tj. uczciwości, automatycznych decyzji, czy prywatności danych⁸⁹.

5.4. Wielka Brytania: podsumowanie

Zagadnienia etyczne związane z użyciem AI w sektorze bankowym, a szerzej w sektorze finansowym, w Wielkiej Brytanii są przedmiotem stałych badań. **Do tej pory nie powstały żadne akty prawne ani dokumenty, które moglibyśmy określić jako tzw. *soft***

85 Zob. *FS 2/33 – Artificial Intelligence and Machine Learning*, 2023, www.bankofengland.co.uk/prudential-regulation/publication/2023/october/artificial-intelligence-and-machine-learning [dostęp: 17 października 2024 r.].

86 Wcześniej jako *British Bankers' Association (BBA)* – od 1 lipca 2017 r. zostało przekształcone w *UK Finance*.

87 Zob. np. ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/; ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/; www.ukfinance.org.uk/events-training/free-webinar-ethical-principles-advanced-analytics-and-artificial-intelligence-aaai/; www.ukfinance.org.uk/blogs/artificial-intelligence-finance-sector/; www.ukfinance.org.uk/system/files/2019_FINAL_ONLINE.pdf [dostęp: 17 października 2024 r.].

88 Zob. www.turing.ac.uk/research/research-programmes/artificial-intelligence-ai/safe-and-ethical [dostęp: 17 października 2024 r.].

89 Zob.: view.officeapps.live.com/office/view.aspx?src=https%3A%2F%2Fwww.turing.ac.uk%2Fsites%2Fdefault%2Ffiles%2F2023-12%2Fpolice_frt_workshop_report.docx&wdOrigin=BROWSELINK; www.turing.ac.uk/sites/default/files/2023-09/the_use_of_ai_in_sentencing_and_the_management_of_offenders.pdf; Szerzej o relacji prawa antydyskryminacyjnego UE a AI w: S. Wachter, B. Mittelstadt, C. Russel, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, „Computer Law & Security Review”, 41/2021, papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922 [dostęp: 17 października 2024 r.].



law, które w sposób kompleksowy i bezpośredni regulowałyby tę kwestię.

Niemniej, liczba oraz różnorodność działań do tej pory podjętych – w tym zwłaszcza opublikowanych raportów, wypowiedzi, opracowań naukowych lub eksperckich, a także Biała Księga – kształtuje pewne ramy, w których możliwe jest zidentyfikowanie kluczowych obszarów będących szczególnie narażonymi w kontekście stosowania AI w sektorze bankowym, tj.:

- **ochrona konsumentów,**
- **ochrona konkurencji,**
- **prywatność danych,**
- **uczciwość, w szczególności w kontekście automatycznych decyzji,**

- **równość = brak dyskryminacji,**
- **bezpieczeństwo wewnętrzne i zewnętrzne,**
- **transparentność,**
- **zrozumiałość i czytelność.**

Należy pamiętać, że korzystanie z AI ma charakter uniwersalny w tym sensie, że co do zasady nie różni się (oprócz ewentualnie kwestii językowych) ze względu na terytorium, w którym ma miejsce jej użycie. Dlatego też, pomimo różnic prawnych, ustrojowych, ekonomiczno-społecznych i gospodarczych, dotychczasowe badania i konkluzje dotyczące etyki AI w Wielkiej Brytanii mogą posłużyć ze cenne wskazówki dla polskiego sektora finansowego, a w szczególności sektora bankowego.

6. Francja

AI

Republika Francuska jest jednym z państw UE, które poczyniło duże nakłady na badania AI (i planuje je dalej zwiększać), a także sformułowało państwową strategię i politykę w tym zakresie. Ponadto Francja jest również zaangażowana w tworzenie polityk dotyczących AI na poziomie unijnym, odgrywając tam jedną z pierwszoplanowych ról. Prezydent Emmanuel Macron podkreśla również chęć bliskiej współpracy w tym zakresie z USA czy z Kanadą oraz to, że **potrzebna jest regulacja AI na poziomie ogólnoświatowym**⁹⁰.

Na wstępie należy zaznaczyć, że z racji pozostawiania w strefie wpływów prawa UE, rozwiązania na poziomie legislacyjnym nie wykazują się znaczącymi różnicami od tych w Polsce. Dlatego też w Raporcie skupiono uwagę na tzw. *soft law*, które zawiera pewne ramy i rekomendacje dla działań czy to instytucji publicznych czy też podmiotów prywatnych, a także ogólnej polityki państwa.

6.1. Strategia i polityka państwa w zakresie rozwoju AI

Za początek działań podejmowanych w kontekście regulacji i dalszego rozwoju AI uznać można rozpoczęcie przez rząd debaty na ten temat w 2017 r.⁹¹. W raporcie przygotowanym przez ówczesnego posła do Parlamentu UE Cédrica Villaniego oraz jego zespół (dalej jako: „**Raport Villaniego**”)⁹² omówiono kluczowe aspekty dla AI, tj.:

- politykę gospodarczą,
- badania naukowe,
- zatrudnienie,
- etykę,
- spójność społeczną.

Do raportu dołączono również pięć załączników, które określały obszary szczególnego zainteresowania, tj.:

- edukacja,
- zdrowie,
- rolnictwo,
- transport,
- obronność i bezpieczeństwo.

Uwagę przy tworzeniu Raportu Villaniego skupiono głównie wokół algorytmów „codziennych”, a nie na czysto spekulatywnych rozważaniach na temat egzystencjalnych zagrożeń ze strony AI. Przede wszystkim, można z niego wywnioskować, że etyczne zastosowanie AI uzależnione jest w głównej mierze nie od samego modelu AI, a od człowieka, który go stworzył albo który następnie wprowadza do niego dane lub inne modyfikacje. Podkreślono także, że kluczowa w kontekście etyki AI jest **wyjaśnialność/zrozumiałość modeli AI, stanowi to bowiem warunek akceptowalności dla jego odbiorców, czyli przede wszystkim społeczeństwa**. W innym wypadku możliwe będzie ukrywanie (nawet nieświadome) źródeł np. wydawanych automatycznie decyzji, które są niesprawiedliwe, stronnictwe czy też mają charakter dyskryminacyjny. Duży nacisk położono również na udział społeczeństwa w zakresie rozwoju AI i jego stosowania. W szczególności **społeczeństwo powinno być dopuszczone do procesów oceny, np. audytów, zastosowania AI w danym obszarze** – w tym względzie to właśnie państwo powinno aktywnie działać na rzecz zaangażowania społeczeństwa.

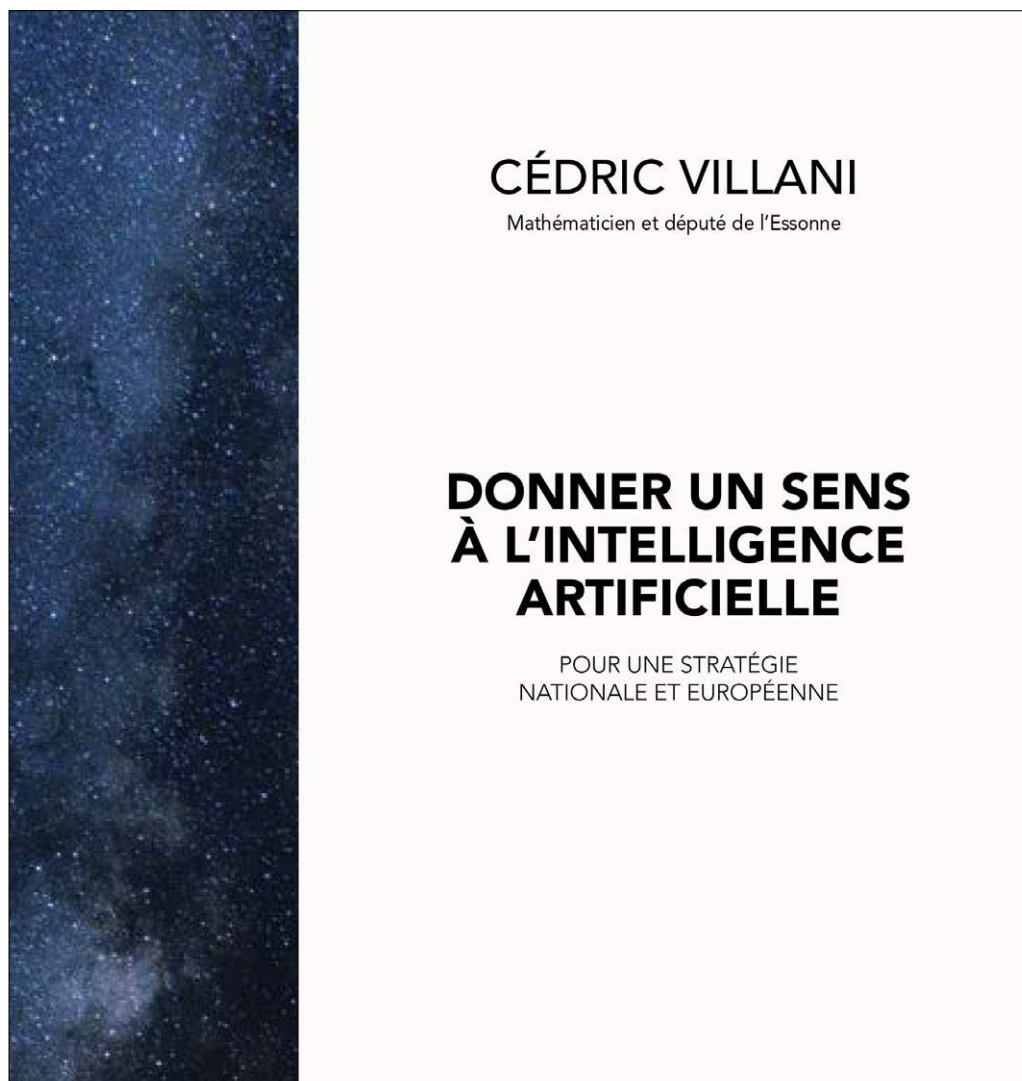
Jak podkreślono w Raporcie Villaniego, jego misją było zaproponowanie rekomendacji umożliwiających stworzenie podstaw etycznych dla rozwoju AI oraz podtrzymania dyskusji na ten temat na odpowiednim poziomie. Rekomendacje te oparto na pięciu założeniach:

- zwiększenia możliwości i przejrzystości audytu systemów autonomicznych;
- odpowiedniego dostosowania ochrony praw i wolności w odniesieniu do potencjalnych nadużyć związanych z wykorzystaniem AI i ML;
- organizacje wdrażające i korzystające z modeli AI muszą ponosić odpowiedzialność prawną za ewentualne szkody przez nie wyrządzone;
- „architekci” społeczeństwa cyfrowego – badacze, inżynierowie i programiści – którzy projek-

90 Zob. A. Kharpal, *France sees global A.I. regulation ideas by the end of this year, wants to work with U.S. on tech laws*, www.cnbc.com/2023/06/15/ai-regulation-france-sees-ideas-on-global-laws-by-end-of-year.html [dostęp: 17 października 2024 r.].

91 Zob. www.economie.gouv.fr/strategie-nationale-intelligence-artificielle [dostęp: 17 października 2024 r.].

92 Zob. C. Villani, *Donner un sens à l'intelligence artificielle*, 2018, www.vie-publique.fr/files/rapport/pdf/184000159.pdf [dostęp: 18 października 2024 r.].



Źródło: www.vie-publique.fr/files/rapport/pdf/184000159.pdf.

tują i wprowadzają te technologie na rynek muszą czynić to w sposób świadomy pod kątem potencjalnych negatywnych skutków swoich technologii na społeczeństwo i aktywnie działać na rzecz ich ograniczania;

- demokratyzacji debaty nad stosowaniem i rozwojem AI, w szczególności poprzez zapewnienie różnorodności i otwartości na społeczeństwo, jako jeden z przykładów podjętych do tej pory w tym zakresie wskazano plany powołania tzw. Izby Przyszłości⁹³.

Na podstawie m.in. Raportu Villaniego ogłoszono w marcu 2018 r. krajową strategię na rzecz AI (*La stratégie nationale pour l'intelligence artificielle*).

Pierwsza jej faza (2018-2022) miała na celu zapewnienie Francji pozycji jednego ze światowych liderów w zakresie AI. W tym zakresie przeznaczono prawie 1,5 mld euro na m.in. promowanie tworzenia i rozwoju sieci interdyscyplinarnych instytutów AI, wspieranie katedr doskonałości w dziedzinie AI, finansowanie programów doktoranckich, wzmacnianie zdolności badawczych, np. dalsze inwestycje w superkalkulator Jean Zay⁹⁴.

⁹³ Celem Izby Przyszłości jest opiniowanie aktów prawnych z określonych obszarów, np. ochrony klimatu, prawa nowych technologii, czy właśnie AI. Założeniem jest to, żeby przedstawiciele społeczeństwa, tj. osoby z różnych grup społecznych spotykali się dwa razy w miesiącu i dyskutowali nad wybranymi zmianami prawa.

⁹⁴ Zob. www.cnrs.fr/fr/presse/jean-zay-le-supercalculateur-le-plus-puissant-de-france-pour-la-recherche [dostęp: 17 października 2024 r.].

Drugi etap (2021-2025) koncentruje się na rozpowszechnianiu AI w gospodarce oraz wparciu rozwoju i innowacji w niektórych wybranych obszarach kluczowych, tj.:

- *embedded AI*, czyli AI wbudowanej w „serca” urządzeń lub komponentów;
- **zaufanej AI, czyli niezawodnej, wysokowydajnej AI, która spełnia standardy przejrzystości i poufności;**
- **oszczędnej AI, ze wsparciem dla rozwoju i wdrażania niskoenergetycznej i efektywnej pod względem wykorzystania danych AI, aby służyć celom dekarbonizacji, transformacji ekologicznej i energetycznej;**
- generatywnej AI i dużych modeli językowych (LLM).

W ramach drugiego etapu uruchomiono np. Komitet ds. Generatywnej AI⁹⁵. Skupia on podmioty z różnych sektorów (kulturalnego, gospodarczego, technologicznego i badawczego) w celu wsparcia rządu w podejmowaniu decyzji i ustalaniu priorytetów w dalszym rozwoju AI.

Warto również wspomnieć o planie inwestycyjnym dla Francji, tzw. *France 2030*. Został on określony na lata 2021-2030 i składa się łącznie na kwotę 54 mld euro. Jednym z jego głównych celów jest zdobycie przez Francję pozycji pioniera innowacji, a więc jest on kompatybilny z oraz uwzględnia wydatki czynione w ramach krajowej strategii na rzecz AI. Ponadto Francja współpracuje również z innymi państwami unijnymi w zakresie rozwoju AI⁹⁶.

6.2. Działalność i rekomendacje francuskich instytucji

Kluczowymi instytucjami regulującymi rynek finansowy we Francji są Bank Centralny Francji (*Banque de France*, dalej jako: „**Bank Francji**”), francuski Urząd

Rynków Finansowych⁹⁷ (*Autorité des Marchés Financiers*, dalej jako: „**AMF**”), francuski Urząd Nadzoru Bankowego i Ubezpieczeniowego (*Autorité de contrôle prudentiel et de résolution*, dalej jako: „**ACPR**”).

Ze względu na realny wpływ tych instytucji na rynek finansowy, w tym sektor bankowy, czy to poprzez wykonywanie swoich obowiązków i kompetencji ustanowionych przez prawo, czy też poprzez podejmowanie własnych inicjatyw w postaci „miękkiego” oddziaływania na rynek, np. publikowanie raportów, organizacja konferencji, wydawanie rekomendacji, ich działalność w obszarze AI może posłużyć za źródło cennych wskazówek dla etycznego postępowania z AI w Polsce.

Poczynając od Banku Francji należy wskazać, że głównym obszarem działań podejmowanych przez tę instytucję w zakresie zastosowania AI są różnego rodzaju aktywności o charakterze informacyjnym, np. wystąpienia przedstawicieli banku⁹⁸, artykuły⁹⁹ i dokumenty robocze¹⁰⁰ przygotowywane przez ekspertów.

W swojej wypowiedzi z czerwca br. Pierwszy Zastępca Dyrektora Denis Beau wskazywał na następujące zagrożenia etyczne związane ze stosowaniem AI, w szczególności w wyniku złego zaprogramowania modeli wyceny¹⁰¹:

- **niewłaściwe, a nawet dyskryminacyjne traktowanie klientów,**
- **naruszenie prywatności klientów,**
- **ryzyko dezinformacji lub manipulacji klientów.**

W odpowiedzi na te zagrożenia, a także wiele innych, których dzisiaj nawet nie jesteśmy w stanie

95 Zob. www.economie.gouv.fr/comite-intelligence-artificielle-generative [dostęp: 17 października 2024 r.].

96 Zob. np. umowę zawartą przez Niemcy, Włochy i Francję wyrażającą chęć współpracy pomiędzy tymi państwami w zakresie regulacji AI na poziomie unijnym - www.bakermckenzie.com/en/insight/publications/2023/11/ai-regulation-agreement-germany-france-italy [dostęp: 17 października 2024 r.].

97 Stanowiący odpowiednik polskiej Komisji Nadzoru Finansowego.

98 Zob. np. *The financial supervisor in the age of AI*, www.banque-france.fr/en/governors-interventions/financial-supervisor-age-ai [dostęp: 17 października 2024 r.].

99 Zob. np. M. Laporte, H. Fraise, H. Buisson, *AI and banks' own funds: a new determination?*, www.banque-france.fr/en/publications-and-statistics/publications/ai-and-banks-own-funds-new-determination [dostęp: 17 października 2024 r.].

100 Zob. np. H. Fraise, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, www.banque-france.fr/en/publications-and-statistics/publications/return-investment-ai-case-capital-requirement [dostęp: 17 października 2024 r.].

101 Zob. *The financial supervisor in the age of AI*, www.banque-france.fr/en/governors-interventions/financial-supervisor-age-ai [dostęp: 17 października 2024 r.].

przewidzieć, Denis Beau podkreślił konieczność pracy nad dalszymi regulacjami. W jego ocenie, pomimo pionierskiego charakteru i określenia „wysokich ryzyk” również w sektorze finansowym m.in. w ocenie zdolności kredytowej osób fizycznych oraz ocenie i wycenie ubezpieczeń zdrowotnych i na życie, AI Act stanowi jedynie początek prawodawstwa w zakresie AI. Wskazał także, że w celu efektywnego zastosowania tych regulacji oraz ich dalszego rozwoju, **niezbędna jest współpraca pomiędzy organami regulacyjnymi w ramach jednego państwa, ale również (a może przede wszystkim) współpraca państw na poziomie unijnym, a docelowo również ogólnosiwiatowym.**

W samych jednak dokumentach publikowanych przez Bank Francji, w szczególności w kodeksie etyki do tej pory nie pojawiły się żadne postanowienia bezpośrednio dotyczące etycznego zastosowania AI czy to na użytek wewnętrzny banku, czy w relacjach zewnętrznych. Zastosowanie mają więc dotychczasowe reguły działania banków, które powinny być odpowiednio stosowane w kontekście użycia AI.

Jeśli chodzi natomiast o działalność AMF, to – podobnie jak w przypadku Banku Francji – jest ona skupiona na publikacjach materiałów informacyjnych¹⁰², mających na celu popularyzację wiedzy o AI¹⁰³, a także wskazywania obszarów, które mogą być zagrożone przez jej nieetyczne lub niewłaściwe zastosowanie.

Podobnie jak w przypadku analogicznej instytucji w Wielkiej Brytanii, AMF również prowadzi aktywną **walkę z podmiotami, które korzystając z popularności AI prowadzą działalność na rynku finansowym nie będąc przy tym autoryzowanymi**¹⁰⁴.

W zakresie działalności edukacyjnej należy wskazać jedną z nowszych inicjatyw AMF podjętą w ramach planu strategicznego „Impact 2027”. Jest nią seria warsztatów (w czerwcu br. odbył się pierwszy z nich) nt. innowacyjności w zakresie AI w usługach finansowych¹⁰⁵. Tego typu inicjatywy są warte śledzenia, ponieważ możliwe, że można przełożyć ich treści/cele również na polskie warunki. Jak już zostało niejednokrotnie podkreślone rozwój AI jest niezwykle dynamiczny, i można zakładać, że tendencja ta się utrzyma, a nawet zwiększy. Dlatego też, szczególnie w kontekście etycznego stosowania AI, niezbędne jest ciągle doszkalanie się i zrozumienie coraz to nowszych rozwiązań, które z pewnością będą bardziej zaawansowane technologicznie.

AMF brała również czynny udział przy regulacji AI na poziomie unijnym, m. in. poprzez uczestnictwo w konsultacjach społecznych odnośnie Białej Księgi dotyczącej AI¹⁰⁶. Ponadto, w swoim rocznym raporcie z 2023 r.¹⁰⁷ wskazano, że prawie 100 tys. dokumentów zostało przetworzonych przy użyciu AI, co umożliwiło znacznie odciążyć pracę urzędników w zadaniach, które są powtarzalne. AMF używa również AI w walce z nadużyciami na rynku finansowym – i jak zostało stwierdzone w raporcie – z obiecującymi wynikami.

Głównym przedsięwzięciem ACPR (podejmowanym wspólnie z Bankiem Francji) była publikacja w czerwcu 2020 r. dokumentu nadającego ramy debaty nad zarządzaniem AI w sektorze finansowym¹⁰⁸. Była to kontynuacja pracy ACPR nad stosowaniem AI, która rozpoczęła się w 2018 r. Dla oceny algorytmów AI i innych narzędzi w sektorze finansowym zidentyfikowano cztery współzależne kryteria:

- **właściwe zarządzanie danymi** – fundamentalna kwestia dla każdego algorytmu, ponieważ od tego zależy jego wydajność i zgodność z przepisami prawa. Należy przy tym uwzględnić także kwestie etyczne jak sprawiedliwość przetwarzania danych oraz brak dyskryminacji;

102 Zob. np. *Artificial intelligence: the AMF encourages market stakeholders to take part in two European Commission initiatives on issues specific to the financial sector*,

www.amf-france.org/en/news-publications/news/artificial-intelligence-amf-encourages-market-stakeholders-take-part-two-european-commission-artificial-intelligence-towards-new-contributions-for-regulators, www.amf-france.org/en/news-publications/news/artificial-intelligence-towards-new-contributions-regulators; Artificial intelligence and Big Data are now a reality for the AMF, www.amf-france.org/en/news-publications/news/artificial-intelligence-and-big-data-are-now-reality-amf [dostęp: 17 października 2024 r.].

103 Zob. np. *AI enthusiasts, the AMF challenges you*, www.amf-france.org/en/news-publications/news/ai-enthusiasts-amf-challenges-you [dostęp: 17 października 2024 r.].

104 Zob. np. www.amf-france.org/en/node/73913; www.amf-france.org/en/node/71938; www.amf-france.org/en/node/71127; www.amf-france.org/en/node/73922 [dostęp: 17 października 2024 r.].

105 Zob. *AMF launches its “innovation workshops”*, www.amf-france.org/en/news-publications/news/amf-launches-its-innovation-workshops [dostęp: 17 października 2024 r.].

106 Zob. <https://www.amf-france.org/sites/institutionnel/files/2020-07/reponse-amf-artificial-intelligence.pdf> [dostęp: 17 października 2024 r.].

107 *AMF Annual Report*, www.amf-france.org/sites/institutionnel/files/private/2024-08/2023_amf_annual_report.pdf [dostęp: 18 października 2024 r.].

108 Zob. L. Dupont, O. Fliche, Su Yang Fintech-Innovation Hub, ACPR, *Governance of Artificial Intelligence in Finance*, 2020, acpr.banque-france.fr/sites/default/files/medias/documents/20200612_ai_governance_finance.pdf [dostęp: 18 października 2024 r.].

MODERNISING ITS SUPERVISION TOOLS: EXPLOITING DATA, USING AI AND REFOCUSING SUPERVISION ON HIGH ADDED-VALUE ACTIVITIES

Exploiting data and deploying natural language processing

In 2023, the AMF continued to use various artificial intelligence (AI) techniques to support its teams in their core tasks. Nearly 100,000 documents were pre-processed using AI. This automated reading and analysis work has not only enabled our teams to save time and increase the efficiency of tedious and repetitive tasks, but also to make more intensive use of the collected data. The scope of application of these techniques, such as natural language processing and image recognition, aligned with the institution's priorities. In 2023, there was an emphasis on improving the detection and analysis of websites offering financial investments, as well as on developing suitable tools to facilitate the supervision of asset management market participants in relation to the new reporting required by the SFDR Regulation⁴. Additionally, as part of the development of the key information documents for funds (see page 93), AI has enabled the adaptation of tools to automatically extract information on fees and risks from the new documents.

The AMF is also capitalising on AI to improve its internal system for detecting market abuse, in particular using supervised and semi-supervised learning techniques (see box below). Some of this work is being carried out in partnership with the *École Normale Supérieure (ENS)* Cachan. Furthermore, 2023 saw the first generative AI tests, as part of experiments on the classification of the ESG issues addressed in issuers' press releases.

At this stage, the results of these tests are promising and will be the subject of future work.

Lastly, in 2023, the AMF played an active role in ESMA's strategic European discussions on data and innovation. These dialogues represent an opportunity to share best practice in this area and to pool as far as is possible the development costs of innovative projects that may be useful to all the AMF's counterparts.

Źródło: www.amf-france.org/sites/institutionnel/files/private/2024-08/2023_amf_annual_report.pdf.

- **wydajność** – może być ona określana na różne sposoby, przy czym powinny być one adekwatne do celu;
- **stabilność** – opisuje, jak stabilne i odporne jest działanie algorytmu w całym cyklu jego „życia”;
- **wyjaśnialność** – czyli przejrzystość i interpretowalność danych. Co ważne, dla prawidłowego zastosowania tego kryterium, konieczne jest umieszczenie danych w określonym kontekście. Owe wyjaśnienie może być konieczne dla użytkowników końcowych (zarówno klientów, jak i wewnętrznych użytkowników). Jego głównym celem jest poinformowanie klienta, zapewnienie spójności procesów decyzyjnych oraz ułatwienie walidacji i monitorowania modeli AI i ML.

W przygotowanym dokumencie wskazano, że kwestie etyczne leżą u podstaw coraz szerszego zastosowania AI w biznesie, co wpływa zarówno na jednostki, jak i na grupy, w szczególności w kontekście sprawiedliwości tych procesów. Powołano się w tym względzie na wytyczne unijne przygotowane

przez grupę ekspertów wysokiego szczebla ds. AI¹⁰⁹, w których wskazano zestaw siedmiu kluczowych wymagań, które powinny spełniać systemy AI, aby mogły zostać uznane za godne zaufania, tj.:

- **Przewodnia i nadzorcza rola człowieka** – systemy AI powinny wzmacniać pozycję ludzi, umożliwiając im podejmowanie świadomych decyzji i wspierając ich prawa podstawowe. Jednocześnie należy zapewnić odpowiednie mechanizmy nadzoru, które można osiągnąć poprzez podejście oparte na człowieku w pętli, człowiek-w-pętli i człowiek-dowództwo.
- **Pewność i bezpieczeństwo techniczne** – systemy AI muszą być odporne i bezpieczne. System można uznać za odporny i bezpieczny, jeżeli m.in. zapewniony jest plan awaryjny w przypadku problemów, a także precyzję, rzetelność i od-

109 Zob. Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji, www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_PL.pdf [dostęp: 18 października 2024 r.].

tworzalność. Jest to jedyny sposób na zapewnienie zminimalizowania i zapobieżenia również niezamierzonym szkodom.

- **Ochrona prywatności i danych** – oprócz zapewnienia pełnego poszanowania prywatności i ochrony danych należy również zapewnić odpowiednie mechanizmy zarządzania danymi, z uwzględnieniem jakości i integralności danych oraz przy zapewnieniu usprawiedliwionego dostępu do danych.
- **Przejrzystość** – modele biznesowe oparte na danych, systemach i AI powinny być przejrzyste. Jako pomoc w osiągnięciu tego celu można wskazać mechanizmy identyfikowalności, tzn. ustalone reguły, które umożliwią prawidłowy odczyt danych. Ponadto systemy AI i ich decyzje należy wyjaśnić w sposób dostosowany do zainteresowanych stron. Osoby mające skorzystać z danego systemu AI muszą być świadomi, że wchodzi w interakcję właśnie z systemem AI (a nie np. z człowiekiem), oraz muszą zostać poinformowani o możliwościach i ograniczeniach tego systemu.
- **Różnorodność, niedyskryminacja i sprawiedliwość** – należy unikać niesprawiedliwej tendencji, ponieważ może ona mieć wiele negatywnych skutków, począwszy od marginalizacji słabszych grup społecznych po nasilenie uprzedzeń i dyskryminacji. Wspierając różnorodność, systemy sztucznej inteligencji powinny być dostępne dla wszystkich, niezależnie od wszelkich niepełnosprawności, i angażować odpowiednie zainteresowane strony w całym cyklu życiowym.
- **Dobrostan społeczny i środowiskowy** – systemy AI powinny przynosić korzyści wszystkim ludziom, w tym przyszłym pokoleniom. Należy zatem zadbać o to, by były one zrównoważone i przyjazne dla środowiska, czyli uwzględniać

konkretne uwarunkowania społeczne, kulturowe i środowiskowe.

- **Rozliczalność** – należy wprowadzić mechanizmy zapewniające odpowiedzialność konkretnych podmiotów (np. spółek) za systemy AI oraz ich wyniki. Możliwość kontroli, która umożliwia ocenę algorytmów, danych i procesów projektowania, odgrywa kluczową rolę w tym procesie, zwłaszcza w zastosowaniach o krytycznym znaczeniu. Ponadto należy zapewnić odpowiednie dostępne środki odwoławcze.

ACPR wskazało, że uprzedzenia, zwłaszcza te o charakterze dyskryminacyjnym, są przedmiotem ciągłych badań, w ramach których wyróżniono następujące etapy walki z nimi:

- 1) staranne określenie, co stanowi problematyczną stronniczość,
- 2) ustalenie, w jakim stopniu algorytmy AI odzwierciedlają lub wzmacniają uprzedzenia obecne w danych,
- 3) łagodzenie tych uprzedzeń, jeśli jest to możliwe na poziomie danych lub algorytmu.

6.3. Podsumowanie

Francja jest jednym z państw przodujących, jeśli chodzi o badania i rozwój AI w UE. Świadczy o tym nie tylko to, że angażuje w to duże zasoby kapitału zarówno finansowego, jak i ludzkiego, ale także to, że jest aktywnym uczestnikiem zmian prawnych. **Aktualnie żadne instytucje państwowe nie przewidują jednak wytycznych czy rekomendacji, które kompleksowo opisywałyby kwestie etyczne w zastosowaniu AI w sektorze bankowym. Na ten moment przy stosowaniu AI powinny być więc brane pod uwagę obowiązujące akty prawne, takie jak AI Act czy RODO.**

7. Hiszpania





7.1. Uwagi wprowadzające

W Hiszpanii (dalej jako: „Hiszpania” lub „Królestwo Hiszpanii”) sektor bankowy **traktowany jest jako wyjątkowo zaawansowana technologicznie sfera życia społeczno-gospodarczego, a prawo bankowe jako jedna z najbardziej rozległych gałęzi prawa.**

Działalność banków regulują przede wszystkim trzy kluczowe akty prawne, tj.: ustawa o regulacji bankowości z 2014 r. (*Ley de 26 de junio, de ordenación, supervisión y solvencia de entidades de crédito*)¹¹⁰, ustawa o marketingu usług finansowych¹¹¹ (*Ley de 25 de febrero, de protección de los consumidores y usuarios frente a situaciones de vulnerabilidad social y económica*) oraz dekret królewski z mocą ustawy nr 16/2011 z dnia 14 października 2011 r. ustanawiający Fundusz Gwarancji Depozytów dla Instytucji Kredytowych (*Real Decreto-ley de 14 de octubre, por el que se crea el Fondo de Garantía de Depósitos de Entidades de Crédito*)¹¹².

Podkreślić należy, że od dnia 4 listopada 2014 r. nadzór nad hiszpańskimi instytucjami kredytowymi i pozostałymi instytucjami kredytowymi strefy euro sprawuje **Jednolity Mechanizm Nadzorczy – system kierowany przez Europejski Bank Centralny**, w którym uczestniczą krajowe organy nadzoru, w tym również podmioty hiszpańskie.

Jednym z tych organów jest **Bank Hiszpanii**¹¹³ (*Banco de España*). Jest to podmiot, który reguluje działalność bankową poprzez system nadzoru ostrożnościowego, który obejmuje monitorowanie bilansów i ocenę ryzyka każdego podmiotu (w tym nadzór nad wypłacalnością, wynikami i zgodnością ze szczegółowymi regulacjami instytucji kredytowych). Ponadto Bank Hiszpanii wydaje zalecenia

i wytyczne dla banków w celu zapewnienia odpowiedniego zarządzania ryzykiem.

W kontekście innych instytucji prawa bankowe należy również wyróżnić:

- **Fundusz Gwarantowania Depozytów (IOS)** – jest instytucją utworzoną w 2003 r. w celu ochrony depozytów klientów banków;
- **Komitet Nadzoru Bankowego (CSB)** – ma na celu wspieranie prawidłowego funkcjonowania instytucji bankowych;
- **Instytut Rachunkowości i Audytu Rachunków (ICAC)** – jest instytucją odpowiedzialną za nadzór nad rachunkowością podmiotów bankowych.

7.2. Hiszpańska Karta Praw Cyfrowych w kontekście AI

Istotnych regulacji w kontekście stosowania sztucznej inteligencji, dostarcza w pierwszej kolejności Hiszpańska Karta Praw Cyfrowych¹¹⁴, zatwierdzona przez Rząd Hiszpanii w dniu 14 lipca 2021 r. (*Carta de Derechos Digitales*). We fragmencie nr XXV zatytułowanym „Prawa przed sztuczną inteligencją” (*Derechos ante la inteligencia artificial*) odniesiono się do praw użytkowników i zasad funkcjonowania sztucznej inteligencji oraz jej stosowania. Zgodnie z jego treścią:

Sztuczna inteligencja zapewnia podejście skoncentrowane na osobie i jej niezbywalnej godności, dąży do wspólnego dobra i zapewnia zgodność z zasadą non-maleficence (tzw. zasada nieszkodzenia) [wyróżnienie i tłumaczenie własne].

Co więcej przepis ten zapewnia osoby korzystające ze sztucznej inteligencji:

- prawo do niedyskryminacji, niezależnie od jej pochodzenia, przyczyny lub charakteru, w odniesieniu do decyzji, wykorzystania danych i procesów opartych na sztucznej inteligencji;
- ustanowienie warunków przejrzystości, audytowalności, wyjaśnialności, identyfikowalności,

¹¹⁰ Ley 10/2014, de 26 de junio, de ordenación, supervisión y solvencia de entidades de crédito. Disposición 6726 del BOE núm. 156 de 2014 [dostęp: 21 września 2024 r.].

¹¹¹ Ley 4/2022, de 25 de febrero, de protección de los consumidores y usuarios frente a situaciones de vulnerabilidad social y económica. Ley 4/2022, de 25 de febrero, de protección de los consumidores y usuarios frente a situaciones de vulnerabilidad social y económica. (boe.es) [dostęp: 21 września 2024 r.].

¹¹² real Decreto-ley 16/2011, de 14 de octubre, por el que se crea el Fondo de Garantía de Depósitos de Entidades de Crédito. Real Decreto-ley 16/2011, de 14 de octubre, por el que se crea el Fondo de Garantía de Depósitos de Entidades de Crédito. (boe.es) [dostęp: 21 września 2024 r.].

¹¹³ <https://www.bde.es/wbe/es/areas-actuacion/supervision-entidades-financieras/> [dostęp: 15 października 2024 r.].

¹¹⁴ La Carta de Derechos Digitales https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carda_Derechos_Digitales_RedEs.pdf [dostęp: 10 października 2024 r.].

nadzoru ludzkiego i zarządzania. **W każdym przypadku przekazywane informacje muszą być dostępne i zrozumiałe dla jednostki;**

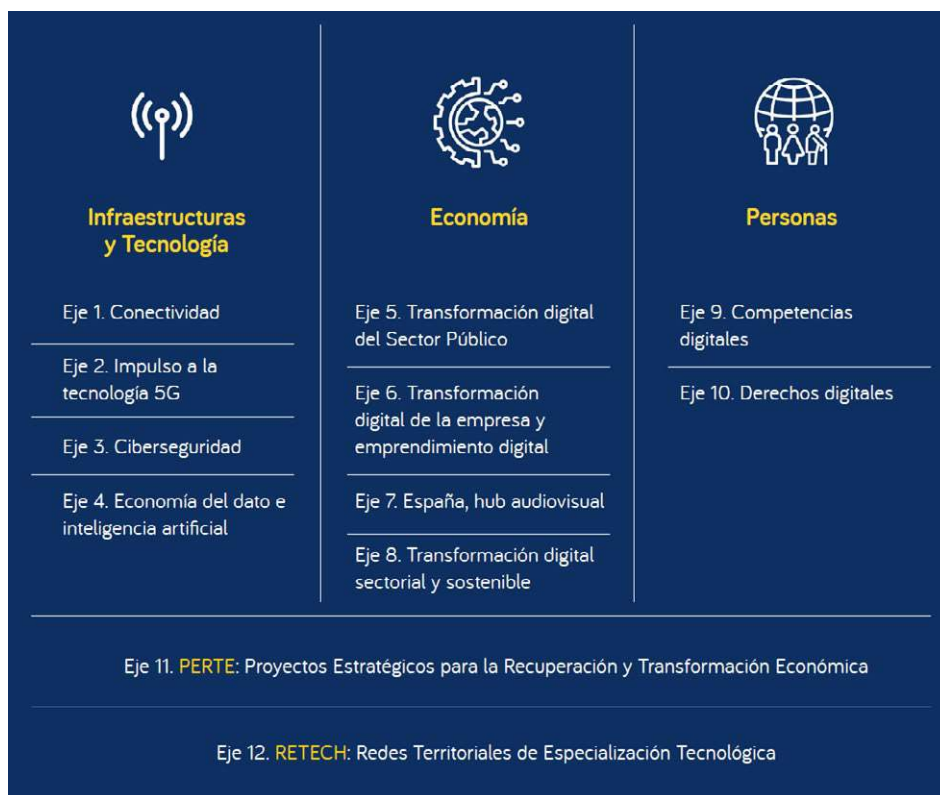
- dostępność, użyteczność i niezawodność.

Ponadto regulacja ta zapewnia również mechanizm kontroli nad sztuczną inteligencją. Przyznaje ona bowiem osobom fizycznym prawo do żądania nadzoru, podjęcia interwencji oraz kwestionowania zautomatyzowanych decyzji podejmowanych przez AI, **w sytuacji, w której mają one wpływ na ich sferę majątkową, osobistą lub na obie jednocześnie.**

7.3. Hiszpańska Agencja Nadzoru nad Sztuczną Inteligencją

W zakresie sztucznej inteligencji kluczową rolę w Hiszpanii odgrywać będzie również AESIA – Hiszpańska Agencja Nadzoru nad Sztuczną Inteligencją (*Agencia Española de Supervisión de la Inteligencia Artificial*), czyli pierwszy taki powołany w Europie organ do nadzoru nad AI¹¹⁵.

Institucja ta będzie działać w ramach ministerstwa, które odpowiedzialne będzie m.in. za transformację cyfrową w Hiszpanii (*Ministerio de Asuntos Económicos y Transformación Digital*). W dniu 15 maja 2024 r. hiszpański rząd zatwierdził strategię w zakresie sztucznej inteligencji, która zapewnia ciągłość i wzmacnia strategię w zakresie sztucznej inteligencji (ENIA) opublikowaną w 2020 r., zgodnie z planem odbudowy, transformacji i zwiększania odporności (RTRP), dostosowując ją do nowych osiągnięć technologicznych. Promowanie sztucznej inteligencji jest również rozważane w hiszpańskiej agendzie cyfrowej pt. „Cyfrowa Hiszpania 2026” (*España Digital 2026*) (strona agendy: <https://espanadigital.gob.es/>) jako kluczowy element przekrojowy służący przekształceniu modelu produkcji i pobudzeniu wzrostu hiszpańskiej gospodarki.



Źródło: strategia „Cyfrowa Hiszpania 2026”¹¹⁶.

¹¹⁵ Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial. Disposición 18911 del BOE núm. 210 de 2023 [dostęp: 21 września 2024 r.].

¹¹⁶ España Digital 2026 Informe de ejecución https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2023/Informe_ejecucion_V7.pdf [dostęp: 16 października 2024 r.].

DIGITALIZACIÓN EN ESPAÑA



Źródło: strategia „Cyfrowa Hiszpania 2026”¹¹⁷.

7.4. Raport Hiszpańskiego Stowarzyszenia Bankowego

Jednym z czołowych sektorów gospodarki krajowej Królestwa Hiszpanii, który w dużym stopniu wykorzystał dotychczasowe możliwości sztucznej inteligencji jest **sektor bankowy** (*El sector bancario*).

Jak wynika z raportu przeprowadzonego przez Hiszpańskie Stowarzyszenie Bankowe (*Asociación Española de Banca*) pt. „Wykorzystanie sztucznej inteligencji w sektorze bankowym” (*El uso de la Inteligencia Artificial en el sector bancario*)¹¹⁸:

„Technologie AI między innymi poprawiają komfort użytkowania, zwiększają wydajność procesów wewnętrznych i zwiększają bezpieczeństwo”;

„Przyszłym wyzwaniem dla sztucznej inteligencji jest poprawa zdolności rozumienia i reagowania na środowisko na głębszym poziomie.

Ostatecznym celem tych technologii jest **zwiększenie zdolności technologicznych i przemysłowych gospodarek, służąc przy tym ludziom i poprawiając ich dobrobyt** [wyróżnienie i tłumaczenie własne].

Co więcej raport dzieli zastosowanie AI w sektorze bankowym na dwie zasadnicze kategorie. Pierwsza z nich dotyczy **ulepszania i usprawnienia procesów związanych z obsługą klienta banku**. Podkreśla się, że sztuczna inteligencja w ciągu najbliższych lat umożliwi nie tylko przetwarzanie języków obcych na język macierzysty, ale i prowadzenie zaawansowanych rozmów głosowych lub czatów z klientami odpowiadając przy tym na bardzo złożone pytania. Przełoży się to z kolei na automatyzację niektórych procesów w działach obsługi klienta, zapewniając klientom większą wygodę i sprawność podczas interakcji z bankiem. Drugą kategorię stanowią **ulepszenia procesów wewnętrznych i bezpieczeństwa**. Raport wskazuje przede wszystkim na wykorzystanie AI poprzez robodoradców implementowanych do stron internetowych oraz aplikacji mobilnych poszczególnych banków. Największą obiecującą prognozą w ocenie raportu jest zastosowanie **AI modelach oceny ryzyka**, co z kolei ma usprawnić proces profilowania kredytowego i generować dokładniejsze informacje o ryzyku związanym z transakcją.

¹¹⁷ España Digital 2026, <https://www.indesia.org/espana-digital-2026/> [dostęp: 16 października 2024 r.].

¹¹⁸ El uso de la Inteligencia Artificial en el sector bancario, „Asociación Española de Banca” *la-inteligencia-artificial-en-el-sector-bancario.pdf* (aebanca.es) [dostęp: 21 sierpnia 2024 r.].

7.5. Kontekst stosowania AI w bankowości w hiszpańskiej literaturze

Również w literaturze podkreśla się, że AI odgrywa istotną rolę w tamtejszym systemie bankowości. Jak trafnie zauważa Teresa Rodríguez de las Heras Ballell (profesor prawa handlowego na Uniwersytecie Karola III w Madrycie):

„z różną intensywnością, systemy AI są wykorzystywane w sektorze bankowym **wzdłuż całego łańcucha wartości front-line-mid-office-back-office**. W zakresie obsługi klienta i interakcji, systemy AI oferują niezwykle możliwości personalizacji, rekomendacji i profilowania, zarządzania kontem, doradztwa inwestycyjnego (trading) i finansowego (robo-doradcy), ciągłej obsługi za pośrednictwem chatbotów i wirtualnych asystentów, a także rozwiązań Know Your Customer (KYC)”¹¹⁹;

„W wewnętrznym zarządzaniu operacjami rozwiązania AI są stosowane w **automatyzacji procesów korporacyjnych, administracyjnych i transakcyjnych, w optymalizacji różnych działań lub zarządzaniu zgodnością**. W przypadku zarządzania ryzykiem, przewiduje się, że rozwiązania AI **usprawnią systemy zapobiegania oszustwom, wczesnego ostrzegania i cyberbezpieczeństwa, a także zostaną włączone do predykcyjnych modeli rekrutacji i promocji**”¹²⁰ [wyróżnienie i tłumaczenie własne].

Z kolei w ocenie Andrésa Alonso-Robisco (starszego ekonomisty w Wydziale Innowacji Finansowych Banku Hiszpanii) oraz Jose Manuel Carbo Martinez (ekonomisty w Wydziale Innowacji Finansowych Banku Hiszpanii), AI wpłynie istotnie na obliczanie i określanie kapitału regulacyjnego banków:

„Innym interesującym zastosowaniem zaawansowanych modeli analitycznych z uczeniem maszynowym jest **obliczanie i określanie kapitału regulacyjnego**”. Szacuje się znaczne oszczędności, jeśli modele te zostaną wykorzystane do obliczenia aktywów ważonych ryzykiem”¹²¹ [wyróżnienie i tłumaczenie własne].

Istotnych informacji w kontekście wykorzystania AI w sektorze bankowym dostarcza również artykuł naukowy pt. „Artificial intelligence in financial services” autorstwa Any Fernandez zamieszczony na stronie internetowej *Banco de España* zawierający raport w zakresie wykorzystania AI w sektorze finansowym, w tym w sektorze bankowym¹²²:

„Wykorzystanie technik sztucznej inteligencji w świadczeniu usług finansowych może **zwiększyć wydajność, obniżyć koszty, poprawić jakość, podnieść poziom zadowolenia klientów i zwiększyć integrację finansową, głównie dzięki możliwościom automatyzacji procesów operacyjnych i zwiększenia zdolności analitycznych**” [wyróżnienie i tłumaczenie własne].

W ocenie cytowanej autorki zastosowania AI w sektorze Bankowym przejawiają się przede wszystkim w takich aspektach jak:

- **chatboty oraz wirtualni asystenci**, którzy umożliwiają użytkownikom rozstrzyganie typowych wątpliwości, w niektórych przypadkach, zabezpieczają lub wykonują określone transakcje (np. zlecenie przelewów, otwieranie rachunków);
- **personalizacja produktów i usług końcowych**. W ocenie autorki instytucje finansowe są w stanie zaoferować spersonalizowane doświadczenie użytkownika, być może nawet wykraczające poza usługi bankowe w najściślejszym tego słowa znaczeniu (np. wysyłanie alertów o stanie konta etc.);
- **kontrola przeciwdziałania praniu pieniędzy (AML) i zapobiegania oszustwom**. Autorka powołuje się w tym zakresie na pogląd Barta Vana Liebergena (kierownika ds. inwestycji w holenderskim bank rozwoju przedsiębiorczości FMO), zgodnie z którym: „Zdolność do analizowania większej ilości danych i łączenia tych danych z nowymi źródłami informacji **umożliwia wykrywanie anomalii lub wzorców, które w przeciwnym razie pozostałyby niezauważone**”¹²³ [wyróżnienie własne].

119 T.R. de las Heras Ballell, *Inteligencia Artificial en el sector bancario: reflexiones sobre su régimen jurídico en la Unión Europea*. ICE, „Revista de Economía”, 926/2022, s. 99.

120 Ibidem.

121 A. Alonso, J.M. Carbó, *Understanding the performance of Machine Learning Models to predict credit default: A novel approach for supervisory evaluation*, „Documentos Ocasionales del Banco de España”, 2105/2021, s. 23-24.

122 A. Fernández, *Artificial intelligence in financial services*, „Banco de España” Article, 3, 2019, s. 19.

123 B. Van Liebergen, *Machine Learning: A Revolution in Risk Management and Compliance?*, „The CAPCO Institute Journal”, No 45, 2017, s. 66.

- **scoring kredytowy.** Zwiększone możliwości analityczne mogą poprawić ocenę kredytową i przyspieszyć udzielanie pożyczek;
- **zgodność z przepisami (tzw. *regulatory compliance*).** Większe możliwości analityczne zapewniane przez narzędzia sztucznej inteligencji ułatwiają przestrzeganie niektórych wymogów regulacyjnych (np. zarządzanie ryzykiem, obowiązki sprawozdawcze), a także monitorowanie zmian regulacyjnych.

Zauważyć należy, że wskazane przez A. Fernández zastosowania znajdują również swoje odzwierciedlenie w aktualnym sektorze bankowym. Tytułem przykładu w Banku Hiszpanii **opracowano narzędzie sztucznej inteligencji przeznaczone do sortowania banknotów na banknoty nadające się i nienadające się do obiegu**¹²⁴.

7.6. Rządowa Strategia na rzecz sztucznej inteligencji 2024

W kontekście etyki AI w zakresie bankowości istotnych wytycznych dostarcza nam również zaprezentowana przez Rząd Królestwa Hiszpanii, Strategia na rzecz sztucznej inteligencji 2024 (*La Estrategia de Inteligencia Artificial 2024*)¹²⁵. Pomimo, że nie odnosi się ona bezpośrednio i wyłącznie do bankowości to zawiera szereg rozwiązań ogólnych, istotnych również dla tego sektora.

Strategia na rzecz sztucznej inteligencji 2024 opiera się na trzech głównych osiach, do których należą kolejno:

- wzmocnienie wdrażania sztucznej inteligencji w całej gospodarce,
- **ułatwianie stosowania sztucznej inteligencji w sektorze publicznym i prywatnym oraz**
- **promowanie przejrzystej, odpowiedzialnej i humanistycznej sztucznej inteligencji.**

W odniesieniu do sektora bankowego, istotną rolę odgrywa oś nr 2. **Koncentruje się ona na promo-**

waniu i wdrażania sztucznej inteligencji w sektorze ogólnej administracji państwowej (*Administración General del Estado*). W ramach niej m.in. bank centralny odpowiedzialny będzie za opracowanie projektów pilotażowych (eksperymentalnych) dotyczących zastosowania sztucznej inteligencji dla państwowych jednostek z sektora publicznego. Ponadto w ramach tej osi, opracowany zostanie wspólny model zarządzania danymi i korpusem ogólnej administracji państwowej Królestwa Hiszpanii, którego celem jest zagwarantowanie standardów bezpieczeństwa – anonimowości, jakości, interoperacyjności i ponownego wykorzystania danych.

Z drugiej strony oś zakłada również wsparcie sektora prywatnego (w tym również banków). Wsparcie to zakłada w szczególności poszerzenie dostępu do usług doradczych w zakresie wdrażania sztucznej inteligencji, dla mniejszych projektów biznesowych.

Co więcej oś tą uzupełni przyszły akt prawny o cyberbezpieczeństwie, który ustanowi jasne i kompleksowe ramy rozwoju krajowego cyberbezpieczeństwa i poprawy ochrony systemów informacyjnych, sieciowych i danych, **a także promowanie ekosystemu biznesowego cyberbezpieczeństwa AI, który będzie promował innowacje, współpracę i przyjęcie technologii sztucznej inteligencji w dziedzinie cyberbezpieczeństwa**¹²⁶.

Niebagatelne znaczenie dla sektora bankowego, może mieć również oś nr 3, której realizacja będzie ściśle powiązana ze wspomnianą wcześniej Hiszpańską Agencją Nadzoru nad Sztuczną Inteligencją (AESIA). Realizacja celów tej osi będzie odbywać się za pośrednictwem tejże agencji, której działanie będzie trzytorowe tj. jako:

- **centrum analiz w zakresie sztucznej inteligencji** – w tym zakresie do jej zadań będzie należało: analizowanie trendów, generowanie debaty społecznej oraz identyfikowanie dobrych praktyk i pojawiających się zagrożeń;
- **organ nadzorczy odpowiedzialny za odpowiedzialne wdrażanie sztucznej inteligencji**

124 J. Fernández-Villaverde, S. Hurtado, G. Nuño, *Financial Frictions and the Wealth Distribution*, „Econometrica” 2019, s. 869-900.

125 La Estrategia de Inteligencia Artificial 2024, https://portal.mineco.gob.es/es-es/digitalizacionIA/Documents/Estrategia_IA_2024.pdf [dostęp: 10 października 2024 r.].

126 El Gobierno aprueba la Estrategia de Inteligencia Artificial 2024, s. 3 https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2024/20240514_NdP_EstrategiaIA_2024.pdf [dostęp: 10 października 2024 r.].

– w tym zakresie do jej zadań będzie należało: wdrażanie sztucznej inteligencji: certyfikacja systemów sztucznej inteligencji zgodnie z europejskim rozporządzeniem w sprawie sztucznej inteligencji oraz ustanawianie dobrych praktyk w celu promowania przejrzystych i otwartych modeli;

- **międzynarodowy punkt odniesienia:** uczestnictwo w europejskich i światowych instytucjach zarządzających sztuczną inteligencją.

Przedstawiona powyżej strategia (w tym podejmowane w jej ramach inicjatywy) będą koordynowane nie tylko przez Sekretarza Stanu ds. Cyfryzacji i Sztucznej Inteligencji, ale **ze względu na jej zakres i wpływ wezmą w niej udział wszystkie ministerstwa**. Potwierdza to tylko, fakt istotnej roli sztucznej inteligencji jaką odgrywa w kontekście gospodarki krajowej Królestwa Hiszpanii.

7.7. Współpraca Hiszpanii z KE (pierwsze kontrolowane środowisko testowe dot. zastosowania sztucznej inteligencji wysokiego ryzyka)

Uwydatnić należy również, że rząd Hiszpanii, we współpracy z Komisją Europejską, uruchamia pierwsze kontrolowane środowisko testowe w ramach europejskiego rozporządzenia w sprawie sztucznej inteligencji (AI Act). Inicjatywa ta jest częścią hiszpańskiej agendy cyfrowej pt. „Cyfrowa Hiszpania 2026”, a jej celem jest sprawdzenie sposobu wdrożenia wymogów, o których mowa w AI Act w zakresie systemów sztucznej inteligencji wysokiego ryzyka.

Co więcej, w ramach omawianej inicjatywy mają powstać przewodniki (podręczniki) dobrych praktyk, które mając na celu przygotowanie i poszerzenie świadomości wśród przedsiębiorstw, zwłaszcza w sektorze małych i średnich przedsiębiorstw i *start-upów*, aby ułatwić im wdrożenie AI Act. Mają one przede wszystkim: zapewnić jasność co do wymogów, które zostały określone w AI Act, ułatwić transfer *know-how* w zakresie wdrażania przepisów dotyczących sztucznej inteligencji, zachęcić do innowacji i promowania rozwoju AI oraz umożliwić im testowanie obowiązków i wymagań w kontrolowanym środowisku, a następnie opracowania wytycznych technicznych i gromadzenie niezbędnych informacji. Niewykluczone, że **w ramach podręczników po-**

ruszone zostaną konkretne kwestie etyczne sztucznej inteligencji¹²⁷.

Jak wskazuje się na stronie Rządu Hiszpanii *Ministerio de Economía, Comercio y Empresa*:

*„projekt pilotażowy prowadzony przez Hiszpanię i wspierany przez Komisję Europejską **zaowocuje wytycznymi i przewodnikami dotyczącymi dobrych praktyk, które przygotują i podniosą świadomość wśród przedsiębiorstw, zwłaszcza MŚP i start-upów, w celu ułatwienia ich wdrożenia**”¹²⁸ [wyróżnienie i tłumaczenie własne].*

7.8. Działania hiszpańskiego banku centralnego *Banco de España*

Równie istotnym podmiotem działającym na rzecz rozwoju sztucznej inteligencji w hiszpańskim sektorze bankowym, jest Centralny Bank Królestwa Hiszpanii (*Banco de España*), który zdecydowanie angażuje się w tym zakresie, ujmując omawiane kwestie sztucznej inteligencji w swoim Planie Strategicznym na rok 2024 (z *hiszp. Plan Estratégico 2024*). Uwzględniono w nim m.in. plan promowania innowacji technologicznych jako centralny element modernizacji bankowości:

*„Jednym z głównych zidentyfikowanych obecnie trendów jest zmiana technologiczna, która wyróżnia się różnorodnością na tle innych technologii, które ewoluują i łączą się ze sobą, aby rozszerzyć i zwiększyć swój potencjał. Nie wyczerpując tematu, możemy podkreślić istnienie zestawu nowych technologii z potencjałem do przekształcenia środowiska: masowe przetwarzanie danych (*Big Data*), **nowe narzędzia do analizy danych (*Data Analytics*)**, **sztuczna inteligencja, przechowywanie i przetwarzanie danych w chmurze (*Cloud*)**, **nowe narzędzia do rejestrowania transakcji (*Blockchain*) i robotyzacja. Wszystkie te nowe technologie służą jako **dźwignia cyfrowej transformacji****”¹²⁹ [wyróżnienie i tłumaczenie własne].*

¹²⁷ España, pionera en el impulso de la regulación de Inteligencia Artificial, pone en marcha el entorno controlado de pruebas del Reglamento Europeo de Inteligencia Artificial, <https://portal.mineco.gob.es/es-es/comunicacion/Paginas/entorno-controlado-de-pruebas-sandbox-del-Reglamento-Europeo-de-Inteligencia-Artificial.aspx> [dostęp: 21 sierpnia 2024 r.].

¹²⁸ *Ibidem*.

¹²⁹ PlanEstratégico2024, <https://www.bde.es/f/webbde/SSICOM/20200115/planestrategico.pdf> [dostęp: 9 października 2024 r.]

7.9. Podsumowanie

Królestwo Hiszpanii jest jednym z pionierów w Europie, jeśli chodzi o rozwój i badania poświęcone w kierunku sztucznej inteligencji (w tym również w zakresie AI w sektorze bankowym). Potwierdzają to m.in. raporty *Banco de España*, Hiszpańskiego Stowarzyszenia Bankowego, Rządowa Strategia na rzecz sztucznej inteligencji 2024 czy Plan Strategiczny Banku Centralnego na rok 2024.

Co więcej Hiszpania, jako pierwsza w Europie, powołała organ do nadzoru AI (EASIA). Tym samym, kraj ten nie tylko angażuje w rozwój AI duże ilości kapitału zarówno finansowego, jak i ludzkiego, ale także jest aktywnym uczestnikiem zmian prawnych.

Nie ulega również wątpliwości, że wprowadzona w 2021 r. Karta Praw Cyfrowych potwierdziła tylko jak istotnym elementem rozwoju gospodarki krajowej tego Państwa jest transformacja cyfrowa (w tym wszelkie działania mające na celu promowanie oraz rozpowszechnianie AI we wszystkich sferach gospodarki, w tym również w sektorze bankowym).

Aktualnie jednak żadne instytucje państwowe **nie przewidują wytycznych czy rekomendacji (lub in-**

nych aktów prawnych typu *soft law* dedykowanych właśnie stosowaniu AI w sektorze bankowym/finansowym), które kompleksowo normowałyby kwestie etyczne w zastosowaniu AI w sektorze bankowym.

Na ten moment więc przy stosowaniu AI powinny być brane pod uwagę obowiązujące akty prawne, takie jak AI Act czy RODO, wspierane wewnętrznym (krajowym) prawem oraz ogólnym rekomendacjami dotyczącymi stosowania AI np. hiszpańską Kartą Praw Cyfrowych czy Rządową Strategią na rzecz sztucznej inteligencji.

Nie można jednak wykluczyć, że w niedalekiej przyszłości takie wytyczne się pojawią. Bankowość jest jednym z kluczowych systemów gospodarki krajowej Królestwa Hiszpanii, w którym również AI zaczyna odgrywać istotną rolę. Podejrzewać można, że prędzej czy później stopień rozwoju sztucznej inteligencji, który z dnia na dzień znacząco wzrasta, spowoduje, że w pewnym momencie, stopień rozwoju AI będzie na tak wysokim poziomie, że konieczne będzie uregulowanie kwestii etycznych.

Na chwilę obecną Królestwo Hiszpanii, nie skupia się jednak na kwestiach etycznych, a na rozwoju AI i kolportowaniu go do wszystkich gałęzi gospodarki krajowej (w tym również podmiotów publicznych i prywatnych z sektora bankowości).

8. Dania



Zanim przejdziemy do poszczególnych rozwiązań wprowadzonych w zakresie przestrzegania etyki AI w Danii, należy wspomnieć, że jest to kraj niezwykle w tym zakresie rozwinięty.

W celu kontroli prawidłowego przestrzegania etyki danych powołano Radę ds. Etyki danych (dalej jako: „Rada”). Obszar jej działalności wydzielono z ogólnej Rady Etyki funkcjonującej w Danii i składa się ona z 15 członków:

- Przewodniczącego, powoływanego przez Ministra Sprawiedliwości w porozumieniu z Ministrem Przemysłu i Ministrem Finansów;
- 3 członków mianowanych przez Ministra Sprawiedliwości;
- 3 członków mianowanych przez Ministra Przedsiębiorczości i Przemysłu;
- 3 członków mianowanych przez Ministra Finansów;
- 3 członków rekomendowanych przez Komisję Prawną i mianowanych przez Ministra Sprawiedliwości po konsultacji z Ministrem Handlu i Przemysłu oraz Ministrem Finansów;
- 2 członków mianowanych przez Ministra Finansów, Ministra Sprawiedliwości oraz Ministra Handlu i Przemysłu.

Podczas powoływania członków Rady szczególny nacisk kładzie się na znajomość problematyki nabytej w drodze doświadczenia praktycznego, w tym „*wiedzę z zakresu praktycznych i technicznych aspektów związanych z wykorzystywaniem technologii w administracji publicznej lub biznesie*” oraz interdyscyplinarny skład, co gwarantować ma wymóg „*doświadczenia badawczego i zainteresowania prawem, filozofią, etyką i technologią*”¹³⁰. Kadencja członków Rady jest czteroletnia z możliwością powołania na kolejną kadencję. Jest ona niezależna od rządu, jednakże musi działać we współpracy z odpowiednimi ministrami. W razie konieczności Rada może powoływać odrębne grupy monitorujące lub robocze.

Do pięciu głównych obszarów działalności Rady zalicza się:

- **pełnienie funkcji konsultacyjno-doradczej**, w tym doradzanie duńskiemu parlamentowi, ministrom oraz innym organom władzy publicznej w kwestiach etyki danych związanych z wykorzystywaniem danych i nowymi technologiami, w tym kierując się z jednej strony korzyściami, jakie może przynieść wykorzystanie danych i nowych technologii, jak i mając na uwadze towarzyszące temu ryzyka i zagrożenia. Rada ma obowiązek monitorować, informować, a także sygnalizować potencjalne problemy lub konieczność podjęcia działań przez dany organ w ramach jego kompetencji;
- **pełnienie funkcji opiniodawczej**;
- **funkcję monitorowania**, tj. permanentnego bieżącego obserwowania zmian społecznych w zakresie wykorzystywania nowych technologii oraz wydawania w związku z tym: zaleceń, sprawozdań i oświadczeń dotyczących ogólnych oraz podstawowych kwestii związanych z etyką danych;
- **funkcję informacyjno-popularyzatorską**, tj. prowadzenie działań informacyjnych oraz inicjujących debaty na temat kwestii etyki danych i wyzwań stojących przed społeczeństwem. Rada zapewnia, aby opinia publiczna była na bieżąco informowana o tym fakcie;
- **sprawozdawanie**, tj. **przedkładanie rządowi rocznego sprawozdania z aktualnej sytuacji, realizacji zaleceń, podejmowanych lub planowanych działań**. Rada wspiera kulturę w przedsiębiorstwach i sektorze publicznym, która opiera się na świadomych decyzjach dotyczących zrównoważonego wykorzystywania danych;
- **zapewnienie wsparcia odpowiednim organom oraz inicjowanie dialogu społecznego z interesariuszami, w tym podmiotami prywatnymi** – na przykład w odniesieniu do zasad etycznych, które stanowią inicjatywę w strategii wykorzystania AI, oraz w odniesieniu do inicjatyw w zakresie etyki danych, za które odpowiada Urząd ds. Przedsiębiorczości, takich jak dynamiczny zastaw narzędzi na rzecz etyki danych¹³¹.

¹³⁰ <https://dataetiskraad.dk/om-dataetisk-raad/baggrund-og-formaal> [dostęp: 18 października 2024 r.].

¹³¹ <https://dataetiskraad.dk/om-dataetisk-raad/baggrund-og-formaal> [dostęp: 18 października 2024 r.].

Rada zobowiązana jest również do **śledzenia trendów międzynarodowych oraz wsparcia instytucji duńskich w dostosowaniu się do nich** w zakresie wykonywaniem swoich zadań.

„Celem duńskiej Rady ds. Etyki Danych jest stworzenie forum, na którym będzie można dyskutować o kwestiach etycznych dotyczących związku między korzyściami wynikającymi ze stosowania nowych technologii z jednej strony, a uwzględnieniem praw podstawowych obywateli, pewności prawa i podstawowych wartości społecznych z drugiej strony, a także na którym można zakotwiczyć trwałe wysiłki na rzecz wspierania odpowiedzialnego korzystania z danych. Należy wziąć pod uwagę i szanować społeczną potrzebę korzystania z danych i możliwości technologicznych duńska Rada ds. Etyki Danych może przyczynić się do zapewnienia, aby korzystanie z danych było etycznie odpowiedzialne, a tym samym przyczynić się do pozytywnego rozwoju w zakresie wykorzystania danych”¹³².

W praktyce Rada skupiona jest na wypełnianiu dwóch głównych kwestii:

- kreowaniu debat na temat etyki danych i zwiększania świadomości społecznej na temat AI;
- aktywności na rzecz odpowiedzialnego i zrównoważonego wykorzystania danych w biznesie i sektorze publicznym.

Kreowanie debaty na temat etyki danych i zwiększaniu świadomości społecznej

„Poprzez zalecenia, raporty, oświadczenia i działania generujące debaty. (...). Jednocześnie duńska Rada ds. Etyki musi stale pracować nad skoncentrowaniem się na odpowiedniej równowadze między rozwojem technologicznym i wykorzystaniem danych z jednej strony, a **poszanowaniem praw podstawowych jednostki i pewnością prawa, a także skuteczną administracją publiczną, która przyczyni się do zwiększenia zaufania obywateli do wykorzystywania danych przez sektor publiczny oraz przedsiębiorstwa.**

Zostanie to zrealizowane w ramach ogólnego nacisku na **ochronę danych oraz bezpieczeństwo cybernetyczne i bezpieczeństwo informacji w sektorze publicznym i biznesowym**, a także w ramach skoncentrowania się na tym, jak zapewnić dobre ramy dla przedsiębiorstw, tak aby Dania miała konkurencyjny i cyfrowy sektor prywatny, który jest przygotowany na zmiany przyszłości.”¹³³.

Aktywności na rzecz odpowiedzialnego i zrównoważonego wykorzystania danych w biznesie i sektorze publicznym:

„w radzie zakotwiczone są trwałe wysiłki na rzecz wspierania odpowiedzialnego i zrównoważonego wykorzystywania danych. Rada będzie działać na rzecz zapewnienia, aby odpowiedzialne podejście do wykorzystywania danych mogło stać się parametrem konkurencyjności zarówno na szczeblu krajowym, jak i międzynarodowym, np. poprzez formułowanie zaleceń i oświadczeń (...).

Duńska Rada ds. Etyki musi wspierać odpowiedzialne i zrównoważone wykorzystanie danych w sektorze publicznym, aby zwiększyć zaufanie obywateli”¹³⁴.

Rada ds. Etyki stworzyła również specjalny formularz, który może wypełnić każdy przedsiębiorca.

„Pomaga on zrozumieć zasady etyki danych (...) Celem formularza oceny nie jest wyprowadzenie jednoznacznego wniosku dotyczącego etyki danych w odniesieniu do danego projektu, w który firma jest zaangażowana. Formularz ma na celu zwiększenie wiedzy i świadomości na temat etyki danych oraz umożliwienie dokonywania szerszych rozważań etycznych dotyczących danych i podejmowania decyzji, które są etycznie uzasadnione.”¹³⁵.

132 <https://dataetiskraad.dk/om-dataetisk-raad/baggrund-og-formaal> [dostęp: 18 października 2024 r.].

133 <https://dataetiskraad.dk/om-dataetisk-raad/baggrund-og-formaal> [dostęp: 18 października 2024 r.].

134 <https://dataetiskraad.dk/om-dataetisk-raad/baggrund-og-formaal> [dostęp: 18 października 2024 r.].

135 <https://dataetiskraad.dk/alle-udgivelser/dataetisk-vurderingsskema-2023> [dostęp: 18 października 2024 r.].



DATAETISK VURDERINGSSKEMA

NULSTIL ALLE SVAR



Indledende overvejelser

Som det første, er det relevant at du danner dig et overblik. Det gør du ved at finde ud af, hvad formålet med samkøringen af data er, hvor i organisationen datasamkøringen foretages, og hvor stor en del af befolkningen der berøres af samkøringen.

Hvilken del af myndigheden foretager datasamkøringen?

Er der taget højde for alle de lovgivningsmæssige krav, som skal iagttages ved samkøringen?

☐

JA

☐

NEJ

☐

DE RETLIGE VURDERINGER ER UNDER UDARBEJDELSE

Hvilke overordnede formål har samkøringen? (sæt gerne flere kryds)

☐

FORSKNING

☐

FOLKESUNDHED

☐

SUNDHED

☐KLIMA/
MILJØ☐KONTROL/
TILSYN☐

UDDANNELSE

☐ADMINISTRATION/
SAGSOPTIMERING☐

ANDET

Hvor stor en del af befolkningen berøres ca. af samkøringen? (sæt ét kryds)

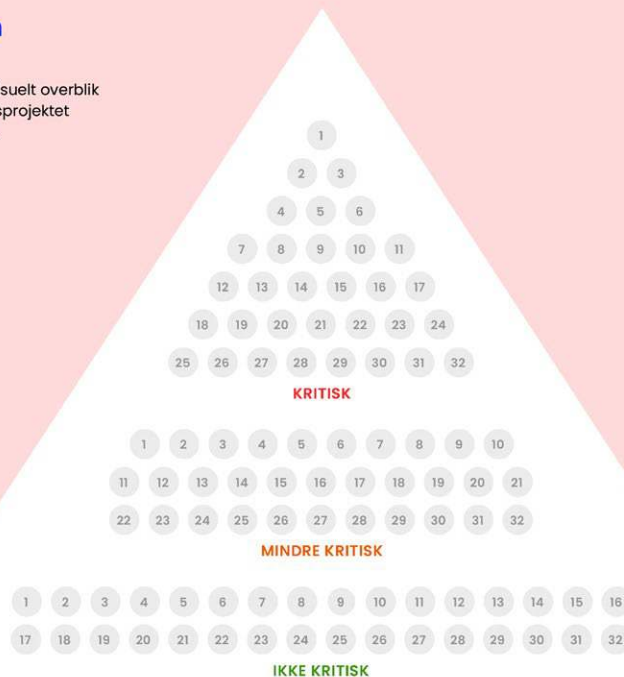
☐MINDRE
END 25 %☐ML.
25-50 %☐ML.
50-75 %☐OVER
75 %

NÆSTE

DATAETISK VURDERINGSSKEMA

Risikopyramiden

Risikopyramiden giver dig et visuelt overblik over, hvordan datasamkøringsprojektet samlet set vurderes dataetisk.



Rada stworzyła także definicję etyki danych w sztucznej inteligencji, zgodnie z którą: „etyka danych jest ogólnie rozumiana jako etyczny wymiar relacji między technologią z jednej strony, a prawami podstawowymi obywateli, pewnością prawa i podstawowymi wartościami społecznymi z drugiej, które rozwój technologiczny skłania do rozważenia. Koncepcja ta obejmuje kwestie etyczne związane z wykorzystaniem danych”¹³⁶.

Również Bank Centralny Danii (*Denmarks Nationalbank*) przywiązuje dużą uwagę do etyki w AI. Podkreśla, że sektor finansowy musi być regularnie monitorowany. **Nie może bowiem dojść do dyskryminacji ze względu na rasę czy płeć.** Istotną obawą jest również możliwość nałożenia przez Unię Europejską na instytucje finansowe kar, sięgających nawet 7% rocznych obrotów firm¹³⁷.

Z tych względów *Denmarks Nationalbank* opublikował dokument pt. „*Five suggestions for banks that want to use artificial intelligence*”¹³⁸, zawierający pięć propozycji rozwiązań, które powinny stosować banki, aby monitorować oraz udoskonalać etykę w bankowości.

Pierwsza z nich opiera się na przygotowaniu i utrzymywaniu katalogów, przeglądów różnych modeli systemu AI wykorzystywanego przez daną organizację. Ma to na celu **zwiększenie wewnątrzorganizacyjnej świadomości, ale też skupienie uwagi na miejscach, w których istnieje duże prawdopodobieństwo wystąpienia ryzyka.**

Druga z nich polega na **stałej ocenie systemu.** Nie tylko na etapie powstawania, ale również bieżącego monitorowania, co jest niezwykle istotne przy stałym funkcjonowaniu systemu. Zawsze bowiem, wprowadzając zmiany, mogą pojawić się nowe błędy, nawet tam, gdzie wcześniej one nie występowały.

Trzecia polega na zatrudnianiu ekspertów, którzy będą na bieżąco kontrolować system. Najistotniejsze nie jest bowiem stworzenie nowego systemu, a **ciągła praca nad nim, kontrola i permanentne ulepszanie**, w celu wyeliminowania wszystkich błędów.

Jako czwartą z nich *Denmarks Nationalbank* podał **krytyczne spojrzenie na wytwarzany projekt.** Sugeruje również tworzenie łatwiejszych systemów, ponieważ w momencie awarii łatwiej będzie go naprawić.

Piątą z nich to zachęta do utworzenia **wspólnej przestrzeni komunikacyjnej dla osób z sektora finansowego, na której będą mogli wymieniać się spostrzeżeniami i sugestiami.** Ma to również zachęcić do szerokiego i jednolitego przestrzegania przepisów prawa w tym obszarze¹³⁹.

Duński Urząd Nadzoru Finansowego (*Finanstilsynet*) przekazał natomiast do konsultacji projekt dokumentu w sprawie dobrych praktyk w zakresie wykorzystania sztucznej inteligencji w sektorze finansowym pt. „*God praksis: Dataetik ved brug af AI i den finansielle sektor*”. Dokument ten został sporządzony w oparciu o konsultacje z innymi organizacjami finansowymi w Danii¹⁴⁰. Są to swego rodzaju rekomendacje dla instytucji finansowych. Nie są one formalnie wiążące jako źródło „twardego prawa”. Dokument zawiera przykłady złych lub wątpliwych etycznie praktyk i stara się je ograniczać.

„Etyka danych w mniejszym stopniu dotyczy przepisów, a bardziej **wartości społecznych dyalematów i rozważania sprzecznych interesów.** Chodzi o to, za co firma może ręczyć. Ważne jest, aby firmy **szanowały godność jednostki.** Za każdym numerem klienta kryją się jednostki. Każda osoba ma **prawo do prywatności i dokonywania własnych wyborów.** Moralność i normy różnią się w zależności od osoby, firmy, sytuacji i kultury. Dlatego, to co jest uważane za etycznie poprawne, również się różni”¹⁴¹.

Dokument zakłada przede wszystkim skupienie się na „identyfikacji, ale także właściwym zarządzaniem ryzykiem związanym z wykorzystywaniem sztucznej inteligencji”¹⁴².

¹³⁶ <https://erhvervsstyrelsen.dk/vejledning-vejledning-om-lovpligtig-redegoerelse-dataetik> [dostęp: 18 października 2024 r.].

¹³⁷ <https://eur-lex.europa.eu/legal-content/PL/TXT/?uri=CELEX:32024R1689> [dostęp: 18 października 2024 r.].

¹³⁸ [news-no-1-five-suggestions-for-banks-that-want-to-use-artificial-intelligence.pdf](https://www.nationalbanken.dk/media/52952/god-praksis_dataetik_AI.pdf) (*nationalbanken.dk*) [dostęp: 18 października 2024 r.].

¹³⁹ <https://www.nationalbanken.dk/media/05jg4iw1/news-no-1-five-suggestions-for-banks-that-want-to-use-artificial-intelligence.pdf>, s. 2 [dostęp: 18 października 2024 r.].

¹⁴⁰ https://www.finanstilsynet.dk/nyheder-og-presse/nyheder-og-pressemeddelelser/2023/maj/hoering_dataetik_ai_170523 [dostęp: 18 października 2024 r.].

¹⁴¹ https://www.finanstilsynet.dk/media/52952/god-praksis_dataetik_AI.pdf, s. 8 [dostęp: 18 października 2024 r.].

¹⁴² https://www.finanstilsynet.dk/media/52952/god-praksis_dataetik_AI.pdf, s. 3-4 [dostęp: 18 października 2024 r.].

„Firmy mogą rozważyć, w jaki sposób oceniają i utrzymują swoją pracę z etyką danych wewnętrznie. Ważne jest, aby etyka danych nadążała za rozwojem społecznym i technologicznym”¹⁴³.

Dokument stawia nacisk na to, by przedsiębiorcy skrupulatnie wybierali swoje systemy i na bieżąco je monitorowali. W dokumencie podkreśla się również, że **etyka danych w AI – to jeszcze nie są w pełni rozwinięte regulacje prawne. Aktualnie są to dopiero wartości społeczne**, które powinny zostać powszechnie przyjęte i uniwersalnie respektowane.

Ponadto, opisane zostało jak podmioty rynku finansowego i instytucje finansowe powinny wprowadzać etykę danych w AI. Duży nacisk kładziony jest na określony sposób myślenia oraz „zdrową kulturę korporacyjną”. Polityka podmiotów rynku finansowego w zakresie etyki danych powinna zawierać opis standardów etycznych. Istotnym też jest prowadzenie szkoleń dla pracowników, aby każdy z pracowników potrafił odnaleźć się w coraz to nowszych zmianach. Dokument podkreśla także, że w każdym przedsiębiorstwie warto wyznaczyć po jednym pracowniku z każdego działu w celu zapewnienia jak najlepszej wymiany informacji w firmie. Oprócz tego prezes powinien wyznaczyć osobę odpowiedzialną za kierowanie całym sektorem etyki danych w AI, natomiast do wprowadzania etyki w politykę firmy warto zaangażować wszystkich pracowników.

„Jeśli firma aktywnie decyduje się na pracę z etyką danych, ważne jest, aby zaangażować wszystkich pracowników w celu stworzenia wspólnej, zdrowej kultury postępowania z danymi. Sposób myślenia o etyce danych nie jest rozwijany i wzmacniany podczas jednego spotkania, ale raczej z czasem poprzez dialog oraz dyskusję wśród szerokiego grona pracowników o wielu różnych perspektywach, obszarach pracy i umiejętnościach. Uwzględnienie wielu punktów widzenia w tym procesie najlepiej służy jego wynikom”¹⁴⁴.

W dokumencie wskazuje się także obawy, które wiążą się z niewłaściwym wykorzystywaniem AI. Jednym z nich jest gromadzenie i wykorzystywanie danych wyłącznie w interesie przedsiębiorstwa, które

łączy się bezpośrednio z prowadzeniem statystyk odnośnie pochodzenia klientów. Niektóre z nich wprowadziły takie rozwiązanie, aby pokazać, iż nie dyskryminują osób ze względu na inną narodowość. Jest to jednak rozwiązanie nieetyczne, ponieważ **narazi klientów na nieprzyjemne i obraźliwe sytuacje chociażby przez pracowników**. Kolejnym **etycznie wątpliwym zachowaniem jest śledzenie bieżącej lokalizacji klientów**. Mimo, iż niektórzy świadomie wyrazili zgodę na to rozwiązanie, ciągłe monitorowanie klientów nie jest etyczne, zwłaszcza gdy przedsiębiorstwa wykorzystują te informacje w celu zwiększenia możliwości sprzedażowych. Kolejnym **dyskusyjnym etycznie rozwiązaniem jest wprowadzenie algorytmów odpowiadających za kredyty**. Dla firm stanowi to wygodne rozwiązanie, klienci jednak mogą czuć się dyskryminowani. W razie wadliwości algorytmu zasady przyznawania kredytu nie będą do końca znane. Może to doprowadzić do nieprawidłowego udzielania kredytów, a istotą powstania dokumentu jest **uczciwe i sprawiedliwe traktowanie wszystkich klientów**¹⁴⁵.

Oprócz tego, w dokumencie wyraźnie podkreślona jest istota stosowania wytycznych, wskazanych przez Radę ds. Etyki danych. Stworzyła ona pięciostopniowy przewodnik dotyczący pracy z etyką danych. Pierwszym krokiem jest określenie postępowania z danymi w projektach oraz jaki jest ogólny cel ich przetwarzania. Następnie przedsiębiorstwa i instytucje finansowe powinny przeanalizować plusy i minusy. W trzecim kroku wszystkie aspekty z poprzedniej części powinny zostać sobie przeciwstawione. Po dokładnej analizie poprzednich punktów, instytucje powinny wybrać, które aspekty mają największe znaczenie i wprowadzić je w życie. Ostatnim krokiem jest bieżąca oraz konsekwentna ocena wybranych rozwiązań.

Rada ds. etyki danych wskazuje również 10 kluczowych zasad i wartości, których przedsiębiorstwa powinny przestrzegać. Są to:

- I. **Godność każdego człowieka** – człowiek i życie ludzkie powinny być stawiane „na pierwszym miejscu”, a przetwarzanie danych w żaden sposób nie powinno szkodzić jednostce.
- II. **Prywatność** – przetwarzanie danych musi odbywać się z poszanowaniem prywatności i ochro-

¹⁴³ https://www.finanstilsynet.dk/media/52952/god_praksis_dataetik_AI.pdf, s. 20 [dostęp: 18 października 2024 r.].

¹⁴⁴ https://www.finanstilsynet.dk/media/52952/god_praksis_dataetik_AI.pdf, s. 17 [dostęp: 18 października 2024 r.].

¹⁴⁵ https://www.finanstilsynet.dk/media/52952/god_praksis_dataetik_AI.pdf [dostęp: 18 października 2024 r.].

ną danych osobowych. Ich zakres powinien być ograniczony do niezbędnych informacji.

III. **Indywidualna wolność, samostanowienie oraz racjonalność** – przetwarzanie danych ma pomagać ludziom podejmować świadome i niezależne decyzje, a nie ich ograniczać.

IV. **Równość, inkluzywność i niedyskryminacja** – przetwarzanie danych nie powinno dyskryminować ze względu na pochodzenie etniczne, orientację seksualną, płeć, status społeczno-ekonomiczny, przekonania polityczne, religię, przynależność związkową, dane genetyczne, dane biometryczne, niepełnosprawność lub inne dane zdrowotne. Trzeba dążyć do zapewnienia różnorodności i inkluzywności.

V. **Prawa podstawowe i wolności jednostki** – przetwarzanie danych powinno odbywać się z poszanowaniem podstawowych praw i wolności obywatelskich.

VI. **Prawo oraz sprawiedliwość** – przetwarzanie danych powinno odbywać się z poszanowaniem podstawowych gwarancji prawnych oraz poziomu bezpieczeństwa prawnego w społeczeństwie.

VII. **Przejrzystość** – każdy powinien mieć możliwość wglądu do swoich danych, a także do informacji z nimi związanych, m.in. celu gromadzenia, funkcji, czy bezpieczeństwa.

VIII. **Bezpieczeństwo** – należy zapewnić bezpieczeństwo w przechowywaniu i udostępnianiu danych, aby zapobiec ich niezamierzonemu ujawnieniu osobom nieupoważnionym.

IX. **Odpowiedzialność** – zastosowanie nowoczesnych technologii, w tym zwłaszcza AI, nie może zwalniać konkretnych osób od odpowiedzialności. Powinno być jasno określone, kto jest odpowiedzialny za skutki wynikające z rozwoju i wykorzystania danych na każdym etapie¹⁴⁶.

Podsumowując, należy podkreślić, że Dania jest niezwykle rozwinięta w zakresie etyki danych w AI. Wprowadza zarówno swoje rozwiązania i regulacje w tym zakresie, jak również wzoruje się na aktach wydawanych przez Unię Europejską, np. AI Act. Powołanie Rady ds. etyki danych pomaga sprawować kontrolę i stale polepszać jakość, co znacząco pomaga w rozwoju tego sektora.

146 <https://admin.nationaltcenterforetik.dk/Media/637847650997152568/Dataetik%20-%20S%C3%A5dan%20g%C3%B8r%20du.pdf> [dostęp: 18 października 2024 r.].

9. Włochy



Przechodząc do Włoch, zauważyć należy, że instytucjonalna etyka w sektorze finansowym dopiero się rozwija. Państwo to powoli zaczyna wprowadzać poszczególne rozwiązania mające zapewnić przejrzystość sztucznej inteligencji oraz bezpieczeństwo danych osobowych.

Banki zaczynają wprowadzać zasady etycznego korzystania ze sztucznej inteligencji. Przykład może stanowić Banca Intesa Sanpaolo. Bank ten przyjął zasady, które stanowią fundament wykorzystywania AI w bezpieczny sposób.

Pierwszą z nich jest **wytłumaczalność**. Bank musi tworzyć modele, które będą zrozumiałe dla użytkowników. Kolejną zasadą jest **uczciwość**. Ma ona na celu zapewnienie przestrzegania kodeksu etycznego w modelach sztucznej inteligencji. Trzecią z nich jest **jakość danych**, która ma być zapewniona na najwyższym poziomie z dbałością o wiarygodność wyników. Ostatnią, ale jednocześnie najważniejszą zasadą, jest **ciągły nadzór człowieka nad decyzjami podejmowanymi przez AI**, który zapewni bezpieczeństwo w stosowaniu wszystkich powyższych zasad¹⁴⁷. Oprócz tego, Intesa Sanpaolo wykorzystuje włoski start-up Aptus.AI¹⁴⁸, który regularnie bada zgodność polityki firmy i ich rozwiązań z aktualnymi rozwiązaniami prawnymi.

Aptus.AI zajmuje się głównie kontrolowaniem zgodności zasad wewnętrznych w różnych instytucjach z obowiązującymi aktami prawnymi, zarówno włoskimi jak i tymi, wydanymi przez Unię Europejską. Jest to istotne, ponieważ we Włoszech kładziony jest duży nacisk na stosowanie regulacji unijnych. Aktualnie kluczową rolę odgrywa oczywiście AI Act, który stanowi punkt wyjścia dla wyznaczenia ram etyki AI. To właśnie na jego podstawie Włochy zaczynają i w przyszłości będą tworzyć regulacje krajowe.

Poszczególne regiony zaczynają tworzyć organizacje zajmujące się tylko etyką w sztucznej inteligencji. Pierwszym z nich jest Turyn, który 31 października 2023 r. powołał komisję etyczną ds. sztucznej inteligencji, która składa się z multidyscyplinarnego zespołu ekspertów. Ma ona na celu spojenie projektów związanych z innowacjami w AI¹⁴⁹.

Ponadto Bank Auriga wskazuje na dwie główne potrzeby sztucznej inteligencji „*utrzymanie wysokiego poziomu konkurencyjności nawet poza znanymi granicami geograficznymi oraz nawiązanie relacji z konsumentami, którzy wymagają coraz prostszych, a jednocześnie skutecznych usług, bezproblemowych, ale coraz bardziej spersonalizowanych*”¹⁵⁰.

Bank wskazuje również na korzyści, które może przynieść korzystanie z AI. Są to:

- spersonalizowane podejście do klienta, które nie wyklucza relacji międzyludzkich;
- zachowywanie inwestycji dzięki natychmiastowej integracji z technologią;
- korzystanie z transformacji cyfrowej, która obejmuje wszystkich interesariuszy banku, od kierownictwa po klientów;
- generowanie coraz większego zaufania do marki, które odnawia się dążąc do mądrego, choć prostego podejścia¹⁵¹.

W kontekście doświadczeń z Półwyspu Apenińskiego, wspomnieć trzeba, że papież Franciszek wraz z innymi przywódcami religii monoteistycznych, podpisał rzymski apel o etyczną sztuczną inteligencję „*Mission report Oct 2022 – Oct 2023*”. Jego głównymi celami są:

- wspólna refleksja i dyskusja różnych środowisk w celu zidentyfikowania różnych zasad etycznych, które mają być stosowane przy rozwoju AI;
- poczucie wspólnej odpowiedzialności;
- refleksja naukowa¹⁵².

Także Premier Włoch, Giorgia Meloni podkreśliła, że „*priorytetem jest zapewnienie, aby sztuczna inteligencja pozostała skoncentrowana na człowieku i kontrolowana przez człowieka*”¹⁵³.

147 <https://group.intesasanpaolo.com/en/newsroom/all-news/news/2024/artificial-intelligence-impact-in-finance> [dostęp: 18 października 2024 r.].

148 <https://www.aptus.ai/en/platform-and-tech/platform/> [dostęp: 18 października 2024 r.].

149 http://www.comune.torino.it/ucstampa/2023/article_583.shtml#:~:text=A%20pochi%20giorni%20dall%E2%80%99annuncio%20del%20Governo%20della%20scelta%20di [dostęp: 18 października 2024 r.].

150 <https://www.aurigaspaspa.com/banking/esigenze/intelligenza-artificiale-banche/> [dostęp: 18 października 2024 r.].

151 <https://www.aurigaspaspa.com/banking/esigenze/intelligenza-artificiale-banche/> [dostęp: 18 października 2024 r.].

152 https://www.romecall.org/wp-content/uploads/2024/02/RomeCall_report-web.pdf, s. 5 [dostęp: 18 października 2024 r.].

153 Meloni vows to address AI ethics as Italy takes G-7 lead, <https://decode39.com/8227/meloni-ai-ethics-italy-g7-lead/> [dostęp: 18 października 2024 r.].

Ponadto, Włoska Agencja ds. Cyfryzacji (*Agenzia per l'Italia Digitale*) (dalej jako: „**AGID**”) opublikowała dokument opisujący wyzwania, strategie i cele dla sztucznej inteligencji, które zakładają:

- uproszczenie rozwoju i korzystania z usług cyfrowych;
- wdrożenie krajowych platform w celu napędzenia transformacji cyfrowej;
- zwiększenie wydajności i bezpieczeństwa krajowej infrastruktury fizycznej;
- podniesienie poziomu świadomości zagrożeń, cyberzagrożeń i przeciwdziałaniu im;
- kierowanie ewolucją usług publicznych.¹⁵⁴

AGID opublikowała również strategię Włoch na lata 2024-2026 („*Italian strategy for artificial intelligence 2024-2026*”). Jest ona skierowana do różnych sektorów: badawczego, administracji publicznej, przedsiębiorstw, czy szkoleń. Zakłada ona m.in.:

- wsparcie dla rozwoju i przyjęcia rozwiązań AI;

- laboratoria dla rozwoju aplikacji AI w kontekście przemysłowym;
- powszechne nauczanie sztucznej inteligencji na uniwersytetach;
- rozwój *startupów* AI¹⁵⁵.

Niestety, tak jak zostało wspomniane na początku tego rozdziału, doniosłość etyki sztucznej inteligencji we włoskim sektorze bankowym dopiero zaczyna być dostrzegana. Banca d'Italia porusza ten temat w raporcie z 2023 r. *Livello di rilevanza criticità attese - IA generativa* (Poziom istotności przewidywanych krytycznych zagadnień – generatywna Sztuczna Inteligencja). Etyka znajduje się w nim dopiero na szóstym miejscu, co nadal czyni ją tematem dosyć pomijanym¹⁵⁶.

Podsumowując, Włochy dopiero zaczynają rozwój w zakresie etyki danych w bankowości. Prowadzą działania mające na celu doprowadzenie do zgodności z regulacjami unijnymi, a także tworzą własne, oparte o już istniejące akty prawne. W perspektywie czasu, wszystkie wyżej wymienione działania mogą mieć istotny wpływ na etykę sztucznej inteligencji także w bankowości.

¹⁵⁴ https://www.agid.gov.it/sites/default/files/repository_files/3_anni_di_agid05giugno.pdf, s. 4, [dostęp: 18 października 2024 r.].

¹⁵⁵ https://www.agid.gov.it/sites/agid/files/2024-07/Italian_strategy_for_artificial_intelligence_2024-2026.pdf, s. 25 [dostęp: 18 października 2024 r.].

¹⁵⁶ https://www.cipa.it/rilevazioni/tecnologiche/2023/Rilevazione_tecnologica_2023.pdf, s. 48 [dostęp: 17 października 2024 r.].

10. Wnioski i rekomendacje





Biorąc pod uwagę przeprowadzoną analizę, należy stwierdzić, że:

- aktualnie prace nad ramami prawnymi – w tym zarówno „twardym” prawem powszechnie obowiązującym, jak i *soft law* – dotyczącymi etyki AI w sektorze bankowym nie są w państwach członkowskich Unii Europejskiej zaawansowane, przeciwnie: we wszystkich uwzględnionych w analizie krajach prace dopiero się toczą, pozostając na ogół na poziomie koncepcyjnym, projektów legislacyjnych, debaty eksperckiej, konsultacji publicznych, a nawet zaledwie analizy unijnego AI Act;
- legislacja krajowa w obszarze AI jest obecnie w relatywnie wczesnym stadium rozwoju – Respondenci podkreślali, że wobec niedawnego przyjęcia AI Act oraz braku stosownego orzecznictwa TSUE lub sądów krajowych aktualny stan prawny określić można mianem „stanu przejściowego” lub „regulacji eksperymentalnej”;
- silne jest przekonanie, że – wobec „delikatności” oraz powagi materii, jaką jest etyka AI – regulacja typu „*hard law*” (precyzyjne przepisy o charakterze *ius cogens*) mają przewagę nad wytycznymi, rekomendacjami, zaleceniami, kodeksami etycznymi i innymi aktami o charakterze *soft law*;
- żaden z uwzględnionych w badaniu krajów nie wprowadził jednego uniwersalnego aktu prawnego, który regulowałby etykę AI dla sektora finansowego/bankowego;
- brak jest obecnie systemowego, zamkniętego i koherentnego zestawu takich norm – zarówno w „prawie twardym”, jak i w „prawie miękkim”;
- mało prawdopodobne wydaje się, aby w najbliższym czasie (szacowanym na rok-trzy lata) udało się wypracować jeden europejski (ogólnounijny) kodeks etyki AI w sektorze bankowym, choć nie oznacza to, że nie ma takiej potrzeby, a tym bardziej, że organy unijne, krajowe oraz banki powinny ze sobą współpracować w tym zakresie, prowadzić dyskusję, utworzyć forum (płaszczyzny) współpracy, dialog w celu wymiany doświadczeń itd.;
- aktualnie regulacja AI, w tym również dotycząca etyki AI, jest rozdrobniona: pojawia się w wielu różnych odrębnych aktach (dotyczących danych

osobowych, praw konsumentów, kredytów bankowych, płatności *online* itp.);

- poszczególne organy i instytucje zajmujące się rynkiem finansowym (krajowe banki centralne, organy nadzoru finansowego, organizacje bankowe) przyjmują różne akty *quasi-prawne* lub prawo miękkie (a także wydają publikacje/stanowiska/komentarze eksperckie itp.) dotyczące etyki AI, ale mają one charakter ogólny, wskazujący zasady i wartości (a nie konkretne reguły);
- rozwiązania AI w bankowości zmieniają się i rozwijają tak szybko, że trudno wyobrazić sobie jeden spójny i precyzyjny kodeks etyki – raczej powinien on bazować i wyrażać pewne ogólne zasady i wartości, jak np.: odpowiedzialności i rozliczalności (ang. *responsibility and accountability*), uczciwości oraz wyjaśnialności (*bias, fairness and explainability*) itp.;
- dla etyki AI kluczowe są – jako kryteria oceny etyczności AI – pewne uniwersalne ogólne zasady i wartości, takie jak: wyjaśnialność, poszanowanie godności i praw człowieka (w tym prywatności i danych osobowych), równość, niedyskryminacja, nadzór ludzki itp.;
- doświadczenia zagraniczne państw zachodniej Europy oraz standardy na szczeblu ponadnarodowym pozwalają sformułować przypuszczenie, że nie należy spodziewać w najbliższej przyszłości wydania koherentnego, precyzyjnego i uniwersalnego kodeksu etycznego dla AI, który stanowiłby wzorzec dla innych dokumentów lub analogicznych aktów poszczególnych państw europejskich, w tym Polski;
- Związek Banków Polskich, Komisja Etyki Bankowej, banki oraz organy władzy publicznej powinny monitorować prawodawstwo unijne i polskie, obserwować praktykę korzystania z AI oraz standardy, zasady i wartości, a także upowszechniać wiedzę o etyce AI oraz promować jak najwyższe normy w tym zakresie, m.in. poprzez: szkolenia, konferencje, seminaria, debaty, konkursy, publikacje naukowe oraz raporty.

Rekomenduje się:

- aktywny udział polskiego sektora bankowego w inicjatywach na poziomie unijnym, prowadzących do wypracowania standardów w zakresie etyki AI na szczeblu ponadnarodowym;



- respektowanie zasad, wartości i standardów wskazanych w raporcie, w tym m.in.: zaangażowania człowieka w decyzje z użyciem systemów AI (ang. *human in the loop*) w krytycznych dla konsumentów obszarach, takich jak ocena wiarygodności kredytowej i decyzja o udzieleniu kredytu, odpowiedzialność, rozliczalność, przejrzystość, wyjaśnialność, ochrona prywatności i równe traktowanie;
- informowanie klientów banków na temat zastosowania systemów SI w ważnych decyzjach podejmowanych przez bank, w tym zwłaszcza uwzględnianych w nim parametrów, kryteriów i okoliczności;
- proaktywne edukowanie konsumentów na temat roli systemów SI w sektorze bankowym;
- jasne i precyzyjne określanie odpowiedzialności za systemy SI wewnątrz banków oraz ich transparentne komunikowanie;
- regularne audyty sprawdzające – zarówno od strony technicznej, prawnej, jak i etycznej – systemy AI wykorzystywane przez podmioty sektora bankowego;
- przeprowadzenia analogicznego badania (w celu weryfikacji aktualności ustaleń i tez przedstawionych w raporcie) w przyszłości, za dwa-trzy lata.



PROGRAM
ANALITYCZNO
BADAWCZY

Raport przygotowany na zlecenie
Programu Analityczno-Badawczego
Fundacji Warszawski Instytut Bankowości