ZASTOSOWANIE SZTUCZNEJ INTELIGENCJI W BANKOWOŚCI – SZANSE ORAZ ZAGROŻENIA

ANALIZA PRAWNO-REGULACYJNA WPYWU TECHNOLOGII UCZENIA MASZYNOWEGO I POKREWNYCH NA OBOWIĄZKI SEKTORA BANKOWEGO Z ZAKRESU ZAPEWNIENIA ZGODNOŚCI (COMPLIANCE) ORAZ ZARZĄDZANIA RYZYKIEM.

SYGN. WIB PAB 12/2022

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

PROGRAM ANALITYCZNO BADAWCZY

Warszawa, sierpień 2022

O Raporcie

Raport **Zastosowanie sztucznej inteligencji w bankowości – szanse oraz zagrożenia** został opracowany na zlecenie Programu Analityczno-Badawczego Fundacji Warszawski Instytut Bankowości.
Projekt badawczy przygotowany zgodnie z umową pomiędzy Uniwersytetem Śląskim a Warszawskim Instytutem Bankowości.

Autorzy:

Raport przygotowany przez zespół w składzie:

Dr hab. Dariusz Szostek – kierownik projektu

– profesor Wydziału Prawa i Administracji Uniwersytetu Śląskiego, Dyrektor Śląskiego Centrum Inżynierii Prawa, Techniki i Kompetencji Cyfrowych Cyber Science Partner i założyciel Kancelarii Szostek-Bar i Partnerzy, Członek European Law Institute w Wiedniu, Ekspert Obserwatorium Sztucznej Inteligencji Parlamentu Europejskiego w Brukseli; Ekspert w EUiPO w zakresie blockchain, współautor koncepcji e-sądu, pomysłodawca elektronicznego potwierdzenia odbioru (zwrotki elektronicznej), autor koncepcji elektronicznego biura podawczego i zmian w kodeksie postępowania cywilnego, wykładowca, autor kilkudziesięciu publikacji (w tym monografii oraz publikacji zagranicznych m.in współautor *Cyber Law* – wydanie w New York, Tokio, Sydney, Amsterdam, London). Autor monografii *Blockchain i Prawo* 2018, i jej wersji angielskiej *Blockchain and Law* (Nomos Germany 2019), *Lega Tech* (Nomos Germany 2021), *Internet and New Technologies* (współredaktor) (Nomos Geremny 2021). Członek szeregu zespołów naukowych w Polsce oraz zagranicą.

Dr Gabriela Bar

– doktor nauk prawnych, radca prawny, współpracownik Cyber Science, partnerka zarządzająca w kancelarii prawnej Szostek_Bar i Partnerzy. Entuzjastka nowych technologii, ze szczególnym uwzględnieniem sztucznej inteligencji. Doświadczony ekspert w zakresie umów elektronicznych, e-commerce, aspektów prawnych wdrożeń systemów IT, ochrony danych osobowych i bezpieczeństwa informacji. TOP100 kobiet w sztucznej inteligencji w Polsce (Fundacja Perspektywy – Women in Tech). Członkini Komitetu Prawnego IEEE w projekcie IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems; członkini Stowarzyszenia Prawa Nowych Technologii oraz Women in AI. Adiunkt na Wydziale Prawa i Administracji Uniwersytetu Opolskiego – odpowiada za wsparcie prawne zespołu projektowego w ramach wdrożenia Digital Manufacturing Platforms for Connected Smart Factories (Industry 4.0). Adiunkt na Wydziale Prawa i Administracji Uniwersytetu w Katowicach – zapewniający wsparcie prawne projektu Multi-Agent Systems for Pervasive Artificial Intelligence for assisting Humans in Modular Production Environments (Horizon 2020). Absolwentka Wydziału Prawa, Administracji i Ekonomii Uniwersytetu Wrocławskiego oraz Akademii Leona Koźmińskiego. Autorka licznych publikacji z zakresu prawa nowych technologii.

Dr inż. Rafał Tomasz Prabucki

– doktor nauk prawnych i inżynier. Wydział Prawa i Administracji Uniwersytetu Śląskiego. Członek CYBER SCIENCE i zespołu naukowego Legalengineering. Asystent w projektach dotyczących wykorzystania nowych technologii w przemyśle: MAS4AI na Uniwersytecie Śląskim i SHOP4CF na Uniwersytecie Opolskim. Alumn stypendiów w Polsce, Hiszpani i Niemczech. Prowadzący wykłady na uczelniach w Polsce i we Włoszech. Członek komitetu sterującego Youth Internet Governance Forum Poland przy ONZ. Kierownik studiów podyplomowych „Tokenizacja i automatyzacja procesów w gospodarce cyfrowej" na Uniwersytecie Śląskim.

Dr Michal Nowakowski

– radca prawny, pełni funkcję Head of NewTech w NGL Advisory oraz Counsela w NGL Legal w obszarze fintech oraz newtech, a także jest co-founderem w start-upie Ceforai zajmującym się implementacją etycznej sztucznej inteligencji. Ekspert fundacji AILAWTECH i FinTech Poland oraz współpracownik Cyber Science Śląskiego Centrum Inżynierii Prawa, Technologii i Kompetencji Cyfrowych – Cyber Science. Współautor opracowania dla Parlamentu Europejskiego pt. „An Analysis of the Selected Aspects of the Draft Artificial Intelligence Act" prezentującego krytyczne uwagi do projektu rozporządzenia w sprawie sztucznej inteligencji. Redaktor naukowy komentarza do ustawy o przeciwdziałaniu praniu pieniędzy i finansowania terroryzmu wydawnictwa Wolters Kluwer, który ukaże się pod koniec 2022 r. Autor książki *Fintech. Technologia, Finanse, Regulacje. Przewodnik praktyczny dla sektora innowacji finansowych*, która otrzymała w 2021 r. prestiżową nagrodę eDukat za najlepsze opracowanie naukowe. Wykładowca na studiach podyplomowych, w tym na Politechnice Warszawskiej oraz Szkole Głównej Handlowej, a także Uniwersytecie SWPS, w obszarze sztucznej inteligencji oraz uczenia maszynowego. Autor publikacji naukowych oraz wystąpień na konferencjach i warsztatach oraz szkoleniach, w szczególności poświęconych zagadnieniom sztucznej inteligencji, etyki oraz walut cyfrowych banków centralnych.

PROGRAM ANALITYCZNO BADAWCZY

O Programie

Program Analityczno-Badawczy przy Fundacji WIB powstał w 2019 roku jako odpowiedź na potrzeby sektora bankowego w zakresie analizy zjawisk, tworzenia opracowań i porządkowania wiedzy w obszarach 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 ich wyników w celu rozwoju sektora bankowego, podnoszenia poziomu bezpieczeństwa usług oraz kreowania wartości dla klientów bankowości. PAB WIB kładzie duży nacisk na rozwój współpracy ze środowiskami akademickimi i eksperckimi, poszukując synergii w zakresie zainteresowań badawczych autorów oraz potrzeb rozwojowych sektora bankowego.

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

- **Nowe technologie i cyberbezpieczeństwo**
- **Zdolność banków do finansowania gospodarki**
- **Bankowość spółdzielcza**
- **Rynek nieruchomości**
- **Zielony ład i finansowanie energetyki odnawialnej**
- **Finansowanie projektów innowacyjnych**

Więcej na temat działalności programu na stronie www.pab.wib.org.pl

Kontakt:

dr Andrzej Banasiak
Koordynator Programu
m: 696 405 104
e: abanasiak@wib.org.pl

Jacek Gieorgica
m: 603 626 254
e: jgieorgica@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 Tomasz Pawlonka
m: 505 917 778
e: tpawlonka@wib.org.pl

Spis treści

Wstęp

Na poziomie Unii Europejskiej są obecnie prowadzone intensywne prace nad sfinalizowaniem projektu *Akt w sprawie sztucznej inteligencji* (publikacja 21 kwietnia 2021 r.), którego jednym z założeń jest wprowadzenie rozbudowanych wymogów dla tzw. systemów sztucznej inteligencji wysokiego ryzyka, do których zaliczają się m.in. systemy oceny zdolności kredytowej dla osób fizycznych, ale również inne systemy, które mogą być stosowane w sektorze bankowym. Projekt zakłada też wprowadzenie pewnych wymagań w odniesieniu do systemów o podwyższonym ryzyku. Ostateczny kształt przepisów nie jest jeszcze przesądzony, jednak wyraźnie widać, że jest to jeden z priorytetów dla Unii Europejskiej. Na poziomie krajowym od 2020 r. realizowana z kolei jest *Polityka rozwoju sztucznej inteligencji dla Polski*, w ramach której powołano m.in. grupę ds. sztucznej inteligencji w sektorze finansowym. Jednocześnie Urząd Komisji Nadzoru Finansowego podejmuje działania zmierzające do poprawy efektywności nadzoru poprzez m.in. wykorzystanie analityki danych oraz Big Data, jak również monitoruje działania sektora bankowego (i szerzej finansowego) w zakresie wykorzystania AI oraz przygotowania na nowe wyzwania w tym zakresie.

Zastosowanie sztucznej inteligencji (w praktycznie każdej konfiguracji) do realizacji kluczowych lub istotnych obszarów działalności banków może generować zróżnicowane ryzyka opisane w skrócie w niniejszym dokumencie. Nowe standardy Unii Europejskiej (nie tylko w obszarze AI, ale i szerzej – w cyfrowej odporności operacyjnej) będą wiązały się nie tylko ze znacznymi kosztami implementacji wymogów prawnych, ale także z wyzwaniami o charakterze infrastrukturalnym i operacyjnym, a także wyzwaniami w kontekście pozyskania zasobów osobowych. Projektowany akt w sprawie AI zakłada wprawdzie możliwość „wykorzystania" rozwiązań wynikających już z implementacji pakietu CRD/CRR, jednakże poprawna diagnoza potencjalnego wpływu zarówno rozwoju sztucznej inteligencji, jak i projektowanych przepisów na sektor bankowy, pozwoli na bardziej efektywną transformację oraz implementację. Nie jest także jasne, w jakiej postaci projektowane przepisy zostaną przyjęte, gdyż projekt budzi spore emocje, a na poziomie prawodawców tworzone są jego kolejne iteracje.

W praktyce może okazać się, że na bazie już istniejących przepisów pokryta jest część wymogów, które mogą pojawić się w związku z nowymi regulacjami prawnymi. Ustalenie *status quo* stanowi więc równie ważny element niniejszego opracowania.

Na marginesie warto zwrócić uwagę, że *Strategia (UE) dla cyfrowych finansów* opracowana przez Komisję Europejską zakłada możliwość wydania przez Europejski Urząd Nadzoru Bankowego wytycznych sektorowych dla sztucznej inteligencji. Takie wytyczne mogą również zostać opracowane na poziomie krajowym, m.in. w związku z wynikami ankiety przeprowadzonej przez UKNF.

ROZDZIAŁ I

Sztuczna inteligencja a systemy sztucznej inteligencji

Jednym z kluczowych pojęć mających znaczenie dla dalszej analizy jest „sztuczna inteligencja" czy też – uprzedzając dalsze wnioski – systemy sztucznej inteligencji. Kwalifikacja określonych rozwiązań jako będących częścią tzw. sztucznej inteligencji będzie bowiem determinowała zastosowanie (lub brak obowiązku stosowania) określonych aktów prawnych i regulacji, toteż właściwa kwalifikacja stanowi punkt wyjścia dla dalszej oceny. Jest to szczególnie istotne w kontekście rozwoju równoległego trendu, tzw. *Robotic Process Automation* (RPA), czyli zrobotyzowanej automatyzacji procesów, która w niewielkim stopniu pokrywa się z cechami charakterystycznymi dla modeli samouczących.

Pojęcie sztucznej inteligencji jest rozumiane niejednolicie[1] i ewoluuje ono wraz z rozwojem różnych technik i podejść opartych m.in. o uczenie maszynowe, uczenie głębokie czy przetwarzanie języka naturalnego, które pozwalają m.in. na tłumaczenie tekstów, kierowanie dronami czy przygotowywanie dokumentów prawnych[2]. Rozwiązania te znajdują również szerokie zastosowanie w sektorze finansowym, czemu poświęcona jest niniejsza analiza.

Samo zdefiniowanie sztucznej inteligencji może stanowić wyzwanie, bowiem zarówno w nauce, jak i prawodawstwie czy naukach politycznych nie ma w tej chwili konsensusu co do definicji, a poza tym rozwój sztucznej inteligencji jest dynamiczny, co może skutkować dezaktualizacją ukutych pojęć.

Obecny stan nauki i rozwoju inżynierii oprogramowania (oraz dziedzin pokrewnych) nie pozwala na jednoznaczne określenie kierunku zmian dla sztucznej inteligencji, w tym możliwości stworzenia tzw. *General Artificial Intelligence*[3], czyli sztucznej inteligencji odwzorowującej – przynajmniej w pewnym stopniu – działanie ludzkiego umysłu[4]. Sama tematyka wykracza jednak daleko poza zakres przedmiotowy niniejszego opracowania i w praktyce nie ma większego znaczenia dla dalszych rozważań. Z tego względu akcent zostanie położony na koncepcję tzw. systemów sztucznej inteligencji wykorzystujących zróżnicowane podejścia i techniki, jak np. uczenie maszynowe, uczenie głębokie, metody statystyczne i pokrewne czy wspomniane już przetwarzanie języka naturalnego.

Pojęcie systemów sztucznej inteligencji pojawiło się już kilka lat temu w opracowaniach organizacji międzynarodowych, jak OECD[5] czy UNESCO[6], a jako propozycja definicji legalnej po raz pierwszy pojawiła się we wniosku Komisji – projekcie rozporządzenia Parlamentu Europejskiego i Rady (UE) ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (Akt w sprawie sztucznej inteligencji) i zmieniającego niektóre akty ustawodawcze Unii[7] („AIA"), gdzie wskazano, że: *„Systemem sztucznej inteligencji jest oprogramowanie opracowane przy użyciu co najmniej jednej spośród technik i podejść wymienionych w załączniku I* [obejmuje on m.in. uczenie maszynowe, głębokie, ale także metody statystyczne – przyp. autora]*, które może – dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję" (art. 3 pkt 1 projektu AIA)".*

Przyjęcie powyższej definicji jako punktu wyjścia ma to znaczenie, że AIA zakłada poddanie wybranych systemów (m.in. odpowiedzialnych za ocenę zdolności kredytowej) specyficznym wymaganiom w zakresie m.in. zarządzania ryzykiem czy zarządzania danymi, co będzie miało istotne znaczenie dla wielu instytucji finansowych, w szczególności banków.

Jednocześnie należy zwrócić uwagę, że w sektorze finansowym nie w każdym przypadku stosowane jest pojęcie sztucznej inteligencji. Przykładowo Europejski Urząd Nadzoru Bankowego (EUNB) w swoich opracowaniach posługuje się raczej pojęciem „zaawansowanej analityki danych"[8] czy po prostu „uczeniem maszynowym"[9] jako lepiej oddających specyfikę stosowanych rozwiązań. Również Turing Institute, który przygotował obszerny raport nt. wykorzystania sztucznej inteligencji w sektorze finansowym, pomimo przybliżenia pojęć dla różnych kategorii sztucznej inteligencji[10], wyraźnie podkreśla, że największe znaczenie i zastosowanie znajduje uczenie maszynowe i jego różne odmiany.

[1] D.S. Grewal, *A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering*, IOSR Journal of Computer Engineering (IOSR-JCE) 2014, Vol. 16, Issue 2, s. 3.
[2] S. M. Liao, *A Short Introduction to the Ethics of Artificial Intelligence* [w:] S. M. Liao (red.), Ethics of Artificial Intelligence, Oxford 2020, s. 1.
[3] R. Fjelland, *Why general artificial intelligence will not be realized*, Humanities and Social Sciences Communications, 2020, 7:10.
[4] M. Świerczyński, Z. Więckowski, *Sztuczna inteligencja w prawie międzynarodowym. Rekomendacje wybranych rozwiązań*, Warszawa 2021, s. 39.
[5] https://oecd.ai/en/ai-principles (dostęp: 16.03.2022 r.).
[6] https://unesdoc.unesco.org/ark:/48223/pf0000380455 (dostęp: 16.03.2022 r.).
[7] https://eur-lex.europa.eu/legal-content/PL/TXT/?uri=CELEX:52021PC0206 (dostęp: 20.06.2022 r.).
[8] https://www.eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf (dostęp: 16.03.2022 r.).
[9] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Discussions/2022/Discussion%20on%20machine%20learning%20for%20IRB%20models/1023883/Discussion%20paper%20on%20machine%20learning%20for%20IRB%20models.pdf (dostęp: 16.03.2022 r.).
[10] Autorzy wskazują tutaj m.in. na symboliczną sztuczną inteligencję, statystyczną sztuczną inteligencję, ogólną sztuczną inteligencję oraz wąską sztuczną inteligencję. F. Ostmann, and C. Dorobantu, *AI in financial services*, *The Alan Turing Institute*, 2021, https://doi.org/10.5281/zenodo.4916041, s. 10.

Na koniec warto wskazać, że jedną z metod klasyfikacji modeli sztucznej inteligencji – jakkolwiek byśmy ich nie definiowali – jest podział na modele oparte o techniki uczenia się przez dany system i modele oparte o rozwiązania pozbawione tego elementu, np. proste drzewa decyzyjne. W tym drugim przypadku powstają wątpliwości, czy rzeczywiście powinny być one traktowane w kategoriach tzw. sztucznej inteligencji. Ma to praktyczne znaczenie w kontekście m.in. zagadnień przejrzystości (*transparency*) oraz wyjaśnialności (*explainability*), które stanowią znaczące wyzwanie zarówno dla podmiotów stosujących modele samouczące, jak i organów nadzoru.

ROZDZIAŁ II

Definicja systemów sztucznej inteligencji

AIA wprowadza pojęcie „systemów sztucznej inteligencji", mające znaczenie dla ewentualnego stosowania przepisów tego rozporządzenia w sektorze bankowym. Przedstawienie tej definicji – z zastrzeżeniem możliwych zmian na etapie prac legislacyjnych – wydaje się więc kluczowe.

Jak już wskazano wyżej, projektowane rozporządzenie w art. 3 pkt 1) definiuje „system sztucznej inteligencji" jako oprogramowanie opracowane przy użyciu co najmniej jednej spośród technik i podejść wymienionych w załączniku I, które może – dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję.

Zakres tych technik i podejść określa wspomniany załącznik I, który wymienia następujące pozycje (zastrzeżenia wymaga, że Komisja Europejska ma być upoważniona do dokonywania zmian w treści tego załącznika):

a) mechanizmy uczenia maszynowego, w tym uczenie nadzorowane, uczenie się maszyn bez nadzoru i uczenie przez wzmacnianie, z wykorzystaniem szerokiej gamy metod, w tym uczenia głębokiego;

b) metody oparte na logice i wiedzy, w tym reprezentacja wiedzy, indukcyjne programowanie (logiczne), bazy wiedzy, silniki inferencyjne i dedukcyjne, rozumowanie (symboliczne) i systemy ekspertowe;

c) podejścia statystyczne, estymacja bayesowska, metody wyszukiwania i optymalizacji.

Bazując na tej definicji możemy wskazać, że aby zakwalifikować dane rozwiązanie jako system sztucznej inteligencji w rozumieniu projektowanego rozporządzenia, spełnione muszą być następujące warunki:

a) rozwiązanie jest oprogramowaniem (brak przy tym wskazówki, jak należy rozumieć „oprogramowanie");

b) oprogramowanie jest opracowane z użyciem co najmniej jednej z technik i podejść określonych w załączniku I;

c) działa ono w celu określonym przez człowieka;

d) generuje wyniki takie jak treści, przewidywania, zalecenia lub decyzje;

e) wyniki te wpływają na środowiska, z którymi te systemy wchodzą w interakcję.

Systemami sztucznej inteligencji mogą być zatem modele stosowane w bankowości, np. na potrzeby oceny ryzyka i określenia kwoty kapitału do pokrycia ekspozycji, jednakże nie musi to – przynajmniej w kontekście projektowanego rozporządzenia – automatycznie rodzić obowiązku spełnienia szeregu dodatkowych wymogów prawnych tam określonych. Takie obowiązki mogłyby powstać w przypadku systemów sztucznej inteligencji wysokiego ryzyka, które zostały zdefiniowane w art. 6 projektowanego rozporządzenia, a które należy rozumieć m.in. jako explicite wskazane w załączniku III do projektowanego rozporządzenia.

Na dzień sporządzania niniejszego raportu nie było jasne czy (i w jakim zakresie) definicja systemów sztucznej inteligencji wysokiego ryzyka mogłaby objąć także rozwiązania szerzej związane z sektorem bankowym, wyłączając te rozwiązania, które zostały ujęte w Załączniku III do AIA, jednakże istotne jest monitorowanie tego obszaru zarówno przez banki, jak i Związek Banków Polskich.

ROZDZIAŁ III

PROGRAM ANALITYCZNO BADAWCZY

Wybrane metody i podejścia stosowane w bankowości

Jak zostało to podkreślone, kategoria systemów sztucznej inteligencji może mieć bardzo szeroki zakres. Załącznik I do AIA zawierający listę technik i podejść z zakresu sztucznej inteligencji wydaje się być listą zamkniętą, jednak rozwój technologiczny szybko może doprowadzić do powstania nowych technik i podejść, które będą zaliczane do szeroko rozumianej sztucznej inteligencji. Niemniej opisane w Załączniku I do AIA rozwiązania stanowią obecnie trzon rozwoju tej dziedziny.

Niemiecki organ nadzoru nad rynkiem finansowym w jednym ze swoich opracowań wskazał, że sama sztuczna inteligencja to nic innego, jak połączenie tzw. Big Data, zasobów komputerowych oraz uczenia maszynowego[11].

Uczenie maszynowe można definiować na wiele sposobów, choć pewne elementy pozostają aktualnie niezależnie od przyjętego kierunku. Za J. Alzubim (i in.) można przyjąć, że uczenie maszynowe jest kategorią „sztucznej inteligencji", które umożliwia komputerom „myślenie" i „uczenie się" we własnym zakresie[12]. Innymi słowy, jest to metoda pozwalająca komputerowi na przetwarzanie danych i wyciąganie z nich określonych wniosków, a także generowanie określonych (przez człowieka) rezultatów.

Zasadniczo wyróżnia się kilka modeli uczenia maszynowego, które w znacznej mierze różnią się poziomem zaangażowania człowieka w proces uczenia. Wyróżnia się jego następujące rodzaje:

1. **Uczenie nadzorowane** (*supervised learning*) polega na pozyskiwaniu przez algorytm „wiedzy" pochodzącej z konkretnych przykładów, które zostają uprzednio opisywane przez specjalistów. Na bazie danych wejściowych algorytm uczenia nadzorowanego „trenuje", dostrzegając pewne zależności (korelacje), które następnie wykorzystuje przy kolejnych przykładach. Uczenie nadzorowane jest obecne m.in. w obszarze zarządzania ryzykiem.

2. **Uczenie nienadzorowane** (*unsupervised learning*) jest rozwiązaniem przeciwnym do uczenia nadzorowanego. W tym przypadku algorytm nie posiada opisanych danych wejściowych, a poszukiwanie korelacji odbywa się na zasadzie „obserwacji" tych danych.

3. **Uczenie częściowo nadzorowane** (*semi-supervised learning*) można sprowadzić do połączenia obu powyższych modeli, co oznacza, że dane wejściowe są częściowo opisane i przyporządkowane (np. do danej kategorii), a częściowo nie. Zadaniem algorytmu jest więc ich odpowiednie przyporządkowanie i odnalezienie korelacji.

4. **Uczenie ze wzmocnieniem** (*reinforcement learning*) jest modelem uczenia maszynowego, który można określić jako uczenie metodą prób i błędów. Algorytm nie otrzymuje klucza, według którego mógłby przyporządkować określone dane, ale zestaw pewnych reguł i oczekiwanych rezultatów. Algorytm taki podejmuje więc działania w celu znalezienia „poprawnej odpowiedzi", czyli osiągnięcia pożądanego celu. W modelu uczenia ze wzmocnieniem można posługiwać się systemem nagród, które są przyznawane za prawidłowe działanie algorytmu.

Modele uczenia maszynowego wykorzystują dwa rodzaje danych – dane wejściowe i wyjściowe, a proces uczenia nazywany jest trenowaniem, którego efektem jest powstanie modelu uwzględniającego analizę danych w procesie trenowania. Jakość i efektywność modeli uczenia maszynowego zależy od wielu czynników, w tym przede wszystkim dostępności danych o określonej jakości oraz ilości, choć znaczna ilość danych niekoniecznie musi być gwarantem skuteczności danego modelu. Modele uczenia maszynowego mogą być podatne na ryzyka związane przykładowo ze stronniczością algorytmiczną, która może być efektem zarówno celowego działania człowieka, jak i niewłaściwego opisania danych.

Uczenie głębokie (*deep learning*) jest swoistą podkategorią uczenia maszynowego, która wykorzystuje jednak nieco inne techniki niż klasyczne uczenie maszynowe. Obecne są tutaj bowiem tzw. sieci neuronowe składające się z połączonych ze sobą warstw (wejściowych, wyjściowych i ukrytych), które przekazują sobie informacje wykorzystywane przykładowo do predykcji. W tym typie uczenia maszynowego funkcjonowanie całego mechanizmu ma w znacznej mierze odtworzyć schemat działania ludzkiego mózgu. Uczenie głębokie jest najczęściej wykorzystywane do przetwarzania języka naturalnego oraz rozpoznawania obrazów.

Metody statystyczne są często wyodrębniane z katalogu szeroko rozumianej sztucznej inteligencji i najczęściej znajdują zastosowanie w ekonometrii. Zasadniczo można przyjąć, że modele statystyczne są sposobem układu pewnych hipotez wyrażonych w konkretnej formule matematycznej. Metody i modele statystyczne pozwalają więc na wykrywanie pewnych korelacji czy dokonywanie predykcji, ale bez komponentu uczenia, jak ma to miejsce w przypadku uczenia maszynowego. Można wyróżnić wiele modeli z tego obszaru, jak modele regresji liniowej.

[11] BaFin, *Big data and artificial intelligence: Principles for the use of algorithms in decision-making processes*, s. 3. Dokument dostępny pod adresem: https://www.bafin.de/SharedDocs/Downloads/EN/Aufsichtsrecht/dl_Prinzipienpapier_BDAI_en.html?nn=9866146 (dostęp: 17.03.2022 r.).

[12] J. Alzubi, A. Nayyar, A. Kumar, *Machine Learning from Theory to Algorithms: An Overview*, IOP Conf. Series: Journal of Physics: Conf. Series 1142, 2018, 012012, s. 1.

Wyżej wskazane techniki i podejścia z szeroko rozumianego obszaru sztucznej inteligencji nie stanowią wszystkich dostępnych i wykorzystywanych rozwiązań. Warto zwrócić uwagę, że Europejski Urząd Nadzoru Bankowego w jednym ze swoich opracowań wskazał, że całokształt rozwiązań wykorzystujących dane w sposób „autonomiczny" lub „półautonomiczny", do których zaliczają się wskazane wyżej rozwiązania, stanowią większy podzbiór **zaawansowanej analityki**[13] (*advanced analytics*) wykorzystującej często tzw. duże zbiory danych (Big Data). Poza uczeniem maszynowym EUNB wyróżnia tutaj m.in. *data mining*, analizę sentymentu, analizę graficzną czy statystyki wielowymiarowe.

W sektorze finansowym, w tym bankowym, najczęściej wykorzystywane są metody statystyczne oraz uczenie maszynowe, a także procesy typu RPA, które nie należą do zbioru sztucznej inteligencji. Warto w tym miejscu zwrócić uwagę, że utożsamianie powyższych procesów zautomatyzowanej robotyzacji ze „sztuczną inteligencją" może prowadzić do tworzenia mylnego obrazu zaawansowania sektora bankowego w zakresie wykorzystania nowych technologii. Wprowadzenie stosownego rozróżnienia wydaje się więc w tym kontekście istotne.

3.1. Zastosowanie sztucznej inteligencji w sektorze bankowym

Wykorzystanie rozwiązań określanych mianem sztucznej inteligencji jest w sektorze bankowym nadal na relatywnie niskim poziomie[14], choć wyraźne przyspieszenie można dostrzec na całym świecie, szczególnie w związku z pandemią Covid-Sars-19[15]. Z raportu Europejskiego Urzędu Nadzoru Bankowego w zakresie wykorzystania rozwiązań typu RegTech (*Regulatory Technology*) wynika, że wykorzystanie uczenia maszynowego i głębokiego, przetwarzania języka naturalnego, a także analityki predykcyjnej, stanowi istotną część nowoczesnych rozwiązań stosowanych w takich obszarach, jak przeciwdziałanie praniu pieniędzy, wykrywanie transakcji oszukańczych (fraudowych), raportowanie nadzorcze, bezpieczeństwo ICT czy ocena zdolności kredytowej[16], ale także w relacji z klientami, zarówno w kontekście personalizacji produktów, jak i obsługi bezpośredniej, np. chatboty.

Warto podkreślić, że w jednym z badań przeprowadzonych w 2021 r.[17] wykazano, że w Polsce nadal niewiele firm zajmuje się rozwojem sztucznej inteligencji (102), a firm z obszaru Fintech-AI jest jeszcze mniej – zaledwie 11. Dane obejmują 2020 r.

Według niektórych szacunków, do 2023 r. potencjalne oszczędności dla globalnego sektora bankowego związane ze stosowaniem sztucznej inteligencji mogą osiągnąć blisko 450 mld dolarów[18], przy czy realne oszczędności w dużej mierze mogą wymagać istotnych zmian po stronie banków, również w zakresie wewnętrznej organizacji oraz procesów, a także kultury organizacyjnej[19]. Jeżeli dodatkowo uwzględnić potencjalne oszczędności związane z uproszczeniem obowiązków raportowania, które są przewidywane przez EUNB[20], to korzyści dla sektora bankowego związane z ograniczeniem kosztów operacyjnych mogą być znaczące.

Zasadniczo wykorzystanie sztucznej inteligencji w sektorze bankowym można podzielić na trzy główne obszary, tj.

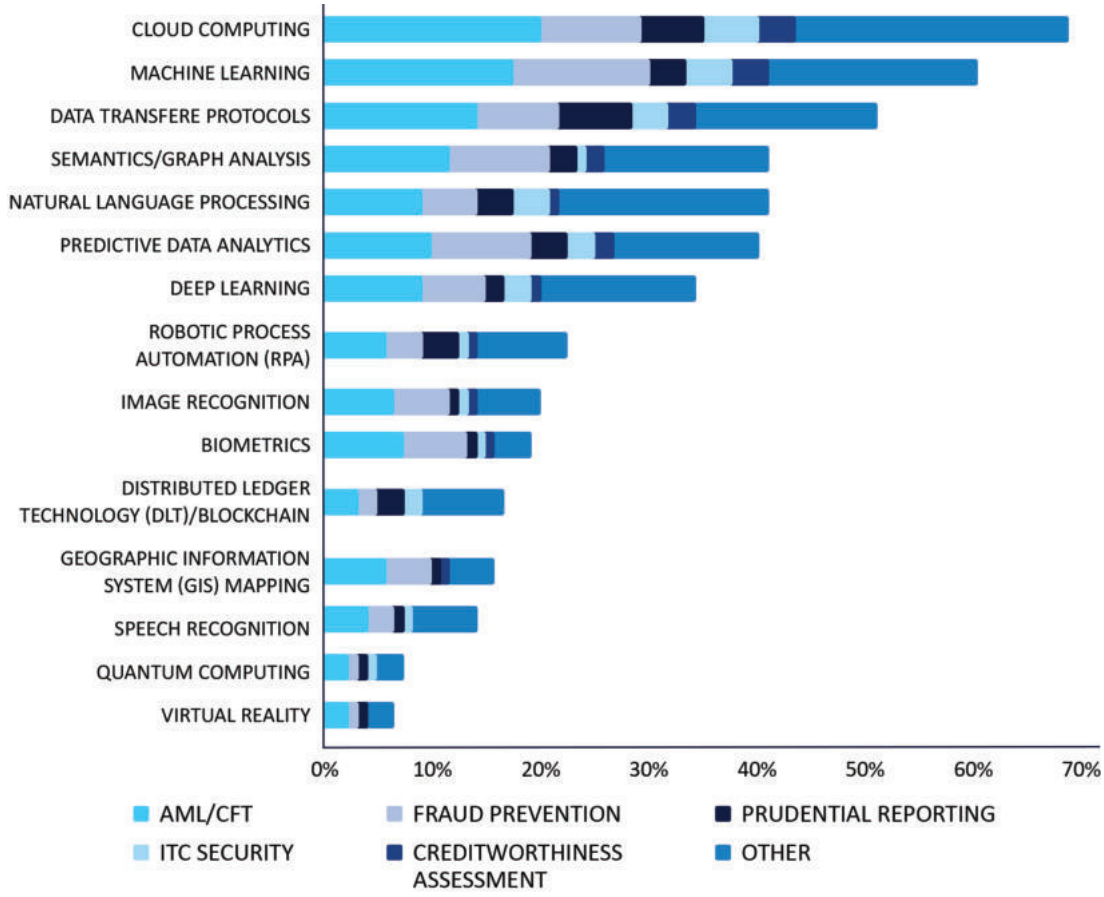
1. Relacje banków z ich klientami, zarówno detalicznymi, jak i korporacyjnymi.

2. Relacje banków z innymi bankami, jak również niebankowymi dostawcami usług płatniczych i szerzej – finansowych, np. w zakresie relacji korespondenckich.

3. Zastosowania wewnętrzne niebędące działalnością regulowaną banków.

W praktyce można także wyodrębnić dodatkową kategorię, tj. relacje pomiędzy bankami a organami regulacyjnymi oraz nadzorczymi, np. z wykorzystaniem tzw. SupTech (*Supervisory Technology*), jednakże obszar ten jest jeszcze bardzo słabo rozwinięty[21], w szczególności w Polsce, co jestem efektem niedostatecznego rozwoju technologii wspierających, w tym interfejsów dostępowych – API (*Application Programming Interfaces*) oraz braku standaryzacji w zakresie oczekiwań nadzorczych[22].

[13] https://www.eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf, s. 13 (dostęp: 21.03.2022 r.).
[14] https://alebank.pl/wp-content/uploads/2020/06/Raport-SZTUCZNA-INTELIGENCJA.pdf (dostęp: 22.03.2022 r.).
[15] Financial Stability Board, *FinTech and Market Structure in the Covid-19 Pandemic. Implications for financial stability*, 21 March 2022, s. 4.
[16] European Banking Authority, *Analysis of RegTech in the EU Financial Sector*, June 2021, EBA/REP/2021/17, s. 20.
[17] B. G. Buchanan, D. Wright, *The impact of machine learning on UK financial services*, Oxford Review of Economic Policy, Volume 37, Number 3, 2021, s. 542.
[18] V. Mahalakshmi, N. Kulkarni K.V. Pradeep Kumar et. al., *The Role of implementing Artificial Intelligence and Machine Learning Technologies in the financial services Industry for creating Competitive Intelligence*, Elsevier 2021, s. 2.
[19] L. Kruse, N. Wunderlich, R. Beck, *Artificial Intelligence for the Financial Services Industry: What Challenges Organizations to Succeed*, Proceedings of the 52nd Hawaii International Conference on System Sciences, 2019, s. 6409.
[20] European Banking Authority, *Study of the cost of compliance with supervisory reporting requirements*. *Report EBA/Rep/2021/15*, July 2021.
[21] Financial Stability Board, *The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions: Market developments and financial stability implications*, 9 October 2020.
[22] Warto jednocześnie, że Urząd Komisji Nadzoru podejmuje pewne działania – będące konsekwencją realizacji Cyfrowej Agendy Nadzoru – w zakresie stworzenia warunków dla rozwoju SupTech: https://us.edu.pl/porozumienie-o-wspolpracy-urzedu-knf-z-uczelniami-ze-slaska/ (dostęp: 22.03.2022 r.).



Rysunek 1. Wykorzystanie wybranych nowoczesnych technologii z podziałem na konkretne segmenty[23].

W jednym z opracowań[24] Organizacji Współpracy Gospodarczej i Rozwoju (OECD) zaproponowano listę przykładowych zastosowań tzw. sztucznej inteligencji w sektorze finansowym, przy czym nie jest to lista wyczerpująca wszystkie dostępne kierunki rozwoju. Lista ta zawiera następujące przykłady:

1. Obszary wspierania tzw. *back office*:
   a. Procesowanie potransakcyjne
   b. Analiza zysków i strat, rekoncyliacja
   c. Raportowanie i zarządzanie dokumentacją
   d. Analityka danych
   e. Ocena punktowa
   f. Infrastruktura i rozwiązania IT
2. Obszary wspierania tzw. *middle office*:
   a. Zarządzanie ryzykiem
   b. Procesy KYC (*Know-Your-Customer*)
   c. Obszar zgodności (*compliance*)
   d. Funkcje kontrolne i procesy
   e. Przeciwdziałanie praniu pieniędzy i finansowaniu terroryzmu
   f. Przeciwdziałanie oszustwom
3. Obszary wspierania tzw. *front office*:
   a. Alokacja aktywów
   b. Robo-doradztwo, chatboty
   c. Uwierzytelnienia z użyciem biometrii.
   d. Realizacja transakcji
   e. Spersonalizowane rekomendacje
   f. Usługi dla klientów

[23] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Reports/2021/1015484/EBA%20analysis%20of%20RegTech%20in%20the%20EU%20financial%20sector.pdf (dostęp: 04.06.2022 r.).
[24] https://www.oecd.org/competition/ai-in-business-and-finance.htm (dostęp: 29.03.2022 r.).

Dodatkowo OECD wyróżnia tutaj cztery ogólne kategorie:

1. Zarządzanie aktywami
2. Handel algorytmiczny
3. Pośrednictwo kredytowe
4. Finanse oparte o technologię łańcucha bloków

Zauważyć należy, że klasyfikacja do jednej z kategorii, tj. wsparcia procesów back, middle oraz front office jest płynna, bowiem systemy sztucznej inteligencji mogą wspierać różne etapy realizacji usług czy procesów.

Jednocześnie rozwiązania oparte o tzw. sztuczną inteligencję mogą w tych relacjach generować zróżnicowane ryzyka, które mogą mieć wpływ zarówno na bank, jak i jego klientów oraz kontrahentów. Część z tych ryzyk może być trudna do uchwycenia, bowiem ramy prawno-regulacyjne dla systemów sztucznej inteligencji, w tym w sektorze finansowym, są jeszcze na wczesnym etapie projektowania, co utrudnia jednoznaczną klasyfikację jako działań potencjalnie generujących np. ryzyko braku zgodności.

3.2. Zastosowanie sztucznej inteligencji w relacji bank-klient

Na wstępie należy zaznaczyć, że niektóre zastosowania systemów sztucznej inteligencji w relacji pomiędzy bankiem a klientem mogą mieć charakter zarówno bezpośredni, jak i pośredni. Przykładowo modele uczenia maszynowego stosowane na potrzeby oceny zdolności kredytowej (w tym zautomatyzowane rozwiązania, o których mowa w art. 105a ust. 1 Prawa bankowego) mogą mieć znaczenie zarówno dla „wyceny" ryzyka kredytowego po stronie instytucji finansowej, jak i samego klienta, np. gdy takie narzędzie jest niejako „wystawiane" klientowi w ramach interaktywnego interfejsu. Odrębną kategorią są też te rozwiązania, które niejako „towarzyszą" realizacji usług na rzecz klienta. Należą do nich m.in. systemy rekomendacyjne dla obszaru przeciwdziałania praniu pieniędzy i finansowania terroryzmu, systemy wykrywania transakcji potencjalnie nieautoryzowanych czy rozwiązania wspierające procesy uwierzytelniania użytkownika, np. wykorzystujące biometrię behawioralną.

Obecnie można wymienić następujące zastosowania szeroko rozumianej sztucznej inteligencji, które mają zastosowanie w relacji bank-klient, które zostaną omówione bardziej szczegółowo w dalszej części podrozdziału:

1. Chatboty oraz wirtualni asystenci, w tym rozwiązania wykorzystujące rozpoznawanie mowy
2. Aplikacje mobilne i desktopowe do zarządzania budżetem i pokrewne
3. Weryfikacja danych identyfikacyjnych użytkowników
4. Rozwiązania usprawniające procesy rozpatrywania reklamacji
5. Rozwiązania wspierające działania o charakterze marketingowym
6. Rozwiązania umożliwiające personalizację produktów oraz wydawanie rekomendacji produktowe
7. Zautomatyzowane systemy doradztwa, w szczególności inwestycyjnego (robo-advisory)
8. Systemy oceny zdolności kredytowej oraz ryzyka kredytowego

Należy przy tym podkreślić, że nie jest to lista wszystkich możliwych zastosowań sztucznej inteligencji, bowiem – przynajmniej teoretycznie – można wyobrazić sobie także zastosowania łączące rozwiązania oparte o tzw. internet rzeczy (*Internet of Things*), np. generujący wirtualną rzeczywistość oraz systemy sztucznej inteligencji.

Systemy sztucznej inteligencji, które znajdują zastosowanie w relacjach z klientami, zarówno osobami fizycznymi, jak i prawnymi, mogą generować znaczne ryzyka dla banków, a także powinny być projektowane z dołożeniem najwyższej staranności oraz poszanowaniem zasady *privacy and data protection by design* oraz *privacy and data protection by default,* czyli zasad, w myśl których ochrona prywatności oraz danych osobowych powinna być stosowana na każdym etapie projektowania jako obowiązkowy aspekt tworzenia nowych rozwiązań[25] oraz jako „ustawienie" domyślne w każdym procesie lub systemie. Systemy sztucznej inteligencji powinny także spełniać wymogi w zakresie przejrzystości (*transparency*) oraz wyjaśnialności (*explainability*), na co wskazują projektowane przepisy AIA, m.in.:

– art. 13, który stanowi w ust. 1: „*Systemy sztucznej inteligencji wysokiego ryzyka projektuje się i opracowuje się w sposób zapewniający wystarczającą przejrzystość ich działania, umożliwiającą użytkownikom interpretację wyników działania systemu i ich właściwe wykorzystanie. Zapewnia się*

[25] Zasadę tę przywołuje m.in. ENISA, *Good Practices for Security of IoT – Secure Software Development Lifecycle*, November 19, 2019.

odpowiedni rodzaj i stopień przejrzystości w celu osiągnięcia zgodności z odpowiednimi obowiązkami użytkownika i dostawcy, określonymi w rozdziale 3 niniejszego tytułu" oraz w ust. 2: *„Do systemów sztucznej inteligencji wysokiego ryzyka dołącza się instrukcję obsługi w odpowiednim formacie cyfrowym lub innym formacie zawierającą zwięzłe, kompletne, poprawne i jasne informacje, które są istotne, dostępne i zrozumiałe dla użytkowników"*.

– art. 52 ust. 1, zgodnie z którym *„[d]ostawcy zapewniają, aby systemy sztucznej inteligencji przeznaczone do wchodzenia w interakcję z osobami fizycznymi projektowano i opracowywano w taki sposób, aby osoby fizyczne były informowane o tym, że prowadzą interakcję z systemem sztucznej inteligencji, chyba że okoliczności i kontekst korzystania z systemu jednoznacznie na to wskazują"*.

Podobnie, stanowisko UKNF w sprawie świadczenia usługi robo-doradztwa[26] podkreśla znaczenie zasady przejrzystości względem klientów.

Nie bez znaczenia pozostają także przepisy prawa oraz regulacje, które mogą odnosić się do reklamowania i sprzedaży produktów oraz usług finansowych, np. o charakterze inwestycyjnym, które mogą wprowadzać specyficzne wymagania względem instytucji je stosujących, co może być niekiedy problematyczne w odniesieniu do zautomatyzowanych systemów sztucznej inteligencji, np. chatbotów o funkcji sprzedażowej. Z tego względu przy projektowaniu tego typu rozwiązań szczególnie istotny jest udział pracowników komórek ds. zgodności, którzy posiadają przynajmniej podstawową wiedzę w zakresie działania algorytmów i modeli sztucznej inteligencji, np. wzorem wymogów określonych w Rozporządzeniu 2017/589 w sprawie handlu algorytmicznego[27].

Poniżej zaprezentowane zostaną wybrane rozwiązania oparte o systemy sztucznej inteligencji, które są lub mogą być wykorzystywane w relacji bank-klient.

3.3. Chatboty oraz wirtualni asystenci

Chatboty są programami komputerowymi, które potrafią komunikować się „jak człowiek" z użyciem języka naturalnego[28] i wykorzystują różne techniki i podejścia z szeroko rozumianego obszaru sztucznej inteligencji, w tym przede wszystkim uczenie maszynowe i głębokie oraz przetwarzanie języka naturalnego. Bardziej zaawansowane rozwiązania mogą także wykorzystywać analizę obrazu oraz sentymentu, a także – w szczególnych przypadkach – doradzać użytkownikom, ze zrozumieniem kontekstu wypowiedzi. Mają one zróżnicowane zastosowanie, a jednym z celów ich wykorzystania – poza obniżeniem kosztów działalności operacyjnej – jest poprawa doświadczeń użytkowników (klientów) banków, którzy dzięki takim narzędziom są w stanie szybciej dotrzeć do niezbędnych informacji. Implementacja chatbotów w sektorze finansowym może wiązać się jednak z konkretnymi ograniczeniami oraz wyzwaniami[29], co jest pochodną faktu, że jest to sektor regulowany poddany szczególnym wymogom, np. w zakresie tajemnicy bankowej czy sprzedaży produktów i usług.

Z tego względu wykorzystanie chatbotów oraz wirtualnych asystentów może wymagać nie tylko szczegółowej analizy pod kątem ochrony danych osobowych oraz ryzyk ICT i cyberbezpieczeństwa (w szczególności, jeżeli za pomocą takiego kanału komunikacji przekazywane są informacje wrażliwe), ale także szczegółowych wymogów i ograniczeń związanych z marketingiem oraz dystrybucją produktów i usług finansowych.

Przykładowo, Rozporządzenie Ministra Finansów z dnia 30 maja 2018 r. w sprawie trybu i warunków postępowania firm inwestycyjnych, banków, o których mowa w art. 70 ust. 2 ustawy o obrocie instrumentami finansowymi, oraz banków powierniczych[30] oraz Rozporządzenie delegowane Komisji (UE) 2017/565 z dnia 25 kwietnia 2016 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2014/65/UE w odniesieniu do wymogów organizacyjnych i warunków prowadzenia działalności przez firmy inwestycyjne oraz pojęć zdefiniowanych na potrzeby tej dyrektywy[31] określają pewne specyficzne wymogi w zakresie reklamy i sprzedaży wybranych produktów inwestycyjnych, w tym także wymogi względem osób odpowiedzialnych za takie działania. Wydaje się, że – pomimo braku wyraźnego odniesienia w tych

[26] https://www.knf.gov.pl/knf/pl/komponenty/img/Stanowisko_UKNF_ws_swiadczenia_uslugi_robo_doradztwa_71303.pdf (dostęp: 23.03.2022 r.).
[27] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2017/589 z dnia 19 lipca 2016 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2014/65/UE w odniesieniu do regulacyjnych standardów technicznych określających wymogi organizacyjne dla firm inwestycyjnych prowadzących handel algorytmiczny, Dz. Urz. UE z 2017 r., L-87/417.
[28] J. Ortiz, *Chatbots and Finance*, June 15, 2020, s. 5. Opracowanie dostępne pod adresem: https://escholarship.org/uc/item/0dw3h6h1 (dostęp: 23.03.2022 r.).
[29] M. Sugumar, S. Chandra, *Do I Desire Chatbots to be like Humans? Exploring Factors for Adoption of Chatbots for Financial Services*, Journal of International Technology and Information Management, Volume 30, Issue 3, 2021, s. 40.
[30] Dz. U. z 2018 r., poz. 1112.
[31] Dz. Urz. UE z 2017 r., L-87/1.

przepisach – wymogi te należy (odpowiednio) stosować także do rozwiązań wykorzystujących systemy sztucznej inteligencji, także w zakresie nadzoru nad nimi, który powinien być sprawowany przez wykwalifikowaną kadrę, m.in. z obszaru zgodności z prawem (*compliance*). Jednocześnie podkreślić należy, że obecnie w niewielkim stopniu widać zainteresowanie banków stosowaniem bardziej rozbudowanych rozwiązań, np. o funkcjonalności sprzedażowej, a znaczna część chatbotów i wirtualnych asystentów jest „delegowana" do doradzania przy mniej skomplikowanych sprawach, w tym do przekazywania dokumentów informacyjnych czy odnośników do odpowiednich sekcji na stronie internetowej banku czy w aplikacji.

Odrębnym zagadnieniem, któremu poświęcony będzie także oddzielny rozdział, jest kwestia przejrzystości w odniesieniu do tych systemów. Pozostaje pytaniem otwartym – przynajmniej do czasu uchwalenia stosownych rozwiązań prawnych – na ile i w jakim zakresie chatboty oraz wirtualni asystenci (w tym te, które wykorzystują metody głosowe) powinny spełniać wymogi w zakresie przejrzystości rozumianej jako obowiązek informowania o fakcie, że kontakt klienta odbywa się z udziałem „robota", a nie człowieka. Choć brak w tym zakresie jasnych i wiążących wytycznych[32], to przyjąć należy, że takie działanie jest pożądane i rekomendowane, w szczególności, jeżeli rozwiązania te nie podlegają szczególnemu nadzorowi ze strony pracowników banku.

Należy w tym miejscu także wskazać, że nie każdy chatbot oraz wirtualny asystent musi wykorzystywać bardziej zaawansowane podejścia i techniki jak np. uczenie maszynowe czy głębokie, co może determinować konieczność zastosowania (lub nie) specyficznych wymagań dla tzw. sztucznej inteligencji. Z tego względu należy dokonać weryfikacji rozwiązania również pod kątem wymagań z szeroko rozumianego obszaru *compliance*, w szczególności pod kątem nadzoru i kontroli działalności tego typu rozwiązań ze wskazanymi wyżej wymogami prawnymi i regulacyjnymi.

Na marginesie należy wskazać, że modele sztucznej inteligencji, które wykorzystują techniki samouczenia, jak np. uczenie maszynowe i głębokie, podatne są na zagrożenia związane z wpływaniem przez ich użytkowników na ich „zachowania", poprzez stosowanie wyrażeń obraźliwych czy krzywdzących, które – jeżeli nie zostaną w odpowiedni sposób wyeliminowane – mogą stanowić dla banku istotne ryzyko reputacyjne.

[32] Wyjątek stanowią tutaj wytyczne ekspertów Komisji Europejskiej w sprawie sztucznej inteligencji godnej zaufania: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60436 (dostęp: 27.06.2022 r.).

Na marginesie można wskazać, że Europejska Rada Ochrony Danych wydała w 2021 r. Wytyczne 02/2021 w sprawie wirtualnych asystentów głosowych[33], które mają istotne znaczenie dla rozwiązań opisanych w niniejszym podrozdziale, gdyż odnoszą się przede wszystkim do wykorzystywania takich rozwiązań, jak przetwarzanie języka naturalnego (*natural languange processing*) oraz rozpoznawanie głosu (*voice recognition*).

Należy także zwrócić uwagę na art. 52 projektowanego rozporządzenia w sprawie sztucznej inteligencji, który wskazuje, że dostawcy zapewniają, aby systemy sztucznej inteligencji przeznaczone do wchodzenia w interakcję z osobami fizycznymi projektowano i opracowywano w taki sposób, aby osoby fizyczne były informowane o tym, że prowadzą interakcję z systemem sztucznej inteligencji, chyba że okoliczności i kontekst korzystania z systemu jednoznacznie na to wskazują. Wirtualni asystenci i chatboty niewątpliwie mogą być zaliczone do tej kategorii, a więc również banki będą obowiązane do informowania, np. w formie stosownych komunikatów głosowych, że rozmowa odbywać się będzie bez udziału człowieka.

3.4. Aplikacje mobilne do zarządzania budżetem

Ten obszar działalności banków jest ściśle powiązany z rozwojem szeroko rozumianej otwartej bankowości, która została wprowadzona do porządku prawnego w Unii Europejskiej poprzez pakiet PSD2, do którego należą dyrektywa[34] oraz rozporządzenie delegowane Komisji (UE)[35], a w Polsce ustawa o usługach płatniczych[36]. Wspomniany pakiet wprowadził dla dostawców usług płatniczych możliwość świadczenia usługi dostępu do informacji o rachunku płatniczym, która w dużym uproszczeniu polega na pozyskiwaniu zagregowanych informacji z takich rachunków (np. z innych banków) i prezentowaniu ich użytkownikowi, który wyraził na to zgodę. Obecnie na poziomie Unii Europejskiej w ramach realizacji

[33] https://edpb.europa.eu/system/files/2022-02/edpb_guidelines_202102_on_vva_v2.0_adopted_pl.pdf (dostęp: 12.05.2022 r.).
[34] Dyrektywa Parlamentu Europejskiego i Rady (UE) 2015/2366 z dnia 25 listopada 2015 r. w sprawie usług płatniczych w ramach rynku wewnętrznego, zmieniająca dyrektywy 2002/65/WE, 2009/110/WE, 2013/36/UE i rozporządzenie (UE) nr 1093/2010 oraz uchylająca dyrektywę 2007/64/WE, Dz. Urz. UE z 2015 r., L 337/35.
[35] Rozporządzenie delegowane Komisji (UE) 2018/389 z dnia 27 listopada 2017 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady (UE) 2015/2366 w odniesieniu do regulacyjnych standardów technicznych dotyczących silnego uwierzytelniania klienta i wspólnych i bezpiecznych otwartych standardów komunikacji, Dz. Urz. UE z 2017 r., L 69/23.
[36] Ustawa z dnia 19 sierpnia 2011 r. o usługach płatniczych, Dz. U. 2011, Nr 199 poz. 1175 (z późn. zm.).

Strategii dla UE w sprawie cyfrowych finansów[37] rozważana jest możliwość wprowadzenia szerszej koncepcji – otwartych finansów – która ma stanowić kolejny krok w możliwości tworzenia bardziej spersonalizowanych i dostosowanych do potrzeb klientów rozwiązań pozwalających na agregację szeregu informacji nie tylko z rachunku płatniczego, ale np. rachunku inwestycyjnego.

Wspomniana usługa dostępu do informacji o rachunku płatniczym otworzyła szereg możliwości w zakresie nowych produktów oferowanych zarówno klientom detalicznym, jak i korporacyjnym, w szczególności ułatwiających zarządzanie budżetem czy posiadanymi środkami zgromadzonymi na wielu rachunkach w różnych instytucjach. Wartość dodana takich rozwiązań jest widoczna dla użytkownika jedynie wtedy, gdy wykracza poza wyłącznie prezentację salda rachunku, a więc np. z pozyskanych informacji transakcyjnych generowane są informacje nt. rzeczywistego stanu finansów, w tym wydatków, a nawet rekomendacji, które umożliwią optymalizację budżetu. Tworzenie tego typu prezentacji i rekomendacji wymaga jednak zastosowania bardziej zaawansowanych narzędzi analitycznych, w szczególności wykorzystujących uczenie maszynowe. Takie rozwiązania często są kwalifikowane jako aplikacje do zarządzania budżetem/stanem finansów.

W tym kontekście pojawiają się więc wyzwania oraz ryzyka nie tylko w kontekście bezpieczeństwa (w tym zabezpieczenia indywidualnych danych uwierzytelniających i danych wrażliwych), ale także związane ze stosowaniem wspomnianych rozwiązań opartych o szeroko rozumianą sztuczną inteligencję. Wykorzystanie danych wrażliwych, jak dane transakcyjne, może prowadzić w pewnych sytuacjach do wygenerowania danych o charakterze danych osobowych podlegających szczególnej ochronie. Już w wytycznych Europejskiej Rady Ochrony Danych w zakresie zależności pomiędzy pakietem PSD2 a Rozporządzeniem 2016/679[38] wskazano, że pewne dane pochodzące z transakcji mogą posłużyć do identyfikacji pewnych cech, jak np. przynależność do grupy religijnej, preferencji seksualnych czy politycznych i innych danych, które objęte są zakazem przetwarzania określonym w art. 9 ust. 1 Rozporządzenia 2016/679 (RODO). Oznacza to, że ich przetwarzanie również z użyciem uczenia maszynowego może podlegać warunkom określonym w art. 9 RODO dla danych szczególnych kategorii.

O ile nie jest to ryzyko inherentnie związane ze stosowaniem systemów sztucznej inteligencji, to pozyskiwanie tych danych i ich przetwarzanie np. na potrzeby trenowania modeli może być źródłem konkretnych zagrożeń, bowiem dane te – wykorzystane w nieodpowiedni sposób, w tym jako element Big Data – mogą być źródłem algorytmicznego *bias'u,* czyli skrzywienia (przechyłu) algorytmicznego, który może skutkować dyskryminacją określonych grup. Z tego względu niezwykle istotne jest, aby zapewnić, że dane te – o ile w ogóle są wykorzystywane np. do trenowania modeli sztucznej inteligencji – podlegały szczególnemu nadzorowi i kontroli. W tym zakresie również projektowane rozporządzenie w sprawie sztucznej inteligencji w art. 10 wskazuje na konkretne wymagania względem danych wykorzystywanych w modelach AI, np. w kontekście jakości, dokładności etc. Przedmiotowy przepis będzie wymagał również stosowania odpowiednich rozwiązań w zakresie trenowania, walidacji i testowania modeli[39].

W kontekście stosowania tych rozwiązań mogą również pojawić się wyzwania związane z koniecznością wyjaśnienia działania algorytmu – choć na dzień sporządzania dokumentu nie funkcjonowały ramy prawne nakazujące ujawnienie takich danych, to obowiązek wykazania podstaw, np. kategoryzacji czy rekomendacji w ramach udostępnionej aplikacji, powinien być elementem szeroko pojętej przejrzystości, również w kontekście profilowania i zautomatyzowanego przetwarzania danych, o których mowa w art. 22 Rozporządzenia 2016/679 (o ile dotyczy).

W pozostałym zakresie stosowanie rozwiązań opartych o uczenie maszynowe i techniki pokrewne w ramach usługi dostępu do informacji o rachunku powinno podlegać „klasycznym" wymogom prawnym i regulacyjnym dla świadczenia usług płatniczych i przetwarzania danych, w tym także Rekomendacji D KNF[40].

[37] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0591 (dostęp: 27.03.2022 r.).

[38] https://edpb.europa.eu/system/files/2021-06/edpb_guidelines_202006_psd2_afterpublicconsultation_pl.pdf (dostęp: 27.03.2022 r.).

[39] Na marginesie warto zauważyć, że projekt AIA w art. 10 ust. 5 przewiduje nową podstawę prawną do przetwarzania danych osobowych szczególnych kategorii, wskazując, iż *„w zakresie, w jakim jest to ściśle niezbędne do celów zapewnienia monitorowania, wykrywania i korygowania tendencyjności systemów sztucznej inteligencji wysokiego ryzyka, dostawcy takich systemów mogą przetwarzać szczególne kategorie danych osobowych, o których mowa w art. 9 ust. 1 rozporządzenia (UE) 2016/679, art. 10 dyrektywy (UE) 2016/680 i art. 10 ust. 1 rozporządzenia (UE) 2018/1725, pod warunkiem stosowania odpowiednich zabezpieczeń gwarantujących ochronę podstawowych praw i wolności osób fizycznych, w tym środków technicznych ograniczających ponowne wykorzystanie tych danych i najnowocześniejszych środków służących zapewnieniu bezpieczeństwa i ochrony prywatności, takich jak pseudonimizacja lub – w przypadku gdy anonimizacja może znacząco wpłynąć na możliwość realizacji zakładanego celu – szyfrowanie".*

[40] https://www.knf.gov.pl/knf/pl/komponenty/img/Rekomendacja_D_8_01_13_uchwala_7_33016.pdf (dostęp: 27.03.2022 r.).

3.5. Weryfikacja danych identyfikacyjnych użytkowników

Systemy sztucznej inteligencji połączone z pozyskiwaniem i przetwarzaniem danych biometrycznych mogą znacznie usprawnić i zabezpieczyć procesy uwierzytelniania użytkowników korzystających z kanałów cyfrowych. Znaczna część analiz wykazuje, że stosowanie rozwiązań opartych o biometrię, także behawioralną, przyczynia się do znacznej poprawy bezpieczeństwa i z powodzeniem mogą one zastępować dotąd najbardziej rozpowszechnione metody ochrony danych[41] czy środków finansowych, jak np. hasła czy kody PIN[42]. Zastosowanie biometrii wymaga jednak stosowania też równolegle odpowiednich rozwiązań wykorzystujących szeroko rozumianą sztuczną inteligencję, która umożliwia efektywne i natychmiastowe potwierdzenie określonych cech użytkownika. Z tego względu zagadnienie wykorzystania biometrii w bankowości nie powinno być analizowane w oderwaniu od kwestii wykorzystania sztucznej inteligencji, tym bardziej, że połączenie tych dwóch obszarów może znacznie przyczynić się do poprawy bezpieczeństwa[43]. Choć nie jest to przedmiotem niniejszego opracowania, to warto w tym miejscu zwrócić uwagę, że na poziomie Unii Europejskiej prowadzone są obecnie prace legislacyjne zmierzające do stworzenia uniwersalnej metody potwierdzania tożsamości w ramach tzw. Europejskich Portfeli Identyfikacji Cyfrowej[44].

Wykorzystanie biometrii może mieć zastosowanie przykładowo w systemach dokonujących uwierzytelnienia użytkownika, w szczególności z użyciem tzw. silnego uwierzytelniania użytkownika (*Strong Customer Authentication* – SCA), które zostało „wprowadzone" w ramach pakietu PSD2. Cechy biometryczne mogą stanowić jeden z elementów SCA pod warunkiem spełnienia wymogów określonych w Rozporządzeniu 2018/389. Jednocześnie wspomniane rozporządzenie dopuszcza możliwość stosowania licznych wyłączeń, np. w oparciu o niski poziom oszustw związanych z realizowanymi transakcjami płatniczymi. Systemy sztucznej inteligencji mogą więc – w czasie rzeczywistym – dokonywać analizy określonych cech biometrycznych w celu dokonania uwierzytelnienia użytkownika. Sposób, w jaki dokonywana jest wspomniana analiza, jest uzależniony od dostępności samych danych biometrycznych, bowiem w znaczącej większości przypadków banki nie wchodzą w „posiadanie" wzorca biometrycznego, a całość analizy dokonywana jest na urządzeniu użytkownika, zaś bank otrzymuje jedynie informację o poprawności porównania pierwotnego wzorca z konkretnymi działaniem.

W przypadku, gdy to bank dokonuje jednak analizy cech biometrycznych, np. w kontekście analizy transakcji potencjalnie nieautoryzowanych, to na banku ciążyć będą obowiązki związane ze stosowaniem tego typu rozwiązań, w tym w szczególności w kontekście Rozporządzenia 2016/679 i ograniczeń związanych z przetwarzaniem danych wrażliwych, o czym mowa w art. 9 RODO. Stawia to przed bankiem liczne wyzwania, w szczególności, jeżeli przedmiotem analizy są dane zaliczane do kategorii biometrii behawioralnej.

Jest to o tyle istotne, że dane wykorzystywane przez systemy samouczące się dla danego użytkownika mogą być wykorzystywane jako dane do trenowania algorytmów i modeli sztucznej inteligencji. Jeżeli jakość tych danych nie będzie na odpowiednio wysokim poziomie, to istnieje ryzyko, że również docelowy model nie będzie efektywny i skuteczny. Podobną uwagę należy poczynić w kontekście wykorzystania tych danych – pod warunkiem wyrażenia zgody przez użytkownika – na potrzeby trenowania modeli stosowanych w innych obszarach niż wskazane powyżej.

W kontekście wykorzystania biometrii można więc wskazać na szereg ryzyk związanych zarówno z samym faktem wykorzystania danych wrażliwych, jak i ryzyk charakterystycznych dla stosowania systemów sztucznej inteligencji, przy czym należy zwrócić uwagę, że obecnie proponowane zmiany do AIA zmierzają w niektórych miejscach do całkowitego zakazu stosowania identyfikacji biometrycznej. Wydaje się jednak, że są to propozycje, które nie mają szansy się ziścić[45].

Należy tutaj podkreślić, że obszar identyfikacji z użyciem danych biometrycznych jest jednym z obszarów, któremu poświęcono wiele miejsca w projektowanym rozporządzeniu w sprawie sztucznej inteligencji, poczynając od samej definicji, poprzez praktyki zakazane, jak i wymagania względem systemów sztucznej inteligencji wysokiego ryzyka.

[41] A. Botana Lopez, *Deep Learning in Biometrics: A Survey*, Advances in Distributed Computing and Artificial Intelligence Journal, vol. 8, no. 4, 2019, s. 31.
[42] J. S. Wang, *Exploring biometric identification in FinTech applications based on the modified TAM*, Financial Innovation 7:42. 2021, s. 2.
[43] B. Purgason, D. Hibler, *Security Through Behavioral Biometrics and Artificial Intelligence*, Procedia Computer Science 12, 2012, s. 403.
[44] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0281 (dostęp: 29.03.2022 r.).
[45] EROD i EIOD wzywają do wprowadzenia zakazu wykorzystywania sztucznej inteligencji do automatycznego rozpoznawania cech ludzkich w przestrzeniach dostępnych publicznie oraz niektórych innych zastosowań sztucznej inteligencji, które mogą prowadzić do niesprawiedliwej dyskryminacji: https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_pl (dostęp: 20.06.2022 r.).

Rozważania w tym zakresie należy rozpocząć od próby zdefiniowania samych danych biometrycznych. Definicja taka została przewidziana art. 3 pkt 33) projektowanego rozporządzenia, zgodnie z którym są to dane osobowe będące wynikiem specjalnego przetwarzania technicznego, które dotyczą cech fizycznych, fizjologicznych lub behawioralnych osoby fizycznej oraz umożliwiają lub potwierdzają jednoznaczną identyfikację tej osoby, takie jak wizerunek twarzy lub dane daktyloskopijne. Takie dane mogą być też wykorzystywane do kategoryzacji biometrycznej, którą zdefiniowano jako (art. 3 pkt 35) systemy sztucznej inteligencji służące do przypisywania osób fizycznych do określonych kategorii, takich jak płeć, wiek, kolor włosów, kolor oczu, tatuaże, pochodzenie etniczne lub orientacja seksualna bądź polityczna, na podstawie ich danych biometrycznych. Takie systemy mogą być wykorzystywane także w systemach bankowych, np. na etapie wideoweryfikacji klienta.

Jednocześnie art. 3 pkt 36-38 projektu AIA wprowadzają kolejne pojęcia dotyczące systemów zdalnej identyfikacji biometrycznej (w czasie rzeczywistym i post factum). Takim systemem jest system sztucznej inteligencji służący do identyfikacji osób fizycznych na odległość poprzez porównanie danych biometrycznych danej osoby z danymi biometrycznymi zawartymi w referencyjnej bazie danych, bez uprzedniej wiedzy użytkownika systemu sztucznej inteligencji, czy dana osoba będzie w nimi figurować i czy może zostać zidentyfikowana. Kwalifikacja określonego rozwiązania może mieć doniosłe konsekwencje na gruncie projektowanego rozporządzenia w sprawie sztucznej inteligencji, bowiem w załączniku III wskazano, że systemy przeznaczone w celu zdalnej identyfikacji biometrycznej osób fizycznych w czasie rzeczywistym i post factum będą mogły być traktowane jako systemy sztucznej inteligencji wysokiego ryzyka.

Na dzień sporządzania niniejszego raportu nie było do końca jasne czy systemy stosowane np. w celu dokonania wideoweryfikacji klienta będą traktowane jako takie systemy, jednakże takiej możliwości nie można wykluczyć.

3.6. Procesy rozpatrywania reklamacji

Jednym z obszarów, który stosunkowo łatwo poddaje się automatyzacji w sektorze finansowym, jest proces rozpatrywania (prostszych) reklamacji składanych przez klientów banków. W znacznej większości przypadków reklamacje składane przez klientów dotyczą nieskomplikowanych spraw obejmujących takie zagadnienia, jak nieautoryzowane transakcje czy nieuprawniony dostęp do konta, choć oczywiście zdarzają się też przypadki bardziej skomplikowanych zagadnień.

W odniesieniu do tematyki nieautoryzowanych transakcji można wyodrębnić pewne elementy, których spełnienie lub niespełnienie będzie warunkowało (niezasadność) określonego roszczenia. Innymi słowy są to przesłanki, które można łatwo skwantyfikować, bazując na obowiązujących przepisach prawa (*vide* ustawa o usługach płatniczych). Pozwala to więc na przynajmniej częściową automatyzację procesów reklamacyjnych, które stanowią kosztowe obciążenie operacyjne dla banków.

W celu usprawnienia i zautomatyzowania takich procesów można wykorzystywać zarówno wskazane już powyżej chatboty oraz wirtualnych asystentów, którzy przeprowadzają przez proces reklamacyjny, w tym przyjmują zgłoszenie, jak również implementować rozwiązania oparte o uczenie maszynowe i przetwarzanie języka naturalnego w celu dokonania analizy zgłoszenia reklamacyjnego i ewentualnego przygotowania odpowiedzi lub skierowania go do procesu manualnego sprawowanego przez człowieka. Zaawansowane rozwiązania wykorzystujące wspomniane wyżej techniki stają się coraz bardziej efektywne i są w stanie analizować również kontekst określonej wypowiedzi czy tekstu z powodzeniem zastępując człowieka przy mniej skomplikowanych przypadkach. Jednocześnie nie można zapominać, że rekomendowanym rozwiązaniem przy interakcji z użytkownikami końcowymi jest udział człowieka (analityka) na każdym etapie działania systemu sztucznej inteligencji.

Wykorzystanie tego typu rozwiązań dla rozpatrywania reklamacji wymaga jednak aktywnego działania ze strony na banku, zarówno na etapie koncepcyjnym, testowania, jak i wdrożenia, bowiem trenowanie wykorzystywanych rozwiązań wymaga m.in. weryfikacji treści i pozyskiwania tzw. *feedbacku*. Nawet częściowa automatyzacja procesu wymagać może także istotnych zmian organizacyjnych i proceduralnych, a także ustalenia procesu nadzoru w ramach kontroli wewnętrznej. Podobnie jak w przypadku innych rozwiązań komunikujących się z użytkownikiem końcowym, niezbędne jest zapewnienie odpowiedniego poziomu przejrzystości, aby użytkownik ten miał świadomość w jaki sposób realizowane są jego prawa związane z procesem reklamacji. Jednocześnie bank może rozważyć proces półautomatyczny, tj. system rekomendacyjny dla osoby odpowiedzialnej za udzielanie odpowiedzi na reklamacje, natomiast sama decyzja co do kierunku pozostaje w gestii tej osoby.

Pamiętać jednocześnie należy, że rozwiązania automatyczne podatne są nie tylko na błędy niewłaściwej

interpretacji zgłoszenia, ale także ewentualnego skrzywienia, które może być skutkiem zastosowania danych treningowych o niskiej jakości. W konsekwencji system odpowiedzialny za rozpatrywanie reklamacji może narazić bank na ryzyko zarówno prawne, jak i regulacyjne.

3.7. Działania o charakterze marketingowym

Działalność marketingowa banków podlega szczególnym wymogom prawnym i regulacyjnym, a także etycznym. Jednocześnie jest to obszar, który łatwo poddaje się automatyzacji z wykorzystaniem systemów sztucznej inteligencji. Jednym z obszarów jest także personalizacja produktów i usług, także z perspektywy cenowej, co jednocześnie może być przedmiotem tzw. dyskryminacji cenowej, która w przypadku wykorzystania AI może występować znacznie częściej.

Banki znajdują się w posiadaniu znacznej ilości danych (finansowych i niefinansowych) dotyczących obecnych, jak i potencjalnych klientów, co może stanowić źródło znacznej przewagi konkurencyjnej, jeżeli dane te zostaną zastosowane w odpowiedni sposób. Może to być jednocześnie źródło rozlicznych ryzyk związanych z wykorzystaniem danych, w tym biometrii behawioralnej i manipulacji, a także naruszeniem „tradycyjnych" przepisów w zakresie sprzedaży produktów.

Dane dotyczące oczekiwań, potrzeb (także tych nieuświadomionych) oraz możliwości klientów połączone ze stosowaniem systemów sztucznej inteligencji oraz Big Data, pozwalają na dostosowywanie oferty do poszczególnych grup klientów lub tworzenia spersonalizowanych treści oraz ofert. Dane sprzedażowe mogą z kolei być wykorzystywane do przewidywania określonych zachowań klientów i prawdopodobnego wzrostu lub obniżenia sprzedaży poszczególnych produktów i usług finansowych[46]. Pozwala to na tworzenie bardziej efektywnych kampanii marketingowych i uzyskiwanie lepszych (bardziej optymalnych) wyników sprzedażowych. Praktycznych zastosowań systemów sztucznej inteligencji w marketingu jest jednak znacznie więcej, na co wskazuje m.in. E. Hermann:

1. Zbieranie danych marketingowych, segmentacja i standaryzacja,

2. Analiza rynkowa i personalizacja,

3. Lepsze zrozumienie klienta, pozycjonowanie oraz egzekucja strategii[47].

Podobnie jak w przypadku pozostałych zastosowań systemów sztucznej inteligencji, w których dużą rolę odgrywają dane (osobowe), również w przypadku wykorzystania AI dla działań marketingowych banków istotne będą więc odpowiednie rozwiązania organizacyjno-techniczne zapewniające wysoki poziom prywatności oraz etyczne[48] działanie zarówno „ze strony" algorytmów i modeli sztucznej inteligencji, jak i osób odpowiedzialnych za działanie tych rozwiązań. W tym kontekście na bankach ciąży szczególna odpowiedzialność, a niewłaściwe zarządzanie procesami związanymi z taką formą działań marketingowych może skutkować nie tylko ryzykami prawnymi, w tym na gruncie przepisów o ochronie danych osobowych czy regulacyjnymi, ale także reputacyjnymi.

3.8. Personalizacja produktów i rekomendacje produktowe

Obszar ten można zasadniczo łączyć z podrozdziałem 3.1.5, bowiem są to niewątpliwie obszary ze sobą współistniejące. Zaawansowana analiza danych, w tym Big Data, pozwala na dokonywanie podziału klientów oraz rynków na segmenty czy obszary, a pozyskane w ten sposób klasyfikacje mogą posłużyć następnie do bardziej spersonalizowanego doboru produktów i usług pod te konkretne kategorie. Zakres danych pozyskiwanych od klientów jest szeroki i obejmuje nie tylko dane transakcyjne, ale także inne informacje pozyskiwane w trakcie interakcji z klientem[49], co dotyczy nie tylko klienta indywidualnego, ale także korporacyjnego[50]. Warto w tym miejscu zaznaczyć, że choć jednym z głównych motywów personalizacji produktów i usług finansowych zazwyczaj będzie maksymalizacja

[46] R. Tiwari, S. Srivastava, R. Gera, *Investigation of Artificial Intelligence Techniques in Finance and Marketing*, Procedia Computer Science 173, 2020, s. 155.

[47] E. Hermann, *Leveraging Artifcial Intelligence in Marketing for Social Good—An Ethical Perspective*, Journal of Business Ethics, s. 3. Artykuł dostępny pod adresem: https://link.springer.com/content/pdf/10.1007/s10551-021-04843-y.pdf (dostęp: 01.04.2022 r.).

[48] P. K. Kopalle, M. Gangwar, A. Kaplan et. al., *Examining artificial intelligence (AI) technologies in marketing via a global lens: Current trends and future research opportunities*, International Journal of Research in Marketing [w opracowaniu], 2021, s. 5. Artykuł dostępny pod adresem: https://reader.elsevier.com/reader/sd/pii/S016781162100094X?token=806330FB098B824585F497DE85C05B23BAE071411D2AD0F1965394805CACD3A41F343931E2C9F729E342440F1EB56E17&originRegion=eu-west-1&originCreation=20220401170250 (dostęp: 01.04.2022 r.).

[49] M. Nowakowski, K. Waliszewski, *Artificial intelligence and algorithms assisting personal finance. A legal and economic perspective*, Przegląd Ustawodawstwa Gospodarczego nr 8, 2021.

[50] L. Cao, *AI in Finance: A Review*, July 2020, s. 25. Opracowanie dostępne pod adresem: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3969480_code3810363.pdf?abstractid=3647625&mirid=1 (dostęp: 02.04.2022 r.).

zysków, to obszar ten może także istotnie przyczyniać się do zwiększenia włączenia finansowego[51], a więc wprowadzenie tego typu rozwiązań należy także rozważać z perspektywy społecznej, gdyż nieodpowiednie zarządzanie personalizacją może także przynieść odwrotny skutek.

Przykładowo, produkty o pewnych cechach, jak i cenie, dostosowanych do określonych kategorii klientów mogą zwiększyć ich dostępność i umożliwić korzystanie z produktów bankowych osobom, które dotąd nie mogły sobie na to pozwolić, chociażby ze względu na barierę kosztów. Jednocześnie możliwość dostosowywania produktów i ich indywidualizacji może być źródłem ryzyk, które zostały już przedstawione w podrozdziale 3.1.5. Ryzyka te mają zarówno charakter prawny, jak i etyczny, w tym odnoszą się do stosowania nieuczciwych praktyk cenowych[52].

Zauważyć należy, że w tym kontekście zastosowanie mają także przepisy Rozporządzenia 2016/679 odnoszące się m.in. do zautomatyzowanego podejmowania decyzji w indywidualnych przypadkach, w tym profilowania (art. 22 RODO), które przewidują liczne obowiązki po stronie administratorów danych osobowych. W szczególności ma to znaczenie w kontekście systemów sztucznej inteligencji i obowiązku wdrażania środków ochrony praw, wolności i prawnie uzasadnionych interesów osoby, której dane dotyczą. Przykładowo, art. 22 ust. 3 RODO wskazuje na konieczność zapewnienia prawa do uzyskania interwencji ludzkiej, co zobowiązuje administratora do wprowadzenia stosownych rozwiązań organizacyjno-technicznych, które zapewnią realizację tego prawa, także w przypadku niewłaściwego działania systemu sztucznej inteligencji.

3.9. Zautomatyzowane systemy doradztwa, w szczególności inwestycyjnego

Ze względu na znaczenie zagadnienia robo-doradztwa dla niniejszego opracowania, zostanie mu poświęcony odrębny rozdział. W tym miejscu należy jedynie wskazać, że zautomatyzowane systemy doradztwa mogą mieć zastosowanie nie tylko w obszarze doradztwa inwestycyjnego, ale także w innych obszarach związanych ze świadczeniem usług finansowych, np. w obszarze tzw. *wealth management*, czyli zarządzania majątkiem.

3.10. Zastosowanie sztucznej inteligencji w relacji bank-bank oraz bank-infrastruktura rynku finansowego

Systemy sztucznej inteligencji mogą mieć zastosowanie także w relacjach pomiędzy bankami, np. w obszarze bankowości korespodenckiej, ale także na linii bank – podmioty należące do szeroko rozumianej infrastruktury rynkowej. Jest to jednak obszar, który dopiero się rozwija i w praktyce najczęściej znajduje odzwierciedlenie w wewnętrznych systemach poszczególnych banków, a nie są to rozwiązania o charakterze systemowym.

Przykładowo, banki świadczące usługi bankowości korespodenckiej mogą wykorzystywać systemy sztucznej inteligencji do dokonywania bardziej dokładnych i efektywnych procesów KYC (Know Your Customer[53]), a także wykrywać ewentualne nieprawidłowości w zlecanych przelewach, w tym identyfikować beneficjentów rzeczywistych tych przelewów. W praktyce oznacza to wykorzystywanie tego typu rozwiązań w systemach przeciwdziałania praniu pieniędzy i finansowania terroryzmu i systemach transakcyjnych. Podobne instrumenty mogą znaleźć zastosowanie w przypadku interakcji banków z instytucjami świadczącymi usług w zakresie infrastruktury rynkowej, w tym także systemami płatności czy rozliczeń.

Oczywistymi przykładami zagrożeń i ryzyk dla podmiotów wykorzystujących takie rozwiązania są:

1. Niewłaściwe funkcjonowanie systemów;
2. Błędy dokonane w trakcie trenowania modelu sztucznej inteligencji, w tym w związku z wykorzystaniem niewłaściwych danych, i nieszczelność systemu;
3. Nieodpowiedni nadzór człowieka nad systemami i przekazanie procesu decyzyjnego w znacznej mierze na rzecz takiego systemu AI.

Decyzja o wprowadzeniu takiego rozwiązania powinna być więc poprzedzona analizą korzyści, możliwości i zagrożeń dla instytucji, bowiem materializacja wspomnianych ryzyk może skutkować odpowiedzialnością regulacyjną, prawną i odszkodowawczą banku.

[51] D. Mhlanga, *Industry 4.0 in Finance: The Impact of Artificial Intelligence (AI) on Digital Financial Inclusion*, Int. J. Financial Stud. 2020, 8, 45, s. 3.

[52] J. A. Gerlick, S. M. Liozu, E*thical and legal considerations of artifcial intelligence and algorithmic decision-making in personalized pricing*, Journal of Revenue and Pricing Management, 19/2020, s. 88 i następne.

[53] J. Han, Y. Huang, S. Liu, K. Towey, *Artifcial intelligence for anti-money laundering: a review and extension*, Digital Finance, 2/2020, s. 214.

3.11. Zastosowania wewnętrzne niebędące działalnością regulowaną banków

Instytucje finansowe, w tym banki, mogą z powodzeniem wykorzystywać systemy sztucznej inteligencji nie tylko w prowadzeniu działalności regulowanej, ale także stosować je do poprawy własnych procesów niezwiązanych z głównym obszarem, ale istotnym dla funkcjonowania instytucji. Przykładem mogą tutaj być obszary wsparcia działalności operacyjnej, systemy wykorzystywane w szeroko rozumianym obszarze zatrudnienia czy rozwiązania wspierające zarządzanie dokumentacją lub zakupami. Choć w tych obszarach – co do zasady – nie występuje ryzyko regulacyjne, to nadal mogą one generować dla banku ryzyka prawne czy ryzyka charakterystyczne dla systemów sztucznej inteligencji. W tym miejscu należy odnotować, że AIA zakłada poddanie stosunkowo szerokiego katalogu systemów AI szczególnym wymaganiom dla tzw. systemów wysokiego ryzyka. Do takiego katalogu – prawdopodobnie – zaliczać się będą przykładowo systemy odpowiedzialne za zatrudnienie, zarządzanie pracownikami i dostęp do samozatrudnienia[54].

Wskazać należy także, że nieodpowiednie stosowanie tego typu rozwiązań może mieć negatywny wpływ na stabilne zarządzanie instytucją, w szczególności, jeżeli proces nie jest poddany odpowiedniemu procesowi nadzoru i monitorowania. Stosowanie takich rozwiązań powinno się również oceniać z perspektywy wymogów w zakresie outsourcingu, w tym outsourcingu bankowego przewidzianego w ustawie Prawo bankowe.

Typowymi ryzykami, na które może być narażony bank wykorzystujący „niebankowe" systemy sztucznej inteligencji, są te odnoszące się do przetwarzania danych osobowych, ryzyk w zakresie cyberbezpieczeństwa i bezpieczeństwa IT oraz ryzyka operacyjne. Zwrócić uwagę należy, że jedną z technik manipulacji systemami sztucznej inteligencji jest tzw. data poisoning, czyli zatruwanie danych sprowadzające się do takiego „pokierowania" danymi, aby wytrenowany model uczenia maszynowego lub głębokiego wygenerował inny rezultat niż ten założony przez jego twórcę[55]. Niewykrycie stosowania przez osobę trzecią takich technik może mieć negatywne konsekwencje dla instytucji oraz osób tam zatrudnionych. Ponieważ kategoria ta jest jednak bardzo szeroka i obejmuje wiele zastosowań o charakterze „ogólnym", tematyka ta nie będzie przedmiotem dalszej analizy.

3.12. Zastosowania wewnętrzne zaliczane do działalności regulowanej

Kolejna kategoria zastosowań systemów sztucznej inteligencji to wszelkie obszary wykorzystywane wewnętrznie przez banki do realizacji szeroko rozumianej działalności regulowanej, w tym:

1. Obszary zarządzania ryzykiem i modele wewnętrzne,
2. Rozwiązania w zakresie zapewnienia zgodności (compliance),
3. Systemy przeciwdziałania oszustwom,
4. Systemy przeciwdziałania praniu pieniędzy i finansowania terroryzmu,
5. Rozwiązania wspierające procesy decyzyjne, w tym zarządu.

Należy przy tym zauważyć, że obszarów zastosowania można wskazać więcej, w tym w szczególności w obszarze szeroko rozumianej działalności inwestycyjnej banków czy skarbowości banków, m.in. z użyciem tzw. HFT (*High-frequency trading*) i handlu algorytmicznego. Zagadnienia te były i są sygnalizowane w całym opracowaniu, jednakże ze względu na ograniczone ramy, jak i tematykę, zostały one opisane jedynie w podstawowym zakresie.

Zwrócić uwagę należy tutaj, że zastosowanie systemów sztucznej inteligencji w wewnętrznych obszarach poddanych regulacji może wiązać się z podwyższonym ryzykiem dla instytucji – a także ze zwiększonymi kosztami zarówno implementacji, jak i utrzymania – a planowane zmiany prawne mogą to pogłębić. Wymagać to będzie po stronie instytucji wprowadzenia szeregu zmian zarówno w istniejących procesach, jak i stworzenia nowych, a także modyfikacji wielu obszarów organizacyjno-technicznych czy edukacji pracowników oraz – z dużą dozą prawdopodobieństwa – zmiany kulturowej, o czym w dalszej części opracowania.

Poniżej zaprezentowane zostaną wybrane rozwiązania, których znaczenie dla sektora bankowego wydaje się – przynajmniej na dzień sporządzania raportu – największe. Nie oznacza to, że wraz z rozwojem metod i podejść opartych o tzw. sztuczną inteligencję (np. przetwarzanie języka naturalnego) inne obszary nie staną się kluczowe z perspektywy instytucji finansowych. Trudno jednak przewidzieć dalszy kierunek tych ewentualnych zmian w związku z niepewnością technologiczną, ale także ewoluującymi potrzebami banków i ich klientów.

[54] Załącznik III do projektu AIA.

[55] Więcej [w:] L. Verde, F. Marulli, S. Marrone, *Exploring the Impact of Data Poisoning Attacks on Machine Learning Model Reliability*, Procedia Computer Science 192, 2021, s. 2625.

W niniejszym podrozdziale zaprezentowane zostaną także dwa zagadnienia z pogranicza działalności banków oraz firm inwestycyjnych, tj. handel algorytmiczny oraz zautomatyzowane doradztwo (*robo-advisory*).

3.13. Obszar zarządzania ryzykiem i modele wewnętrzne

Systemy sztucznej inteligencji, a także modele statystyczne, na potrzeby zarządzania ryzykiem[56], stanowią jeden z kluczowych obszarów wykorzystania przez instytucje finansowe, w tym banki, choć wykorzystywane są nie tylko do predykcji[57]. Istotność samego obszaru zarządzania ryzykiem dla banków nie wymaga dodatkowego wyjaśnienia[58], dlatego w dalszej części akcent zostanie położony na obszar wykorzystania systemów sztucznej inteligencji w tym obszarze[59], tym bardziej że niektóre dane wskazują wyraźnie na przewagę modeli uczenia maszynowego nad dotychczas stosowanymi rozwiązaniami[60]. Warto podkreślić, że skuteczność modeli uczenia maszynowego w obszarze oceny zdolności kredytowej w relacji do np. modeli regresji liniowej, może być znaczna[61].

Na wstępie należy wskazać, że obecnie na poziomie Unii Europejskiej, jak i krajowym nie funkcjonują regulacje, które specyficznie (i wprost) odnoszą się do kwestii zastosowania np. uczenia maszynowego do obszaru zarządzania ryzykiem czy modeli wewnętrznych, choć Europejski Urząd Nadzoru Bankowego opublikował pod koniec 2021 r. dokument „EBA Discussion Paper on Machine Learning for IRB Models"[62], który ma stanowić punkt wyjścia do stworzenia ewentualnych wytycznych sektorowych w tym obszarze. Warto w tym miejscu nadmienić, że także wspomniany już w niniejszym opracowaniu niemiecki organ nadzoru nad rynkiem finansowym BaFin przygotował ogólne wytyczne w zakresie wykorzystania AI w sektorze finansowym[63].

W prawie krajowym oraz unijnym (art. 22 Rozporządzenia 2016/679) można znaleźć wprawdzie pośrednie odniesienia do takich zastosowań, jak chociażby w art. 105a ust. 1a Prawa bankowego, który stanowi, że:

„Banki, inne instytucje ustawowo upoważnione do udzielania kredytów, instytucje pożyczkowe oraz podmioty, o których mowa w art. 59d ustawy z dnia 12 maja 2011 r. o kredycie konsumenckim, a także instytucje utworzone na podstawie art. 105 ust. 4, mogą w celu oceny zdolności kredytowej i analizy ryzyka kredytowego podejmować decyzje, opierając się wyłącznie na zautomatyzowanym przetwarzaniu, w tym profilowaniu, danych osobowych – również stanowiących tajemnicę bankową – pod warunkiem zapewnienia osobie, której dotyczy decyzja podejmowana w sposób zautomatyzowany, prawa do otrzymania stosownych wyjaśnień co do podstaw podjętej decyzji, do uzyskania interwencji ludzkiej w celu podjęcia ponownej decyzji oraz do wyrażenia własnego stanowiska".

Bez wątpienia można przyjąć, że wspominane tutaj „zautomatyzowane przetwarzanie" odnosi się także do szerokiego zastosowania systemów sztucznej inteligencji, nawet jeżeli nie zostało to wypowiedziane wprost. W takim przypadku systemy stosowane są w dwóch przypadkach:

1. do oceny zdolności kredytowej oraz
2. analizy ryzyka kredytowego.

Wykorzystanie tego typu rozwiązań wiąże się dla banków nie tylko z wyzwaniami związanymi z implementacją nowego rozwiązania z szeroko rozumianego ICT, ale także koniecznością spełnienia dodatkowych wymogów określonych we wspomnianym art. 105a ust. 1a, tj. obowiązkowemu zapewnieniu osobie, której dotyczy decyzja, prawa:

1. do otrzymania stosownych wyjaśnień co do podstaw podjętej decyzji,
2. do uzyskania interwencji ludzkiej w celu podjęcia ponownej decyzji,
3. do wyrażenia własnego stanowiska.

[56] Ch. L. Dunis, P. W. Middleton et. al., *Artificial Intelligence in Financial Markets. Cutting-edge Applications for Risk Management, Portfolio Optimization and Economics*, Palgrave/MacMillan 2016.
[57] B. V. Liebergen, *Machine Learning: A Revolution in Risk Management and Compliance?*, The Capco Institute Journal of Financial Transformation, April 27, 2017, s. 64.
[58] M. Hellwig, *Systemic Aspects of Risk Management in Banking and Finance*, Swiss Journal of Economics and Statistics, 1995, Vol. 131 (4/2), s. 724.
[59] M. Nowakowski, K. Waliszewski, *Sztuczna inteligencja w problematyce modeli oceny ryzyka w instytucjach finansowych z perspektywy prawno-regulacyjnej*, Finanse i Prawo Finansowe, *1*(33), 2022.
[60] L. Gambacorta, Y. Huang, H. Qiu, J. Wang, *How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm*, BIS Working Papers No. 834, December 2019, s. 20.
[61] M. C. Aniceto, F. Barboza, H. Kimura, *Machine learning predictivity applied to consumer creditworthiness*, Future Business Journal, 6(1):37, 2020, s. 13.
[62] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Discussions/2022/Discussion%20on%20machine%20learning%20for%20IRB%20models/1023883/Discussion%20paper%20on%20machine%20learning%20for%20IRB%20models.pdf (dostęp: 05.04.2022 r.).
[63] https://www.bafin.de/SharedDocs/Downloads/EN/Aufsichtsrecht/dl_Prinzipienpapier_BDAI_en.pdf?__blob=publicationFile&v=2 (dostęp: 05.04.2022 r.).

Jednocześnie decyzje takie nie mogą – zgodnie z art. 105a ust. 1c Prawo bankowego – być podejmowane w oparciu o dane szczególnych kategorii, które zostały określone w art. 9 Rozporządzenia 2016/679. Powyższe obowiązki nie były jako dotąd przedmiotem wyjaśnień ze strony Urzędu Komisji Nadzoru Finansowego, choć w lipcu 2020 r. został opublikowany „Komunikat Urzędu Komisji Nadzoru Finansowego w sprawie realizacji przez banki i inne instytucje ustawowo upoważnione do udzielania kredytów prawa wnioskującego o kredyt do uzyskania wyjaśnień na temat dokonanej oceny zdolności kredytowej”[64], który można stosować posiłkowo do zautomatyzowanych systemów w powyższym zakresie. Sam UKNF wskazuje, że celem udzielenia odpowiedzi jest przekazanie klientowi (lub innej osobie) zindywidualizowanej i szczegółowej informacji, w tym informacji na temat środków, które powinna taka osoba przedsięwziąć w celu usunięcia negatywnych skutków determinujących decyzję kredytodawcy o nieprzyznaniu kredytu.

Oznacza, to że przykładowo realizacja prawa do otrzymania stosownych wyjaśnień co do podstaw decyzji kredytowej nie jest tożsama z koniecznością wskazania dokładnych informacji na temat działania algorytmu (modelu), ale wskazania czynników, które taki algorytm wykorzystywał przy opracowywaniu wyniku. Nie ma wątpliwości także, że powyższy obowiązek nie jest tożsamy z koniecznością ujawniania jakichkolwiek informacji stanowiących tajemnicę przedsiębiorstwa, tym bardziej, że znajomość dokładnego mechanizmu funkcjonowania modelu może posłużyć do jego obejścia. Innymi słowy, w takiej sytuacji bank powinien przekazać takie informacje, które pozwolą konkretnej osobie na zrozumienie podstaw decyzji, np. relacji do grupy, do której została ona przypisana, oraz ewentualne usunięcie przeszkód.

Problem, na który mogą napotkać banki korzystające ze zautomatyzowanych systemów przetwarzania to niewątpliwie (nie)możliwość ewentualnego wyjaśnienia, w jaki sposób działa określony model sztucznej inteligencji. Często wskazywaną zależnością jest ta, zgodnie z którą im bardziej zaawansowany model sztucznej inteligencji (i zasadniczo bardziej skuteczny), tym większe trudności w możliwości dokonania wyjaśnienia jego działania. Jednocześnie organy regulacyjne i nadzorcze zasadniczo zawsze będą dążyły do zapewnienia, aby model podlegał wyjaśnieniu, zarówno na potrzeby wewnętrzne instytucji, jak i w ramach Badania i Oceny Nadzorczej czy inspekcji *ad hoc*. Pomocne w tym zakresie są przepisy art. 174-177 Rozporządzenia Parlamentu Europejskiego i rady (UE) nr 575/2013 z dnia 26 czerwca 2013 r. w sprawie wymogów ostrożnościowych dla instytucji kredytowych i firm inwestycyjnych, zmieniające rozporządzenie (UE) nr 648/2012[65] („CRR”), jak i rekomendacji Komisji Nadzoru Finansowego dotyczącej zarządzania ryzykiem modeli w bankach[66], m.in. rekomendacja nr 11[67].

Oczywistymi zagrożeniami związanymi ze stosowaniem modeli automatycznych, zarówno w obszarze zarządzania ryzykiem, jak i oceną zdolności kredytowej, są te dotyczące jakości danych wykorzystywanych zarówno do trenowania modelu, jak i jego bieżącego zasilania. Nieodpowiedni dobór tych danych może skutkować nieefektywnością modelu i skutkować negatywnymi skutkami dla instytucji, w tym w zakresie pokrycia ekspozycji odpowiednim kapitałem. Wymaga to, aby w przypadku stosowania takich rozwiązań w instytucji funkcjonowały specyficzne rozwiązania dotyczące szeroko rozumianego *data governance*, rozumianego jako całokształt rozwiązań organizacyjno-technicznych w obszarze zarządzania danymi. Jednocześnie mogą one być częścią rozwiązań określonych w art. 176 CRR, a także spełniać wymogi określone w Wytycznych EUNB w sprawie zarządzania ryzykiem związanym z technologiami i bezpieczeństwem ICT[68]. Nieodpowiednie ułożenie procesów w tym zakresie, a także brak stosowania odpowiednich procedur wewnętrznych może generować znaczne ryzyka dla instytucji, zarówno w sferze finansowej, jak i odpowiedzialności administracyjnej.

Nadmienić należy, że również projektowane rozporządzenie w sprawie sztucznej inteligencji przewiduje specyficzne wymogi dla obszaru zarządzania danym art. 10), choć jednocześnie nie zawiera konkretnych rozwiązań w odniesieniu do samych instytucji finansowych. W przypadku przyjęcia aktu – zakładając jednak dalsze zmiany na etapie legislacyjnym – zastosowanie tych wymagań powinno stanowić punkt wyjścia dla banków. Na marginesie należy wskazać – gdyż jest to etap dyskusji na poziomie Unii Europejskiej i ostateczny kształt jest trudny do przewidzenia – że zgodnie z treścią załącznika III do wspomnianego rozporządzenia, systemy sztucznej inteligencji przeznaczone do wykorzystania w celu oceny zdolności kredytowej osób fizycznych lub ustalenia ich punktowej oceny kredytowej, z wyjątkiem systemów sztucznej inteligencji oddawanych do użytku przez drobnych dostawców na ich własny użytek, mają stanowić systemy

[64] https://www.knf.gov.pl/knf/pl/komponenty/img/Komunikat_UKNF_ws_prawa_do_uzyskania_wyjasnien_nt_oceny_zdolnosci_kredytowej_wersja_szczegolowa_70332.pdf (dostęp: 06.04.2022 r.).

[65] Dz. Urz. UE z 2013 r., L.176/1, z późn. zm.

[66] https://www.knf.gov.pl/knf/pl/komponenty/img/knf_137749_Rekomendacja_W_42219.pdf (dostęp: 09.04.2022 r.).

[67] Więcej w M. Nowakowski, K. Waliszewski, *Sztuczna inteligencja w problematyce modeli oceny ryzyka w instytucjach finansowych z perspektywy prawno-regulacyjnej*, Finanse i Prawo Finansowe, Tom I/2022.

[68] EBA/GL/2019/04.

wysokiego ryzyka, które podlegać będą specyficznym wymogom, m.in. w zakresie zarządzania ryzykiem. Nie można wykluczyć, że na etapie dalszych prac również systemy zarządzania ryzykiem zostaną poddane takiej kwalifikacji.

Poza obszarem zarządzania danymi, w przypadku stosowania przez bank zautomatyzowanych systemów oceny ryzyka (a także oceny zdolności kredytowej) powinny funkcjonować odpowiednie rozwiązania organizacyjno-techniczne pozwalające na bieżącą weryfikację działania modelu, także w obszarze walidacji i testowania, jak również sprawowanie efektywnego nadzoru nad ich działaniem, co obejmuje także obowiązek wykrywania i reagowania na incydenty występujące w związku z niewłaściwym funkcjonowaniem modelu, a także odpowiednie rozwiązania infrastrukturalne[69]. Ponieważ jest to obszar podlegający ryzykom charakterystycznym dla rozwiązań cyfrowych, należy pamiętać także o obowiązkach związanych z ryzykami typu ICT oraz bezpieczeństwa.

Rekomendowanym rozwiązaniem jest wyznaczenie członka zarządu oraz osób z wyższego szczebla zarządzającego instytucji, którzy będą odpowiedzialni za ten specyficzny obszar działania instytucji, a także gromadzenie i systematyzowanie wiedzy w tym zakresie, także w kontekście stosowanych metodologii. Ułożenie rozwiązań organizacyjnych, w tym kwestii odpowiedzialności oraz raportowania stanowi bardzo istotny element funkcjonowania tego obszaru w banku.

Tematyka ta jest przynajmniej częściowo adresowana w projektowanym rozporządzeniu w sprawie sztucznej inteligencji w zakresie, w jakim odnosi się do systemów sztucznej inteligencji przeznaczonych do wykorzystania w celu oceny zdolności kredytowej osób fizycznych lub ustalenia ich punktowej oceny kredytowej, z wyjątkiem systemów sztucznej inteligencji oddawanych do użytku przez drobnych dostawców na ich własny użytek, czyli tych określonych w załączniku III do projektowanego rozporządzenia. Takie systemy będą traktowane jako systemy wysokiego ryzyka, a więc podlegać będą wymogom określonym w projektowanym rozporządzeniu, choć pamiętać należy, że w wielu miejscach zawiera ono odesłanie do odpowiednich przepisów CRD, a w praktyce aktów je implementujących, jak prawo bankowe.

Po stronie banków zasadnym wydaje się przeprowadzenie mapowania procesów, które odnoszą się do oceny zdolności kredytowej i weryfikacji potrzeby przeprowadzenia stosowanych dostosowań, w szczególności w kontekście danych (art. 10 projektowanego rozporządzenia) i nadzoru człowieka nad funkcjonowaniem modeli (art. 14). Jednocześnie nie można wykluczyć, że również „klasyczne" modele oceny ryzyka występujące w bankach poddane zostaną przedmiotowej regulacji ze względu na istotność predykcji dla stabilności systemu finansowego.

Podobnie, art. 9 projektu AIA odnosi się do systemów zarządzania ryzykiem, które będą miały obligatoryjny charakter dla systemów wysokiego ryzyka. Przepis ten jednocześnie dopuszcza możliwość „wykorzystania" istniejących systemów zarządzania ryzykiem w instytucjach kredytowych do zapewnienia zgodności z przepisami projektowanego rozporządzenia.

3.14. Rozwiązania w zakresie zapewnienia zgodności (compliance)

Obszar zapewnienia zgodności z wymogami prawa i regulacjami jest obszarem, który powinien podlegać dynamicznemu rozwojowi ze względu na rosnącą liczbę obowiązków regulacyjnych, w tym raportowych, nakładanych na banki. Na konieczność obniżenia tych kosztów, przynajmniej w odniesieniu do wybranych typów instytucji finansowych, wskazuje m.in. Europejski Urząd Nadzoru Bankowego w swoim opracowaniu z drugiej połowy 2021 r.[70] i jednocześnie w innym raporcie na temat rozwoju obszaru *Regulatory Technology* („RegTech") podkreśla znaczenie wykorzystania zaawansowanej analityki danych w tym obszarze[71]. Rozwój narzędzi wspierających organizację w zarządzaniu ryzykiem zgodności zależy w dużej mierze od równoległego rozwoju instrumentów nadzorczych wykorzystujących nowe technologie – Supervisiory Technology[72] („SupTech") – które mają stanowić „punkt odbioru" danych przekazywanych przez instytucje nadzorowane.

Szeroko rozumiana sztuczna inteligencja może znaleźć szerokie zastosowanie w obszarze zapewnienia zgodności, choć dalszy rozwój uzależniony jest od spełnienia kilku warunków, które zostaną poruszone w dalszej części opracowania. Można wskazać następujące przykłady:

[69] EUNB w dokumencie dot. stosowania uczenia maszynowego na potrzeby modeli wewnętrznych wskazuje na cztery podstawowe zasady, które powinny znaleźć zastosowanie w bankach: (i) zarządzanie danymi; (ii) infrastruktura techniczna; (iii) ład korporacyjny oraz organizacja (iv) odpowiednia metodologia analityczna. EBA, Discussion Paper on Machine Learning for IRB Models, EBA/DP/2021/04, 11 November 2021, s. 23.

[70] EBA, *Study of the cost of compliance with supervisory reporting requirements*, EBA/Rep/2021/15, 2021.

[71] EBA, *Analysis of RegTech in the EU Financial Sector,* EBA/REP/2021/17, June 2021.

[72] Financial Stability Board, *The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions Market developments and financial stability implications*, 9 October 2020.

1. Wsparcie analityki danych raportowych przekazywanych organom nadzorczym i regulacyjnym, jak również realizacja tych obowiązków.

2. Wykorzystanie technik przetwarzania języka naturalnego na potrzeby gromadzenia i aktualizacji wiedzy w zakresie prawa i regulacji, a także obowiązków o takim charakterze.

3. Wspieranie procesów regulacyjnych, np. w zakresie zarządzania ryzykiem, o czym była mowa w rozdziale 3.4.1.

4. Wspieranie procesów z zakresu bezpieczeństwa i ryzyk operacyjnych.

5. Wspieranie procesów z zakresu przeciwdziałania oszustwom oraz praniu pieniędzy i finansowania terroryzmu (te obszary zostaną opisane w dalszej części opracowania).

6. Wspieranie osób zatrudnionych w instytucji finansowej w zakresie spełniania wymogów prawnych i regulacyjnych, w tym z udziałem tzw. chatbotów.

7. Kontrola korespondencji, np. w zakresie działalności inwestycyjnej banku.

Nie są to oczywiście wszystkie przykłady zastosowania szeroko rozumianej sztucznej inteligencji w obszarze compliance, jednak wskazane powyżej zastosowania znajdują najczęściej zastosowanie w instytucjach. Należy jednocześnie zwrócić uwagę, że pojęcie RegTech jest stosunkowo szerokie, co powoduje, że jednoznaczna kwalifikacja określonych pojęć jako służących zapewnieniu zgodności może być nieco utrudniona.

Za przykład może posłużyć obszar tzw. handlu algorytmicznego, który podlega Rozporządzeniu (UE) 2017/589, gdzie art. 3 ust. 4 wskazuje, że *„[f]irma inwestycyjna zapewnia, by pracownicy odpowiedzialni za zarządzanie ryzykiem i zgodność z przepisami w zakresie handlu algorytmicznego posiadali:*

a) wystarczającą wiedzę na temat handlu algorytmicznego i strategii handlowych;

b) wystarczające umiejętności w zakresie reagowania na informacje pochodzące z automatycznych ostrzeżeń;

c) wystarczające uprawnienia, aby zastosować środki dyscyplinarne wobec pracowników odpowiedzialnych za handel algorytmiczny w przypadkach, w których taki handel powoduje zakłócenie obrotu lub prowadzi do podejrzenia wystąpienia nadużycia na rynku”.

Pojawia się więc pytanie, czy rozwiązania technologiczne/techniczne zapewniające spełnienie tych wymogów będą kwalifikowane jako RegTech czy też „produktowe” odnoszące się do samego handlu algorytmicznego. Wydaje się jednak, że wszelkie rozwiązania techniczne wspierające spełnienie obowiązków prawnych i regulacyjnych należy traktować jako będące częścią szeroko rozumianego obszaru zgodności z przepisami.

Ryzyka związane ze stosowaniem szeroko rozumianej sztucznej inteligencji w obszarze zarządzania ryzykiem zgodności mogą mieć zróżnicowany charakter i mogą być powiązane zarówno z kwestiami dotyczącymi outsourcingu (w tym braku możliwości outsourcingu niektórych czynności, w tym zarządzania ryzykiem), jak i IT oraz bezpieczeństwa, czy też ryzykami związanymi ze stosowaniem samej sztucznej inteligencji.

Te szczególne ryzyka polegać mogą przykładowo na:

1. Niedoskonałości modeli i systemów sztucznej inteligencji, np. w zakresie przetwarzania języka naturalnego.

2. Nadmiernym poleganiu na tych systemach przez inspektorów compliance i innych osób korzystających z tych rozwiązań i zatraceniem ludzkiego osądu czy podejmowaniu decyzji wyłącznie w oparciu o wyniki systemu.

3. Braku dostatecznego nadzoru na funkcjonowaniem systemów i niedostatecznym udziale człowieka w całym procesie.

Efektem zmaterializowania się takich ryzyk może być m.in. odpowiedzialność regulacyjna, prawna, jak i reputacyjna względem np. klientów, toteż zarówno decyzja o stosowaniu tego typu rozwiązań, jak i jego stosowanie powinno podlegać specyficznym zasadom i podejściu opartym na analizie ryzyk wewnątrz organizacji.

Warto nadmienić, że rozwój RegTech i SupTech z wykorzystaniem różnych podejść i technik z obszaru sztucznej inteligencji w znacznej mierze uzależniony jest od równoległego rozwoju m.in. aktów prawnych i regulacji tworzonych w formie do odczytu maszynowego, dostępu do baz danych i rejestrów z użyciem interfejsów dostępowych (API), jak również jasnego określenia wymogów dla dostawców tego typu rozwiązań z perspektywy organów nadzoru oraz regulacji.

3.15. Systemy przeciwdziałaniu transakcjom oszukańczym (fraudowym)

Banki mają szczególne obowiązki w zakresie przeciwdziałania transakcjom nieautoryzowanym, co obejmuje nie tylko obowiązki o charakterze informacyjnym czy edukacyjnym, ale także technologicznym i operacyjnym. Wysoki poziom transakcji nieautoryzowanych w rozumieniu ustawy o usługach płatniczych może być źródłem zarówno odpowiedzialności odszkodowawczej (i tym samym słabszych wyników finansowych), ale także regulacyjnej, jeżeli bank nie dołożył starań w celu mitygacji ryzyk w tym obszarze. Niski poziom transakcji oszukańczych może warunkować także (nie)stosowanie tzw. silnego uwierzytelniania klienta zgodnie z art. 18 i następne Rozporządzenia (UE) 2018/389.

W bankach stosowane są różne rozwiązania mające na celu przeciwdziałanie transakcjom oszukańczym, przy czym nie wszystkie korzystają z bardziej zaawansowanych rozwiązań opartych np. na uczeniu maszynowym lub głębokim i przykładowo biometrii behawioralnej, ale wykorzystują mniej skomplikowane rozwiązania oparte np. bieżącym monitorowaniu określonych wskaźników, np. położenia geograficznego płatnika.

Niewątpliwie jednak systemy przeciwdziałania oszustwom płatniczym, także z wykorzystaniem mniej standardowych kanałów, jak np. IVR (*Interactive Voice Response*), są obszarem, w którym szeroko rozumiana sztuczna inteligencja ma spory potencjał rozwoju i zwiększenia efektywności systemów antyfraudowych.

Systemy sztucznej inteligencji mogą przykładowo „uczyć się" pewnych określonych zachowań czy przyzwyczajeń lub nawyków (np. zakupowych) i porównywać w czasie rzeczywistym te elementy z daną transakcją. Jeżeli model dostrzeże istotne (określenie tego poziomu będzie uzależnione od wielu czynników) różnice może albo automatycznie zablokować próbę realizacji transakcji, albo przekazać użytkownikowi informację o podejrzeniu. Może to istotnie przyczynić się do obniżenia wskaźnika transakcji nieautoryzowanych, a więc uchronić użytkowników, jak i same instytucje przed wskazanymi wyżej ryzykami.

Powyższy przykład stanowi dość zaawansowany sposób przeciwdziałania transakcjom oszukańczym, jednakże mniej skomplikowane rozwiązania oparte o bieżącą analitykę danych również mogą okazać się skuteczne. Dodatkowo systemy oparte o tzw. *voice recognition* (rozpoznawanie głosu) oraz *image recognition* (rozpoznawanie obrazu) połączone z modelami sztucznej inteligencji mogą wspierać analityków oraz decydentów w tym zakresie (np. na etapie *onboardingu* klienta z użyciem zdalnej weryfikacji klienta, o czym szerzej w podrozdziale dotyczącym przeciwdziałania praniu pieniędzy i finansowania terroryzmu).

Ryzyka związane ze stosowaniem tych rozwiązań są zasadniczo zbieżne z tymi, które zostały opisane w podrozdziale 3.4.2, przy czym w tym miejscu należy zwrócić uwagę, że szczególnego znaczenia nabiera kwestia monitoringu człowieka i prawidłowego określenia wskaźników referencyjnych (a także weryfikacja m.in. *false-positive*), jak również walidacji modeli.

3.16. Systemy przeciwdziałania praniu pieniędzy i finansowania terroryzmu

Postępująca cyfryzacja finansów powoduje, że ryzyka związane z praniem pieniędzy i finansowaniem terroryzmu („AML" – *Anti-money laundering*) rosną, podobnie jak koszt obsługi tych procesów w instytucjach finansowych[73]. Zwiększone zapotrzebowanie na automatyzację procesów AML będzie z pewnością widoczne w związku z dynamicznym rozwojem rynku kryptoaktywów, które za sprawą planowanych i już obowiązujących rozwiązań prawnych[74] i regulacyjnych[75], podlegać będą znacznie wyższym standardom w zakresie przeciwdziałania praniu pieniędzy i finansowania terroryzmu.

Jednocześnie zastosowanie tzw. sztucznej inteligencji w obszarze AML jest wskazywane jako mające największy potencjał w zakresie zwiększenia efektywności, w tym kosztowej, jak również skuteczności wykrywania transakcji podejrzanych. Hong Kong Monetary Authority – organ odpowiedzialny za nadzór nad sektorem finansowym w Hong Kongu – opracował nawet swoisty poradnik w zakresie wdrażania takich rozwiązań[76]. Jednocześnie jednak – ze względu na swoją specyfikę, jak i wysoki poziom ryzyka regulacyjnego – obszar AML z użyciem cyfrowych rozwiązań może być poddany pewnym ograniczeniom wynikającym chociażby z zasad dotyczących outsourcingu (powierzenia).

Szeroko rozumiana sztuczna inteligencja w kontekście AML może znaleźć zastosowanie m.in. w następujących obszarach:

[73] FATF, *Opportunities and Challenges of New Technologies for AML/CFT*, July 2021, s. 36.
[74] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0422 (dostęp: 19.04.2022 r.).
[75] https://www.fatf-gafi.org/media/fatf/documents/recommendations/Updated-Guidance-VA-VASP.pdf (dostęp: 19.04.2022 r.).
[76] https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210121e1a1.pdf (dostęp: 19.04.2022 r.).

1. Systemy wykorzystujące przetwarzanie języka naturalnego oraz uczenia maszynowego w celu opracowania rekomendacji dla analityka co do statusu transakcji.

2. Systemy analityczne dokonujące analizy dużych zbiorów danych oraz danych niestandardowych w celu wykrycia transakcji podejrzanej.

3. Modele oceny ryzyka klienta lub transakcji.

4. Wykorzystanie biometrii behawioralnej do oceny ryzyka transakcji.

W praktyce powyższe rozwiązania mogą być rozszerzane lub zawężane w zakresie funkcjonalności w zależności od indywidualnych potrzeb instytucji finansowej (instytucji obowiązanej). Zakres tych rozwiązań można także rozszerzyć o obszar wideoweryfikacji, czyli weryfikacji klienta z użyciem kanałów zdalnych, gdzie zastosowanie mogą mieć przykładowo systemy analizujące zachowanie użytkownika, porównujące obraz twarzy z dokumentami identyfikacyjnymi i podobne.

Zastosowanie uczenia maszynowego czy głębokiego w obszarze AML wymaga jednak zapewnienia odpowiednich rozwiązań organizacyjnych i technicznych, jak również dobrego zarządzania danymi, które służą zarówno do trenowania modelu, jak i zastosowań produkcyjnych. Niewłaściwe wykonywanie obowiązków w zakresie przeciwdziałania praniu pieniędzy i finansowania terroryzmu może być źródłem zarówno odpowiedzialności prawnej, jak i regulacyjnej, a także kryzysu wizerunkowego, jeżeli system okaże się nieszczelny.

Zwrócić należy tutaj także uwagę na projektowane przez Europejski Urząd Nadzoru Bankowego wytyczne w sprawie roli inspektorów AML, jak również szeroko rozumianej zgodności z wymogami w zakresie AML[77]. W punkcie 74 wytycznych EUNB znajdziemy wyraźne wskazanie, że strategiczne decyzje w zakresie przeciwdziałania praniu pieniędzy i finansowaniu terroryzmu nie powinny być powierzane podmiotom trzecim (w tym dokumencie znajduje się także niewyczerpujący katalog czynności operacyjnych, które również nie powinny podlegać powierzeniu). Istotne jest więc wyraźne określenie, jaki jest zakres czynności generowanych przez systemy oparte o sztuczną inteligencję, jeżeli są one dostarczane przez podmioty trzecie.

W tym kontekście również należy zwrócić uwagę na szczególną wagę obowiązków ludzi w kontekście tych systemów. Dotyczy to nie tylko sprawowania efektywnego i skutecznego nadzoru nad funkcjonowaniem systemów, ale także ryzyko tzw. *overreliance*, czyli nadmiernego polegania przez analityków na systemach sztucznej inteligencji (lub automatyzacji) w zakresie podejmowania decyzji. Podkreślić należy, że to człowiek ostatecznie powinien podejmować decyzję co do zaklasyfikowania określonej transakcji lub relacji jako podejrzanej, a także podjąć stosowne działania w zakresie ewentualnego zgłoszenia. Nie wyklucza to jednocześnie możliwości stosowania systemów rekomendacyjnych w tym zakresie. Jednocześnie ze względu na liczne obowiązki banków (ale i ograniczenia) w zakresie dopuszczalności możliwości blokowania środków czy transakcji, konieczne jest zwrócenie uwagi na efektywność stosowanych modeli.

3.17. Rozwiązania wspierające procesy decyzyjne, w tym zarządu

Kolejnym obszarem, w którym rozwiązania oparte o sztuczną inteligencję znajdą zastosowanie, jest wsparcie procesów decyzyjnych w instytucji. Systemy rekomendacyjne czy predykcyjne mogą tworzyć zarówno konkretne rekomendacje co do działań, które powinien podjąć jej adresat, np. w zakresie pokrycia dodatkowego kapitału lub zmniejszenia skali sprzedaży określonych produktów, jak i wskazywać na przewidywane skutki określonych decyzji, np. wpływ decyzji na model biznesowy instytucji.

Obszar ten jest jeszcze w stosunkowo wczesnej fazie rozwoju, co w znacznej mierze jest też konsekwencją niedostatecznego rozpowszechnia tego typu rozwiązań, ich nie zawsze zadowalającej skuteczności, jak i obaw co do stosowania tzw. sztucznej inteligencji w obszarze zarządczym. Niewątpliwie jednak jest to obszar, który ma szansę dynamicznie się rozwijać i podejmować bardziej efektywne decyzje zarządcze, m.in. w takich obszarach, jak ryzyko operacyjne, sprzedaż produktów i usług czy inwestycji.

Jednocześnie umożliwienie osobom decyzyjnym podejmowania decyzji przy wsparciu systemu rekomendacyjnych czy predykcyjnych może wiązać się z określonymi ryzykami o zróżnicowanym charakterze. Poza oczywistymi ryzykami związanymi z nieefektywnością lub nieskutecznością modeli, należy wskazać tutaj na ryzyko nadmiernego polegania na takich systemach przez osoby podejmujące decyzje[78]. Jest to o tyle istotne, że decyzje podejmo-

[77] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Consultations/2021/Consultation%20on%20draft%20Guidelines%20on%20the%20role%2C%20tasks%20and%20responsibilities%20AML-CFT%20compliance%20officers/1018277/CP%20GLs%20on%20AMLCFT%20compliance%20officer.pdf (dostęp: 19.04.2022 r.).

[78] Na konieczność stosowania środków przeciwdziałających tzw. overreliance wskazuje m.in. dokument ekspertów Komisji Europejskiej w sprawie sztucznej inteligencji godnej zaufania. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419, s. 26.

wane przez te osoby są też ściśle powiązane z ich kompetencjami, wiedzą i doświadczeniem, a także bezpośrednią odpowiedzialnością. Zapewnienie, że tego typu systemy są jedynie swoistymi dopełnieniem, jest więc w tym przypadku kluczowe.

3.19. Handel algorytmiczny

Zagadnienie handlu algorytmicznego nie zostało tutaj przedstawione w sposób wyczerpujący, a jedynie w odniesieniu do kluczowych obszarów powiązanych ze stosowaniem tzw. systemów sztucznej inteligencji.

Handel algorytmiczny został zdefiniowany zasadniczo w dwóch aktach prawnych na poziomie Unii Europejskiej, tj.:

1. W dyrektywie Parlamentu Europejskiego i Rady (UE) 2014/65/UE z dnia 15 maja 2014 r. w sprawie rynków instrumentów finansowych oraz zmieniająca dyrektywę 2002/92/WE i dyrektywę 2011/61/UE oraz

2. Rozporządzeniu delegowanym Komisji (UE) 2017/565 z dnia 25 kwietnia 2016 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2014/65/UE w odniesieniu do wymogów organizacyjnych i warunków prowadzenia działalności przez firmy inwestycyjne oraz pojęć zdefiniowanych na potrzeby tej dyrektywy.

Z kolei warunki dla prowadzenia handlu algorytmicznego dla firm inwestycyjnych określa Rozporządzenie 2017/589 z dnia 19 lipca 2016 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2014/65/UE w odniesieniu do regulacyjnych standardów technicznych określających wymogi organizacyjne dla firm inwestycyjnych prowadzących handel algorytmiczny mające kluczowe znaczenie z perspektywy zagadnienia stosowania tzw. sztucznej inteligencji w tym obszarze[79].

Sama definicja handlu algorytmicznego jest dość rozbudowana. Zgodnie z przywołanymi aktami prawnymi handel algorytmiczny oznacza obrót na instrumentach finansowych, w którym algorytm komputerowy automatycznie ustala indywidualne parametry zleceń, takie jak warunki uruchomienia zlecenia, moment jego realizacji, cenę lub ilość instrumentów będących przedmiotem zlecenia lub sposób zarządzania zleceniem po jego złożeniu, przy ograniczonym lub zerowym udziale człowieka i nie obejmuje jakichkolwiek systemów wykorzystywanych wyłącznie do celu przekierowywania zleceń z jednego systemu obrotu do innego lub do celu przetwarzania zleceń nieobejmującego określania jakichkolwiek parametrów transakcji lub potwierdzania zleceń lub przetwarzania potransakcyjnego zawartych transakcji. Dodatkowo system uznaje się za działający przy ograniczonym lub zerowym udziale człowieka, jeżeli w odniesieniu do każdego procesu zlecenia lub opracowania wyceny lub każdego procesu służącego optymalizacji wykonywania zleceń zautomatyzowany system podejmuje decyzje na dowolnym z etapów generowania, tworzenia, przekierowywania lub wykonywania zleceń lub wycen, zgodnie z wcześniej określonymi parametrami.

Kluczowym elementem jest tutaj więc zautomatyzowany proces, który odbywa się z niewielkim lub żadnym udziałem człowieka. Nie oznacza to jednak, że jest on pozbawiony stosownej kontroli, a przepisy wspomnianego rozporządzenia 2017/589 są w tym zakresie wręcz restrykcyjne, szczególnie w kontekście kwalifikacji pracowników odpowiedzialnych za ten obszar działalności firmy inwestycyjnej. Zauważyć jednocześnie należy, że kluczowymi obszarami wskazanymi m.in. w art. 1 przedmiotowego rozporządzenia są:

1. Wyraźnie sprecyzowane zakresy odpowiedzialności, w tym procedury zatwierdzania w zakresie tworzenia, wdrażania i późniejszego aktualizowania algorytmów handlowych oraz procedury rozwiązywania problemów wykrytych podczas monitorowania algorytmów handlowych.

2. Skuteczne procedury przekazywania informacji wewnątrz firmy inwestycyjnej – tak, aby instrukcje były opracowywane i wdrażane w sposób wydajny i terminowy.

3. Rozdzielenie zadań i obowiązków jednostek odpowiadających za handel oraz funkcji wsparcia, w tym funkcji kontroli ryzyka i komórki ds. nadzoru zgodności z prawem w celu zapewnienia, by nie można było zataić nieuprawnionej działalności handlowej.

Innymi słowy, kluczowe w wykorzystywaniu algorytmów handlowych oraz ich monitorowaniu jest odpowiednie ułożenie procesów, które w samym rozporządzeniu zostały doprecyzowane w kolejnych artykułach. Wyraźny akcent jest w tym przypadku położony na szeroko rozumianą odporność systemów transakcyjnych, w tym poprzez stosowanie zautomatyzowanych systemów nadzoru służących wykrywaniu manipulacji na rynku.

Jak zostało to już wskazane w niniejszym opracowaniu, szczególną rolę w kontekście stosowania handlu algorytmicznego w firmach inwestycyjnych przypisuje się komórce ds. nadzoru zgodności z prawem

[79] Warto jednocześnie podkreślić, że zagadnienie handlu algorytmicznego jest również poruszane w opracowaniach ESMA, m.in.: https://www.esma.europa.eu/file/121044/download?token=3CWH3QwW (dostęp: 21.04.2022 r.).

(compliance). Przepisy rozporządzenia wymagają przykładowo, aby pracownicy tej jednostki, jeżeli odpowiadają za ten obszar, *„posiadali przynajmniej ogólne pojęcie na temat działania systemów handlu algorytmicznego oraz algorytmów handlowych”*. Muszą oni także utrzymywać stały kontakt z osobami wewnątrz firmy, które posiadają bardziej szczegółową wiedzę w tym zakresie oraz tych osób, które posiadają dostęp do tzw. funkcji awaryjnej uruchamianej na wypadek awarii systemów transakcyjnych. Choć nie wynika to wprost ze wskazanych przepisów[80], to rekomendowanym kierunkiem jest zapewnienie co najmniej podstawowego szkolenia w zakresie działania systemów handlu algorytmicznego, jak również szkoleń okresowych. Istotne jest także zastrzeżenie, że art. 3 ust. 4 przewiduje także dodatkowe wymogi dla pracowników zarządzania ryzykiem i zgodności w zakresie handlu algorytmicznego, które można sprowadzić zasadniczo do umiejętności sprawnego reagowania na ewentualne zagrożenia i ryzyka, zarówno o charakterze technicznym, jak i osobowym.

Rozporządzenie 2017/589 wymaga również, aby sam personel odpowiedzialny za działanie tych rozwiązań posiadał stosowne umiejętności oraz wiedzę techniczną, a także aby ilość osób była adekwatna do skali zastosowania handlu algorytmicznego. Istotne jest przy tym to, że takie osoby – niezależnie od osób zatrudnionych w komórkach compliance posiadały także kompetencje i wiedzę dotyczące:

1. funkcjonowania systemów handlu algorytmicznego,
2. monitorowania i testowania takich systemów i algorytmów (szczegółowe warunki określa tutaj m.in. art. 6-8 Rozporządzenia 2017/589),
3. strategii handlowych wykorzystywanych do tworzenia algorytmów,
4. zobowiązań prawnych (wynikających nie tylko z przedmiotowego rozporządzenia, ale także Rozporządzenia 2017/565 i innych aktów oraz regulacji).

Minimalny zakres wiedzy i umiejętności powinien być dostosowany do zakresu działania firmy inwestycyjnej i jej potrzeb, a także na bieżąco monitorowany i poddawany rozwojowi w formie szkoleń.

Pozostałe wymogi w zakresie nadzoru i kontroli nad funkcjonowaniem handlu algorytmicznego odnoszą się do szerokiego spektrum zagadnień, takich jak dostęp do danych, zarządzania ryzykami, regularnych przeglądów czy obowiązków w zakresie testowania i walidacji. Zauważyć należy, że w przypadku przyjęcia przez Unię Europejską rozporządzenia w sprawie sztucznej inteligencji wiele z dotychczasowych obowiązków należało będzie powiązać z tymi określonymi w AIA, m.in. w kontekście zarządzania ryzykiem. Jest to istotna kwestia, w szczególności w kontekście projektowanego art. 10 AIA dotyczącego danych, w tym ich jakości.

3.19. Robo-doradztwo, czyli doradztwo inwestycyjne z wykorzystaniem algorytmów i modeli sztucznej inteligencji

Doradztwo inwestycyjne zostało zdefiniowane w art. 76 ustawy o obrocie instrumentami finansowymi i polega ono na *„przygotowaniu, z inicjatywy firmy inwestycyjnej lub na wniosek klienta, oraz przekazywaniu klientowi, określonej w art. 9 rozporządzenia 2017/565, pisemnej, ustnej lub w innej formie, w szczególności elektronicznej, spełniającej wymóg trwałego nośnika, przygotowanej w oparciu o potrzeby i sytuację klienta rekomendacji, dotyczącej nabycia lub zbycia jednego instrumentu finansowego lub większej ich liczby, albo dokonania innej czynności wywołującej równoważne skutki, której przedmiotem są instrumenty finansowe, albo rekomendacji dotyczącej powstrzymania się od wykonania takiej czynności”*.

Usługa doradztwa inwestycyjnego wiąże się z licznymi obowiązkami po stronie firmy inwestycyjnej oraz jej pracowników, w szczególności w kontekście przejrzystości względem klientów oraz posiadaniem odpowiednich rozwiązań organizacyjno-technicznych[81]. Ze względu na szczególną rolę tej usługi na rynku kapitałowym, Urząd Komisji Nadzoru Finansowego opublikował w 2020 r. Stanowisko UKNF w sprawie świadczenia usługi robo-doradztwa[82], które określa oczekiwania organu nadzoru w zakresie zautomatyzowanej (lub częściowo zautomatyzowanej) usługi doradztwa inwestycyjnego. Już na wstępie należy zauważyć, że stosowanie samego stanowiska nie oznacza „łagodniejszego” podejścia do przepisów

[80] Abstrahując od wymogów specyficznych dla komórek ds. nadzoru zgodności z prawem dla firm inwestycyjnych i banków, które mogą wynikać z aktów wykonawczych do ustawy o obrocie instrumentami finansowymi.

[81] Aktami kształtującymi te wymogi są nie tylko Rozporządzenie 2017/565 czy ustawa o obrocie instrumentami finansowymi, ale także Rozporządzenie Ministra Finansów z dnia 29 maja 2018 r. w sprawie szczegółowych warunków technicznych i organizacyjnych dla firm inwestycyjnych, banków, o których mowa w art. 70 ust. 2 ustawy o obrocie instrumentami finansowymi, i banków powierniczych (Dz. U. z 2018 r., poz. 1111) oraz Rozporządzenie Ministra Finansów z dnia 30 maja 2018 r. w sprawie trybu i warunków postępowania firm inwestycyjnych, banków, o których mowa w art. 70 ust. 2 ustawy o obrocie instrumentami finansowymi, oraz banków powierniczych (Dz. U. z 2018 r., poz. 1112).

[82] https://www.knf.gov.pl/knf/pl/komponenty/img/Stanowisko_UKNF_ws_swiadczenia_uslugi_robo_doradztwa_71303.pdf (dostęp: 22.04.2022 r.).

obowiązujących banki świadczące usługi wspomniane powyżej.

UKNF definiuje robo-doradztwo (*robo-advisory*) jako szczególną formę wykonywania usługi doradztwa inwestycyjnego, którą można rozumieć jako proces, w ramach którego udzielanie i przekazywanie rekomendacji odbywa się z wykorzystaniem algorytmów, systemów automatycznych i półautomatycznych. Urząd wskazuje tutaj, że kluczowe jest, aby to algorytm (model) był wykorzystywany do analizy i przyporządkowania instrumentów finansowych do profilu klienta.

Wykorzystanie usługi robo-doradztwa może mieć niewątpliwie duże znaczenie w rozwoju usług finansowych na rynku kapitałowym[83], a także zmierzać do ograniczenia niektórych kosztów operacyjnych po stronie banków oraz firm inwestycyjnych, poprzez ograniczenie czynnika ludzkiego w świadczeniu tego typu usługi. Jednocześnie jednak można wyodrębnić zróżnicowane modele jej świadczenia, także takie, które służą jedynie jako system rekomendacyjny dla ludzkich doradców banków lub przynajmniej teoretycznie – odbywają się bez jego udziału. Każde z rozwiązań może wiązać się z określonymi ryzykami[84], które zostaną zarysowane poniżej.

Stanowisko UKNF zwraca uwagę, że robo-doradztwo składa się zasadniczo z trzech etapów:

1. Zebrania informacji od klienta, które posłużą następnie do dokonania oceny odpowiedniości niezbędnej do przygotowania rekomendacji (oraz samej kwalifikacji klienta).

2. Dokonania analizy (zautomatyzowanej) instrumentów finansowych objętych usługą doradztwa inwestycyjnego.

3. Zestawienia klienta z konkretnymi rekomendacjami z użyciem odpowiedniego algorytmu lub modelu.

Każdy z etapów może wiązać się dla banku z innymi wyzwaniami, które mogą mieć charakter organizacyjno-techniczny, jak również osobowy czy regulacyjny. Z tego względu przed podjęciem decyzji o wdrożeniu rozwiązania opartego o robo-doradztwo zaleca się dokonanie analizy SWOT, która pozwoli na ocenę potrzeb, możliwości i wyzwań związanych z jej wdrożeniem. Sam UKNF zwraca uwagę, że przy przygotowywaniu firmy inwestycyjnej do świadczenia takiej usługi powinny brać „rzeczywisty i czynny udział" zróżnicowane jednostki, w tym co najmniej te odpowiedzialne za obszary doradztwa inwestycyjnego, obsługi klienta, IT, compliance (tutaj UKNF przewiduje szczególne wymagania) oraz zarządzania ryzykiem. Jednocześnie jednak w odniesieniu do pewnej specyficznej grupy, tzw. pracowników autoryzowanych, nie wymaga się wiedzy i kompetencji w zakresie IT, a jedynie takiego poziomu zrozumienia, który pozwoli na aktywne uczestnictwo w tym procesie. Jednocześnie tacy pracownicy nie są obowiązani do uczestnictwa w procesie technicznym związanym ze świadczeniem usługi robo-doradztwa.

UKNF wskazuje także w swoim Stanowisku na szczególną rolę komórek ds. nadzoru zgodności działalności z prawem (compliance). Zauważyć można tutaj analogię do wymogów określonych w Rozporządzeniu (UE) 2017/589 odnoszącym się do handlu algorytmicznego, gdzie komórka taka ma również istotne znaczenie w kontekście całokształtu procesu dla tej usługi. Urząd podkreśla tutaj, że zarząd lub członek zarządu (przy czym UKNF nie precyzuje obszaru nadzorowanego) powinien włączyć komórkę compliance zarówno w tworzenie, wdrożenie, jak i działalność produkcyjną (operacyjną) usługi robo-doradztwa (z tym zastrzeżeniem, że komórka compliance nie musi być zaangażowana w cały proces świadczenia usługi).

Oznacza to, że inspektorzy compliance powinni także uczestniczyć w pracach nad stosownymi politykami, procedurami czy dokumentami klientowskimi, w tym także o charakterze informacyjnym i marketingowym. W konsekwencji na zarządzie banku (firmy inwestycyjnej) ciąży obowiązek zapewnienia rozwiązań organizacyjnych istotnych dla prawidłowego sprawowania nadzoru, w tym odpowiedzialność za współpracę z właściwymi jednostkami czy też weryfikację poprawności funkcjonowania algorytmów robo-doradztwa. Sprowadza to na inspektorów tych jednostek obowiązek posiadania stosownych kompetencji oraz wiedzy, choć UKNF podkreśla, że *„(...) nie jest konieczne posiadanie przez nich wiedzy i kompetencji na poziomie równoważnym do posiadanego przez pracowników komórki zajmującej się doradztwem inwestycyjnym czy służb IT, lecz takim, który pozwoli na zrozumienie istoty procesu (...) oraz dokonanie jego oceny"*. Weryfikacja oraz podnoszenie kwalifikacji podlega zasadniczo ogólnym zasadom określonym w odrębnych przepisach.

Jednocześnie w ocenie Urzędu odpowiednie procedury i rozwiązania organizacyjne, techniczne i dotyczące stosunków umownych z klientami powinny uwzględniać wszystkie etapy (cykle) związane z usługą robo-doradztwa, co obejmuje zarówno

[83] F. D'Acunto, N. Prabhala, A. G. Rossi, *The Promises and Pitfalls of Robo-Advising*, The Review of Financial Studies, v. 32, no. 5, 2019, s. 2017-2018.

[84] Ryzyka typowe dla samych modeli robo-doradztwa opisuje m.in. J. W. Lam, *Robo-Advisors: A Portfolio Management Perspective*, April 4, 2016, s. 25. Artykuł dostępny pod adresem: http://economics.yale.edu/sites/default/files/files/Undergraduate/Nominated%20Senior%20Essays/2015-16/Jonathan_Lam_Senior%20Essay%20Revised.pdf (dostęp: 24.04.2022 r.).

etap koncepcyjny, jak i testowanie (oraz walidację), wdrożenie oraz monitorowanie działania rozwiązań. Wszystkie te elementy powinny zostać w odpowiedni sposób udokumentowany zgodnie z wewnętrznymi procedurami. Zwrócić tutaj należy szczególną uwagę na kwestię dokumentacji technicznej, która powinna obejmować nie tylko samą instrukcję użytkowania, ale wszystkie zmiany techniczne, w tym aktualizacje oprogramowania. Warto nadmienić, że choć nie jest to (na dzień sporządzania niniejszej opinii) wymóg prawny, to zasadnym jest kierowanie się w tym zakresie wymogami określonymi w art. 13 projektu rozporządzenia w sprawie sztucznej inteligencji oraz w załączniku IV do tego projektu, przy czym należy podkreślić, że te przepisy odnoszą się przede wszystkim do systemów wysokiego ryzyka oraz że mogą ulec istotnej zmianie w toku prac nad aktem. Nie stoi to oczywiście w opozycji do innych wymagań, które mogą wynikać z aktów i regulacji specyficznych dla doradztwa inwestycyjnego.

Należy podkreślić, że UKNF wymaga, aby „(...) została przygotowana odpowiednia dokumentacja (...) procesów. Dokumentacja powinna określać uzasadnienie projektowania, rozwoju i modyfikacji, a także strukturę, zamierzone wyniki, cele i zakres algorytmów – s. 10 Stanowiska. Warto także zaznaczyć, że UKNF zaleca stosowanie rozwiązań automatycznych pozwalających na wykrywanie sprzeczności w udzielanych przez użytkowników odpowiedziach.

UKNF stawia pewne specyficzne wymogi w zakresie stosowanych algorytmów robo-doradztwa (vide pkt 5.9 Stanowiska), jednakże nie są to wymogi zasadniczo odbiegające od „standardowych" wymogów dla stosowania tzw. systemów uczenia maszynowego (dokumentacja procesów testowania, walidacji i wdrażania czy wykrywania błędów) oraz wymogów specyficznych dla doradztwa inwestycyjnego.

Urząd rekomenduje także, aby algorytmy były poddane odpowiedniemu testowaniu, które powinno obejmować nie tylko aspekty techniczne, ale również, czy zaprojektowana przez bank metodologia, także w zakresie oceny odpowiedniości klientów, jest „(...) odpowiednia, algorytm jest prawidłowo skonstruowany, a wykorzystywane dane są wiarygodne i rzetelne". Należy wprowadzić także środki bezpieczeństwa uniemożliwiające np. manipulację algorytmem czy wprowadzanie zmian. Istotny będzie tutaj obszar zarządzania ryzykiem, a więc m.in. odpowiednie stosowanie Rekomendacji D KNF, jak również aktów wykonawczych do ustawy Prawo bankowe. Przegląd i weryfikacja rozwiązań powinny odbywać się z częstotliwością dostosowaną do poziomu ryzyk, jak i zaawansowania rozwiązań, a także być określone w stosownych politykach i procedurach.

Ważne jest tutaj także zastrzeżenie dotyczące outsourcingu (choć nie jest to przedmiotem bardziej szczegółowej analizy), który może mieć kluczowe znaczenie dla prawidłowego funkcjonowania usługi robo-doradztwa, w szczególności w kontekście zasilania modelu odpowiednimi danymi. Zapewnienie spełnienia wszystkich wymogów, w tym np. w kontekście wykorzystania chmury obliczeniowej oraz ciągłości działania i bezpieczeństwa danych jest kluczowe zarówno w kontekście odpowiedzialności regulacyjnej, ale także odszkodowawczej względem klienta za ewentualne szkody.

Istotnym elementem – który jest również adresowany w przepisach art. 13 i art. 52 projektowanego rozporządzenia w sprawie sztucznej inteligencji – jest kwestia przejrzystości względem (potencjalnego) klienta. Oprócz ogólnych wymogów, które są charakterystyczne dla usługi doradztwa inwestycyjnego (katalog został wskazany w pkt 7.2 Stanowiska), UKNF wskazuje na pewne istotne z punktu widzenia zautomatyzowanego doradztwa elementy. Przekaz kierowany do (potencjalnego) klienta tej usługi powinien być skonstruowany w taki sposób, aby *„(...) klient mógł w pełni zrozumieć charakter usługi robo-doradztwa, ryzyka z nim związane (...) oraz podjąć świadomą i przemyślaną decyzję co do zawarcia umowy z firmą inwestycyjną i korzystania z przekazywanych mu rekomendacji"*. UKNF przywołał także w swoim Stanowisku elementy wskazywane przez ESMA (pkt 7.4 ostatni akapit Stanowiska), który wymaga m.in. wyraźnego określenia stopnia udziału człowieka w całym procesie.

Istotne jest przy tym takie sformułowanie przekazu, aby był on zrozumiały dla osoby, która w założeniu nie posiada ani wiedzy, ani kompetencji w zakresie szeroko rozumianej sztucznej inteligencji, a więc aspekty techniczne rozwiązania należy ograniczyć do minimum. Jednocześnie UKNF wymaga, aby w regulaminie i/lub umowie o świadczenie usługi znalazły się pewne treści charakterystyczne dla usługi robo-doradztwa (vide pkt 9.4 Stanowiska), do których należy (poniżej zaprezentowane zostały główne elementy istotne z punktu widzenia wykorzystania tzw. systemów sztucznej inteligencji):

1. charakterystyka usługi robo-doradztwa ze wskazaniem skali zaangażowania człowieka i algorytmów,

2. informacja o rodzaju doradztwa inwestycyjnego i częstotliwość udzielania rekomendacji, a także wskazanie elementów immanentnych dla samej usługi doradztwa inwestycyjnego, które zostały określone w pkt 9.4,

3. opis wykorzystywanych algorytmów, ograniczenia i ryzyka związane z algorytmami,

4. określenie źródeł i rodzajów danych wykorzystywanych przez algorytm,

5. wskazanie sposobu postępowania w przypadku awarii czy błędów w działaniu algorytmów.

UKNF nie wskazał jednocześnie szczegółowych oczekiwań w tym zakresie, pozostawiając decyzję co do zakresu i sposobu prezentacji podmiotowi świadczącemu usługę. Należy jednak poczynić pewne uwagi na tym tle.

Wydaje się zasadnym przyjąć, że w tym przypadku nie chodzi o ujawnienie informacji stanowiących tajemnicę przedsiębiorstwa[85], a jedynie takich informacji, które mają znaczenie z perspektywy odbiorcy, a więc klienta banku (firmy inwestycyjnej). W takiej sytuacji chodzi więc o ogólne określenie, że rozwiązanie opiera się o zautomatyzowane przetwarzanie danych wykorzystujące przykładowo uczenie maszynowe oraz wskazanie w jaki sposób (i z jakich źródeł) pozyskiwane są dane i jak ich przetwarzanie przekłada się na konkretne rekomendacje. Określić należy też rolę, jaką pełni człowiek w tym procesie, a także jakie możliwości w zakresie interakcji z człowiekiem przysługują w tym kontekście.

Z powyższej analizy wynika, że świadczenie usługi robo-doradztwo zasadniczo podlega obowiązkom i wymogom dla „klasycznego" doradztwa inwestycyjnego, choć oczywiście niektóre elementy zostały rozszerzone o dodatkowe wymogi związane ze zautomatyzowanym świadczeniem robo-doradztwa, w tym w zakresie spełnienia wymogów szeroko pojętego IT, którym banki i tak podlegają na mocy odrębnych przepisów i regulacji, jak np. Rekomendacja D UKNF.

Jednocześnie należy zwrócić uwagę na pewne charakterystyczne dla tej usługi ryzyka, które niekoniecznie muszą pojawić się w przypadku tradycyjnego świadczenia usługi doradztwa inwestycyjnego[86]. Przede wszystkim chodzi tutaj o kwestię dostępu do danych, jak również właściwego łączenia przez algorytm indywidualnych potrzeb klienta (zebranych na podstawie ankiety odpowiedniości) z konkretnymi instrumentami finansowymi. Brak należytego nadzoru nad tym obszarem, jak również korygowanie ewentualnych niedopasowań czy też nawet pewnej stronniczości może skutkować dotkliwymi stratami po stronie klienta, jak również być źródłem odpowiedzialności regulacyjnej, jeżeli nie zostaną zastosowane odpowiednie rozwiązania w tym zakresie. Ważna jest też tutaj analiza tego, w jaki sposób algorytm dokonuje oceny odpowiedniości na bazie posiadanych przez bank informacji, co ma szczególne znaczenie w przypadku algorytmów i modeli uczących się w czasie rzeczywistym. Nie można także zapominać o zastosowaniu odpowiednich rozwiązań w zakresie tzw. *kill switch*, czyli możliwości wyłączenia systemu i przekierowania usługi na model manualny.

[85] S. Chesterman, *We, The Robots? Regulating Artificial Intelligence and the Limits of Law*, Cambridge 2021, s. 144 i następne.

[86] L. Riulin, *Research on Financial Risks of Robo-Advisor Platforms*, E3S Web of Conferences 218, 01035, 2020, s. 2-3.

ROZDZIAŁ IV

PROGRAM ANALITYCZNO BADAWCZY

Zagadnienia szczegółowe związane z wykorzystaniem systemów sztucznej inteligencji w sektorze bankowym

4.1. Strategia w zakresie sztucznej inteligencji lub danych

Jednym z kluczowych dokumentów mających znaczenie z perspektywy instytucji zamierzającej stosować rozwiązania oparte o tzw. sztuczną inteligencję oraz dane jest odpowiednia strategia[87] zawierające cele krótko-, średnio– i długoterminowe. Wymóg posiadania strategii w zakresie ICT, której częścią niewątpliwie są wskazane rozwiązania, wynika chociażby z Wytycznych EUNB w sprawie zarządzania ryzykiem związanym z technologiami i bezpieczeństwem ICT[88] (vide pkt. 1.2.2). Na potrzebę opracowania takiej strategii wskazuje także BaFin (niemiecki organ nadzoru nad rynkiem finansowym) w swoich zasadach dotyczących stosowania wybranych rozwiązań w zakresie tzw. sztucznej inteligencji[89].

Zakres informacji określonych w strategii będzie uzależniony zarówno od skali działalności instytucji finansowej, jak również swoistego „apetytu" na wdrażanie takich rozwiązań, ale niewątpliwie powinna ona zawierać wizję, cele oraz instrumenty (sposoby jej realizacji), a także zasoby, które pozwolą na jej realizację. Taka strategia powinna być w odpowiedni sposób komunikowana wewnątrz banku i przyczyniać się przez to do tworzenia innowacyjnej kultury opartej o (bezpieczne) wykorzystywanie danych.

W celu realizacji takiej strategii, w instytucji warto rozważyć również powołanie odpowiedniego zespołu projektowego zgodnie z rekomendacjami EBA określonymi w przytoczonych już wytycznych (pkt 1.6.1). Rekomendowane jest, aby takie zespoły miały charakter interdyscyplinarny, a więc uwzględniały zróżnicowane kompetencje, wiedzę i doświadczenie, a także różnorodność.

Strategia powinna być dokumentem przyjętym przez organ zarządzający banku i poddawana regularnej ewaluacji oraz ewentualnej aktualizacji zgodnie z przyjętymi w banku rozwiązaniami.

[87] P. Kirby et. al., *Designing Data Strategies: A Playbook for Action*, October 2020, Development Gateway: Washington, DC, s. 11.
[88] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Guidelines/2020/GLs%20on%20ICT%20and%20security%20risk%20management/Updated%20Translations/880823/Final%20draft%20Guidelines%20on%20ICT%20and%20security%20risk%20management_COR_PL.pdf (dostęp: 26.04.2022 r.).
[89] BaFin, *Big Data and Artificial Intelligence: Principles for the Use of Algorithms in Decision-Making Processes*, 2021, s. 8.

4.2. Odpowiedzialność banku za wykorzystywane rozwiązania z obszaru sztucznej inteligencji

4.2.1. Odpowiedzialność wewnętrzna

Jednym z kluczowych zagadnień w kontekście wykorzystania systemów sztucznej inteligencji jest kwestia odpowiedzialności (*accountability*), która ma wymiar zarówno zewnętrzny, jak i wewnętrzny, a prawidłowe ułożenie zasad w tym zakresie jest istotne dla prawidłowego funkcjonowania rozwiązań w banku. Odpowiedzialność w tym kontekście odnosi się do możliwości ustalenia, czy dana decyzja została podjęta zgodnie z odpowiednimi wymogami oraz do przypisania odpowiedzialności, jeśli te standardy nie zostały spełnione[90]. Jest to o tyle istotne, że obecnie nie ma przesłanek do tego, aby przypisywać odpowiedzialność algorytmom czy modelom sztucznej inteligencji[91], a więc to człowiek odpowiada za naruszenia, których takie rozwiązanie mogło „się dopuścić".

Oznacza to, że w ramach instytucji finansowej wdrażającej rozwiązania oparte o AI muszą funkcjonować odpowiednie rozwiązania organizacyjne i techniczne, w tym w zakresie raportowania i komunikacji, które zapewnią, że ewentualne błędy czy incydenty zostaną w efektywny sposób rozwiązane, a także że będzie możliwe jednoznaczne ustalenie odpowiedzialności za ewentualne naruszenia. Jednocześnie zasada nie ma na celu „wskazywania winnych", a ma raczej charakter prewencyjny, zapewniający, że nawet w przypadku do zmaterializowania się określonego ryzyka możliwe będzie skuteczne rozwiązanie napotkanego problemu, np. niewłaściwego działania systemu sztucznej inteligencji.

Warto jednocześnie zwrócić uwagę, że projekt rozporządzenia w sprawie sztucznej inteligencji w pewnym stopniu odnosi się do kwestii odpowiedzialności wewnętrznej, m.in. w art. 14, który odnosi się do nadzoru człowieka nad działaniem systemów, np. w kontekście konieczności posiadania tzw. „stop button", czyli rozwiązania pozwalającego na zatrzymanie systemu sztucznej inteligencji, jeżeli zajdzie taka potrzeba (należy przy tym zaznaczyć, że konieczne jest też wdrożenie odpowiednich środków awaryjnych pozwalających na przełączenie się na procesy manualne lub zapasowe).

Sam wymiar odpowiedzialności wewnętrznej jest zagadnieniem wielowątkowym, który można podzielić na następujące obszary:

[90] F. Doshi-Velez, M. Kortz, R. Budish et. al., *Accountability of AI Under the Law: The Role of Explanation*, 2019. Artykuł dostępny pod adresem: https://arxiv.org/pdf/1711.01134 (dostęp: 26.04.2022 r.).
[91] M. Coeckelbergh, *AI Ethics*, MIT 2020, s. 109 i następne.

1. rozwiązania organizacyjne,

2. rozwiązania techniczne,

3. kultura organizacji.

Pierwszy punkt odnosi się do konieczności takiego ułożenia struktury banku, aby odzwierciedlała ona wszystkie aspekty odpowiedzialności za obszar systemów sztucznej inteligencji oraz danych, zarówno na poziomie zarządu, jak i pozostałych poziomów w organizacji. Wymaga to po pierwsze wyznaczenia członka zarządu odpowiedzialnego, np. CIO, CDO etc., jak również stworzenia odpowiedniej struktury departamentalnej, w tym wyznaczenia pracowników odpowiedzialnych za specyficzne obszary związane z wykorzystaniem AI. Zauważyć przy tym należy, że stosowanie tych rozwiązań ma charakter złożony, obejmujący nie tylko ściśle techniczne wątki, ale także obszar etyki, prawa i regulacji, które powinny być uwzględnione w tym kontekście, także kontroli wewnętrznej. Zmiany te powinny być oczywiście w odpowiedni sposób udokumentowane zarówno na poziomie struktury organizacyjnej, jak i regulaminów (organizacyjnych) oraz polityk i procedur. Wyznaczenie właścicieli poszczególnych procesów powinno stanowić dla instytucji kluczowy element procesu wdrażania, toteż opracowanie stosownej mapy procesów również powinno zostać dokonane. Ważne jest również ułożenie odpowiednich kanałów komunikacji i raportowania, w tym do jednostek odpowiedzialnych za raportowanie zewnętrzne, np. do CSIRT czy UKNF.

Warto zwrócić uwagę na jedną istotną kwestię często pomijaną w kontekście aspektu odpowiedzialności. Wykorzystanie systemów sztucznej inteligencji wiąże się zazwyczaj z wykorzystaniem danych osobowych oraz zastosowaniem innowacyjnych rozwiązań i może wiązać się z koniecznością przeprowadzenia oceny skutków dla ochrony danych[92], o którym mowa w Rozporządzeniu 2016/679. Zaangażowanie jednostek odpowiedzialnych za obszar ochrony danych i prywatności jest więc absolutną koniecznością po stronie instytucji.

W aspekcie technicznym należy zwrócić uwagę na dostarczenie osobom odpowiedzialnym – wyznaczonym w ramach punktu 1 – narzędzi, które umożliwią im realizację ich zadań, jak np., wdrożenie systemów automatycznego zbierania informacji i przekazywania alertów czy też możliwości składania raportów na wyższy poziom.

Swoistą klamrą spinającą powyższe wątki jest kwestia budowania w banku kultury innowacyjności oraz odpowiedzialności. Jest to wątek wykraczający daleko poza zakres niniejszego opracowania, jednakże warto podkreślić, że zapewnienie ram kulturowych dla rozwoju i promowania innowacyjności oraz zrozumienia istoty przetwarzania danych z użyciem nowych rozwiązań może być kluczowe dla realizacji postulatu odpowiedzialności wewnętrznej. Taka kultura powinna być dostosowana do pewnych podstawowych zasad, na których opiera się działanie danej instytucji i powinna być promowana przez kadrę kierowniczą oraz zarząd na niższych poziomach organizacji. W tym zakresie pomocna może być również wspomniana już strategia dla danych lub sztucznej inteligencji, która określa także oczekiwania względem pracowników instytucji.

Na koniec należy wskazać, że odpowiednie ułożenie odpowiedzialności wewnętrznej w organizacji będzie miało znaczenie także w aspekcie zewnętrznym, który został opisany w kolejnym podrozdziale.

4.2.2. Odpowiedzialność zewnętrzna

Kwestia odpowiedzialności zewnętrznej jest zagadnieniem wielowątkowym i nie dość rozstrzygniętym na poziomie przepisów prawa. Ramy odpowiedzialności specyficzne dla systemów sztucznej inteligencji podlegają obecnie analizom i dyskusjom na poziomie Unii Europejskiej, a przyjęcie określonych rozwiązań będzie miało także znaczenie w kontekście krajowych porządków prawnych[93]. W obecnym stanie prawnym nie funkcjonują przepisy, które dostatecznie regulowałyby tę tematykę.

Pisząc o odpowiedzialności zewnętrznej za systemy sztucznej inteligencji mamy na myśli odpowiedzialność (producenta, operatora, użytkownika) takich systemów względem użytkowników końcowych (oraz organów nadzoru), na których może ono wpływać zarówno w pozytywny, jak i negatywny sposób, uaktywniając niejako odpowiedzialność odszkodowawczą, choć odpowiedzialność ta może przyjąć także inne formy, jak np. konieczność dokonania określonej czynności czy przedstawienia wyjaśnienia. Sprawa komplikuje się także wobec faktu, że banki poddane są szczególnym przepisom prawnym i regulacjom, które mogą wprowadzać zmodyfikowane reżimy odpowiedzialności (zazwyczaj wyższe), jak również zobowiązywać do wyższych standardów staranności. Nawet więc w przypadku braku przepisów charakterystycznych dla sztucznej inteligencji możliwe jest przyjęcie odpowiedzialności na już istniejących podstawach prawnych.

[92] https://monitorpolski.gov.pl/M2018000082701.pdf (dostęp: 26.04.2022 r.).

[93] https://ec.europa.eu/growth/news/commission-collects-views-making-liability-rules-fit-digital-age-artificial-intelligence-and-2021-10-20_en (dostęp: 28.04.2022 r.).

Wspomniane już prace nad stworzeniem ram prawnych specyficznych dla systemów sztucznej inteligencji nie będą przedmiotem niniejszej analizy, gdyż nie jest jasny ostateczny kierunek działań. Można w tym miejscu jedynie wskazać, że wskazuje się na kilka możliwych rozwiązań, w tym konieczność wprowadzenia specyficznej odpowiedzialności solidarnej, odpowiedzialności zbliżonej jak w przypadku produktu niebezpiecznego czy wprowadzenie zróżnicowanych reżimów w zależności od „ryzykowności" określonego systemu.

Wracając jednak na grunt odpowiedzialności zewnętrznej za te systemy, należy wskazać, że bank może występować w bardzo zróżnicowanych rolach i względem różnych podmiotów. Przykładowo bank może być twórcą algorytmu czy modelu sztucznej inteligencji, który wykorzystuje samodzielnie względem swoich klientów albo udostępnia, np. za pomocą cyfrowej platformy czy interfejsu dostępowego API innym podmiotom. Bank może być też użytkownikiem systemu opracowanego przez inny podmiot. Takie rozwiązania mogą z oczywistych względów generować zróżnicowane ryzyka dla użytkowników, np. w kontekście niedozwolonej dyskryminacji czy zawodności systemów transakcyjnych. Stosowane przez banki rozwiązania oparte o tzw. sztuczną inteligencję mogą także generować ryzyko wewnątrz banku, np. w kontekście modeli wewnętrznych odpowiedzialnych za spełnienie wymogów ostrożnościowych. Niewłaściwe ich działanie może narazić bank na odpowiedzialność względem organu nadzorczego i to nie tylko KNF, ale także Prezesa Urzędu Ochrony Danych Osobowych.

Zakres tej odpowiedzialności może być więc zróżnicowany, choć w przypadku tego typu rozwiązań często postuluje się odpowiedzialność na zasadzie ryzyka. Na ocenę samej odpowiedzialności, w tym przyczynienia się banku do określonego niechcianego zdarzenia, wpływ mogą mieć takie kwestie jak nieodpowiednie rozwiązania organizacyjno-techniczne czy też niezastosowanie się do instrukcji stosowania systemu. Wraz z ewentualnym wejściem w życie rozporządzenia w sprawie sztucznej inteligencji obowiązki te będą podlegały także dodatkowym sankcjom administracyjnym, nawet do 30 mln euro.

Dziś, wobec braku specyficznych rozwiązań dla systemów sztucznej inteligencji, przyjąć należy, że zastosowanie będą miały przede wszystkim przepisy Kodeksu cywilnego i ewentualne przepisy specyficzne dla sektora bankowego.

Poniżej prezentujemy bardziej szczegółową analizę zagadnienia.

W pewnym uproszczeniu w ekosystemie AI w praktyce banków możemy rozróżnić cztery podmioty, które ze sobą oddziałują. Mowa tutaj o banku, kliencie banku (konsumencie bądź przedsiębiorcy), organie nadzoru oraz dostawcy usług opartych o sztuczną inteligencję (w skrócie dostawcy AI).

Pierwsza relacja, o charakterze cywilnoprawnym, łączy bank oraz jego klienta. Nie jest ona jednak regulowana tylko i wyłącznie przepisami o charakterze cywilnoprawnym. Na banku ciąży bowiem obowiązek przestrzegania szeregu wymagań prawnych o charakterze administracyjnoprawnym, których egzekwowanie jest nadzorowane przez właściwy organ – przede wszystkim Komisję Nadzoru Finansowego. Nadto istnieje relacja między bankiem a dostawcą AI, co oznacza w większości sytuacji typowych korzystanie z wyspecjalizowanego podmiotu trzeciego. W ten sposób powstaje pomiędzy tym podmiotem a bankiem stosunek cywilnoprawny. Jednocześnie, w obydwu tych relacjach na banku ciąży obowiązek wykazania spełnienia szeregu wymogów prawnych jak np. w zakresie outsourcingu bankowego[94] czy wykorzystania chmury obliczeniowej[95].

W zakresie wspomnianych relacji bank może ponosić dwa rodzaje odpowiedzialności: **cywilnoprawną** lub **administracyjnoprawną**.

4.3. Obecny stan prawny

4.3.1. Rozważania kolizyjnoprawne

W pierwszej kolejności rozważania o cywilnoprawnej odpowiedzialności banków za stosowanie sztucznej inteligencji należy odnieść bezpośrednio do norm prawa prywatnego międzynarodowego obowiązującego w Unii Europejskiej. Relacja pomiędzy bankami a dostawcami usług czy klientami banków, nie jest pozbawiona tzw. **elementu międzynarodowego**, który może decydować o zastosowaniu prawa innego niż prawa państwa, w którym bank prowadzi swoją główną działalność. Najprostszym przykładem może być tutaj sytuacja, w której polski bank korzysta z systemu AI na podstawie umowy zawartej ze spółką posiadającą swoją siedzibę np. w Niemczech, Chinach czy USA. O ile w tym zakresie strony stosunku umownego mają możliwość wyboru prawa właściwego, ze względu na art. 3 Rozporządzenia Rzym I[96], o tyle będzie to już

[94] Zob. art. 6a i n. PrBank oraz Wytyczne EBA z dnia 25 lutego 2019 r. w sprawie outsourcingu, EBA/GL/2019/02.

[95] Zob. Komunikat urzędu Komisji Nadzoru Finansowego z dnia 23 stycznia 2020 r. dotyczący przetwarzania przez podmioty nadzorowane informacji w chmurze obliczeniowej publicznej lub hybrydowej.

96 Rozporządzenie Parlamentu Europejskiego i Rady (WE) nr 593/2008 z dnia 17 czerwca 2008 r. w sprawie prawa właściwego dla zobowiązań umownych (Rzym I) (Dz. U. UE. L. z 2008 r. Nr 177, str. 6 z późn. zm.).

utrudnione w relacjach banku z jego klientem, gdzie w związku z wykorzystaniem wobec niego systemu AI poza ponoszeniem odpowiedzialności kontraktowej, bank może odpowiadać także z deliktu (czynów niedozwolonych).

W pierwszej kolejności trzeba zasygnalizować, że już obecnie, w zasadzie w przededniu ery powszechnego wykorzystania AI, to nie zawsze bank ma silniejszą pozycję negocjacyjną z zagranicznym dostawcą szeroko pojętych usług informatycznych czy oprogramowania. Nierzadko bowiem okazuje się, że bank, by dotrzymać kroku konkurencji, nie ma zbyt wielkiego wyboru w zakresie dostawców usług, co w zasadzie pozwala takiemu dostawcy na dość swobodne kształtowanie warunków umownych, w tym także w zakresie prawa właściwego. Naraża to zatem bank na poruszanie się często w mniej przewidywalnych oraz kosztowniejszych niż prawo polskie systemach prawa. Wydaje się, że w wypadku stosowania systemów AI będziemy świadkiem tego samego problemu.

W drugiej kolejności należy wskazać, że w relacji pomiędzy bankiem a klientem możliwość umownego uregulowania prawa właściwego nie zawsze będzie możliwa. Szczególnie chodzi o sytuację, w której może dojść do zbiegu odpowiedzialności deliktowej i kontraktowej[97]. Będzie to miało miejsce m.in. w przypadku naruszenia dóbr osobistych klienta banku (np. prawa do prywatności). W tym przypadku najczęstszym łącznikiem wskazującym właściwość prawa jest łącznik prawa ojczystego takiej osoby, ewentualnie z możliwością wyboru prawa właściwego państwa, na którego terytorium doszło do naruszenia[98]. Nie można jednak wykluczyć, że wobec wyjścia Wielkiej Brytanii z UE przepisy te zostaną ujednolicone na poziomie unijnym – prawdopodobnie zmieniony zostanie wtedy łącznik z prawa ojczystego na prawo państwa zwykłego pobytu osoby fizycznej[99].

4.3.2. Koncepcje cywilnoprawnej odpowiedzialności za wykorzystanie AI

Na podstawie obecnie obowiązujących przepisów nie jest możliwe ustalenie jednego właściwego reżimu odpowiedzialności cywilnej za szkody wyrządzane przez czy przy użyciu systemów AI. Co więcej, nie jest na ten moment jasne, który z podmiotów uczestniczących w projektowaniu, wdrażaniu i wykorzystywaniu systemu AI miałby zostać pociągnięty do odpowiedzialności w przypadku powstania szkody wynikającej z zastosowania takiego systemu. Pojawia się zatem istotne pytanie, na ile istniejące regulacje prawne są wystarczające dla normowania odpowiedzialności cywilnoprawnej za sztuczną inteligencję.

Podstawową zasadą odpowiedzialności cywilnej m.in. w polskim systemie prawnym jest zasada winy. Odnosi się to zarówno do reżimu deliktowego, jak i kontraktowego[100]. Skuteczne dochodzenie roszczeń wymaga w takim wypadku zarówno wykazania zdarzenia wywołującego szkodę, szkody, związku przyczynowego pomiędzy tymi dwoma oraz winy sprawcy – wynikającej bądź z czynu niedozwolonego, bądź z nienależytego wykonania zobowiązań umownych.

Nie ulega wątpliwości, że Bank jako operator[101] systemu AI może ponosić odpowiedzialność zgodnie z powyższym w razie zawinionego działania lub zaniechania. W reżimie odpowiedzialności kontraktowej będzie to kwestia niezachowania należytej staranności wymaganej od profesjonalisty.

Koncepcja ponoszenia przez operatora AI odpowiedzialności na zasadzie winy jest jednak mało popularna. Mogłoby to doprowadzić do przesadnego zawężenia odpowiedzialności takich podmiotów, można bowiem przypuszczać, że w wielu przypadkach operatorzy systemów AI byliby w stanie wykazać fakt dołożenia należytej staranności w zakresie istniejących standardów technicznych, etycznych lub prawnych, podczas gdy szkoda wynikałaby z nieprzewidzianego (i niemożliwego do przewidzenia) błędu w działaniu systemu AI. W praktyce mogłoby to oznaczać, że operatorzy AI nie ponosiliby odpowiedzialności wcale, co jest trudne do zaakceptowania[102].

Prawo polskie i większość reżimów prawnych w Europie przewiduje także odpowiedzialność na zasadzie ryzyka. Jest to odpowiedzialność za niebezpieczeństwo wynikające z danego typu działalności. Taki reżim przewidziany jest chociażby dla podmiotów prowadzących przedsiębiorstwo lub zakład wprawiany w ruch za pomocą sił przyrody[103]. Zwolnienie z odpowiedzialności na zasadzie ryzyka jest możliwe jedynie na podstawie konkretnych, sformułowanych w przepisach przesłanek egzoneracyjnych. Wydaje

[97] Zob. szczególnie art. 443 KC. Na marginesie należy zauważyć, że często reżim deliktowy może być korzystniejszy ze względu na możliwość żądania zapłaty zadośćuczynienia.
[98] Zob. chociażby art. 16 PPM. Warto podkreślić, że kwestia naruszenia dóbr osobistych nie jest zharmonizowana na poziomie unijnym.
[99] Zob. więcej: P. Mostowik, E. Figura-Góralczyk, *art. 16 PPM* [w:] *Prawo prywatne międzynarodowe. Komentarz*, red. M. Pazdan, Warszawa 2018 r.
[100] Zob. art. 415 oraz art. 471 KC.
[101] Zgodnie z definicją zawartą w projekcie AIA (art. 3 pkt 8): operator oznacza dostawcę systemu AI, jego użytkownika, upoważnionego przedstawiciela, importera i dystrybutora.
[102] M. Jagielska, *Odpowiedzialność za sztuczną inteligencję* [w:] L. Lai, M. Świerczyński (red.), *Prawo sztucznej inteligencji*, Warszawa 2020.
[103] Zob. art. 435 KC. Przykładami takich przedsiębiorstw lub zakładów są gospodarstwo rolne, zakład górniczy, lotnisko czy przedsiębiorstwo budowlane.

się słusznym kierunek postulowany w wielu opracowaniach tego zagadnienia w ramach UE, aby operatorzy AI odpowiadali za szkody spowodowane działaniem tych systemów właśnie na zasadzie ryzyka lub jak za produkt niebezpieczny. Obecnie jednak brak ku temu jednoznacznych podstaw.

Odpowiedzialność za produkt niebezpieczny jest regulowana na poziomie unijnym, zwłaszcza w ramach dyrektywy 85/374 o odpowiedzialności za produkty niebezpieczne[104]. Należy mieć na uwadze, że dyrektywa została przyjęta jeszcze w latach 80. XX w., co utrudnia jej stosowanie w sytuacjach obejmujących wykorzystanie AI. Problemem jest m.in. brak objęcia tymi przepisami oprogramowania oraz fakt, że standardowo producent traci kontrolę nad produktem po jego wprowadzeniu do obrotu. W przypadku systemów AI sytuacja jest o tyle złożona, że z jednej strony system taki, jako rodzaj oprogramowania może podlegać obowiązkowym aktualizacjom zapewnianym przez dostawcę systemu, ale z drugiej strony – ucząc się w środowisku konkretnego użytkownika (np. banku) podlega zmianom.

Podkreślić jednak trzeba, że Komisja Europejska w swoim raporcie w sprawie wpływu AI, Internetu Rzeczy i robotyki na bezpieczeństwo i odpowiedzialność[105], zwraca uwagę na możliwość stosowania przepisów o produkcie niebezpiecznym do systemów AI, a nawet zaleca taki sposób postępowania.

Przyjmując tę koncepcję należałoby uznać, że do odpowiedzialności za AI zastosowanie znajdą przepisy art. 449^1 – 449^{11} KC. Podmiotem odpowiedzialnym jest przede wszystkim producent[106], rozumiany jako producent produktu gotowego, producent każdego surowca lub producent części składowej, jak i każda osoba, która przedstawia się jako producent umieszczając swą nazwę, znak handlowy lub inną wyróżniającą cechę na produkcie[107]. W odniesieniu do AI trudno mówić o jednym tylko producencie. W procesie tworzenia sztucznej inteligencji udział biorą różne podmioty, jak producenci rzeczy, twórcy oprogramowania czy osoby wdrażające system[108]. Podmioty te będą odpowiedzialne, gdy produkt jest wprowadzany do obrotu całościowo, tj. rzecz i oprogramowanie razem. Problematyczna jest sytuacja, w której produkt końcowy jest „składany" z elementów pochodzących od różnych wytwórców – w tym zakresie brak obecnie odpowiednich regulacji. Wydaje się, że należałoby rozważyć wprowadzenie odpowiedzialności solidarnej[109].

Produktem w myśl dyrektywy jest każda rzecz ruchoma, nawet będąca częścią składową innej rzeczy ruchomej lub nieruchomej[110]. Problematyczne może wydawać się uznanie oprogramowania (i tym samym AI) za produkt w rozumieniu tej definicji, ale wydaje się, że istnieją podstawy do rozszerzającej interpretacji, choćby z uwagi na stanowisko Komisji wyrażonej w omawianym raporcie. Produkt jest niebezpieczny, jeżeli nie zapewnia bezpieczeństwa, jakiego osoba ma prawo oczekiwać, biorąc pod uwagę wszystkie okoliczności, w szczególności sposób użycia produktu, którego można rozsądnie oczekiwać czy czas, w którym produkt został wprowadzony do obrotu[111]. Ze względu na możliwość uczenia się przez AI, trudne jest jednoznaczne stwierdzenie, czy można ją uznać za produkt niebezpieczny w kontekście obecnie obowiązujących przepisów. Podobnie niełatwo wyznaczyć odpowiedni standard oczekiwań odnośnie do bezpieczeństwa.

Producent nie odpowiada za produkt niebezpieczny m.in., jeżeli wada, która spowodowała szkodę, nie istniała w momencie wprowadzenia produktu do obrotu lub jeżeli wada powstała później. Zastosowanie tej przesłanki egzoneracyjne w kontekście sztucznej inteligencji jest problematyczne, gdyż system AI z definicji podlega ciągłym zmianom. Problem łatwego zwolnienia się z odpowiedzialności producenta systemu AI pojawia się także w związku z tym, że nie on odpowiada, jeżeli stan wiedzy naukowej i technicznej w momencie wprowadzenia przez niego produktu do obrotu nie pozwalał na wykrycie istnienia wady. Względy dla jej niestosowania są podobne jak w przypadku pierwszej[112].

Reżim odpowiedzialności za produkt niebezpieczny obejmuje szkody spowodowane przez śmierć lub przez uszkodzenie ciała oraz uszkodzenie lub zniszczenie każdej rzeczy innej niż produkt wadliwy o wartości powyżej 500 euro[113]. Należy w związku z tym mieć na uwadze, że nie będzie możliwe dochodzenie

[104] Dyrektywa Rady z 25.07.1985 r. w sprawie zbliżenia przepisów ustawowych, wykonawczych i administracyjnych Państw Członkowskich dotyczących odpowiedzialności za produkty wadliwe, Dz. Urz. L Nr 210, s. 29–33 z 7.8.1985 r.

[105] Report on the Safety and Liability Implication of Artificial Intelligence, the Internet of Thing and Robotics. Opracowanie dostępne pod adresem: https://ec.europa.eu/info/sites/default/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf (dostęp: 04.05.2022 r.).

[106] Kwestię komplikuje fakt, że w odniesieniu do systemów AI, projekt AIA nie używa pojęcia producenta, posługując się w tym kontekście wyłącznie pojęciem dostawcy, rozumianego jako osoba fizyczna lub prawna, organ publiczny, agencja lub inny podmiot, które opracowują system sztucznej inteligencji lub zlecają jego opracowanie w celu wprowadzenia go do obrotu lub oddania go do użytku pod własną nazwą handlową lub własnym znakiem towarowym – odpłatnie lub nieodpłatnie,

[107] Zob. art. 3 dyrektywy 85/374.

[108] M. Jagielska, *op. cit.*

[109] *Ibidem.*

[110] Zob. art. 2 dyrektywy 85/374.

[111] Zob. art. 6 dyrektywy 85/374.

[112] Zob. art. 7 dyrektywy 85/374 i art. 449^3 KC.

[113] O ile rzecz ta jest zwykle przeznaczona do prywatnego użytku lub konsumpcji i była używana przez osobę poszkodowaną głównie dla jej prywatnego użytku lub konsumpcji; zob. art. 9 dyrektywy 85/374 i art. 449^7 KC.

szkód w samym produkcie wyposażonym w sztuczną inteligencję, a także szkód czysto ekonomicznych, takich jak straty na rynkach finansowych czy utrata lub naruszenie integralności danych[114]. Wysoce kłopotliwe jest ustalenie związku przyczynowego między szkodą, a „zachowaniem" systemu AI. Już obecnie w doktrynie prawa wymienia się przykładowe problemy, które mogą pojawić się w procesie ustalania istnienia związku przyczynowego:

- czy dana szkoda została spowodowana przez jedną pierwotną przyczynę, czy przez współdziałanie wielu przyczyn?
- czy szkoda jest efektem uczenia maszynowego bez zmiany oprogramowania?
- czy system AI błędnie odczytał poprawne dane, a może otrzymał nieprawidłowe dane?
- czy aktualizacja oprogramowania wykonana przez pierwotnego producenta lub osobę trzecią była wadliwa?
- czy użytkownik nie zainstalował aktualizacji, która zapobiegłaby szkodzie?[115]

Tytułem uzupełnienia kwestii odpowiedzialności cywilnej za szkody wyrządzone przy użyciu lub przez AI, warto zauważyć, że podstawę do dochodzenia roszczeń zarówno od administratora danych, jak i podmiotu przetwarzającego (procesora) przewiduje również RODO w art. 82 i nast. Odszkodowania może domagać się osoba, która poniosła szkodę (rozumianą szeroko, także jako krzywdę) w wyniku naruszenia przepisów rozporządzenia[116]. Sztuczna inteligencja opiera się na wykorzystaniu danych, także osobowych, więc stosunkowo łatwo wyobrazić sobie sytuację, w której dojdzie do naruszenia praw osoby fizycznej. Nie można również zapomnieć o ochronie dóbr osobistych, przewidzianej w Kodeksie cywilnym[117]. Katalog dóbr osobistych ma charakter otwarty, a do ich naruszenia niewątpliwie może dojść w wyniku zastosowania systemu AI.

4.3.3. Administracyjnoprawna odpowiedzialność za wykorzystanie AI

Banki podlegają nadzorowi Komisji Nadzoru Finansowego, co wiąże się z zagadnieniem odpowiedzialności administracyjnoprawnej. Jeśli wykorzystanie AI odbywa się w ramach działalności bankowej, KNF jest właściwe do podejmowania czynności nadzorczych dotyczących AI. Nadzór odbywa się na podstawie przepisów ustawy – PB[118] oraz ustawy o nadzorze nad rynkiem finansowym[119]. W szczególności KNF może zalecić bankowi podjęcie środków koniecznych do przywrócenia płynności płatniczej lub osiągnięcia i przestrzegania innych norm dopuszczalnego ryzyka w działalności banku, a także opracowanie i stosowanie procedur, które zapewnią funkcjonowanie systemu zarządzania bankiem[120]. Niestosowanie się do zaleceń i nakazów KNF może skutkować nałożeniem dotkliwych sankcji, także finansowych[121].

Odpowiedzialność administracyjną przewiduje także RODO, które umożliwia Prezesowi Urzędu Ochrony Danych Osobowych nałożenie kar za naruszenie przepisów rozporządzenia. Kara pieniężna może w wypadku naruszenia praw podmiotów danych, m.in. w zakresie art. 22 RODO, sięgać nawet 20 000 000 euro lub 4% rocznego obrotu przedsiębiorstwa[122]. Ten reżim odpowiedzialności administracyjnoprawnej może także znaleźć zastosowanie do wykorzystania systemów AI.

4.4. Propozycje przyszłych regulacji

Projektodawcy AIA w zasadzie nie pochylają się nad zagadnieniem odpowiedzialności cywilnej. Producent ma być co prawda odpowiedzialny za zapewnienie zgodności systemu AI z wymogami stawianymi przez rozporządzenie[123], a dystrybutorzy lub importerzy w określonym zakresie (dotyczącym transportu lub przechowywania) również mają odpowiadać za system[124], jednak nie są to kwestie dotyczące bezpośrednio dochodzenia roszczeń.

W projekcie AIA przewidziano utworzenie **systemu nadzoru**. Państwa członkowskie będą musiały ustanowić (wyznaczyć) właściwe organy krajowe na potrzeby zapewnienia stosowania i wdrażania rozporządzenia, a spośród nich zostanie wybrany co najmniej jeden krajowy organ nadzorczy[125]. Ponadto w odniesieniu do systemów AI zastosowanie znajdą przepisy rozporządzenia w sprawie nadzoru rynku i zgodności

[114] M. Jagielska, *op. cit.*
[115] *Ibidem*.
[116] Zob. art. 82 RODO.
[117] Zob. art. 23 i art. 24 KC.
[118] Ustawa z dnia 29 sierpnia 1997 r. Prawo bankowe (t.j. Dz. U. z 2021 r. poz. 2439 z późn. zm.).
[119] Ustawa z dnia 21 lipca 2006 r. o nadzorze nad rynkiem finansowym (t.j. Dz. U. z 2022 r. poz. 660 z późn. zm.).
[120] Zob. art. 138 ust. 1 pkt 1 i pkt 4 PrBank.
[121] Zob. szczególnie art. 138 ust. 3 pkt 2a i pkt 3a PrBank, gdzie mowa kolejno o karze pieniężnej dla osoby fizycznej do wysokości 21 312 000 zł, oraz o karze dla banku w wysokości do 10% przychodu.
[122] Zob. art. 83 RODO.
[123] Zob. art. 24 AIA.
[124] Zob. art. 26 ust. 4 oraz art. 27 ust. 3 AIA.
[125] Zob. art. 59 AIA.

produktów[126]. W myśl tych przepisów AI będzie produktem stwarzającym ryzyko, co zwiększy dodatkowo liczbę obowiązków i uprawnienia organów nadzorczych[127]. Nieprzestrzegającym przepisów grozić będą kary administracyjne. Mają one mieć zróżnicowany, zależny od rodzaju naruszonego obowiązku, charakter. Ich górna granica może wynosić nawet 30 000 000 euro lub 6% całkowitego rocznego obrotu przedsiębiorstwa[128]. Uzupełniająco warto wskazać na projektowane obecnie rozporządzenie w sprawie produktów maszynowych[129]. Stawiać będzie ono kolejne wymogi, które znajdą zastosowanie do produktów maszynowych, w tym takich, które zawierają systemy sztucznej inteligencji.

Chociaż AIA nie zawiera w sobie (na ten moment) postanowień dotyczących odpowiedzialności cywilnej za AI to temat ten jest szeroko dyskutowany w UE.

Grupa Ekspercka Komisji Europejskiej w dokumencie *Liability For Artificial Intelligence And Other Emerging Digital Technologies* z 2019r.[130] wskazała, że odpowiedzialność na zasadzie winy (niezależnie od istnienia domniemania winy) powinna współistnieć z odpowiedzialnością na zasadzie ryzyka – poszkodowany powinien mieć więcej niż jedną podstawę do dochodzenia roszczeń. Wedle autorów raportu producent powinien być ściśle (na zasadzie ryzyka) odpowiedzialny za wady powstających technologii cyfrowych, nawet jeśli wymienione wady pojawiają się po wprowadzeniu produktu do obrotu, o ile producent nadal kontrolował aktualizacje technologii. Na producentach powinien spoczywać obowiązek wyposażania technologii w środki rejestrowania informacji o działaniu technologii (*logging by desing*) dla celów dowodowych. Jeżeli w grę wchodziłaby większa ilość podmiotów, które stworzyły daną technologię, a poszkodowany może wykazać, że co najmniej jeden element spowodował szkodę, ale nie który element, to wszyscy potencjalni sprawcy powinni być solidarnie odpowiedzialni.

W raporcie pt. *Safety and Liability Related Aspects of Software*[131], sporządzonym na zamówienie KE, autorzy zaproponowali następujące ramy prawne dotyczące odpowiedzialności za AI:

- Zaproponowano stworzenie pojęcie „nieuczciwych praktyk algorytmicznych" zabronionych dla wszystkich zastosowań opartych o AI, niezależnie od stwarzanego przez nie ryzyka. Do ich kategorii należałoby zaliczyć m.in. dyskryminację, wykorzystywanie słabych punktów, czy manipulację.

- Zaproponowano także stworzenie listy z praktykami dotyczącymi danych, które byłyby wprost dozwolone przez uczynienie z nich ustawowej podstawy przetwarzania danych w rozumieniu przepisów RODO.

- Szkody wyrządzonej przez AI często nie można przypisać winie ludzkiej, a ustalenie jej defektu jest znacznie trudniejsze, ponieważ może być trudne odróżnienie szkody wyrządzonej przez wadliwą AI od szkody, która była nieunikniona podczas korzystania z określonej AI. Odpowiedzialność na zasadzie ryzyka jest zatem odpowiednia w przypadkach zastosowań wysokiego ryzyka. Definicja tego, co zalicza się do „wysokiego ryzyka", powinna skupiać się nie tylko na zagrożeniach fizycznych, ale także na czysto ekonomicznych i społecznych.

- Niezależnie uznania zastosowania za należące do grupy „wysokiego ryzyka", osoba wdrażająca AI powinna ponosić odpowiedzialność za nieprawidłowe działanie AI w takim samym stopniu, w jakim byłaby odpowiedzialna za działania lub zaniechania pomocnika.

W sprawozdaniu KE na temat wpływu AI, IoT i robotyki na bezpieczeństwo i odpowiedzialność[132] wskazano, że nie ma potrzeby przeprowadzania kompleksowej rewizji rozwiązań prawnych dotyczących odpowiedzialności AI, ale specyfika AI wymaga pewnych dostosowań i wyjaśnień, aby zapewnić w każdym przypadku wskazanie osoby odpowiedzialnej za jej działanie. Podkreślono, że AI nie jest w pełni autonomiczna i zawsze za jej działania odpowiada człowiek. Zasadniczą propozycją w zakresie ukształtowania odpowiedzialności jest zastosowanie reżimu odpowiedzialności jak dla produktu niebezpiecznego. Odpowiedzialność tę ponosiłby tzw. podmiot wdrażający AI (ang. *deployer*), który:

- podejmuje decyzję o wykorzystaniu AI;

- kontroluje ryzyko związane z AI oraz

- czerpie korzyści z AI.

[126] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2019/1020 z dnia 20 czerwca 2019 r. w sprawie nadzoru rynku i zgodności produktów oraz zmieniające dyrektywę 2004/42/WE oraz rozporządzenia (WE) nr 765/2008 i (UE) nr 305/2011.
[127] Zob. art. 63 i art. 65 AIA.
[128] Zob. art. 71 AIA.
[129] Rozporządzenie Parlamentu Europejskiego I Rady w sprawie produktów maszynowych (https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52021PC0202&from=EN), (dostęp: 06.05.2022).
[130] https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en (dostęp: 11.05.2022).
[131] Ch. Wendehorst, Y. Duller, *Safety and Liability Related Aspects of Software*, European Commission 2021.
[132] Report on the Safety and Liability Implication of Artificial Intelligence, *op. cit.*

Podmiot taki ponosiłby odpowiedzialność za wszelkie działania AI, niezależnie od tego, gdzie (fizycznie, wirtualnie) i kiedy miała miejsce szkoda. Przy większej ilości wdrażających przewidziano odpowiedzialność solidarną z prawem regresu. Wskazano jednocześnie, że rodzaj systemu AI może mieć znaczenie dla zakresu odpowiedzialności: w przypadku rozwiązań generujących potencjalnie wysokie ryzyko dla społeczeństwa (ochrona takich wartości jak zdrowie, życie, integralność cielesna czy prawo własności) odpowiedzialność powinna być zaostrzona. Szczególnie dla AI wysokiego ryzyka należałoby przewidzieć oddzielny reżim odpowiedzialności, który powinien uwzględniać to, że AI się uczy i może wygenerować ryzyka, których nie przewidziano w momencie jej wdrożenia.

Parlament Europejski w swojej rezolucji z 20 października 2020 r.[133] przedstawił szereg zaleceń odnośnie przyszłych ram prawnych dotyczących odpowiedzialności za AI. Wyrażono w niej przekonanie, że nie będzie konieczny całkowity przegląd obecnie funkcjonujących systemów odpowiedzialności, a jedynie ich dostosowanie[134]. W dołączonym do rezolucji tekście proponowanego wniosku[135], zaproponowano, by operator[136] systemu AI wysokiego ryzyka ponosił odpowiedzialność na zasadzie ryzyka.

Niezależnie jednak od tego, jaki ostatecznie reżim odpowiedzialności zostanie przyjęty, pojawią się wątpliwości dotyczące jego rozumienia. Każde państwo członkowskie posługuje się własną siatką pojęciową w zakresie odpowiedzialności cywilnej. Oznacza to, że przykładowo „wina" nie jest dokładnie tym samym pojęciem prawnym w dwóch różnych państwach UE. Można założyć, że zasady odpowiedzialności cywilnej za działanie AI będą autonomicznymi pojęciami prawa unijnego, ale to z kolei wywoła potrzebę doprecyzowania i problemy dotyczące relacji z prawem krajowym. Jednocześnie warto nadmienić, że Parlament Europejski uważa za konieczne przyjęcie jednolitego ustawodawstwa w zakresie systemów AI. Z pewnością zatem powstanie na tym tle wiele wątpliwości prawnych, których rozstrzygnięcia będzie się wymagało od TSUE.

Banki mogą spodziewać się więc, że ich sytuacja również w znacznej większości regulowana będzie na zasadach ogólnych.

Nie można także wykluczyć, że kolejne wersje AIA będą jeszcze bardziej restrykcyjne pod względem nakładanych obowiązków[137]. W swojej propozycji z 20 kwietnia 2022 r., Parlament przedstawił szereg poprawek, skupiając się na kwestii praw podstawowych[138].

Dla regulacyjnego otoczenia sztucznej inteligencji, znaczenie będą mieć także przepisy projektowanego obecnie General Product Safety Regulation (GPSR)[139]. Rozporządzenie, zgodnie z zamiarem projektodawców, zastąpić ma obecnie obowiązującą General Product Safety Directive (GPSD; dyrektywa w sprawie ogólnego bezpieczeństwa produktów[140]). Nowa regulacja ma za zadanie m. in. dostosować stan prawny dotyczący bezpieczeństwa produktów związanych z nowymi technologiami, w tym sztuczną inteligencją. GPSR ma doprecyzować środki ochrony prawnej przysługujące konsumentom i zharmonizować maksymalne kary za naruszenia przepisów.

Publiczne konsultacje realizowane przez Komisję Europejską doprowadziły KE do wniosku, że GPSD nie przystaje do obecnych realiów[141]. Nie jest jasne, w jakim stopniu GPSD ma zastosowanie do nowych technologii, takich jak sztuczna inteligencja i Internet of Things (IoT) czy też do aktualizacji oprogramowania i samodzielnego oprogramowania, a także czy zagrożenia związane z bezpieczeństwem cybernetycznym są uwzględnione w definicji bezpieczeństwa. W badaniu GPSD wskazano również na problemy z produktami wykorzystującymi uczenie maszynowe i sztuczną inteligencję, ponieważ mogą one ewoluować w czasie, a tym samym potencjalnie zwiększać ryzyko związane z produktem, który początkowo mógł być bezpieczny.

Istniejące w tym zakresie mankamenty dyrektywy mają zostać wyeliminowane w projektowanym rozporządzeniu. Wskazano choćby na potrzebę ponownego zdefiniowania terminów „produkt" i „bezpieczny produkt", tak, aby w przypadku produktów z wbudowaną sztuczną inteligencją, samodzielne oprogramowanie i aktualizacje oprogramowania prowadzące do znacznych modyfikacji były wyraźnie w definicji uwzględnione.

[133] Rezolucja Parlamentu Europejskiego z dnia 20 października 2020 r. z zaleceniami dla Komisji w sprawie systemu odpowiedzialności cywilnej za sztuczną inteligencję (2020/2014(INL)).
[134] Pkt 6 Rezolucji 2020/2014.
[135] Wniosek dotyczący Rozporządzenia Parlamentu Europejskiego i Rady w sprawie odpowiedzialności za działanie systemów sztucznej inteligencji.
[136] Według propozycji wniosku operator to (w uproszczeniu) osoba fizyczna lub prawna, która do pewnego stopnia kontroluje ryzyko związane z działaniem systemu sztucznej inteligencji i czerpie korzyści z jego działania lub w sposób ciągły określa cechy technologii, dostarcza dane i podstawowe usługi wsparcia.
[137] Zob. Ada Lovelace Institute, *People, risk and the unique requirements of AI. 18 recommendations to strengthen the EU AIA* (2022).
[138] https://www.europarl.europa.eu/doceo/document/CJ-40-PR-731563_EN.pdf, (dostęp: 06.05.2022).
[139] https://ec.europa.eu/info/sites/default/files/proposal_for_a_regulation_on_general_product_safety.pdf (dostęp: 18.05.2022).
[140] Dyrektywa 2001/95/WE Parlamentu Europejskiego i Rady z dnia 3 grudnia 2001 r. w sprawie ogólnego bezpieczeństwa produktów.
[141] https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698028/EPRS_BRI(2021)698028_EN.pdf (dostęp: 18.05.2022).

Tym samym definicja produktu ma zostać rozszerzona. Pojęcie „product" ma zostać zastąpione pojęciem „item" i obejmie także produkty „połączone lub niepołączone z innymi produktami" (interconnected or not to other items)[142]. Celem jest tutaj uwzględnienie pojawiających się zagrożeń bezpieczeństwa związanych z nowymi technologiami. Znaczenie tych ostatnich zostanie też uwzględnione w ramach procedury oceny bezpieczeństwa produktów. Na gruncie GPSR, ocena będzie obejmowała bezpieczeństwo cybernetyczne[143] oraz ewoluujące, uczące się i przewidujące funkcjonalności produktu (*evolving, learning and predictive functionalities of a product*)[144].

Co do zasady, wymagania dotyczące bezpieczeństwa produktów są określone na gruncie GPSD dla producentów, importerów i dystrybutorów. Producenci i importerzy są zobowiązani do wprowadzania na rynek wyłącznie produktów bezpiecznych. Dystrybutorzy są zobowiązani do zachowania należytej staranności, aby nie dostarczać produktów, o których wiedzą lub powinni wiedzieć, że nie spełniają wymagań. Producenci, importerzy i dystrybutorzy są zobowiązani do wycofania produktów wprowadzonych na rynek, jeśli okaże się, że są niebezpieczne lub w ostateczności wycofać je od konsumentów, którzy już je kupili.

Na gruncie GPSR również przede wszystkim te podmioty będą ponosiły odpowiedzialność za spełnienie określonych przepisami wymagań. Banki, w zasadniczej części będą użytkownikami produktów (w tym produktów z AI). Nie oznacza to jednak, że reżim GPSR nie będzie ich dotyczył. Na gruncie rozporządzenia pojawić się ma pojęcie „istotnej modyfikacji". **Osoba fizyczna lub prawna, która w istotny sposób modyfikuje produkt, zostanie uznana za producenta**, co wiąże się ze wszystkimi związanymi z tym obowiązkami. Modyfikację będzie uważać się za istotną, jeżeli spełnione są trzy następujące kryteria:

a) modyfikacja zmienia zamierzone funkcje, rodzaj lub działanie produktu w sposób, który nie został przewidziany w pierwotnej ocenie ryzyka produktu;

b) w wyniku modyfikacji zmienił się charakter zagrożenia lub wzrósł poziom ryzyka;

c) zmiany nie zostały wprowadzone przez konsumenta na jego własny użytek[145].

Wprowadzenie takiego przepisu skutkować będzie koniecznością dopełnienia także przez niektórych użytkowników (np. banki) obowiązków przewidzianych pierwotnie dla producenta[146]. Maksymalna wysokość kary, jaka będzie mogła zostać nałożona za naruszenie rozporządzenia, ma wynieść 4% rocznego obrotu naruszyciela[147].

4.5. Problematyka udziału człowieka w cyklu życia systemu sztucznej inteligencji

Na konieczność udziału człowieka w procesach związanych z wykorzystaniem technologii zwracają uwagę nie tylko regulacje dotyczące bezpośrednio sztucznej inteligencji, ale szerzej obszaru IT[148]. Niewątpliwie jednak – ze względu na specyfikę systemów sztucznej inteligencji, a także ograniczenia wynikające z przepisów o ochronie danych osobowych – nadzór ze strony człowieka staje się szczególnie istotny[149], choć niepozbawiony trudności i wyzwań, które organizacje powinny brać pod uwagę przy projektowaniu odpowiednich rozwiązań[150]. Zaprojektowanie adekwatnych i skutecznych rozwiązań powinno stanowić istotny element tworzenia rozwiązań organizacyjnych poświęconych „zarządzaniu" systemami sztucznej inteligencji. Warto zwrócić uwagę, że wspomniane już wytyczne BaFin stawiają nawet na konieczność wdrożenia koncepcji *human-in-the-loop*[151], wymagającej udziału człowieka na każdym etapie funkcjonowania takich rozwiązań.

Projekt rozporządzenia w sprawie sztucznej inteligencji zakłada w art. 14 szczególne wymagania w tym zakresie. Kluczowy w tym miejscu jest art. 14 ust. 1, który wskazuje, że „[s]ystemy sztucznej inteligencji wysokiego ryzyka projektuje się i opracowuje w taki sposób, w tym poprzez uwzględnienie odpowiednich narzędzi interfejsu człowiek-maszyna, aby w okresie wykorzystywania systemu sztucznej inteligencji wysokiego ryzyka mogły je skutecznie nadzorować osoby fizyczne". Zakres wymagań, które będą stawiane podmiotom będzie szeroki, choć zasadniczo zbieżny z już stawianymi np. instytucjom finansowym w kontekście wykorzystania systemów IT, np. transakcyjnych czy przeciwdziałania praniu pieniędzy i finansowania terroryzmu. Istotne jest jednak, aby osoby odpowiedzialne za poszczególne obszary (etapy) działania miały dostęp do

[142] Zob. art. 3 ust. 1 GPSR.
[143] Zob. art. 7 lit. h GPSR.
[144] Zob. art. 7 lit. i GPSR.
[145] Zob. art. 12 GPSR.
[146] Odnoszą się do tego przede wszystkim art. 5 i art. 8 GPSR.
[147] Zob. art. 40 ust. 4 GPSR.
[148] Również sama Rekomendacja D dotycząca zarządzania obszarami technologii informacyjnej i bezpieczeństwa środowiska teleinformatycznego w bankach KNF wskazuje na konieczność nadzoru nad procesami IT.
[149] A. Simpson Rochwerger, W. Pang, *Real World AI. A Practical Guide for Responsible Machine Learning*, Appen 2021, s. 168.
[150] B. Green, *The flaws of policies requiring human oversight of government algorithms*, Computer Law & Security Review, Volume 45, 2022, s. 11 i następne.
[151] BaFin, *Big Data and..., op.cit.*, s. 12.

odpowiednich narzędzi zapewniających przykładowo możliwość „wyłączenia" systemu sztucznej inteligencji wysokiego ryzyka czy też wprowadzenia odpowiednich zmian, a także raportowania ewentualnych incydentów (szczególnie, że ma to być odrębny wymóg). Szczególnie istotne będzie także „uczulenie" osób odpowiedzialnych za te systemy, że w tym wypadku może wystąpić ryzyko automatycznego polegania lub nadmiernego polegania na wyniku działania systemu sztucznej inteligencji wysokiego ryzyka, czyli tzw. *automation bias*[152], co może mieć szczególne znaczenie w przypadku, gdy analityk podejmuje określoną decyzję bazując na rekomendacjach systemu.

Kwestie związane z nadzorem człowieka nad systemami sztucznej inteligencji należy jednak rozpatrywać nie tylko z perspektywy narzędzi, w które jest on wyposażony, ale także rozwiązań o charakterze organizacyjnym[153], a więc m.in. ułożenia struktury odpowiedzialności (za poszczególne komponenty czy etapy działania systemu), raportowania i komunikacji, a także decyzyjności, np. w kontekście konieczności użycia przycisku „stop" lub podobnej procedury, gdy system generuje nadmierne ryzyko. Powoduje to także, że projektując stosowne rozwiązania należy wziąć pod uwagę także ewentualne wymogi prawne lub regulacyjne, które przykładowo odnoszą się do płynnego przejścia od procesów automatycznych do manualnych. Przykładem takiej sytuacji będzie ta określona w art. 105a ust. 1a Prawa bankowego, gdzie osoba poddana zautomatyzowanej ocenie zdolności kredytowej, ma prawo do żądania otrzymania stosownych wyjaśnień co do podstaw decyzji, do uzyskania interwencji ludzkiej w celu ponownej decyzji oraz wyrażenia własnego stanowiska.

Warto zwrócić uwagę, że niektóre akty prawne, jak wspomniane już rozporządzenie w sprawie handlu algorytmicznego czy stanowisko UKNF w sprawie usługi robo-doradztwa mogą nakładać na banki dodatkowe wymogi w kontekście udziału człowieka w cyklu życia systemu AI. Stworzenie uniwersalnych rozwiązań w zakresie nadzoru nad takimi systemami nie powinno więc skutkować zaniechaniem tworzenia bardziej zindywidualizowanych struktur w przypadku występowania specyficznych wymogów prawno-regulacyjnych.

Jednocześnie podkreślić należy, że osobom oddelegowanym do realizacji obowiązków w tym zakresie należy zapewnić także odpowiednie narzędzia oraz wiedzę i kompetencje, aby były one w stanie realizować te obowiązki bez szkody dla organizacji[154]. Powoduje to, że powinny mieć one dostęp nie tylko do rozwiązań technicznych czy możliwości uzyskania stosownego wsparcia, ale powinny podlegać wstępnym i okresowym szkoleniom, a także należy komunikować im zmiany, które mogą mieć wpływ na funkcjonowanie systemów AI.

Nieoczywistym zagadnieniem, które jest nierzadko pomijane zarówno w praktyce, jak i literaturze, jest też kwestia udziału samych użytkowników w cyklu życia produktu, co ma istotne znaczenie w kontekście tych rozwiązań, które bezpośrednio oddziałują na klientów instytucji. Warto w tym miejscu zwrócić uwagę, że projektowane rozporządzenie w sprawie AI zakłada pewne dodatkowe wymogi w zakresie przejrzystości – informowania o ewentualnych interakcjach pomiędzy człowiekiem a „maszyną". Sam wpływ systemów AI na człowieka i odwrotny kierunek – wpływ człowieka na algorytmy i modele AI również powinny podlegać odpowiedniemu nadzorowi ze strony osób zatrudnionych w banku, w szczególności w kontekście ewentualnego negatywnego wpływu działania systemów na człowieka. Niezwykle istotna pozostaje także rola zespołów odpowiedzialny za szeroko rozumiane *data science,* którego rolą jest m.in. przygotowywanie danych do wykorzystania przez algorytmy i modele sztucznej inteligencji.

Zauważyć też należy, że zgodnie z treścią Rekomendacji D KNF (s. 25) *„[d]o każdej zinwentaryzowanej grupy danych (lub jej podzbioru) powinien zostać przypisany podmiot (jednostka organizacyjna, rola, osoba itp.), który jest ostatecznie odpowiedzialny za jakość tych danych i nadzór nad nimi, w szczególności w zakresie zarządzania związanymi z nimi uprawnieniami i udziału w rozwoju systemów informatycznych, w których są one przetwarzane".*

4.6. Zarządzanie danymi, w tym w zakresie jakości danych wykorzystywanych zarówno do trenowania, jak i stosowania modeli uczenia maszynowego i pokrewnych

Problematyka zarządzania danymi w kontekście sektora finansowego jest bardzo złożona i wielowątkowa[155], jak również dotyka zróżnicowanych aktów prawnych i regulacji, toteż wskazanie wszystkich

[152] K. Goddard, A. Roudsari, J. C. Wyatt, *Automation bias: a systematic review of frequency, effect mediators, and mitigators*, J Am Med Inform Assoc, 2012, 19, s. 121.

[153] M. Mäntymäki, M. Minkkinen1, T. Birkstedt et. al., *Defining organizational AI governance*, AI and Ethics, Springer 2022. Rozdział dostępny pod adresem: https://link.springer.com/content/pdf/10.1007/s43681-022-00143-x.pdf, s.2 (dostęp: 12.05.2022 r.).

[154] A. Simpson Rochwerger, W. Pang, *Real World AI. A Practical Guide for Responsible Machine Learning*, Las Vegas 2021, s. 111 i następne.

[155] Warto przykładowo zwrócić uwagę na Wytyczne 6/2020 w sprawie wzajemnych zależności między dyrektywą PSD2 a RODO opublikowane przez Europejską Radę Ochrony Danych Osobowych w 2020 r., https://edpb.europa.eu/system/files/2021-06/edpb_guidelines_202006_psd2_afterpublicconsultation_pl.pdf (dostęp: 12.05.2022 r.), które przynajmniej pośrednio odnoszą się do tej kwestii.

zagadnień w tym obszarze nie jest możliwe w ramach ograniczonych rozmiarów niniejszego opracowania. Wiele z tych kwestii podlega też obecnie projektowaniu na poziomie Unii Europejskiej, a kształt aktów prawnych i regulacji nie jest do końca rozstrzygnięty.

Przykładem aktu prawnego, który może mieć istotny wpływ na procesy zarządzania danymi w instytucjach finansowych, jest propozycja Dyrektywy Parlamentu Europejskiego i Rady w sprawie kredytów konsumenckich[156], która w wielu miejscach odnosi się do kwestii zarządzania danymi wykorzystywanymi m.in. do oceny zdolności kredytowej czy ryzyka kredytowego[157].

4.6.1. Stan prawny obecnie

Obecnie w polskim systemie prawnym brak jest regulacji prawnych, które odnosiłyby się do wykorzystania systemów sztucznej inteligencji wprost. Problematyka zarządzania danymi, także w związku z AI, mieści się jednak w zakresie problematyki zarządzania wewnętrznego instytucji. Aktualne pozostają więc w tym kontekście obowiązki nakładane przez Dyrektywę CRD IV oraz przepisy ją implementujące[158]. Szczególnie istotne są zasady zarządzania wewnętrznego, z naciskiem położonym na procedury zarządzania ryzykiem[159]. W tym zakresie projektowane rozporządzenie w sprawie AI odsyła właśnie do obowiązujących systemów zarządzania[160].

Wybiegając nieco w przyszłość, warto zwrócić uwagę, że planowany na gruncie AIA system zarządzania jakością, również ma być zgodny z zasadami i procedurami określonymi w CRD IV. Jest to o tyle istotne, że system ten obejmuje także systemy i procedury zarządzania danymi, w tym gromadzenia danych czy analizy danych[161].

Istnieje więc spore prawdopodobieństwo, że dokumentacja związana z AI, będzie w przypadku banków stanowić część szerszej dokumentacji, tworzonej w ramach systemu zarządzania wewnętrznego[162]. Stosowanie się do obowiązujących regulacji „twardych" i „miękkich" powinno zabezpieczać sytuację prawną banku. Projektowanie ewentualnych rozwiązań w zakresie AI należałoby oprzeć na Wytycznych EBA w sprawie zarządzania wewnętrznego[163], które jednak nie skupiają się na technologii i tym samym nie wspominają o samej sztucznej inteligencji.

Prawo bankowe przewiduje obowiązek wprowadzenia systemu zarządzania, tj. co najmniej systemu zarządzania ryzykiem oraz systemu kontroli wewnętrznej. W ramach tych systemów obecnie należy umiejscowić kwestie związane z zarządzaniem danymi. W prawie bankowym brak jest innych szczegółowych regulacji związanych z tym zagadnieniem.

Na osobną uwagę zasługuje za to tematyka ochrony danych osobowych. Ogólna uwaga o nieobecności regulacji dotyczących AI znajduje zastosowanie i tutaj, ale możliwe jest wyróżnienie kilku przepisów, które pośrednio odnoszą się do sztucznej inteligencji. RODO jest neutralne technologicznie, więc należy spojrzeć na jego treść z szerszej perspektywy. W odniesieniu do AI ważne będą przede wszystkich pojęcia zautomatyzowanego przetwarzania danych oraz profilowania. Pierwsze nie wymaga kompleksowych wyjaśnień – chodzi o takie przetwarzanie danych, które nie jest wykonywane w całości w sposób manualny przez człowieka. Z kolei profilowanie, w uproszczeniu, jest formą zautomatyzowanego przetwarzania i polega na wykorzystaniu danych do oceny czynników osoby, np. w celu prognozy jej sytuacji ekonomicznej[164]. W pełni uzasadnione jest przyjęcie, że algorytmy AI analizujące dane osobowe mogą dokonywać profilowania. Przy czym może to prowadzić do podjęcia określonej decyzji przez „maszynę" lub być częścią procesu, w którym bierze udział człowiek. Ocena zdolności kredytowej jest przykładem procesu, który obejmuje profilowanie i zautomatyzowane podejmowanie decyzji w rozumieniu RODO.

Zautomatyzowane podejmowanie decyzji sprowadza się do tego, że podejmowana jest decyzja, która opiera się wyłącznie na zautomatyzowanym przetwarzaniu (w tym profilowaniu) i wywołuje wobec osoby, której dane dotyczą, skutki prawne lub inaczej istotnie na tę osobę wpływa. Przepis art. 22 RODO ustanawia zakaz podejmowania automatycznych decyzji wobec

[156] https://eur-lex.europa.eu/resource.html?uri=cellar:2df39e27-da3e-11eb-895a-01aa75ed71a1.0012.02/DOC_1&format=PDF (dostęp: 12.05.2022 r.).
[157] Warto w tym miejscu zwrócić uwagę na krytyczną opinię Europejskiego Inspektora Ochrony Danych, który negatywnie odniósł się przykładowo co do możliwości wykorzystania niestandardowych danych na potrzeby takiej oceny. Opinia dostępna pod adresem: https://edps.europa.eu/system/files/2021-08/opinion_consumer-credit-final_en.pdf (dostęp: 12.05.0222 r.).
[158] Wdrożenie dokonane przede wszystkim dokonane w formie nowelizacji ustawy – Prawo bankowe, dokonanej ustawą z dnia 5 sierpnia 2015 r. o nadzorze makroostrożnościowym nad systemem finansowym i zarządzaniu kryzysowym w systemie finansowym.
[159] Zob. art. 74 dyrektywy CRD IV.
[160] Zob. szczególnie art. 9 ust. 9 projektu AIA.
[161] Katalog operacji na danych wynika z art. 17 ust. 1 lit. f AIA i obejmuje także: etykietowanie danych, przechowywanie danych, filtrowanie danych, eksplorację danych, agregację danych, zatrzymywanie danych i wszelkie inne operacje dotyczących danych, które przeprowadza się przed wprowadzeniem do obrotu lub oddaniem do użytku systemów sztucznej inteligencji wysokiego ryzyka i do celów wprowadzenia ich do obrotu lub oddania ich do użytku.
[162] Projektowany art. 18 AIA stanowi o tym, że instytucje kredytowe dokumentację techniczną AI wysokiego ryzyka prowadzą jako jeden z elementów dokumentacji z art. 74 CRD IV.
[163] EBA/GL/2017/11.
[164] Zob. art. 4 pkt 4 RODO.

osób, których dane dotyczą. Jednocześnie wymienia przypadki, w których jest to dozwolone. Zakaz nie ma zastosowania, jeśli automatyczna decyzja:

– jest niezbędna do zawarcia lub wykonania umowy,

– jest dozwolona prawem,

– opiera się na wyraźnej zgodzie, osoby, której dane dotyczą[165].

Ocena zdolności kredytowej opiera się na zautomatyzowanym przetwarzaniu i prowadzi do dokonania decyzji na podstawie wyniku dostarczonego przez algorytm. W tym przypadku zachodzi przypadek, o którym mowa w art. 22 ust. 2 lit b) RODO, jako że kwestię tę normuje Prawo bankowe (PB).

PB umożliwia przetwarzanie informacji dotyczących osób fizycznych, stanowiących tajemnicę bankową (i informacji udostępnianych przez określone podmioty) w celu oceny zdolności kredytowej i analizy ryzyka kredytowego[166]. W tych samych celach banki mogą podejmować decyzje, opierając się w całości na zautomatyzowanym przetwarzaniu i profilowaniu danych osobowych. Skorzystanie z takiej możliwości wymaga jednak dopełnienia odpowiednich obowiązków informacyjnych. Bank musi zapewnić osobie, której dane dotyczą:

– prawo do otrzymania stosownych wyjaśnień co do podstaw podjętej decyzji,

– prawo do uzyskania interwencji ludzkiej w celu podjęcia ponownej decyzji,

– prawo do wyrażenia własnego stanowiska[167].

Do tego, zgodnie z RODO, konieczne jest wcześniejsze poinformowanie osoby o samym fakcie zautomatyzowanego podejmowania decyzji oraz informacje o istotnych zasadach ich podejmowania i przewidywanych konsekwencjach[168].

Wątpliwości może budzić szczególnie kwestia opisu zasad podejmowania zautomatyzowanych decyzji i wyjaśnienia podstaw podjętej decyzji, choćby pod kątem tajemnicy przedsiębiorstwa. Wydaje się, że takie obawy mogą powstawać także na innych polach związanych z AI.

Ważnym zagadnieniem, które zostało już zaadresowane w niniejszym opracowaniu, jest kwestia zarządzania danymi w kontekście tzw. modeli wewnętrznych. Zgodnie z art. 176 CRR instytucje gromadzą i przechowują dane dotyczące aspektów przyznawanych wewnętrznych ratingów zgodnie z wymogami określonymi w części ósmej tego aktu. Przepis określa dokładne wymogi w zakresie tego, jakie dane instytucje powinny wykorzystywać i przechowywać. Podobnie rekomendacja T[169] Komisji Nadzoru Finansowego przewiduje pewne specyficzne wymogi w zakresie zarządzania danymi, choć w tym przypadku rekomendacje mają raczej charakter ogólny pozostawiając nieco większą swobodę instytucjom zobowiązanym do jej stosowania.

Równie istotna jest też rekomendacja D – wielokrotnie wskazywana już w niniejszym opracowaniu – która w rekomendacji 8 wskazuje, że *„[b]ank powinien posiadać sformalizowane zasady zarządzania danymi wykorzystywanymi w ramach prowadzonej działalności, obejmujące w szczególności zarządzanie architekturą oraz jakością danych i zapewniające właściwe wsparcie działalności banku”*. Rekomendacja ta ma bardzo istotne znaczenie dla budowania w organizacji swoistej „kultury danych”, której tworzenie jest ważne z punktu widzenia postępującej algorytmizacji i automatyzacji, a także tzw. „danetyzacji” życia (rekomendacja 8.11).

W przypadku zarządzania ryzykiem należy również wskazać, że Rozporządzenie Ministra Finansów, Funduszy i Polityki Regionalnej z dnia 8 czerwca 2021 r. w sprawie systemu zarządzania ryzykiem i systemu kontroli wewnętrznej oraz polityki wynagrodzeń w bankach[170] w par. 19 pkt 4 również przewidziano pewne wymogi w zakresie zarządzania danymi, toteż dokonując całościowej analizy obszaru, należy uwzględniać także ewentualne zależności z innymi aktami prawnymi i regulacjami, które nie odnoszą się wprost do wykorzystywania systemów automatycznych.

Na marginesie należy wskazać także, że w 2022 r. Instytut Inżynierów Elektryków i Elektroników (IEEE) opracowała standard ISO/IEC 38507:2022 odnoszący się m.in. do kwestii zarządzania danymi – *Information Technology – Governance of IT – Governance implications of the use of artificial intelligence by organizations.*

4.6.2. Projektowane regulacje prawne

Będący obecnie w fazie projektowania AIA odnosi się do kwestii związanych z danymi głównie w Rozdziale II Tytułu III. Trzeba zaznaczyć, że sformułowane tam propozycje regulacji dotyczą systemów AI wysokiego ryzyka, tj. m.in. takich, które ujęte będą

[165] Zob. art. 22 ust. 1 i 2 RODO.
[166] Zob. art. 105a ust. 1 pr. Bank.
[167] Zob. art. 105a ust. 1a pr. Bank.
[168] Zob. art. 22 ust. 2 lit. f RODO.
[169] https://www.knf.gov.pl/knf/pl/komponenty/img/Rekomendacja_T_%2814_09_2018%29_63160.pdf (dostęp: 12.05.2022 r.).
[170] Dz. U. z 2021 r., poz. 1045.

w Załączniku III do AIA. Zgodnie z tym załącznikiem systemy sztucznej inteligencji przeznaczone do oceny zdolności kredytowej osób fizycznych lub ustalania ich punktacji kredytowej[171], mają być kwalifikowane jako systemy AI wysokiego ryzyka. Tym samym znajdą do nich zastosowanie wymogi określone we wspomnianym rozdziale.

W projekcie AIA podkreśla się, że obszarem, w którym stosowanie systemów AI zasługuje na szczególną uwagę, jest dostęp do niektórych podstawowych usług i świadczeń prywatnych i publicznych niezbędnych ludziom do pełnego uczestnictwa w życiu społecznym lub do poprawy poziomu życia oraz korzystanie z tych usług i świadczeń. W szczególności systemy AI wykorzystywane do przeprowadzania punktowej oceny kredytowej lub oceny zdolności kredytowej osób fizycznych należy klasyfikować jako systemy wysokiego ryzyka, ponieważ decydują one o dostępie tych osób do zasobów finansowych lub podstawowych usług, takich jak mieszkalnictwo, energia elektryczna i usługi telekomunikacyjne. Systemy sztucznej inteligencji wykorzystywane w tym celu mogą prowadzić do dyskryminacji osób lub grup i utrwalać historyczne wzorce dyskryminacji, na przykład ze względu na pochodzenie rasowe lub etniczne, niepełnosprawność, wiek, orientację seksualną lub powodować powstawanie dyskryminujących skutków w nowej postaci (z motywu 37. preambuły AIA).

Regulacja dotycząca **systemu zarządzania ryzykiem** zawarta w projektowym art. 9 AIA zakłada, że system taki składa się z ciągłego, iteracyjnego procesu prowadzonego przez cały cykl życia AI wysokiego ryzyka, wymagającego regularnej, systematycznej aktualizacji i obejmuje wskazane w projekcie powołanego przepisu etapy. Jednocześnie w ust. 9 art. 9 AIA znajduje się bezpośrednie odesłanie do przepisów dotyczących instytucji kredytowych[172]: *W przypadku instytucji kredytowych regulowanych dyrektywą 2013/36/UE aspekty opisane w ust. 1–8 stanowią część procedur zarządzania ryzykiem ustanowionych przez te instytucje zgodnie z art. 74 tej dyrektywy.*

Jeśli chodzi o jakość danych i zarządzanie danymi, odpowiednie kryteria będą musiały spełnić co najmniej zbiory danych **treningowych, walidacyjnych i testowych** (art. 10 AIA) z zastrzeżeniem, że w przypadku opracowywania systemów AI wysokiego ryzyka niewykorzystujących technik obejmujących uczenie modeli ust. 2–5 art. 10 AIA, stosuje się wyłącznie do testowych zbiorów danych. Zbiory takich danych mają podlegać odpowiednim praktykom w zakresie zarządzania oraz zarzadzania danymi (ang. *appropriate data governance and management practices*), być adekwatne, reprezentatywne, a także – w najszerszym możliwym zakresie (ang. *to the best extent possible*) – wolne od błędów i kompletne.

Proponuje się, aby **praktyki w zakresie zarządzania, w tym zarządzania danymi**, dotyczyły w szczególności:

- odpowiednich decyzji projektowych;
- procesów gromadzenia danych;
- odpowiednich operacji przetwarzania związanych z przygotowywaniem danych, takie jak anotacja, etykietowanie, czyszczenie, wzbogacanie i agregowanie;
- sformułowania odpowiednich założeń, w szczególności w odniesieniu do informacji, do których pomiaru i reprezentowania mają służyć dane;
- uprzedniej oceny dostępności, ilości i przydatności potrzebnych zestawów danych;
- badania pod kątem ewentualnych uprzedzeń, które mogą mieć wpływ na zdrowie i bezpieczeństwo osób lub prowadzić do dyskryminacji zakazanej przez prawo Unii;
- identyfikacji wszelkich możliwych luk lub niedociągnięć w danych oraz sposobu zaradzenia im.

Projektodawca unijny nie poprzestaje jednak na tych wymogach. Konieczne będzie, by dane charakteryzowały się **odpowiednimi właściwościami statystycznymi**, w tym, w stosownych przypadkach, w odniesieniu do osób lub grup osób, wobec których ma być wykorzystywany system AI wysokiego ryzyka[173].

Zbiory danych zgodnie z projektem AIA będą musiały uwzględniać elementy, które są specyficzne dla określonego kontekstu geograficznego, behawioralnego lub funkcjonalnego lub okoliczności, w których ma być wykorzystywany system AI. W zakresie danych osobowych istotne jest projektowane wprowadzenie **nowej podstawy przetwarzania danych osobowych szczególnych kategorii**[174], np. danych biometrycznych. Takie przetwarzanie będzie mogło odbyć się w celu zapewnienia monitorowania, wykrywania

[171] Z wyjątkiem systemów sztucznej inteligencji wprowadzanych do użytku przez usługodawców działających na niewielką skalę na własny użytek; zob. ust. 5 lit. b załącznika III.

[172] Konkretnie do art. 74 dyrektywy 2013/36/UE w sprawie warunków dopuszczenia instytucji kredytowych do działalności oraz nadzoru ostrożnościowego nad instytucjami kredytowymi i firmami inwestycyjnymi.

[173] Te kryteria mogą zostać spełnione na poziomie pojedynczych zbiorów danych lub ich kombinacji.

[174] Chodzi o kategorie danych osobowych określone w art. 9 RODO.

i korygowania tendencyjności systemów AI. W zakresie przetwarzania tych szczególnych kategorii danych, konieczne będzie zapewnienie dodatkowych zabezpieczeń, np. środków ograniczających ponowne wykorzystanie tych danych oraz służących zapewnieniu bezpieczeństwa i ochrony prywatności, takich jak pseudonimizacja.

System zarządzania danymi będzie musiał również uwzględnić przepisy rozporządzenia w sprawie europejskiego zarządzania danymi (DGA)[175]. Akt ten ma przede wszystkim za zadanie ułatwić udostępnianie i wymianę danych na terenie UE. Dotyczy to zarówno uzyskiwania danych od podmiotów publicznych, jak i obiegu danych między podmiotami prywatnymi. W tym drugim przypadku istotne jest rozwiązanie w postaci pośrednika w udostępnianiu danych. Pośredniczenie między posiadaczami danych, polegające np. na udostępnianiu odpowiedniej infrastruktury technicznej do wymiany danych, będzie wymagało spełnienia szeregu warunków[176], co ma w zamierzeniu projektodawców wzmocnić zaufanie do pośredników i usprawnić cyrkulację danych. Komisja Europejska będzie także prowadzić rejestr pośredników, a odpowiednie organy krajowe będą zajmować się nadzorem. Banki będą mogły korzystać z usług pośredników udostępniania danych – zarówno jako podmioty udostępniające, jak i podmioty, którym dane są udostępniane. Zwraca się przy tym uwagę, że istnienie pośredników i regulacja ich obowiązków i uprawnień w rozporządzeniu nie powinny doprowadzić do sytuacji, w której wymiana danych bezpośrednio między przedsiębiorstwami, tj. bez udziału pośrednika, okaże się w praktyce niemożliwa[177].

Niewątpliwie jednak bezpieczna platforma wymiany danych może stanowić cenne źródło danych do trenowania modeli sztucznej inteligencji, co powinno zostać uwzględnione przy planowaniu rozwoju systemów AI w bankowości. Nowe kanały gromadzenia danych wynikające z rozwoju technologicznego oznaczają, że źródła danych i metody ich pozyskiwania stale się zmieniają. Wskazuje się, że może to prowadzić do zmian w świadczeniu niektórych usług finansowych, a także stwarzać ryzyko dla instytucji, stąd konieczność zwrócenia uwagi na ocenę ryzyka związanego z nowymi rozwiązaniami[178]. Jednocześnie zwiększony zakres danych źródłowych może posłużyć do budowania coraz bardziej zaawansowanych modeli, umożliwiających precyzyjniejszą ocenę ryzyka[179].

W nieco bardziej odległej przyszłości istotne w działalności bankowej mogą okazać się także wymogi dotyczące realizacji praw **użytkowników w zakresie dostępu do danych wygenerowanych przez ich urządzenia i usługi sieciowe, zaproponowane w** projekcie rozporządzenia Parlamentu Europejskiego i Rady w sprawie zharmonizowanych przepisów dotyczących sprawiedliwego dostępu do danych i ich wykorzystywania (**Akt w sprawie danych)**[180]**. Projekt aktu** opublikowany został przez Komisję Europejską w dniu 23 lutego 2022r. i wraz z DGA, realizował będzie założenia Europejskiej Strategii Danych[181]. O ile jednak DGA dotyczy przede wszystkim **ram prawnych, procesów i struktur wspierających wymianę danych, to projekt Aktu w sprawie danych kładzie większy nacisk na wyjaśnienie, kto i na jakich warunkach ma prawo korzystać z danych**[182]**.**

[175] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/868 z dnia 30 maja 2022 r. w sprawie europejskiego zarządzania danymi i zmieniające rozporządzenie (UE) 2018/1724 (akt w sprawie zarządzania danymi. PE/85/2021/REV/1, Dz.U. L 152 z 3.6.2022, s. 1. 1-44. Rozporządzenie zacznie obowiązywać 24 września 2023r.

[176] Zob. art. 9-11 DGA.

[177] Zob. EBF Response to the European Commission's consultation on the Data Governance Act, s.2, https://www.ebf.eu/innovation-cybersecurity/european-commissions-consultation-on-the-data-governance-act-ebf-response/ (dostęp: 26.04.2022 r.).

[178] Zob. EBA Report on Big Data and Advanced Analytics, s. 47 i n., https://www.eba.europa.eu/sites/default/documents/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf (dostęp: 04.05.2022 r.).

[179] Firmy z branży FinTech już teraz wykorzystują szeroki zakres danych źródłowych (tzw. Dane alternatywne) do np. oceny zdolności kredytowej; zob. Johnson, Kristin N., *Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit*, Tulane Public Law Research Paper No. 19-7. Artykuł dostępny pod adresem: https://ssrn.com/abstract=3481102 (dostęp: 26.04.2022 r.).

[180] Tekst wniosku dostępny na: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN (dostęp: 24.06.2022).

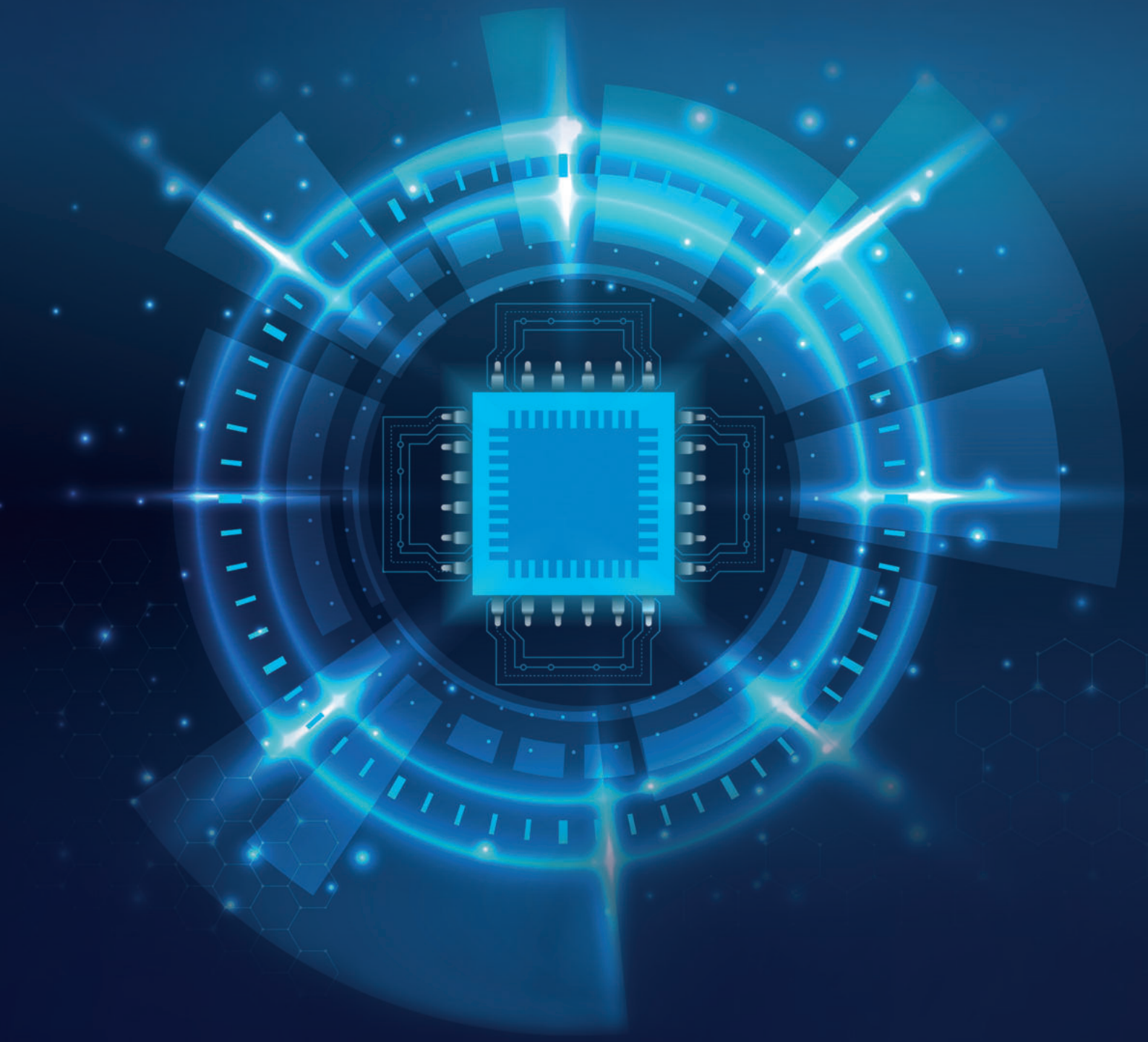
[181] Por. https://ec.europa.eu/commission/presscorner/detail/pl/ip_20_273 (dostęp: 24.06.2022).

[182] Zob.: https://poland.representation.ec.europa.eu/news/akt-w-sprawie-danych-2022-02-23_pl (dostęp: 24.06.2022).

ROZDZIAŁ V

PROGRAM ANALITYCZNO BADAWCZY

Zagadnienie przejrzystości (*transparency*) oraz wyjaśnialności (*explainablity*)



Problematyka przejrzystości i wyjaśnialności algorytmów i modeli sztucznej inteligencji stanowi jedno z kluczowych zagadnień mających wpływ zarówno na stosowanie, jak i rozwijanie tych rozwiązań[183]. Znalezienie odpowiednich rozwiązań w tym obszarze dla systemów sztucznej inteligencji stanowi istotne wyzwanie, bowiem mogą one wpływać także na jakość i skuteczność algorytmów i modeli, o czym w dalszej części opracowania. Sprawiają także pewne trudności interpretacyjne czy nawet pojęciowe, bowiem nie zawsze są rozumiane w sposób jednolity. Z tego względu na wstępie należy dokonać rozróżnienia tych pojęć.

Przejrzystość (*transparency*) może być rozumiana na wiele sposobów, natomiast dość często pojawia się w szerszym kontekście jako całokształt rozwiązań zapewniających informacje i przejrzystość użytkownikom systemów sztucznej inteligencji, na co składać się może wyjaśnialność, identyfikowalność i komunikacja[184]. Z kolei F. Hussain, R. Hussain oraz E. Hossain wskazują na trzy wymiary tzw. *explainable artificial intelligence* (XAI), do których zalicza się: (i) wyjaśnialność; (ii) interpretowalność oraz (iii) przejrzystość. Przy czym ten ostatni element należy rozumieć jako możliwość zrozumienia działania modelu bez użycia dodatkowych komponentów[185]. Na potrzeby niniejszego opracowania posłużymy się jednak propozycją rozumienia przejrzystości zaproponowaną przez A. H. Briggsa (i in.), który wskazał, że jest to możliwość opisania struktury modelu, równań, wartości parametrów i założeń w taki sposób, aby umożliwiało to innym osobom zrozumienie tego modelu[186], co obejmuje także rozumienie przejrzystości w rozumieniu art. 13 i art. 52 projektowanego rozporządzenia w sprawie sztucznej inteligencji.

Z kolei wyjaśnialność (*explainability*) jest pojęciem węższym i będącym „podzbiorem" szerszej przejrzystości lub jej częścią. Odnosi się zasadniczo do możliwości wyjaśnienia podstaw stojących za konkretnym rezultatem działania modelu[187], ale w sposób, który jest zrozumiały dla konkretnego odbiorcy. Z pojęciem wyjaśnialności łączą się dwa pojęcia *„black-box"* i *„glass-box"*[188], które dotyczą możliwości lub niemożliwości odtworzenia sposobu „rozumowania" modelu sztucznej inteligencji.

Przykładowo, art. 105a ust. 1a PB przewiduje obowiązek zapewnienia osobie, wobec której stosowane są zautomatyzowane systemy oceny zdolności kredytowej, stosownych wyjaśnień co do podstaw podjętej decyzji. W przypadku systemu AI niezbędne będzie więc określenie informacji, jakie osoba dotknięta taką decyzją powinna otrzymać. W jednym z komunikatów KNF odnoszącym się do art. 70a PB wskazano, że przedmiotem wyjaśnienia powinna być zindywidualizowana i szczegółowa informacja, w tym informacja na temat środków, które powinien przedsięwziąć wnioskujący, aby usunąć negatywne skutki determinujące decyzję kredytodawcy o nieprzyznaniu kredytu[189]. Oznacza to, że bank powinien mieć na tyle szczegółowe informacje na temat danych i parametrów (oraz ich wagi), które dany model sztucznej inteligencji wziął pod uwagę, aby przekazać potencjalnemu kredytobiorcy zrozumiałą i wartościową informację pozwalającą na usunięcie przeszkód do pozyskania kredytu. Może to być znaczącym wyzwaniem w przypadku modeli, które nie poddają się łatwo interpretacji w tym zakresie.

Warto tutaj także zwrócić uwagę na zbliżone wymogi w zakresie udzielania wyjaśnień odnoszących się do detalicznych ekspozycji kredytowych, które znajdziemy w rekomendacji 19.6 Rekomendacji T dotyczącej dobrych praktyk w zakresie zarządzania ryzykiem detalicznych ekspozycji kredytowych.

W kontekście wyjaśnialności warto także nawiązać do jednego z opracowań Banku Rozrachunków Międzynarodowych, w którym wskazano, że oczekiwania dotyczące wyjaśniania i możliwości kontroli są generalnie takie same zarówno w przypadku modeli AI, jak i tradycyjnych, i dotyczą ujawniania informacji wewnętrznych, zwłaszcza zarządowi i kierownictwu wyższego szczebla, aby mogli oni lepiej zrozumieć ryzyko i konsekwencje stosowania AI. Jednak oczekiwania dotyczące ujawniania informacji na zewnątrz wydają się być specyficzne dla stosowania AI[190].

[183] J. M. Schraagen, S. K. Lopez, C. Schneinder et. al., *The Role of Transparency and Explainability in Automated Systems*, Proceedings of the 2021 HFES 65th International Annual Meeting, s. 31. Artykuł dostępny pod adresem: https://journals.sagepub.com/doi/pdf/10.1177/1071181321651063 (dostęp: 12.05.2022 r.).

[184] Tak rekomendacje ekspertów Komisji Europejskiej w sprawie sztucznej inteligencji godnej zaufania, s. 18. Dokument dostępny pod adresem: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60436 (dostęp: 16.05.2022 r.).

[185] F. Hussain, R. Hussain, E. Hossain, E*xplainable Artificial Intelligence (XAI): An Engineering Perspective*, arXiv: 2101.03613, 2021, s. 1-2.

[186] A. H. Briggs, M. C. Weinstein, E. A. L. Fenwick et. al., *Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6*, Value in Health, No. 15, 2012, s. 835 i następne. Za M. A. Clinciu, H. F. Hastie, *A Survey of Explainable AI Terminology*, Proceedings of the 1st Workshop omn Interactive Natural Language Technology for Explainable Artificial Intelligence, 2019, s. 9.

[187] U. Ehsan, Q. Vera Liao, M. Muller et. al., *Expanding Explainability: Towards Social Transparency in AI systems*, CHI 21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, May 2021, s. 2 i następne.

[188] A. Rai, *Explainable AI: from black box to glass box*, Journal of the Academy of Marketing Science, No. 48, 2020, s. 138.

[189] https://www.knf.gov.pl/knf/pl/komponenty/img/Komunikat_UKNF_ws_prawa_do_uzyskania_wyjasnien_nt_oceny_zdolnosci_kredytowej_wersja_szczegolowa_70332.pdf (dostęp: 17.05.2022 r.).

[190] J. Prenio, J. Yong, *Humans keeping AI in check – emerging regulatory expectations in the financial sector*, FSI Insights on policy implementation, No 35, August 2021, s. 8.

Określenie wymogów w tym zakresie będzie uzależnione więc od kilku elementów, tj. specyficznych wymogów prawno-regulacyjnych (ale także ograniczeń związanych przykładowo z tajemnicą przedsiębiorstwa), wewnętrznych zasad dotyczących przejrzystości czy też konkretnych potrzeb i oczekiwań odbiorców. Inny zakres informacji będzie wymagany względem klientów (użytkowników końcowych), a inny względem inżynierów odpowiedzialnych za prawidłowe funkcjonowanie systemów sztucznej inteligencji. Należy przy tym pamiętać, że specyficzne wymagania mogą się pojawić ze strony organu nadzoru, który również może żądać wyjaśnień co do funkcjonowania określonych rozwiązań.

Innymi słowy, poziom zaawansowania i skomplikowania, a także użyteczności powinien być dostosowany do poszczególnych grup odbiorców[191]. W idealnym scenariuszu bank powinien zwrócić uwagę na konieczność opracowania stosownej dokumentacji, w tym technicznej, dotyczącej kwestii wyjaśnialności.

Jednym z istotnych wątków jest też wspomniana już przejrzystość. Projektowany przepis art. 13 ust. 1 AIA wprost wskazuje, że *„[s]ystemy sztucznej inteligencji wysokiego ryzyka projektuje się i opracowuje się w sposób zapewniający wystarczającą przejrzystość ich działania, umożliwiającą użytkownikom interpretację wyników działania systemu i ich właściwe wykorzystanie. Zapewnia się odpowiedni rodzaj i stopień przejrzystości w celu osiągnięcia zgodności z odpowiednimi obowiązkami użytkownika i dostawcy, określonymi w rozdziale 3 niniejszego tytułu."* Rozdział 3 nosi tytuł „Obowiązki Dostawców i Użytkowników Systemów Sztucznej Inteligencji Wysokiego Ryzyka". Z kolei art. 52 ust. 1 zdanie pierwsze projektowanego rozporządzenia w sprawie sztucznej inteligencji przewiduje, że *„[d]ostawcy zapewniają, aby systemy sztucznej inteligencji przeznaczone do wchodzenia w interakcję z osobami fizycznymi projektowano i opracowywano w taki sposób, aby osoby fizyczne były informowane o tym, że prowadzą interakcję z systemem sztucznej inteligencji, chyba że okoliczności i kontekst korzystania z systemu jednoznacznie na to wskazują"*. Oczywistym przykładem są tutaj czatboty oraz usługi typu IVR (*Interactive Voice Response*), które mogą „wchodzić" w interakcje z klientami banku. Jednocześnie art. 52 ust. 2 zdanie pierwsze wskazuje, że *„[u]żytkownicy systemów rozpoznawania emocji lub systemów kategoryzacji biometrycznej informują osoby fizyczne, wobec których systemy te są stosowane, o fakcie ich stosowania"*, co oznacza, że również w tym przypadku bank powinien spełnić odpowiednie obowiązki informacyjne. Niektóre obowiązki w tym zakresie zostały już opisane w niniejszym opracowaniu, m.in. w kontekście tzw. robo-doradztwa.

Na koniec warto podkreślić, że wymogi w zakresie wyjaśnialności, a także przejrzystości, podlegają znaczącej ewolucji, co jest związane także z rozwijającym się zagadnieniem zależności pomiędzy skutecznością algorytmów i modeli sztucznej inteligencji a poziomem wyjaśnialności, który on przewiduje[192]. Dziś przyjmuje się, że w przypadku wielu systemów tzw. sztucznej inteligencji zapewnienie wysokiego poziomu wyjaśnialności może wpłynąć negatywnie na jego skuteczność i dokładność, co wymaga od podmiotów je stosujących proporcjonalnego podejścia.

5.1. Specyficzne ryzyka dla systemów sztucznej inteligencji

Jak wynika z analizy przedstawionej w niniejszym opracowaniu, systemy sztucznej inteligencji mogą generować ryzyka charakterystyczne zarówno dla szeroko rozumianego obszaru ICT, jak i ryzyka związane ściśle z tą konkretną technologią, np. w kontekście stronniczości algorytmicznej. Powoduje to, że zakres zagadnień, które potencjalnie składają się na obszar zarządzania ryzykami dla systemów sztucznej inteligencji, wykracza znacząco poza ramy opracowania. Poniżej zostaną wskazane więc wybrane – subiektywnie najistotniejsze – zagadnienia.

Analizę warto rozpocząć od podkreślenia, że sam projekt rozporządzenia w sprawie sztucznej inteligencji przewiduje specyficzne wymagania w zakresie systemów zarządzania ryzykiem systemami wysokiego ryzyka, zaś wielokrotnie przywoływane wytyczne niemieckiego organu nadzoru nad rynkiem finansowym wskazują na konieczność powiązania ryzyk inherentnych dla systemów AI z już istniejącym systemem zarządzania ryzykiem obowiązującym w instytucji. W tym przypadku banki znajdują się w stosunkowo dobrej sytuacji, bowiem obowiązujące akty prawne i regulacje niejako narzucają surowe wymogi w tym zakresie[193].

Jednocześnie pierwotne brzmienie wspomnianego projektu rozporządzenia przewiduje w art. 9 ust. 9, że „[w] przypadku instytucji kredytowych podlegających przepisom dyrektywy 2013/36/UE aspekty opisane w ust. 1-8 stanowią część procedur służących zarządzaniu ryzykiem ustanowionych przez te instytucje

[191] G. Bar, *Przejrzystość, w tym wyjaśnialność, jako wymóg prawny dla systemów Sztucznej Inteligencji*, Monitor Prawniczy 20/2020 (dodatek specjalny) s. 70.

[192] M. Kearns, A. Roth, *The Ethical Algorithm. The Science of Socially Aware Algorithm Design*, Oxford 2020, s. 169.

[193] M. Nowakowski, K. Waliszewski, *Sztuczna inteligencja w problematyce modeli oceny ryzyka w instytucjach finansowych z perspektywy prawno-regulacyjnej*, Finanse i prawo finansowe, Vol. 1(33), Marzec 2022, s. 169 i następne.

zgodnie z art. 74 tej dyrektywy". Jest to znaczne ułatwienie, bowiem banki mogą dokonać jedynie modyfikacji istniejących rozwiązań w celu dostosowania się do przyszłych (prawdopodobnych) wymogów w zakresie zarządzania ryzykiem.

Należy w tym miejscu jednocześnie podkreślić, że obecnie na poziomie Unii Europejskiej projektowane są także inne akty prawne, które będą wpływały na zakres dokumentów czy rozwiązań organizacyjno-technicznych, które podmioty wykorzystujące systemy sztucznej inteligencji będą musiały wdrożyć. Przykładem takich aktów są przede wszystkim:

1. Projekt Rozporządzenia Parlamentu Europejskiego i Rady (UE) w sprawie operacyjnej odporności cyfrowej sektora finansowego i zmieniające rozporządzenia (WE) nr 1060/2009, (UE) nr 648/2012, (UE) nr 600/2014 oraz (UE) nr 909/2014 („DORA") oraz

2. Projekt Dyrektywy Parlamentu Europejskiego i Rady (UE) w sprawie środków na rzecz wysokiego wspólnego poziomu cyberbezpieczeństwa na terytorium Unii, uchylająca dyrektywę (UE) 2016/1148 tzw. NIS2.

Wymogi określone w tych aktach będą miały przynajmniej pośrednie przełożenie na systemy zarządzania ryzykiem. Nie można także zapominać o wymogach, które mogą wynikać bezpośrednio z takich aktów jak RODO czy ustawa o przeciwdziałaniu praniu pieniędzy i finansowania terroryzmu, które mogą zawierać specyficzne wymogi dla systemów zarządzania ryzykiem „na styku" z ryzykami inherentnymi dla systemów sztucznej inteligencji.

W dalszej części opracowania przyjmujemy, że wymogi określone w art. 9 projektowanego rozporządzenia w sprawie sztucznej inteligencji (jak również przepisy uzupełniające, w tym określone w stosownych załączniku[194]) będą stanowiły punkt wyjścia dla banków zamierzających wykorzystywać systemy sztucznej inteligencji (wysokiego ryzyka), a przedstawione poniżej rozwiązania stanowią modelowy system zarządzania ryzykiem w tym obszarze. Nie można jednocześnie wykluczyć, że w toku dalszych prac legislacyjnych, wymogi te ulegną zmianie.

Głównym wymogiem określonym w projektowanym art. 9 AIA jest obowiązek ustanowienia, wdrożenia, udokumentowania i utrzymania systemu zarządzania ryzykiem w odniesieniu do systemów sztucznej inteligencji wysokiego ryzyka, przy czym ze względu na ryzyka charakterystyczne dla systemów sztucznej inteligencji banki powinny rozważyć szersze stosowanie przedmiotowych postanowień, tj. wykraczających wyłącznie poza systemy wysokiego ryzyka. Sam system zarządzania ryzykiem powinien składać się z ciągłego, iteracyjnego procesu realizowanego przez cały cykl życia systemu sztucznej inteligencji wysokiego ryzyka[195], który powinien podlegać regularnej aktualizacji – zasady, w tym okresowość lub wskazanie czynników zobowiązujących do podjęcia działań, powinny być określone w stosownych procedurach.

Propozycja Komisji Europejskiej zakłada, że taki system powinien składać się z następujących etapów:

1. Identyfikacji i analizy znanego i dającego się przewidzieć ryzyka związanego z każdym systemem sztucznej inteligencji wysokiego ryzyka.

2. Oszacowania i oceny ryzyka, jakie może wystąpić podczas wykorzystywania systemu zgodnie z jego przeznaczeniem i w warunkach dającego się racjonalnie przewidzieć niewłaściwego wykorzystania (np. niezgodnie z instrukcją przekazywaną użytkownikowi).

3. Oceny innego mogącego wystąpić ryzyka na podstawie analizy danych zebranych z systemu monitorowania po wprowadzenia do obrotu (dla dostawców takich systemów).

4. Przyjęcia odpowiednich środków zarządzania ryzykiem zgodnie z przepisami określonymi w art. 9.

Warto w tym miejscu nadmienić, że bank może występować zasadniczo w dwóch rolach – dostawcy rozwiązania (np. na skutek działań zespołów R&D), jak i użytkownika. Może to więc „aktywować" różne obowiązki po stronie instytucji.

Pewne trudności może sprawiać proces identyfikacji ryzyk charakterystycznych dla systemów sztucznej inteligencji, bowiem te mogą mieć bardzo zróżnicowany charakter, obejmujący nie tylko kwestie techniczne czy związane z bezpieczeństwem, ale także prawne i regulacyjne dotyczące dyskryminacji czy biznesowe, związane z konkretnym komercyjnym wykorzystaniem systemu. Z tego względu rekomendowanym rozwiązaniem jest tworzenie interdyscyplinarnych zespołów, które będą odpowiedzialne za zidentyfikowanie

[194] https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0012.02/DOC_2&format=PDF (dostęp: 19.05.2022 r.).

[195] Przy czym warto w tym miejscu rozważyć stosowanie wytycznych ENISA w sprawie bezpieczeństwa oprogramowania. Wytyczne te odnoszą się wprawdzie do rozwiązań typu IOT, jednakże mają wiele uniwersalnych elementów, które z powodzeniem można stosować w odniesieniu do oprogramowania opartego o uczenie maszynowe czy głębokie. Dokument dostępny pod adresem: https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot-1/@@download/fullReport (dostęp: 19.05.2022 r.).

konkretnych ryzyk. Ważną rolę będą w tym miejscu pełniły także jednostki odpowiedzialne za szeroko rozumiany obszar danych, w tym *data science*, ale także ochrony danych osobowych. Stworzenie dokładnej matrycy ryzyk jest etapem niezwykle istotnym z punktu widzenia instytucji, bowiem ich nieprzewidziana materializacja może mieć doniosłe (negatywne) skutki dla funkcjonowania całej instytucji, także w sferze administracyjno-prawnej czy odszkodowawczej, jak również reputacyjnej. W idealnym scenariuszu bank powinien prześledzić wszystkie procesy związane z wykorzystywaniem lub planowanym wykorzystywaniem systemów sztucznej inteligencji wysokiego ryzyka i zidentyfikować potencjalne ryzyka, które mogą wiązać się z tymi procesami.

Warto tutaj zwrócić uwagę na wspomniane już rekomendacje w ramach normy ISO/IEC 38507:2022, w których wskazuje się na potencjalne źródła ryzyk związanych z wykorzystaniem systemów AI. Należą do nich:

1. nieodpowiednie źródła danych,
2. wadliwy łańcuch dostaw,
3. brak wyjaśnialności algorytmów i modeli,
4. zagrożenia z zakresu cyberbezpieczeństwa,
5. niejasne specyfikacje (techniczne),
6. brak ekspertyzy w zakresie sztucznej inteligencji,
7. stronniczość algorytmiczna.

Jednym z istotnych ryzyk jest także ryzyko nadmiernego polegania przez ludzi na systemach sztucznej inteligencji, w szczególności w przypadku systemów decyzyjnych lub rekomendacyjnych i predykcyjnych, które mogą w sposób istotny wpływać na decyzyjność człowieka.

Materializacja powyższych ryzyk może mieć negatywne skutki dla całej organizacji i dlatego powinny być one uwzględnione w analizie ryzyk charakterystycznych dla tego obszaru.

Kolejnym istotnym elementem jest odpowiednie dopasowanie narzędzi zarządzania tymi ryzykami. Art. 9 ust. 3 projektowanego AIA przewiduje, że stosując te środki należy uwzględniać skutki i możliwe interakcje wynikające z łącznego stosowania wymogów określonych w całym rozporządzeniu, ale w przypadku banków katalog możliwych interakcji jest zdecydowanie szerszy i obejmuje akty prawne i regulacje o charakterze sektorowym. Ważne jest przy tym to, aby stosując te rozwiązania uwzględniać powszechnie uznawany stan techniczny i ewentualne normy zharmonizowane czy wspólne specyfikacje (m.in. wspomniane już normy wydawane przez IEEE, ale także organizacje branżowe). Zwrócić należy także uwagę na fakt, że środki zarządzania ryzykiem powinny być tak dopasowane, aby wszelkie ryzyko szczątkowe związane z każdym zagrożeniem, jak również ogólne ryzyko szczątkowe systemów sztucznej inteligencji wysokiego ryzyka, oceniano jako dopuszczalne, jedynie pod warunkiem, że te systemy wykorzystywane są zgodnie z ich przeznaczeniem lub w warunkach racjonalnie przewidywalnego niewłaściwego wykorzystania (te aspekty powinny znaleźć odzwierciedlenie także w stosownej dokumentacji technicznej oraz instrukcji).

Konstruując zestaw narzędzi do zarządzania tymi ryzykami należy zapewnić (art. 9 ust. 4 AIA):

1. eliminację lub ograniczenie ryzyka w możliwie największym stopniu poprzez odpowiedni projekt systemu i proces jego opracowywania,
2. wdrożenie odpowiednich środków służących ograniczeniu i kontroli ryzyka, którego nie można wyeliminować (np. ryzyko niedokładności, błędów),
3. dostarczenie odpowiednich informacji, w szczególności w odniesieniu do wybranych ryzyk,
4. przeszkolenie użytkowników, jeżeli wynika to z charakteru systemu.

Istotnymi elementami są tutaj także kwestie o charakterze organizacyjnym i osobowym, bowiem system zarządzania ryzykiem dla systemów AI wymaga, aby nadzór nad poszczególnymi jego komponentami sprawowały osoby posiadające odpowiednią wiedzę, w tym także techniczną, wykształcenie czy doświadczenie oraz stosowne certyfikaty, jeżeli jest to wymagane. Ważne jest także zapewnienie, aby całokształt rozwiązań (instrumentów) wykorzystywanych w ramach takiego systemu poddany został uprzedniemu testowaniu, jak również testom okresowym, które pozwolą na ocenę, czy spełniają one swoją rolę.

Podkreślić także należy, że narzędzia do zarządzania ryzykami charakterystycznymi dla systemów sztucznej inteligencji mają zróżnicowany charakter i nie są wyłącznie instrumentami o technicznym charakterze, to z pewnością badanie ich skuteczności może wymagać zastosowania specjalistycznego oprogramowania. Bank powinien podejść indywidualnie i proporcjonalnie do identyfikacji i oceny ryzyk zgodnie z rekomendacjami wskazanymi powyżej.

Zależności pomiędzy systemami sztucznej inteligencji a technologią blockchain oraz komputerami kwantowymi

6.1. Technologia blockchain i rozproszonego rejestru

Rozwiązania w zakresie sztucznej inteligencji bez wątpienia będą integrować się z innymi rozwiązaniami technologicznymi, które obecnie są rozwijane lub wykorzystywane w różnych aspektach w biznesie[196]. W bankowości, czy też w branży FinTech, również można spodziewać się w przyszłości dojrzałych i w pełni funkcjonalnych połączeń AI z rozwiązaniami bazującymi na blockchain[197]. Takie połączenia są bowiem technicznie możliwe[198]. Zainteresowanie nimi będzie tym większe, im więcej pojawi się usług, w których nowoczesne narzędzia będą pozytywnie wpływać na rozwiązania biznesowe. Blockchain oferuje w obszarze finansów następujące usprawnienia[199]:

- większe bezpieczeństwo i zaufanie do danych;
- weryfikację informacji, która obejmuje możliwość potwierdzenia autentyczności dokumentu, dyplomu lub innych informacji;
- rozliczalność w zakresie uczenia sztucznej inteligencji, wymagane przez AIA;
- mniejszą złożoność i większą niezawodność, ponieważ wykorzystanie zdecentralizowanej pamięci masowej zmniejsza prawdopodobieństwo, że wyłączenie serwera spowoduje niedostępność danych[200].

Przekłada się to też na trzy główne płaszczyzny adopcji technologii blockhain, które wedle stanowiska ekspertów OECD wykorzystuje się do budowania[201]:

- zaufania w kwestii transakcji;
- zaufania w kwestiach warunków, które będą automatycznie wyegzekwowane;
- zaufania, że coś pochodzi od kogoś lub z jakiegoś miejsca.

Zapotrzebowanie w aspekcie wykorzystania zabezpieczonej i zaufanej informacji jest ogromne, a uzyskanie tych wartości w Internecie determinuje rozwój takich rozwiązań, jak blockchain[202]. Odpowiednie wykorzystanie kryptografii pozwala uzyskać odpowiednio wartościowy zapis w współdzielonym rejestrze. Tak zgromadzone dane mają olbrzymi potencjał informacyjny. W kontekście AI należy podkreślić też, że potrzebuje ona danych (informacji) do działania[203].

Istotne jest w tym miejscu zasygnalizowanie, że choć w biznesie różnica między DLT a blockchain nie jest wyraźna, o tyle w technice jest ona bardzo mocno akcentowana. Ma to też przełożenie na powstające prawo[204].

Rozróżnienie pojęć DLT i blockchain jest też istotne ze względu na fakt, że rozwiązania blockchain, jako otwarte są często krytykowane, ze względu na ich istotną cechę, jaką jest możliwość przystąpienia każdego człowieka do pewnego systemu. Otwarty blockchain to taki, gdzie każdy może przystąpić do partycypowania w projekcie, zarówno jako posiadacz tokenów, osoba udostępniająca moc obliczeniową, czy też programista, który jako członek społeczności proponuje poprawki do oprogramowania stojącego za pewnym rozwiązaniem blockchain[205]. Otwarte rozwiązania bazują też na specyficznym sposobie uzyskiwania konsensusu, tj. zgodności danych w rejestrze. Osoby przyłączające się do systemu, chcące autoryzować transakcje, muszą oddać moc obliczeniową swojego urządzenia. Powoduje to znaczne zużycie energii elektrycznej, i w związku z tym jest mocno krytykowane, a twórcy tego typu rozwiązań blockchain starają się przejść na nowe metody uzyskiwania konsensusu, bardziej zrównoważone pod kątem energetycznym[206].

[196] V. E. Balas, R. Kumar, R. Srivastava, *Recent Trends and Advances in Artificial Intelligence and Internet of Things* [w:] Intelligent Systems Reference Library, vol. 172, 2020, s. 358.

[197] Należy podkreślić, że samo wykorzystanie rozwiązań AI do np. analizowania cen kryptowalut nie jest niczym nowym i jest wykorzystywane np. naukowo. Zob. J. Kyung-Soo Liew, R. Ziyuan, T. Budavari, A. Sharma, *Cryptocurrency Investing Examined,* The Journal of The British Blockchain Association, 2019. Artykuł dostępny pod adresem: https://jbba.scholasticahq.com/article/8720 (dostęp: 01.05.2022 r.). Tak samo wykorzystywanie AI do analizowanie dokumentów, które zostały zabezpieczone z wykorzystaniem rozwiania blockchain. Zob. LawTechUK, *Smarter Contracts*, 2022, s. 23. Opracowanie dostępne pod adresem: https://resources.lawtechuk.io/files/report_smarter_contracts.pdf (dostęp: 04.04.2022 r.).

[198] Zob. Ganesh P. Kumble, *Practical Artificial Intelligence and Blockhain,* Birmingham 2020; zob. też: EC, *First six Artificial Intelligence and Blockchain Technology funds backed by InnovFin raise a total of EUR 700m.* Opracowanie dostępne pod adresem: https://digital-strategy.ec.europa.eu/en/news/first-six-artificial-intelligence-and-blockchain-technology-funds-backed-innovfin-raise-total-eur (dostęp: 01.05.2022 r.).

[199] The European Union Blockchian Observatory and Forum, *Blockchain in trade finance and supply chain,* 2018.

[200] Zob. D. Szostek, *Blockchain and law,* Baden-Baden 2019, s. 34-36. Książka dostępna pod adresem: https://www.nomos-shop.de/nomos/titel/blockchain-and-the-law-id-95162/ (dostęp: 20.04.2022 r.).

[201] Ch. Pike, A. Capobianco, *Antitrust and trust machine,* 2020. Dokument dostępny pod adresem: https://www.oecd.org/daf/competition/antitrust-and-the-trust-machine-2020.pdf (dostęp: 05.05.2022 r.).

[202] K. Werbach, *The Blockchain and the New Architecture of Trust*, London/Cambridge 2018, s. 2-7.

[203] W. Ertel, *Introduction to Artificial Intelligence*, London 2011, s. 180.

[204] D. Szostek, *Blockchain..., op.cit.,* s. 34-37.

[205] Należy podkreślić, że grupa programistów jest ważna w cyklu życia tego typu projektów. W aktualizacjach bitcoina brali udział akademicy z MIT, co spowodowało, że projekt mógł się dalej rozwijać i stał się bardziej skalowalny. Zob. D. Shrier, *Basic Blockchain: What It is and How It Will Transform the Way We Work and Live,* Londyn 2020.

[206] A. Castro, *Why Ethereum is switching to proof of stake and how it will work*, MIT Technology Review, 2022. Artykuł dostępny pod adresem: https://www.technologyreview.com/2022/03/04/1046636/ethereum-blockchain-proof-of-stake (dostęp: 20.04.2022 r.).

Tabela 1. Podział rozproszonych rejestrów

	Kontrola dostępu	Dostęp publiczny do historii transakcji	Kontrola sposobu uzyskiwania konsensusu
Publiczny	Nie	Tak	Nie
Prywatny	Tak	Nie	Tak
Hybrydowy	Tak	Nie	Nie

Z punktu widzenia wykorzystania AI w rozproszonych rejestrach nie można wykluczyć żadnego rozwiązania. W związku z tym podzielić należy rozproszone rejestry na publiczne (otwarte), prywatne oraz hybrydowe, zgodnie z tabelą 1. Na chwilę sporządzania tego raportu najbardziej znanym blockchainem, który oferuje otwartość i dodatkowe możliwości w postaci wirtualizacji i tworzenia smart kontraktów, jest Ethereum[207]. Należy więc wskazać, że trendem, który będzie się rozwijał, są rozwiązania łączone. O ich potencjale zadecydują przyszłe decyzje w zakresie zarządzania Ethereum, w tym najważniejsza decyzja, dotycząca problemu sposobu autoryzacji transakcji.

Smart Contract wykorzystywany w rozwiązaniach, takich jak np. Ethereum, daje możliwość zautomatyzowania za pomocą kodu pewnych reguł, które umożliwiają egzekucję, gdy spełnione zostaną odpowiednie warunki[208]. Z punktu widzenia unijnego prawa wskazano na dwie wartościowe możliwości wykorzystania tej technologii dla budowania cyfrowej niezależności Uni Europejskiej:

- blockchian i eIDAS2;
- rozpoznanie i uznanie rejestrów w blockchain.

Nowe regulacje spowodowały pojawienie się również pomysłów uwzględniających bardziej zaawansowane rozwiązania, w których to kod regulowałby zasady. Jako przykład należy wskazać tworzenie całych organizacji, gdzie głosowanie odbywałoby się na bazie rozdysponowanych na pewnych ustalonych zasadach tokenów. Tak zaawansowane inteligentne kontrakty skłoniły do sformułowania nurtu, w którym proponuje się umieszczenie odpowiednio zaawansowanej AI w tego typu organizacji, gdzie kod pełniłby rolę ram dla jej funkcjonowania. Pozwoliłoby to też nadzorować lub kontrolować wykorzystanie danych, co w kwestii gospodarki cyfrowej nie pozostaje bez znaczenia[209]. Należy również podkreślić, że w zakresie przemysłu 4.0 znajdują się też aspekty związane z finansami[210].

Postrzeganie rozwiązań na bazie blockchain, jako narzędzia do bezpiecznego i kontrolowanego rozwijania AI widzi też europejska nauka, wskazując na wymierne korzyści nakierowane na[211]:

- przejrzystość;
- odpowiedzialność;
- odpowiednią dynamikę rozwoju.

Takie podejście powoduje też, że konieczne staje się postawienie pytania, kto miałby takie narzędzia do kontroli stworzyć. Niewątpliwie podejście prezentowane w nauce znajduje odzwierciedlenie w raportach, które analizują sam fakt stosowania kodu, jako regulacji[212]. Jest to więc trend, który może ewoluować i dotknąć też branżę finansów.

Odpowiednio przedstawione ramy postępowania spowodują, że AI będzie działała etycznie. Na przykład uniemożliwią AI zbieranie danych, które są zbyteczne, nadmiarowe lub przesyłanie danych poza wyznaczony krąg odbiorców[213].

Nowe rozwiązania na rynku finansowym nie tylko się rozwijają, ale spowodowały również powstanie w prawie nowej gałęzi – prawa kryptoaktywów[214]. Nie jest jednak

[207] V. Butterin, *Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform*, 2014. Opracowanie dostępne pod adresem: https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf (dostęp: 10.04.2022 r.).

[208] D. Szostek, *Blockchain..., op. cit.*, s. 110-111.

[209] A. Sulkowski, *Industry 4.0. Era technology (AI, BIG DATA, BLOCKCHAIN, DAO): Why the Law Needs New Memes,* Kansas Journal of Law & Public Policy vol. XXIX:1, 2019, s. 9-12.

[210] M. Javaid, A. Haleem, R. P. Singh, S. Khan, R. Suman, *Blockchain technology applications for Industry 4.0: A literature-based review,* Blockchain: Research and Applications vol. 2, Issue 4, 2021.

[211] D. Szostek, *Is the Traditional Method of Regulation (the Legislative Act) Sufficient to Regulate Artificial Intelligence, or Should It Also Be Regulated by an Algorithmic Code?,* Bialystok Legal Studies vol. 26, 3, 2020. Artykuł dostępny pod adresem: http://bsp.uwb.edu.pl/wp-content/uploads/2021/10/43_bsp-26-3.pdf (dostęp: 05.05.2022 r.).

[212] T. Barraclough, H. Fraser, C. Barnes, *Legislation as code for New Zealand*, 2021. Dokument dostępny pod adresem: https://www.lawfoundation.org.nz/wp-content/uploads/2021/03/Legislation-as-Code-9-March-2021-for-distribution.pdf (dostęp: 05.05.2022 r.).

[213] Zagadnienie to wykracza jednakże poza ramy niniejszego raportu.

[214] K. Zacharzewski, *Digital Asset Capital Market Law: A New Discipline of Private Law,* Krytyka Prawa t.13, 2, 2021.

celem tego raportu zagłębianie się w zagadnienia dotyczące tokenów. Istotnym elementem natomiast jest zasygnalizowanie ogólnych aspektów prawnych, które są następstwem stosowania rozwiązań blockchain, jak i innych DLT. W tym celu organizacja ITU (ang. *International Telecommunication Union*) przygotowała klucze do analizy wybranych rozwiązań. Między innymi znalazł się wśród nich klucz do analizy prawnej, szczególnie, że rozwiązania DLT są globalne i mogą działać w różnych krajach i systemach prawnych (zob. rys. 2).

Korzystanie ze wszelkich możliwości, jakie oferują rozwiązania w ramach DLT, ale też i udostępnianie wykorzystania tych możliwości innym jest związane z zorganizowaniem i przygotowaniem odpowiedniej infrastruktury. DLT może funkcjonować w trzech modelach: publicznym, prywatnym i hybrydowym. W sytuacji mnogości rozwiązań oraz innych specyfikacji technicznych dla każdego z nich istotne jest zachowanie daleko idącej neutralności technologicznej. Trend technologiczny wskazuje, że kierunkiem dającym obopólne korzyści wraz ze spełnieniem różnych wymogów prawnych jest stosowanie rozwiązania *crosschain*. Oznacza to, że każdy rodzaj DLT będzie się rozwijał w sferze biznesowej, w zależności od zastosowania i potrzeb. Trend regulacyjny, zarówno w UE, jak i innych krajach wskazuje, że konieczne będzie dookreślenie definicji DLT[215].

Należy również zauważyć, że nowo wprowadzane kwalifikowane rejestry elektroniczne w ramach eIDAS2 uwzględniają kwestie związane z rynkiem kryptoaktywów. Zgodnie z preambułą tego aktu: *„W przypadku, gdy rejestry elektroniczne wykorzystuje się w obsłudze emisji obligacji lub obrotu nimi, lub na potrzeby kryptoaktywów, przypadki użycia powinny być zgodne z wszystkimi mającymi zastosowanie przepisami finansowymi, np. z dyrektywą w sprawie rynków instrumentów finansowych, dyrektywą w sprawie usług płatniczych i przyszłym rozporządzeniem w sprawie rynków kryptoaktywów"*[216]. W związku z tym przyjąć należy, że w zgodzie z obowiązującymi przepisami prawa kryptoaktywów, regulacja eIDAS przewiduje domniemania prawne dla kwalifikowanych rejestrów. Mogą nimi być rejestry DLT, jak także i blockchain[217].

Niewykluczone też, że pojawiają się rozwiązania skierowane w kierunku osiągnięcia korzyści społecznych w związku z wykorzystaniem zarówno rozwiązań AI, jak i DLT. Jak wskazano wcześniej, kierunek ten może przybrać też formę nowego podejścia w aspekcie regulowania funkcjonowania rozwiązań wykorzystujących AI.

[215] D. Szostek, *Blockchain...*, *op. cit.*, s. 34-37.
[216] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0281 (dostęp: 05.05.2022 r.).
[217] Zagadnienie to wykracza jednakże poza ramy niniejszego raportu.

6.2. Komputery kwantowe

Dynamiczny rozwój technologii kwantowych sprawia, że na obecny moment trudno jednoznacznie sprecyzować, jak wpłynie on na kształtowanie się regulacji, w tym regulacji odnoszących się do rynku finansowego. Szybkość przetwarzania informacji, w tym informacji, które zawierają dane wymagające szczególnej ochrony jak np. dane osobowe[218], będzie kluczowym czynnikiem wywierającym znaczący wpływ na różne sfery gospodarki, co skutkować będzie koniecznością jego regulacji. Skala nadchodzących zmian oraz ich dynamika powoduje, że twórcy skupieni wokół rozwoju technologii kwantowych już teraz apelują o dyskusję[219], w tym dialog na płaszczyźnie etyczno-regulacyjnej, która może być znacząca dla dalszego kierunku rozwoju regulacji w przedmiotowym obszarze. Chris J. Hoofnagle i Simson L. Garfinkel formułują trzy cele, które na płaszczyźnie regulacyjnej powinny być odpowiednio zabezpieczone normami technicznymi i prawnymi[220]:

1. Inżynieria kwantowa ma potencjał, by przynieść głębokie korzyści dla społeczeństwa ludzkiego, o ile będzie dominować nastawienie na wykorzystanie jej w zakresie cywilnym;

2. Wymagana jest dyskusja publiczna nad zastosowaniem inżynierii kwantowej w sferze bezpieczeństwa. Nowe rozwiązania mogą spowodować bezprecedensowy poziom i skalę inwigilacji i ingerencji w sferę prywatną;

3. Możliwości, jakie przyniosą detekcja kwantowa, obliczenia kwantowe i informatyka kwantowa mogą doprowadzić do niszczącej destabilizacji infrastruktury cywilnej oraz podważyć zaufanie społeczne i mechanizmy integralności wywodzące się z prawa publicznego i prywatnego, a nawet mechanizmy historycznie wypracowane i zaadoptowane przez społeczeństwo obywatelskie.

Oznacza to, że niektóre sektorowe wymagania dotyczące ustalania granic inżynierii kwantowej w branżach wysokiego ryzyka, takich jak zdrowie, żywność, energia, bezpieczeństwo, finanse i obrona, należy uznać za bardziej rygorystyczne niż przyjęte zasady i standardy obowiązujące w branżach kwalifikowanych jako niższego ryzyka, takich jak

[218] Edward J. Swan, *Internet Law. A Concise Guide to Regulation Around the World*, Alphen aan den Rijn 2022, Chapter 21, Section A.
[219] Mauritz Kop, *Establishing a Legal-Ethical Framework for Quantum Technology,* Yale Journal of Law & Technology, March 2021. Artykuł dostępny pod adresem: https://yjolt.org/establishing-legal-ethical-framework-quantum-technology#_ftnref36 (dostęp: 05.05.2022 r.).
220 C. J. Hoofnagle, S. L. Garfinkel, *Law and policy for the quantum age,* Cambridge 2022, s. 375-377.

rozrywka i sztuka itd. Biorąc pod uwagę powyższe należy stwierdzić, że nowoczesna bankowość i branża FinTech niewątpliwie będą musiały zmierzyć się z nowymi obowiązkami, w szczególności w sytuacji, w której będą chciały wdrażać rozwiązania z zakresu inżynierii kwantowej, w celu budowania swojej przewagi biznesowej. W literaturze przedmiotu zostały wstępnie sformułowane również obszary ryzyka, jakie mogą się pojawić w momencie osiągnięcia przez inżynierię kwantową tzw. dojrzałości, tj. [221]:

1. ryzyko wzrostu nierówności w fazie początkowej;
2. ryzyko dla stabilności systemu finansowego;
3. ryzyko związane z prywatnością danych, bezpieczeństwem danych, pewnością prawną i zaufaniem;
4. ryzyko związane z fałszywymi wiadomościami, dezinformacją eoraz ich wpływ na procesy demokratyczne;
5. ryzyko związane z nadzorem i kontrolą państwa;
6. ryzyko związane ze zmianą stosunków geopolitycznych.

Należy zwrócić uwagę, że pomimo wyszczególnienia jako osobnej kategorii ryzyka dla stabilności systemu finansowego, rynek finansowy będzie musiał również się zmierzyć z problematyką związaną z ochroną danych klientów, zapewnieniem etycznych rozwiązań w zakresie wykorzystywania tych danych, a także dalszego rozwoju zabezpieczeń danych i dobrych praktyk, szczególnie w zakresie inżynierii kwantowej, jakim jest kryptoanaliza kwantowa[222].

[221] M. Kop, *Regulating Transformative Technology in The Quantum Age: Intellectual Property*, Standardization & Sustainable Innovation, 2 TTLF Newsletter on Transatlantic Antitrust and IPR Developments Stanford-Vienna Transatlantic Technology Law Forum, Stanford University 2020. Artykuł dostępny pod adresem: https://yjolt.org/establishing-legal-ethical-framework-quantum-technology#_ftnref36 (dostęp: 05.05.2022 r.).

[222] M. Mosca, *Cybersecurity in an Era with Quantum Computers: Will We Be Ready?,* IEEE Security & Privacy, vol. 16, September/October 2018, s. 38-41.

ROZDZIAŁ VII

PROGRAM ANALITYCZNO BADAWCZY

Rekomendacje dla banków oraz Związku Banków Polskich

7.1. Uwagi wstępne

Dalszy rozwój systemów sztucznej inteligencji w sektorze bankowym jest uzależniony od szeregu czynników o zróżnicowanym charakterze, zarówno zewnętrznych, jak i wewnętrznych. Nowe akty prawne i regulacje dotyczące tego obszaru mogą mieć pozytywny wpływ (większa przejrzystość w zakresie wymogów), jak i negatywny, np. ograniczenie innowacyjności i rosnące koszty obsługi. Patrząc przez pryzmat proponowanych rozwiązań prawno-regulacyjnych, banki mogą oczekiwać znacznego wzrostu kosztów operacyjnych w związku ze stosowaniem systemów sztucznej inteligencji.

Jednocześnie trudno jednoznacznie przewidywać dzisiaj, jaki ostateczny kształt przyjmie przykładowo rozporządzenie w sprawie sztucznej inteligencji i jak będą kształtowały się poszczególne wymogi czy nawet sama kategoryzacja systemów sztucznej inteligencji, która będzie miała wpływ na to, w jakim zakresie banki będą zobowiązane do przystosowania się do nowych wyzwań prawno-regulacyjnych.

Jednocześnie jednak banki zdają się być w stosunkowo dobrej sytuacji, bowiem znaczna część wymogów, które prawdopodobnie będą wynikać z rozporządzenia w sprawie sztucznej inteligencji, będzie mogła być „pokryta" już istniejącymi wymogami o charakterze sektorowym, np. wspomniane systemy zarządzania ryzykiem czy wymogi w zakresie zarządzania danymi. Nie oznacza to jednak, że po stronie banków nie będzie konieczne podjęcie aktywności w tym obszarze.

Warto także zwrócić uwagę, że na znaczeniu zyskuje ostatnio problematyka etyki sztucznej inteligencji, która pojawia się przykładowo w wytycznych EIOPA[223], ale także wytycznych sektorowych. Może to wymagać, aby banki oraz organizacje branżowe, w tym Związek Banków Polskich, podjęły decyzję o kierunku rozwoju w tym kontekście. Pożądana będzie też reakcja ze strony organów nadzoru i regulacji.

Wiele z działań na poziomie krajowym, w tym także realizowanych przez Urząd Komisji Nadzoru Finansowego, może być uzależnionych – przynajmniej pośrednio – od inicjatyw podejmowanych na poziomie Unii Europejskiej, zarówno tych o charakterze prawodawczym, jak i regulacyjnym. Jest to także szansa dla polskiego sektora bankowego oraz wsparcia technologicznego, aby stać się jednym z centrów wykorzystania nowych technologii w sektorze finansowym, co jest też wskazywane przez Fundację FinTech Poland[224]. Ważne w tym kontekście jest promowanie rozwiązań o charakterze RegTech[225], które mogą przyczynić się do bardziej dynamicznego wykorzystania tzw. sztucznej inteligencji.

Nie można przy tym zapominać, że ważnym aspektem w kontekście wykorzystania systemów sztucznej inteligencji jest problematyka istniejących systemów (*legacy*) stosowanych przez banki, które mogą być niedostosowane do wymagań stawianych przez nowe regulacje. Jest to jednak zagadnienie znacząco wykraczające poza ramy niniejszego opracowania.

Tym samym, sektor bankowy w Polsce stoi w obliczu znaczących zmian, które determinowane będą jedynie w pewnym stopniu przez akty prawne i regulacje, choć niewątpliwie będą miały one wpływ na wiele kwestii istotnych z perspektywy sektora bankowego. W niniejszym opracowaniu nie dokonano głębszej analizy aspektów kulturowych, w tym tworzenia innowacyjnych organizacji o formule *data-driven*, co jest jednak zagadnieniem ważnym w omawianym w opracowaniu kontekście.

Na koniec warto zwrócić uwagę, że pierwotne brzmienie projektu rozporządzenia w sprawie sztucznej inteligencji przewiduje w wielu miejscach, że wymogi dla systemów sztucznej inteligencji wysokiego ryzyka będą mogły być „realizowane" w ramach istniejących rozwiązań, co wymagać będzie po stronie banków analizy istniejącego modelu zarządzania.

Poniżej zaprezentowane zostały ogólne rekomendacje kierowane zarówno do banków, jak i Związku Banków Polskich, stanowiącego głos banków na arenie krajowej i międzynarodowej. Realizacja tych rekomendacji będzie wymagała jednak ścisłej współpracy pomiędzy instytucjami finansowymi, dostawcami rozwiązań technologicznych oraz organami nadzorczymi i regulacyjnymi, jak i prawodawcami.

7.2. Rekomendacje kierowane do banków:

1. Ustanowienie w ramach organizacji jednostek interdyscyplinarnych odpowiedzialnych za obszar szeroko rozumianej sztucznej inteligencji, jak również wyznaczenie członka zarządu odpowiedzialnego za ten obszar.

[223] https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf (dostęp: 20.05.2022 r.).

[224] http://fintechpoland.com/initiatives2/#ini-nextgen (dostęp: 20.05.2022 r.).

[225] https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Reports/2021/1015484/EBA%20analysis%20of%20RegTech%20in%20the%20EU%20financial%20sector.pdf (dostęp: 20.05.2022 r.).

2. Stworzenie odpowiednich ram organizacyjnych zapewniających odpowiedzialność, raportowanie i kompetencje w zakresie systemów sztucznej inteligencji.

3. Opracowanie i wdrożenie strategii dla danych i/lub sztucznej inteligencji, która pokrywać będzie zarówno perspektywę krótko-, średnio-, jak i długoterminową, a następnie zakomunikowanie jej w ramach organizacji.

4. Tworzenie w organizacji kultury data-driven, która opiera się o wspólne zrozumienie i wykorzystanie danych.

5. Monitorowanie zmian prawnych i regulacyjnych w obszarze tzw. sztucznej inteligencji, jak również reagowanie na te zmiany z odpowiednimi wyprzedzeniem.

6. Ustanowienie w ramach organizacji komitetu odpowiedzialnego za etyczne wykorzystanie systemów sztucznej inteligencji oraz danych.

7. Przeszkolenie pracowników i kontraktorów w zakresie ryzyk, które mogą generować systemy sztucznej inteligencji, jak również ich regularne przeprowadzanie na poziomie organizacji.

8. Przeprowadzenie analizy istniejących procesów, które mogą wykorzystywać systemy sztucznej inteligencji lub podobnych pod kątem podatności na ryzyka charakterystyczne dla AI, a także podjęcie działań w celu ich eliminacji.

9. Dokonanie inwentaryzacji dostępnych rozwiązań infrastrukturalnych oraz oprogramowania i identyfikacja ryzyk.

Rekomendacje kierowane do Związku Banków Polskich:

1. Utworzenie interdyscyplinarnej grupy roboczej lub jednostki odpowiedzialnej za obszar systemów sztucznej inteligencji.

2. Włączenie się przez Związek Banków Polskich w inicjatywy prawne i regulacyjne, a także polityczne, które mogą wpływać na kształt przyszłych aktów prawnych i regulacji.

3. Opracowanie wytycznych w zakresie odpowiedzialnego stosowania systemów sztucznej inteligencji w sektorze bankowym, w tym przykładowo ze względu na obszary działalności, w tym operacyjnej, jak np. przeciwdziałanie praniu pieniędzy i finansowania terroryzmu, modeli wewnętrznych etc.

4. Wypracowanie zmian do Kodeksu Etyki Bankowej, aby ten uwzględniał także wyzwania o charakterze etycznym związanym ze stosowaniem nowych technologii.

5. Rozpoczęcie dialogu z organami nadzoru i regulacji w celu ewentualnego wypracowania wytycznych o charakterze sektorowym.

Wiele wskazuje na to, w tym prace Prezydencji słoweńskiej i francuskiej, iż AIA zostanie w nieodległym czasie przyjęty. Jego zakres oraz treść jest aktualnie przedmiotem intensywnych prac w komisjach Parlamentu Europejskiego. Na poziomie Unii Europejskiej odbywają się burzliwe dyskusje co do obranego kierunku regulacji AI. Powoduje to, że nie jest możliwe wskazanie rekomendacji bazujących wyłącznie na projektowanym akcie prawnym, choć jednocześnie z dużą dozą prawdopodobieństwa można założyć, że wiele ze wskazanych tam przepisów stanie się częścią uchwalonego rozporządzenia.

Jednocześnie nie oznacza to, że banki nie powinny dostosowywać swoich rozwiązań organizacyjno-technicznych do wyzwań i ryzyk związanych ze stosowaniem systemów sztucznej inteligencji. Postępująca algorytmizacja wielu obszarów związanych z szeroko rozumianym sektorem finansowym wymaga, aby już dzisiaj instytucje finansowe działały w sposób etyczny[226] i „wyprzedzający" akty prawne i regulacje. Wiele z rozwiązań już obowiązujących banki, np. w obszarze kontroli wewnętrznej, ma istotne znaczenie dla prawidłowego i bezpiecznego funkcjonowania algorytmów i modeli AI. Z tego względu wdrażanie – na zasadzie dobrych praktyk – wielu z aspektów projektowanego rozporządzenia czy wytycznych unijnych organów regulacyjnych i nadzorczych, wydaje się nie tylko celowe, ale wręcz konieczne, przynajmniej do czasu opracowania przez krajowy organ nadzoru własnych rekomendacji w tym obszarze. Należy przy tym zwrócić uwagę, że brak właściwych rozwiązań, w tym organizacyjno-technicznych, w bankach może mieć negatywne skutki dla całego systemu finansowego, jak i gospodarki.

Warto w tym miejscu podkreślić, że wdrażanie odpowiednich rozwiązań powinno odbywać się z poszanowaniem i pełnym zrozumieniem takich zasad jak:

[226] Więcej na temat etyki sztucznej inteligencji w sektorze finansowym [w:] M. Nowakowski, K. Waliszewski, *Ethics of Artificial Intelligence in the Financial Sector*, Przegląd Ustawodawstwa Gospodarczego 01/2022, s. 2-9. Artykuł dostępny pod adresem: https://www.pwe.com.pl/czasopisma/przeglad-ustawodawstwa-gospodarczego/numery-czasopisma/przeglad-ustawodawstwa-gospodarczego-012022,p1749805129 (dostęp: 05.05.2022 r.).

a. podejście oparte na ryzyku (*risk-based approach),*

b. neutralność technologiczna,

c. zasada proporcjonalności.

Jednocześnie banki powinny bardziej intensywnie wdrażać rekomendacje zawarte w rekomendacji D KNF, ale w kontekście systemów AI i budowania kultury typu *data-driven*, która w założeniu ma pozwolić na bardziej efektywne wykorzystanie danych oraz kompetencji i możliwości instytucji finansowych. Zasadnym przy tym wydaje się, aby banki dokonały mapowania własnych procesów, które mogą obejmować wykorzystanie szeroko rozumianych systemów sztucznej inteligencji, także pod kątem identyfikacji ryzyk charakterystycznych właśnie dla tego typu systemów.

Zauważyć należy, że bardziej efektywne wykorzystanie systemów sztucznej inteligencji może wymagać istotnych zmian infrastrukturalnych i oprogramowania po stronie banków. W szczególności wymiany tzw. zastanych systemów (*legacy systems*), które mogą nie być dostosowane do wymagań nowych technologii, także tych, które towarzyszą wykorzystaniu systemów sztucznej inteligencji. Zasadna byłaby tutaj także reakcja ze strony organu nadzoru w zakresie oczekiwań co do przeprowadzanej transformacji cyfrowej.

Jednym z obszarów, który wymagają szczególnej uwagi po stronie zarówno instytucji finansowych, jak i organów regulacyjnych, jest obszar etycznego i odpowiedzialnego wykorzystania systemów sztucznej inteligencji. Niewłaściwy dobór danych, nieuwzględnianie przy projektowaniu rozwiązań aspektów związanych z różnorodnością czy ochroną prywatności, może generować dla instytucji określone ryzyka, w tym reputacyjne. Zaawansowane rozwiązania związane z wykorzystaniem systemów sztucznej inteligencji mogą generować zróżnicowane ryzyka, które nie są tożsame z tymi, które tradycyjnie przypisuje się obszarowi IT. Z tego względu instytucje powinny rozważyć wprowadzenie odpowiednich rozwiązań, np. komitetów ds. etyki AI (lub włączenia tematyki w już istniejące komitety) czy kodeksów etycznych, które w połączeniu z edukacją mogą przyczynić się do bardziej efektywnego i bezpieczniejszego wykorzystania systemów AI. Jednocześnie Związek Banków Polskich powinien rozważyć możliwość rewizji Kodeksu Etyki Bankowej właśnie pod kątem etycznej i odpowiedzialnego wykorzystania systemów sztucznej inteligencji.

Powyższe rekomendacje mają jednocześnie dość ogólny charakter, a każda z instytucji powinna rozsądnie podejść do ich wdrażania, co powinno odbywać się w zgodzie z poziomie obecnego i planowanego wykorzystania systemów sztucznej inteligencji.

Związek Banków Polski powinien także włączyć się aktywnie w dialog z organami nadzorczymi i regulacyjnymi, a także organizacjami branżowymi (np. PONIP, ZNIP etc.), w celu wypracowania jednolitego podejścia do wymogów w zakresie wdrażania rozwiązań opartych o systemy sztucznej inteligencji.

Raport przygotowany na zlecenie Fundacji
Warszawski Instytut Bankowości

PROGRAM
ANALITYCZNO
BADAWCZY