

AI W FINANSACH

Zastosowanie i ryzyko dla konsumentów
oraz uprawnienia wobec instytucji
finansowych

Warszawa, październik 2025 r.



**Rzecznik
Finansowy**

www.rf.gov.pl

prof. Krzysztof Jajuga
dr Marta Karpińska-Bielec
dr Michał Nowakowski
Michał Sas
Igor Sawczuk
prof. Krzysztof Waliszewski

AI w finansach

Zastosowanie i ryzyko dla konsumentów oraz
uprawnienia wobec instytucji finansowych

Rzecznik Finansowy

Warszawa, październik 2025

Spis treści

Wstęp.....	4
Pojęcie sztucznej inteligencji	5
Definicja AI.....	5
Rodzaje AI.....	9
AI Act – akt w sprawie sztucznej inteligencji	11
Kim jest Rzecznik Finansowy i co ma wspólnego z AI.....	12
Zastosowanie AI w usługach i produktach finansowych.....	14
Chatboty.....	14
Systemy zautomatyzowanej oceny zdolności kredytowej.....	15
Robodoradztwo inwestycyjne.....	15
Personalizacja ofert i dopasowywanie produktów do klienta	17
Zastosowanie AI w przeciwdziałaniu praniu pieniędzy i finansowaniu terroryzmu	19
Szacowanie ryzyka ubezpieczeniowego.....	20
Zautomatyzowane badanie potrzeb klienta.....	21
Likwidacja i wycena szkód ubezpieczeniowych.....	23
Korzyści i zagrożenia wynikające ze stosowania AI	25
Korzyści	25
Zagrożenia – rozważania ogólne	25
Bias i halucynacje – metody przeciwdziałania.....	27
Korzyści i zagrożenia na przykładzie robodoradztwa	29
Zagrożenia na przykładzie chatbotów	30
Zagrożenia związane z hiperpersonalizacją.....	31
AI w finansach a regulacje prawne – co musi przedsiębiorca, a co może konsument.....	32
Zakazane systemy AI w ramach AI Act.....	32
Ochrona danych osobowych (RODO) a wykorzystanie AI.....	33
Prawo bankowe a wykorzystanie AI.....	34
Prawo ubezpieczeniowe a wykorzystanie AI	35
Robodoradztwo	36
AI Act – czyli jak będzie już za nieco mniej niż rok.....	39
Termin wejścia w życie	39
Systemy wysokiego ryzyka.....	39
Prawa osób, na które systemy AI mają wpływ	41
Prawo do wniesienia skargi.....	41
Kodeks postępowania w zakresie AI ogólnego przeznaczenia	42

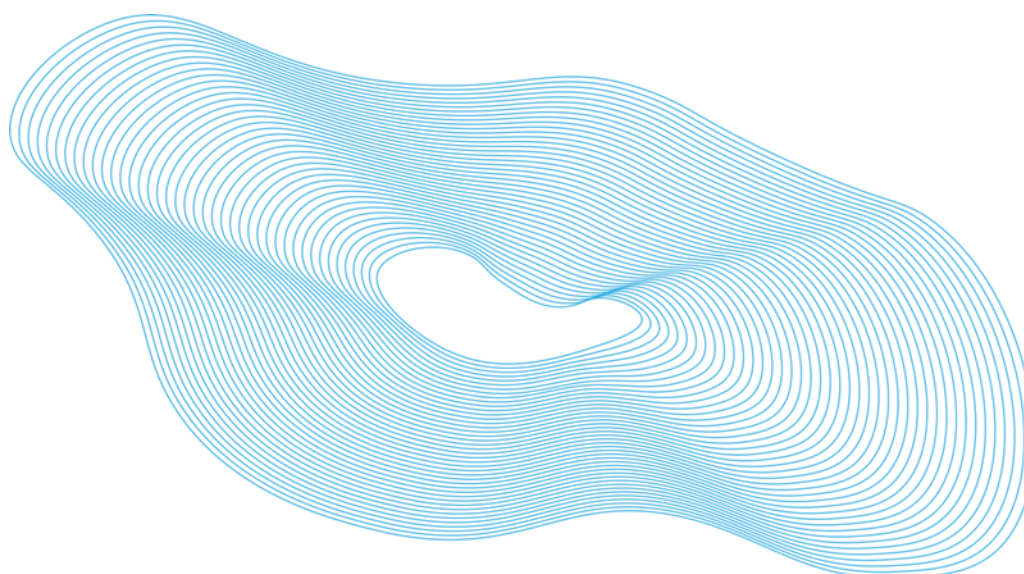
Wstęp

Nie ulega wątpliwości, że termin „sztuczna inteligencja” stał się w ostatnich latach jednym z najczęściej używanych terminów w działalności biznesowej oraz w mediach. Tak dynamiczny przyrost zainteresowania pewnym pojęciem wśród bardzo szerokiego kręgu osób może budzić pewien niepokój. Z jednej strony atakują nas bowiem komunikaty o „AI pozbawiającej ludzi pracy” czy wręcz „realizacji wizji z filmów Terminator”, z drugiej wielu ekspertów próbuje nas przekonać, że wdrożenie sztucznej inteligencji jest ostatnim krokiem do zapewnienia dobrobytu, a wszelkie próby jej regulacji to walka z innowacyjnością.

Tymczasem – jak zwykle – prawda leży gdzieś pośrodku.

Naszą intencją jako autorów tego dokumentu było:

- wyjaśnienie i oswojenie pojęcia sztucznej inteligencji,
- pomoc w zrozumieniu, w jakim zakresie i w jakiej formie sztuczna inteligencja już jest wykorzystywana w sektorze finansowym,
- wskazanie, jakie są obowiązki podmiotów gospodarczych wykorzystujących sztuczną inteligencję i jakie są uprawnienia osób, wobec których stosowana jest ta technologia.



Pojęcie sztucznej inteligencji

Definicja AI

Jednym z pierwszych określeń pojęcia „Sztuczna Inteligencja” jest to, które podał John McCarthy: „Sztuczna inteligencja jest to nauka i inżynieria tworzenia inteligentnych maszyn, zwłaszcza inteligentnych programów komputerowych”.

Sztuczna inteligencja (Artificial Intelligence, dalej określana jako **AI**) to szeroka dziedzina technologii, której celem jest stworzenie programów komputerowych zdolnych do wykonywania zadań, które do tej pory uważano, że wymagają ludzkiego intelektu.

Wyróżnia się trzy główne obszary sztucznej inteligencji, które znacząco zmieniły postrzeganie zdolności komputerów oraz sposób, w jaki postrzega się dziedziny życia zarezerwowane wyłącznie dla ludzkiego intelektu. Jest to uczenie maszynowe, sieci głębokie oraz wielkie modele językowe.

Uczenie maszynowe (Machine Learning) to proces, który polega na znajdowaniu wzorców w dużych ilościach danych. Te wzorce mogą obejmować wszystko, od zależności między różnymi zmiennymi, do powtarzających się sekwencji lub trendów. Na przykład system uczenia maszynowego może nauczyć się, że kiedy temperatura spada, ludzie zaczynają częściej kupować grzejniki – to jest wzorzec, którego system nauczył się z danych. Kiedy system uczenia maszynowego wykryje wzorzec, przekazuje te informacje do sztucznej sieci neuronowej. Sztuczne sieci neuronowe to rodzaj modelu sztucznej inteligencji, który naśladuje sposób, w jaki ludzki mózg przetwarza informacje. Wykorzystują one te wzorce do „uczenia się” i podejmowania decyzji. W przykładzie z grzejnikami sztuczna sieć neuronowa mogłaby nauczyć się, że kiedy temperatura spada, powinna zarekomendować grzejniki do kupienia. W całym procesie budowania wzorców często bierze udział ekspert w danej dziedzinie, tj. osoba posiadająca głęboką wiedzę na temat danego tematu i pomagająca systemowi zrozumieć, jakie wzorce są istotne. Na przykład w odniesieniu do systemu przewidującego sprzedaż grzejników ekspertem mógłby być ktoś, kto ma doświadczenie w sprzedaży grzejników lub ktoś, kto zna się na trendach pogodowych. Ekspert ten przekłada swoje rozumowanie na działanie danej sieci, pomagając systemowi uczenia maszynowego zrozumieć, jakie wzorce są istotne i jakie decyzje powinny być podejmowane na

podstawie tych wzorców. Dzięki temu system staje się coraz lepszy w przewidywaniu i podejmowaniu decyzji.

Głębokie uczenie (Deep Learning) jest bardziej zaawansowaną formą uczenia maszynowego, która polega na pozostawieniu większości procesu decyzyjnego w rękach sztucznej sieci neuronowej. W przeciwieństwie do tradycyjnego uczenia maszynowego, gdzie często potrzeba eksperta do interpretacji wzorców i kierowania procesem uczenia, głębokie uczenie polega na tym, że istnieje tak wiele danych i opisanych wzorców, że można pozwolić modelowi samemu decydować, jakie wzorce są ważne. Można to przedstawić na przykładzie dziecka, któremu da się miliony puzzli i pozwoli mu na złożenie ich razem w dowolny sposób, jakie uważa za stosowne. Z biegiem czasu dziecko (lub w tym przypadku model głębokiego uczenia) zaczyna rozumieć, które kawałki pasują do siebie i jakie wzorce tworzą. W praktyce oznacza to, że sztuczna sieć neuronowa analizuje dane, identyfikuje wzorce, a następnie na podstawie tych wzorców podejmuje decyzje – wszystko to bez wpływu człowieka. To podejście sprawiło, że wiele zadań, które wcześniej były trudne dla komputerów, stało się znacznie prostszych. Na przykład rozpoznawanie obiektów na zdjęciach – coś, co ludzie robią naturalnie i bez wysiłku – stało się dużo łatwiejsze dla komputerów dzięki głębokiemu uczeniu. Aktualnie model uczenia maszynowego może przejrzeć miliony zdjęć, samodzielnie nauczyć się rozpoznawać różne obiekty, a następnie prawidłowo identyfikować te obiekty na nowych zdjęciach.

Wielkie modele językowe (Large Language Models, dalej określane jako **LLM**) są najnowszą innowacją w dziedzinie sztucznej inteligencji. Są to modele oparte na technologii głębokiego uczenia, które zostały wytrenowane na milionach stron tekstów z całego Internetu i wielu księgozbiorów. Pierwotnym celem LLM było nauczenie sztucznej sieci neuronowej kończenia zdań tekstowych. Innymi słowy, zadaniem modelu było przewidzenie, jakie słowo lub fraza najprawdopodobniej pojawi się następnie w danym kontekście. W trakcie tego procesu odkryto jednak, że LLM mają wiele innych, nieoczywistych zalet – w tym, że modele są nie tylko zdolne do analizowania tekstu, ale także do tworzenia nowego tekstu. Na przykład model LLM może generować pełne, składne zdania, które są logiczne i zrozumiałe dla czytelnika. Może także dostosować swój styl pisania do kontekstu, naśladując ton i styl danego tekstu źródłowego. Jednak najbardziej imponującą cechą LLM jest ich zdolność do prowadzenia rozmów

z użytkownikami. Modele są zdolne do zrozumienia pytań postawionych przez użytkowników, do wykorzystania swojego obszernego zbioru wiedzy do formułowania odpowiedzi, a następnie do generowania odpowiedzi, które są zrozumiałe i sensowne. To oznacza, że LLM mogą być wykorzystywane do tworzenia zaawansowanych systemów konwersacyjnych. Te systemy mogą wchodzić w interakcję z użytkownikami, zrozumieć ich pytania, a następnie wykorzystać swoją obszerną wiedzę do generowania odpowiedzi.

Profesor Krzysztof Jajuga proponuje przyjęcie następującej **klasyfikacji sztucznej inteligencji**.

Wąska sztuczna inteligencja (Artificial Narrow Intelligence) obejmuje systemy, które są zaprojektowane tak, aby wykonać jedno (nawet bardzo skomplikowane) zadanie. Znanym przykładem jest system Deep Blue, który w 1997 pokonał mistrza świata w szachach, Garriego Kasparowa. Inny przykład to rekomendowanie konsumentowi produktu. W bankowości od co najmniej końca XX wieku stosowane są systemy oceny ryzyka kredytowego osoby ubiegającej się o kredyt. Bardziej rozwinięte systemy, które też są zaliczane do wąskiej sztucznej inteligencji, to te, w których jest wiele danych wejściowych i wiele możliwych wyników. Przykładem jest system planowania finansowego gospodarstwa domowego, w którym jest wiele danych wejściowych (charakterystyki gospodarstwa domowego i dane otoczenia społeczno-gospodarczego), a konstruowane są możliwe scenariusze realizacji celów finansowych gospodarstwa domowego.

Ogólna sztuczna inteligencja (Artificial General Intelligence) obejmuje systemy, które mają funkcje poznawcze na tym samym poziomie jak człowiek. Są one stosowane w wielu obszarach, takich jak np. rozpoznawanie mowy, rozpoznawanie obrazów, tworzenie tekstu itd. Oczywiście w tych obszarach zastosowań od dawna występują też systemy wąskiej sztucznej inteligencji i trudno stwierdzić, czy mamy do czynienia już z ogólną sztuczną inteligencją, a nie tylko z wąsko rozumianą sztuczną inteligencją. Taka możliwość istniałaby, jeśli byłoby całkowicie znane funkcjonowanie ludzkiego mózgu.

Supersztuczna inteligencja (Artificial Super Intelligence, dalej określana jako **ASI**) obejmuje hipotetycznie te systemy, które przewyższają wszystkie ludzkie możliwości.

Ten typ AI nie istnieje, oznaczałoby to przewagę AI nad ludzką inteligencją, co raczej byłoby niebezpieczne, na przykład w zakresie inteligencji emocjonalnej.

Wiele osób w branży AI uznaje Wielkie Modele Językowe (LLM) za przełomowy krok na drodze do stworzenia ASI. ASI nie tylko rozumie i przetwarza tekst, ale jest także zdolna do przyjmowania i interpretowania danych z różnych czujników, takich jak dźwięk, obraz i inne. Prawdziwa ASI byłaby w stanie uzyskać dostęp do dowolnych danych, przetworzyć je i na ich podstawie przeprowadzać skomplikowane analizy i podejmować decyzje. To oznacza, że mogłaby analizować sytuacje, rozumieć kontekst, a następnie podejmować decyzje na podstawie tej analizy. Innymi słowy, ASI byłaby w stanie symulować ludzki mózg i ludzką naturę w formie programu komputerowego. Istnieje jednak wiele różnych opinii na temat tego, kiedy dokładnie zostanie osiągnięty taki poziom zaawansowania w dziedzinie sztucznej inteligencji. Jedno jest jednak pewne – sztuczna inteligencja będzie jednym z najważniejszych obszarów technologii w najbliższych latach. Wielu uważa, że będzie to gałąź technologii, która zdefiniuje aktualną epokę, podobnie jak komputer osobisty czy Internet zdefiniowały epoki poprzednie. Bez względu na to, kiedy ASI zostanie osiągnięte, nie ma wątpliwości, że sztuczna inteligencja będzie miała ogromny wpływ na przyszłość ludzkości.

Jak wskazuje prof. Krzysztof Jajuga, rozwój systemów AI w XXI wieku wynika z dwóch procesów, są nimi:

- rozwój technologii informatycznych;
- wzrost dostępnych danych.

Rozwój technologii informatycznych to przede wszystkim gigantyczny wzrost prędkości przetwarzania danych, mierzony często liczbą operacji zmiennoprzecinkowych na sekundę (tzw. FLOPS). Najszybszy komputer świata w 1985 roku (superkomputer Cray-2) charakteryzował się prędkością 1,9 GIGAFLOPS (10^9). Obecny najszybszy komputer świata (El Capitan) charakteryzuje się prędkością 1,742 EXAFLOPS (10^{18}).

Wzrost dostępnych danych jest od lat rozważany w ramach Big Data. Są to duże, zmienne i różnorodne zbiory danych, których przetwarzanie i analiza są trudne, ale dają nową wartość, gdy mogą prowadzić do zdobycia nowej wiedzy. Warto zwrócić uwagę zwłaszcza na różnorodność danych, które oprócz klasycznych rodzajów danych, czyli numerycznych i tekstowych, rozważane są pliki audio, pliki video, sygnały generowane

przez sensory, zdjęcia satelitarne itd. Oszacowania wskazują, że w 2020 roku wielkość danych dostępnych w Internecie wynosiła 59 zettabajtów (10^{21} bajtów), a prognoza na 2025 wynosi 170 zettabajtów.

Drugim elementem systemów AI są metody. Te stosowane w AI, w ramach uczenia maszynowego, są co do ich koncepcji znane od wielu lat, większość z nich została stworzona w ramach szeroko rozumianej matematyki stosowanej, zwłaszcza statystyki wielowymiarowej. Połączenie wysiłków naukowców z matematyki stosowanej i statystyki z naukowcami zajmującymi się technologiami informacyjnymi zaowocowało rozwojem w zakresie implementacji tych metod (np. sztucznych sieci neuronowych) w praktyce.

Rodzaje AI

W najprostszym ujęciu metody AI można podzielić na dwie grupy:

1. Metody dyskryminacyjne (Discriminative Methods) – metody te na podstawie znanej klasyfikacji danych charakteryzujących obiekty generują reguły przydziału obiektów do tych klas, co następnie umożliwia przydział nowych obiektów do tych określonych wcześniej klas. Jedną z najprostszych metod dyskryminacyjnych jest analiza dyskryminacyjna, znana od prawie 90 lat, a w finansach stosowana od ponad 50 lat. Obecnie najczęściej stosowane metody dyskryminacyjne to sztuczne sieci neuronowe.
2. Metody generatywne (Generative Methods) – metody te na podstawie analizowanych danych identyfikują wzorce, które są następnie wykorzystywane do tworzenia informacji wyjściowych. Prosty przykładem metod generatywnych jest dowolna metoda prognozowania wielkości ekonomicznych (np. kursu walutowego czy ceny ropy naftowej), w której na podstawie danych historycznych modelowana jest zależność danej wielkości ekonomicznej od innych wielkości. Następnie zależność ta jest wykorzystana do konstrukcji prognozy.

Chociaż metody generatywne znane są od kilkudziesięciu lat, w ostatnich trzech latach zyskały olbrzymią medialną popularność. Stało się to za sprawą m.in. systemu Chat GPT (Generative Pretrained Transformer), który na podstawie dostępnych danych tekstowych przy zastosowaniu modeli przetwarzania języka naturalnego wykrywa wzorce

w tych danych, a następnie generuje nowe informacje tekstowe zgodne z tymi wzorcami. Te systemy są rozwijane w kierunku wykorzystania nie tylko danych tekstowych (i numerycznych), lecz również tzw. alternatywnych danych (plików audio, plików wideo, danych generowanych przez sensory itp.), a zatem również informacje wyjściowe mogą mieć postać danych alternatywnych.

W praktyce konsumenci mają do czynienia ze sztuczną inteligencją od wielu lat. Aplikacje czy systemy wykorzystujące AI są w stanie uczyć się, rozumieć i przetwarzać informacje, podejmować decyzje, rozpoznawać wzorce, a nawet przewidywać przyszłe zdarzenia na podstawie dostępnych danych.

Początki AI są datowane na lata 70. i 80. Pierwsze systemy AI były wykorzystywane głównie do obliczeń matematycznych i statystycznych. W 1987 roku Chase Lincoln First Bank (obecnie część JP Morgan Chase) wprowadził System Planowania Finansowego. Krótco potem, w 1989 roku, został wydany FICO Score, system oceny zdolności kredytowej oparty na podobnym algorytmie. Systemy AI mogły analizować dane klienta, takie jak historia kredytowa, dochód i inne czynniki, aby przewidzieć prawdopodobieństwo, że klient spłaci kredyt.

AI rozwinęła się w latach 90. i 2000., kiedy zaczęła odgrywać jeszcze większą rolę w dziedzinie bankowości i ubezpieczeń. Banki zaczęły wykorzystywać AI do automatycznego wykrywania podejrzanej aktywności, takiej jak oszustwa kredytowe czy pranie brudnych pieniędzy, a zakłady ubezpieczeń m.in. do automatycznego procesowania roszczeń i oceny ryzyka.

Ekspansja AI rozpoczęła się od 2010 roku. W 2014 roku brytyjski zarządca funduszy Man Group rozpoczął wykorzystywanie uczenia maszynowego do inwestowania pieniędzy swoich klientów. W 2016 roku Bank of America wprowadził swojego chatbota Erica, co uważa się za kamień milowy w interakcji z klientem. Obecnie AI jest nieodłącznym elementem bankowości i ubezpieczeń. Wykorzystuje się ją do prawie wszystkich czynności bankowych i ubezpieczeniowych, od personalizacji ofert dla klientów, przez detekcję oszustw, aż po rozwijanie nowych produktów i usług.

Sektor finansowy jest jednym z najbardziej dynamicznych i innowacyjnych obszarów biznesu, który od zawsze szukał sposobów na poprawę efektywności i skuteczności

swoich działań. W kolejnych rozdziałach wskażemy konkretne przykłady faktycznego wykorzystania AI.

AI Act – akt w sprawie sztucznej inteligencji

2 sierpnia 2024 roku weszło w życie unijne rozporządzenie ustanawiające przepisy dotyczące sztucznej inteligencji (Artificial Intelligence Act¹, dalej określane jako **AI Act**). Wejście w życie nie oznacza jednak pełnego obowiązywania – co do zasady przepisy będą obowiązywać od 2 sierpnia 2026 r., z wyjątkiem części, która obowiązuje już od 2 lutego 2025 r.

AI Act ma na celu zapewnienie bezpiecznego i etycznego wykorzystania technologii sztucznej inteligencji, z uwzględnieniem praw obywateli oraz wspieranie innowacji technologicznych. Nowe przepisy obejmują m.in. wymogi dotyczące transparentności algorytmów, uprawnień osób, wobec których jest wykorzystywane AI czy oznaczania treści generowanych przez sztuczną inteligencję.

Systemy AI na europejskim rynku będą musiały spełniać konkretne normy i standardy, co zwiększy poziom ochrony praw obywateli i konsumentów. Regulacje obejmują różne obszary zastosowań AI, takich jak sektor finansowy, opieka zdrowotna, transport, edukacja czy rynek pracy.

AI Act wprowadza podział systemów sztucznej inteligencji według związanego z nimi ryzyka dla zdrowia i bezpieczeństwa lub praw podstawowych człowieka.

- systemy objęte zakazem wykorzystania,
- systemy wysokiego ryzyka,
- systemy z podwyższonym poziomem przejrzystości, a więc tych, które „wchodzi” w bezpośrednią interakcję z użytkownikiem końcowym.

Systemy wysokiego ryzyka będą dopuszczalne pod warunkiem spełnienia określonych wymagań. Za istotne należy również uznać obowiązki w zakresie przejrzystości

¹ Rozporządzenie Parlamentu Europejskiego i Rady (Ue) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji)

w odniesieniu do określonych systemów sztucznej inteligencji. O powyższych kwestiach napiszemy więcej w dalszej części dokumentu.

Kim jest Rzecznik Finansowy i co ma wspólnego z AI

Rzecznik Finansowy (dalej określany jako **RF**) jest organem ochrony prawa powołanym do życia ustawą w 2015 r.² Jego zadaniem jest ochrona klientów podmiotów rynku finansowego, przez co rozumie się osobę fizyczną, w tym również osobę fizyczną prowadzącą jednoosobową działalność gospodarczą, która jest m.in.:

1. ubezpieczonym, ubezpieczającym, uposażonym,
2. klientem banku lub członkiem spółdzielczej kasy oszczędnościowo-kredytowej,
3. klientem instytucji płatniczej lub pożyczkowej,
4. klientem firmy inwestycyjnej.

Termin „podmiot rynku finansowego” także pochodzi z tej ustawy, ale w dalszej części tego dokumentu zastąpimy go bardziej przystępnym terminem „instytucji finansowej”.

Klient instytucji finansowej, zgłaszając się po pomoc do RF w sporze prowadzonym z tą instytucją, może skorzystać z kilku alternatywnych rozwiązań. Może wybrać wypracowanie ugody i polubowne zakończenie sporu lub złożyć wniosek do RF o interwencję w celu zmiany stanowiska instytucji na korzyść klienta.

Z obu wspomnianych trybów klient podmiotu rynku finansowego może skorzystać w sytuacji, zarówno gdy jego roszczenia nie zostały uwzględnione przez instytucję w trybie rozpatrywania reklamacji, jak i wtedy, gdy reklamacja co prawda została rozpatrzona pozytywnie przez instytucję, ale nie wykonała ona czynności wynikających z reklamacji rozpatrzonej zgodnie z wolą klienta w terminie określonym w odpowiedzi na reklamację.

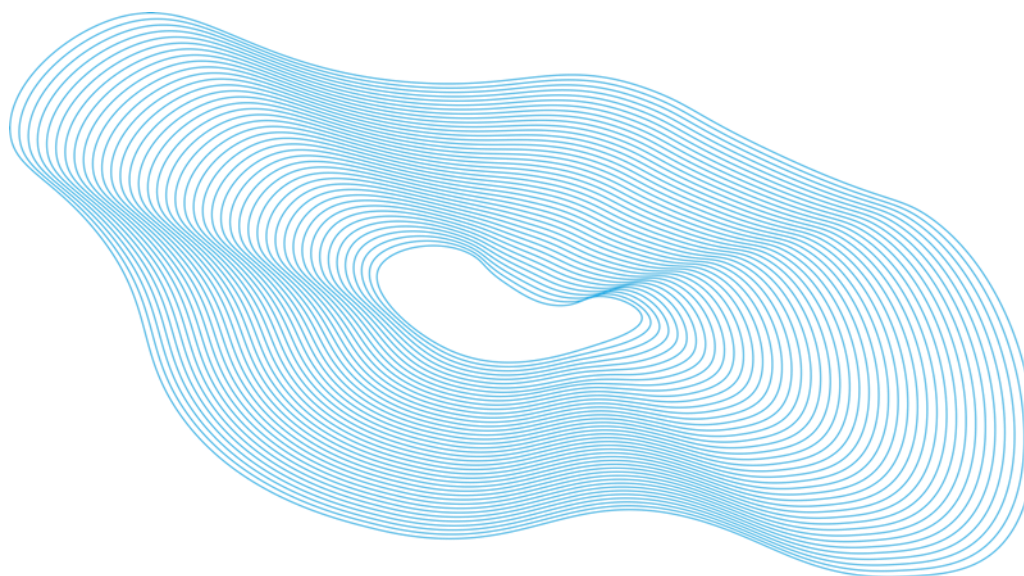
Rzecznik Finansowy może również pomagać klientom instytucji finansowych w trakcie toczących się już postępowań sądowych, zgłaszając istotny pogląd w sprawie. Istotny pogląd to materiał opiniodawczy o charakterze pomocniczym. Nie ma on wiążącej

² Ustawa z dnia 5 sierpnia 2015 r. o rozpatrywaniu reklamacji przez podmioty rynku finansowego, o Rzeczniku Finansowym i o Funduszu Edukacji Finansowej (t.j. Dz. U. z 2024 r. poz. 1109 z późn. zm.)

mocy prawnej dla sądu, jednak sąd uzasadniając orzeczenie, powinien ustosunkować się argumentów w nim przedstawionych (co często robi).

Od 1 stycznia 2023 roku RF może żądać wszczęcia postępowania w sprawach cywilnych na rzecz klientów instytucji finansowych, a także wziąć udział w toczącym się postępowaniu cywilnym, jeżeli według jego oceny wymaga tego ochrona praw.

Jak wskażemy w rozdziale II, instytucje finansowe (banki, zakłady ubezpieczeń itd.) bardzo często wykorzystują AI w swojej działalności, co może prowadzić do pozytywnych lub negatywnych konsekwencji – wskazanych w rozdziale III. Jednocześnie wykorzystanie AI podlega określonym regulacjom (rozdział IV), w niedalekiej przyszłości znacznie rozszerzonym (rozdział V) – zobowiązującym instytucje do określonych zachowań i przyznającym ich klientom odpowiednie uprawnienia. Brak ich odpowiedniej realizacji może stanowić podstawę do interwencji RF lub właściwych organów nadzoru rynku.



Zastosowanie AI w usługach i produktach finansowych

W poprzednim rozdziale wyjaśniliśmy pojęcie sztucznej inteligencji. W tym rozdziale przedstawimy najczęściej stosowane w instytucjach finansowych narzędzia AI.

Chatboty

Chatbot to program komputerowy zaprojektowany do symulowania rozmów tekstowych z ludźmi. Chatboty są najczęściej używane w obszarach obsługi klienta, marketingu i sprzedaży w celu usprawnienia obsługi klienta. Mogą one odpowiadać na proste pytania dotyczące godzin otwarcia sklepu, pomagać w składaniu zamówień online czy udzielać informacji na temat produktów lub usług. Niektóre bardziej zaawansowane chatboty są także w stanie uczyć się i dostosowywać swoje odpowiedzi na podstawie wcześniejszych interakcji.

Chatboty, ze względu na zaawansowanie technologiczne, można podzielić na trzy główne kategorie: klikalne, wykrywające intencje i generatywne.

Klikalne chatboty. To najprostsza forma chatbotów, która polega na interakcji z użytkownikiem za pomocą predefiniowanych opcji, które można kliknąć. Zamiast wpisywać pytania lub komendy, użytkownik wybiera odpowiedzi z określonej listy. Przykładem może być chatbot pomagający w przeprowadzeniu transakcji bankowej, gdzie użytkownik klika na opcje takie jak „przelew”, „sprawdzenie salda”.

Chatboty wykrywające intencje. Są bardziej zaawansowane technologicznie. Wykorzystują one sztuczną inteligencję i uczenie maszynowe do zrozumienia intencji użytkownika. Na przykład jeśli użytkownik napisze „Chcę zrobić przelew”, chatbot rozpoznaje intencję wykonania transakcji i odpowiada odpowiednio, prowadząc użytkownika przez proces.

Generatywne chatboty. To najbardziej zaawansowane chatboty, które nie tylko rozpoznają intencje użytkownika, ale również generują własne, unikalne odpowiedzi, zamiast polegać na predefiniowanych odpowiedziach. Wykorzystują technologie LLM do zrozumienia i odpowiedzi na zapytania w bardziej naturalny i ludzki sposób.

W przypadku bankowości taki chatbot jako źródło wiedzy może mieć tylko regulamin usług i sam na podstawie tego regulaminu generować odpowiedzi.

Chatboty, dzięki dostępności 24/7, są wygodnym narzędziem dla klientów, którzy oczekują natychmiastowej odpowiedzi na swoje pytania, niezależnie od pory dnia czy nocy. Co więcej, mogą one obsługiwać wielu użytkowników jednocześnie, co jest niemożliwe dla ludzkich konsultantów.

Systemy zautomatyzowanej oceny zdolności kredytowej

Bank (lub inna instytucja finansowa) uzależnia przyznanie kredytu (lub pożyczki) od zdolności kredytowej klienta. Rozumie się przez to możliwość spłaty zaciągniętego kredytu wraz z odsetkami w terminach określonych w umowie. Zdolność kredytowa zależna jest od szeregu czynników – w szczególności obecnych dochodów i kosztów życia, potencjalnych zmian tych danych, jak również oceny sumienności danego klienta w spłacaniu dotychczasowych zobowiązań (zarówno kredytów, pożyczek, jak i wykonywania umów).

Podstawą analizy zdolności kredytowej mogą być dane dotyczące potencjalnego kredytobiorcy (w tym m.in. jego wykształcenie, miejsce zamieszkania, zatrudnienie, dochody itd.), jak również historia dotychczas zaciągniętych kredytów i pożyczek, w tym terminowości ich spłaty. Dane te mogą być pozyskiwane bezpośrednio od wnioskodawcy, jak również pozyskiwane (i weryfikowane) z innych źródeł – m.in. publicznych lub prywatnych rejestrów.

Jednym ze sposobów na wykorzystanie systemu AI w bankowości jest poddanie procesów związanych z oceną zdolności kredytowej automatyzacji z użyciem uczenia maszynowego lub głębokiego. Z ich pomocą możliwa jest dokładniejsza ocena „parametrów”, w tym cech wnioskodawcy, a tym samym zmniejszenie ryzyka braku spłaty kredytu w terminie lub nawet całkowitej niewypłacalności.

Robodoradztwo inwestycyjne

Jak wskazuje prof. Krzysztof Waliszewski, robodoradztwo (robo-advice) jest elementem szerszego procesu automatyzacji w usługach finansowych, szczególnie tych kierowanych do konsumentów i nieprofesjonalnych inwestorów. Robodoradztwo jest szczególną formą wykonywania doradztwa inwestycyjnego, jest procesem, w ramach

którego udzielanie i przekazywanie rekomendacji odbywa się z wykorzystaniem algorytmów, systemów automatycznych i półautomatycznych. Istotą tej usługi jest wykorzystanie algorytmu do analizy i przyporządkowania instrumentów finansowych do profilu klienta. Algorytmy wbudowane w tę formę doradztwa mogą wykorzystywać uczenie maszynowe, uczenie głębokie, a także inne narzędzia statystyczne do analizy dużych zbiorów danych.

Klient wprowadza określone dane dotyczące jego sytuacji finansowej, wiedzy finansowej, doświadczeń związanych z inwestowaniem i przyszłych zamierzeń w tym zakresie, a także stopień akceptacji ryzyka i horyzont czasowy inwestycji. Dane te służą do oceny odpowiedniości usługi lub instrumentu finansowego w celu ustalenia profilu klienta poprzez zaprojektowanie odpowiedniego zestawu pytań oraz algorytmu łączącego warianty odpowiedzi z odpowiednią punktacją (tzw. profilowanie). Klient w zależności od wyniku przeprowadzonej ankiety ma do wyboru różne portfele inwestycyjne począwszy od portfeli bezpiecznych z wiodącym udziałem obligacji i instrumentów o stałej stopie dochodu, poprzez portfele zrównoważone z udziałem obligacji i akcji wynoszącym po 50%, aż do portfeli agresywnych, w których udział najbardziej ryzykownych akcji wynosi 100% lub jest jemu bliski (tzw. dopasowanie portfela modelowego).

Można wyróżnić 2 główne modele świadczenia usług robodoradztwa:

- *Model pełnego robodoradztwa (model czysty)*, w którym występuje pełne zautomatyzowanie usługi doradztwa inwestycyjnego poprzez jej całościowe świadczenie klientowi (począwszy od przeprowadzenia testu odpowiedniości, aż po udzielenie rekomendacji inwestycyjnej) za pośrednictwem Internetu. W modelu tym nie dochodzi do interakcji użytkownika i fizycznego doradcy finansowego (*human advisor*).
- *Model półautomatycznego robodoradztwa (model hybrydowy)*, w którym występuje automatyzacja procesu przygotowania rekomendacji inwestycyjnej, gdzie wynikiem zastosowania algorytmu do selekcji i przetworzenia zebranych danych są rekomendacje inwestycyjne, które następnie przekazywane są pracownikom firmy inwestycyjnej obsługującym klientów (bezpośrednio lub poprzez kanały komunikacji) w celu ich dalszego przekazania klientom. W modelu tym

automatyzacji podlegają czasochłonne czynności i samo sformułowanie rekomendacji, która następnie jest przekazywana przez tradycyjnego doradcę.

Robodoradztwo jest przykładem tzw. inwestowania pasywnego, ponieważ nie wymaga codziennego śledzenia trendów rynkowych i podejmowania decyzji, jak ma to miejsce w przypadku inwestowania aktywnego. W tym celu robodoradcy najczęściej wykorzystują tzw. fundusze ETF (*exchange-traded fund*), które są funduszami giełdowymi indeksowymi. Indeks to grupa papierów wartościowych będących przedmiotem obrotu na określonym rynku, która stanowi najbardziej reprezentacyjną próbę. Indeks można opisać jako wzorzec rynku, jego przeciętny skład. Fundusz indeksowy naśladuje indeksy, czyli pozwala inwestować w te same papiery wartościowe, z których złożony został dany indeks, w tych samych proporcjach. Dzięki pasywnemu inwestowaniu i niskim opłatom wyniki funduszy ETF są wyższe od wyników polskich funduszy inwestycyjnych (akcji).

W przypadku robodoradztwa należy pamiętać o okresowym równoważeniu portfela inwestycyjnego, co nazywa się rebalancing. Proporcje aktywów w portfelu pierwotnym z czasem się zmieniają na skutek zmian ich wycen rynkowych, co oznacza, że należy okresowo przywracać pierwotnie określone udziały w portfelu, aby utrzymać charakter ryzyka portfela z punktu widzenia udziału instrumentów bezpiecznych i ryzykownych. Możemy przy tym wyróżnić prosty *rebalancing okresowy* – raz na 3 miesiące bez względu na to, jak zmieniły się proporcje w portfelu, co 3 miesiące są one przywracane do pierwotnego stanu lub *rebalancing warunkowy*, co oznacza, że jest on wykonywany pod warunkiem, że proporcje odchyliły się o określoną wartość.

Rynek robodoradztwa jest najbardziej rozwinięty w jego ojczyźnie, czyli USA. Tam robodoradców jest najwięcej i są najtańsi. Roczne koszty zarządzania wynoszą ok. 0,25%. W Europie takich form jest mniej i są one droższe. Koszty zarządzania wynoszą przeważnie od 0,5 do 1%.

Personalizacja ofert i dopasowywanie produktów do klienta

Jednym z najważniejszych aspektów personalizacji ofert jest analiza danych. Banki gromadzą ogromne ilości danych o swoich klientach, począwszy od podstawowych informacji demograficznych, a kończąc na szczegółach dotyczących ich transakcji i zachowań zakupowych. Te dane są następnie analizowane za pomocą

zaawansowanych algorytmów AI, które pozwalają na identyfikację wzorców i trendów, które mogą być wykorzystane do stworzenia bardziej spersonalizowanych ofert.

Na przykład jeśli analiza danych pokaże, że dany klient regularnie robi zakupy w określonym sklepie, bank może zaoferować mu specjalną ofertę kredytu lub karty kredytowej związanej z tym sklepem. Podobnie jeśli klient często podróżuje za granicę, bank może zaproponować mu produkt ubezpieczeniowy skierowany do osób często podróżujących. Dodatkowo AI pozwala bankom na bieżąco monitorować zachowania swoich klientów i szybko reagować na zmiany. Jeśli na przykład algorytm zauważy, że klient nagle zaczął robić duże zakupy lub zaciągać duże pożyczki, bank może skontaktować się z klientem, aby upewnić się, że nie jest to wynik jakiegoś problemu finansowego.

Jednak personalizacja ofert to nie tylko kwestia dostarczania odpowiednich produktów. To także dostarczanie odpowiedniej komunikacji. AI może pomóc bankom w dostosowaniu tonu i stylu swojej komunikacji do preferencji poszczególnych klientów, co może znacznie zwiększyć skuteczność ich komunikacji marketingowej.

Hiperpersonalizacja to wykorzystanie zaawansowanych technologii, takich jak sztuczna inteligencja, uczenie maszynowe i analiza dużych zbiorów danych (big data), aby dostarczyć jak najbardziej spersonalizowane i dopasowane do indywidualnych potrzeb i preferencji użytkownika treści, oferty produktów czy usług.

Generatywna grafika otwiera zupełnie nowe możliwości w dziedzinie tworzenia treści reklamowych. Jej główna zaleta polega na możliwości generowania obrazów, animacji i innych elementów graficznych na podstawie danych wejściowych, co pozwala na tworzenie niezwykle spersonalizowanych treści. Podobnie jak w przypadku wykorzystania AI do personalizacji ofert bankowych, generatywna grafika może być wykorzystana do tworzenia treści reklamowych dostosowanych do indywidualnych preferencji klienta.

Dodatkowo generatywna grafika może pomóc w tworzeniu bardziej angażujących i interaktywnych reklam. Ostatnim, ale nie mniej ważnym aspektem jest możliwość tworzenia treści w czasie rzeczywistym. Algorytmy generatywnej grafiki mogą na bieżąco dostosowywać treści reklamowe na podstawie najnowszych danych o użytkowniku, takich jak jego lokalizacja, pogoda czy aktualne zainteresowania.

Zastosowanie AI w przeciwdziałaniu praniu pieniędzy i finansowaniu terroryzmu

Wiele instytucji finansowych (m.in. banki, instytucje płatnicze) jest zobowiązanych do przeciwdziałania praniu pieniędzy (dalej określanego jako AML) tj. wykorzystywanie systemu finansowego do ukrywania źródła pochodzenia środków uzyskanych w wyniku przestępstwa i ich legalizacji. Dzięki AI wykorzystanej szeroko w obszarze AML możliwe jest wykrywanie powiązań między różnymi przepływami finansowymi, co może pomóc w identyfikacji podejrzanych transakcji.

Wykrywanie wspólnych cech między transakcjami, które zostały oznaczone jako pranie pieniędzy, a tymi, które nie były jeszcze sprawdzone. To pozwala na skuteczniejsze identyfikowanie potencjalnie nielegalnych działań. Na przykład system AI może być nauczony, że transakcje o dużej wartości przeprowadzane późnym wieczorem są często związane z praniem pieniędzy. System może następnie monitorować transakcje w czasie rzeczywistym i flagować te, które pasują do tego wzorca, nawet jeśli nie zostały jeszcze sprawdzone. Innym przykładem może być wykorzystanie AI do identyfikacji powtarzających się transakcji o niskiej wartości między dwoma kontami. Choć pojedyncza transakcja może wyglądać niewinnie, powtarzalność może sugerować, że jest to tzw. smurfing – technika prania pieniędzy polegająca na podziale dużych kwot na wiele małych transakcji, aby uniknąć wykrycia.

System AI może być używany do skanowania paszportów lub dowodów osobistych. Rozpoznawanie obrazu pozwala na zidentyfikowanie i odczytanie informacji takich jak imię i nazwisko, data urodzenia czy numer dokumentu. Dzięki temu proces weryfikacji tożsamości klienta może być zautomatyzowany i przyspieszony. Innym przykładem jest wykorzystanie technologii OCR (Optical Character Recognition), która umożliwia konwersję różnych typów dokumentów, takich jak skany, PDF-y czy obrazy, na dane, które mogą być łatwo przetwarzane i analizowane. Dzięki OCR sztuczna inteligencja może „czytać” i analizować dane z dokumentów tożsamości, takie jak dowody rejestracyjne firm, umowy czy faktury. AI może również być używana do porównywania danych z dokumentów tożsamości z danymi wprowadzonymi przez klienta podczas

rejestracji, co pozwala na szybkie wykrycie niezgodności i potencjalnych prób oszustwa.

AI jest niezwykle skuteczna w identyfikowaniu nieprawidłowych wzorców i anomalii w danych transakcyjnych. Działa ona poprzez analizę dużej ilości danych historycznych, uczenie się na podstawie tych danych, a następnie zastosowanie nauczonego modelu do nowych danych w celu wykrycia nieprawidłowości. Anomalią może być np. transakcja o wyjątkowej wysokiej kwocie, gdy pozostałe transakcje historycznie były o rząd wielkości mniejsze albo sytuacja, gdy klient, który łączy się zawsze w godzinach porannych, nagle dokonuje transakcji w godzinach nocnych albo łączy się z nowego i nieznanego adresu IP.

Analiza mediów społecznościowych, aby zidentyfikować potencjalne powiązania między osobami przeprowadzającymi transakcje. Jeśli na przykład dwie osoby są powiązane w mediach społecznościowych i przeprowadzają regularnie transakcje o dużej wartości, system AI może to zidentyfikować jako potencjalne pranie pieniędzy.

Szacowanie ryzyka ubezpieczeniowego

Szacowanie ryzyka ubezpieczeniowego jest jedną z najważniejszych kwestii dla ubezpieczyciela, ponieważ w zależności od kalkulacji może on objąć dane ryzyko umową ubezpieczenia lub odmówić klientowi świadczenia ochrony ubezpieczeniowej. W tym celu zakład ubezpieczeń ma obowiązek gromadzenia odpowiednich danych statystycznych, podlegających licznym obowiązkom regulacyjnym. Ubezpieczyciele stoją przed koniecznością szczegółowej analizy ogromnej ilości danych. Aby przyspieszyć proces przetwarzania danych oraz zmniejszyć ilość błędów wynikających z dużej ilości zmiennych, bardzo przydatne staje się zastosowanie narzędzi AI.

Technologia AI umożliwia gromadzenie i przetwarzanie danych z różnych niepowiązanych ze sobą źródeł jak np. portale społecznościowe, historia wizyt na stronach internetowych, historia transakcji płatniczych czy dane medyczne. Dane te są często nieustrukturyzowane lub częściowo ustrukturyzowane i zawierają różne typy danych, np. wideo, dźwięk czy tekst, a zatem wymagają dodatkowego przetworzenia, aby wydobyć ich znaczenie. Z uwagi na ilość danych, przetworzenie ich bez AI byłoby praktycznie niemożliwe lub na tyle czasochłonne, że dane mogłyby utracić ich przydatność, np.

klient zdążyłby wrócić z wycieczki zagranicznej, której ubezpieczenie, chciałby zaofiarować ubezpieczyciel.

Algorytmy na podstawie zebranych danych historycznych, mogą oszacować prawdopodobieństwo wystąpienia danego zdarzenia mającego zostać objętym ochroną ubezpieczeniową, tzw. wypadku ubezpieczeniowego oraz określić jak wysoka powinna być składka ubezpieczeniowa, aby ubezpieczyciel mógł zlikwidować w pełni powstałą szkodę.

Spersonalizowana ocena ryzyka ubezpieczeniowego może wpłynąć również na obniżenie ceny produktu ubezpieczeniowego np. w przypadku osoby nieobciążonej poważnymi chorobami, pomimo kryterium osiągnięcia określonego wieku, zaproponowanie niższej składki ubezpieczeniowej adekwatnej do ryzyka. Jednym z narzędzi personalizacji ryzyka ubezpieczeniowego jest ocena stylu jazdy kierowcy poprzez pomiar niektórych parametrów jazdy jak np. przekraczanie dozwolonej prędkości, otrzymywanie mandatów itp.

Zautomatyzowane badanie potrzeb klienta

Badanie potrzeb klienta przeprowadza, przed zawarciem umowy ubezpieczenia, dystrybutor ubezpieczeń (zakład ubezpieczeń, agent ubezpieczeniowy lub broker ubezpieczeniowy). W ramach badania pozyskuje się od klienta informacje, jego wymagania i potrzeby, a następnie na tej podstawie proponuje adekwatny do potrzeb i wymagań produkt ubezpieczeniowy. Dystrybutor ubezpieczeń ma obowiązek podać informacje o produkcie ubezpieczeniowym w zrozumiałej formie, z uwzględnieniem złożoności produktu oraz rodzaju klienta, aby umożliwić klientowi podjęcie świadomej decyzji.

Aby dystrybutor ubezpieczeń mógł się prawidłowo wywiązać ze wspomnianego obowiązku, musi on pozyskać od klienta adekwatne dane. Można byłoby wyobrazić sobie sytuację, w której dystrybutor ubezpieczeń profilaktycznie pyta klienta poszukującego ochrony ubezpieczeniowej o wszystkie możliwe dane adekwatne dla każdego oferowanego przez dystrybutora produktu ubezpieczeniowego. Jednak takie rozwiązanie byłoby niezwykle czasochłonne, a klient pytany o tak dużą ilość danych w pewnym momencie mógłby zniechęcić się i zrezygnować z poszukiwania umowy ubezpieczenia. Jednocześnie zaniechanie przez dystrybutora ubezpieczeń badania potrzeb klienta poszukującego ochrony ubezpieczeniowej, grozi dystrybutorowi ubezpieczeń

poważnymi sankcjami finansowym oraz administracyjnymi, łącznie z cofnięciem zezwolenia na wykonywanie działalności dystrybucyjnej.

W związku z powyższym gromadzenie danych, następnie ich selekcja z uwzględnieniem preferencji danego klienta, a jednocześnie ograniczeń dla konkretnego produktu ubezpieczeniowego jest niezwykle istotna. Uwzględniając bardzo dużą ilość danych do przetworzenia oraz pożądaną szybkość i dokładność ich przetworzenia narzędzia AI stają się szczególnie przydatne.

Wspomniane narzędzia wykorzystujące biometrikę behawioralną mogą również posłużyć do stworzenia tzw. silników rekomendacyjnych, personalizujących ofertę dla danego klienta, na podstawie jego aktywności w Internecie oraz dokonywanych w czasie rzeczywistym transakcji. Dzięki takim danym bank mający w swojej ofercie ubezpieczenia, korzystając z tzw. real time, może zaproponować ofertę ubezpieczenia samochodu, którego zakup został przed chwilą dokonany i zostało to odnotowane na rachunku bankowym. Pozwala to na kierowanie oferty do klienta, który potencjalnie byłby zainteresowany danym rodzajem umowy ubezpieczenia, eliminując sytuacje, w których klienci czują się nagabywani przez ubezpieczyciela, ciągłymi ofertami produktów, których nie potrzebują. Co więcej, zebrane dane mogą również posłużyć do wygenerowania spersonalizowanej zawartości strony ubezpieczeniowej ubezpieczyciela, oferując każdemu klientowi inne produkty ubezpieczeniowe, które są zgodne z jego wcześniej ustalonymi na podstawie biometriki behawioralnej preferencjami.

W przypadku silników rekomendacyjnych można również wprowadzić zmienne klasyfikujące klientów do określonych grup albo anty grup, w zależności od tego, czy mogą potrzebować danego produktu, czy będzie on dla nich nieodpowiedni. Dzięki temu można przyspieszyć proces przygotowania propozycji ofert produktów, od razu eliminując te, w których dany klient należy do antygrupy.

Automatyzacja procesu badania potrzeb może odbywać się również poprzez wykluczenie decyzji ludzkich i zastąpieniu ich automatycznymi decyzjami, które w odróżnieniu od decyzji człowieka potrzebującego czasu na analizę danej umowy ubezpieczenia, mogą być wydawane w niemal nieograniczonej liczbie w czasie rzeczywistym. Jako przykład można wskazać ustawienie progu wartości transakcji poniżej, którego umowy ubezpieczenia będą zawierane automatycznie po kliknięciu przez klienta

ostatecznej akceptacji umowy, bez oczekiwania na decyzję człowieka. Podobnie można zautomatyzować proces wystawiania dokumentu polisy ubezpieczeniowej, która może być wysłana automatycznie po zaksięgowaniu środków na koncie bankowym ubezpieczyciela.

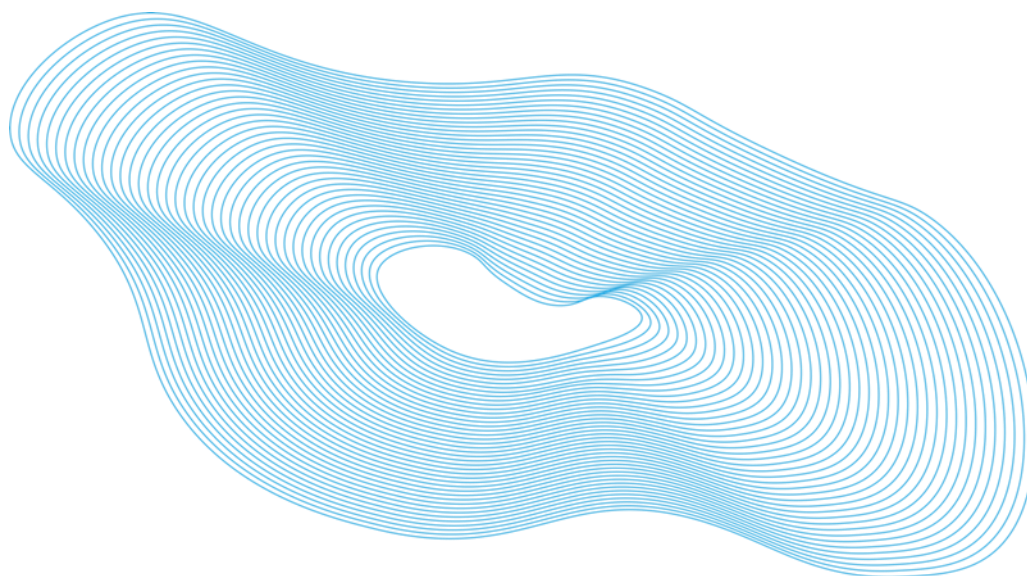
Likwidacja i wycena szkód ubezpieczeniowych

Ocena wysokości szkód jest kluczową częścią procesu roszczeń ubezpieczeniowych. Technologia AI, korzystając z dużych zbiorów danych dotyczących wcześniejszych szkód i decyzji, jest w stanie wykrywać wzory i odchylenia, które mogą sugerować próby nadużyć lub błędy w roszczeniach. Daje to ubezpieczycielom większą pewność co do prawidłowości roszczeń, a jednocześnie pozwala im szybciej identyfikować potencjalne nadużycia. Co więcej, AI może również automatycznie oszacować koszt napraw, co znacznie przyspiesza proces rozpatrywania roszczeń. Dzięki temu ubezpieczyciele mogą szybciej reagować na roszczenia klientów, co z kolei poprawia doświadczenia klientów i zwiększa ich zadowolenie.

Sztuczna inteligencja umożliwia zautomatyzowanie wielu etapów procesu likwidacji szkód, w tym ocenę wniosków. AI analizuje dane klientów, aby szybko i trafnie określić poziom ryzyka i podjąć decyzję o zaakceptowaniu bądź odrzuceniu wniosku.

W kontekście likwidacji szkód chatboty mogą pełnić wiele funkcji. Przede wszystkim, mogą prowadzić rozmowę z klientem w celu pozyskania danych związanych z obsługą roszczenia. Mogą zadawać pytania dotyczące okoliczności zdarzenia, rodzaju i zakresu szkód, a także żądanej formy odszkodowania. Dzięki temu proces zgłaszania roszczeń staje się bardziej wydajny i mniej stresujący dla klienta. Ponadto chatboty mogą odpowiadać na zapytania klientów dotyczące postępu w obsłudze ich roszczenia. Dzięki temu klienci mogą uzyskać szybką i precyzyjną informację bez konieczności oczekiwania na połączenie z konsultantem. Chatboty mogą również samoczynnie aktualizować statusy roszczeń, informując klientów o każdym nowym etapie procesu. Dzięki wykorzystaniu chatbotów proces likwidacji szkód staje się bardziej przejrzysty, efektywny i przyjazny dla klienta. Klienci czują, że mają większą kontrolę nad procesem i otrzymują informacje, które potrzebują, kiedy tylko ich potrzebują. To z kolei przekłada się na większą lojalność wobec zakładu ubezpieczeń.

AI przegląda dane dotyczące roszczeń, aby zidentyfikować podejrzone schematy i nieprawidłowości, które mogą sugerować próby nadużyć. Technologie takie jak profilowanie behawioralne mogą pomóc w identyfikacji klientów o zwiększonym ryzyku oszustwa. AI wykorzystuje zdjęcia uszkodzonych pojazdów lub nieruchomości do automatycznego wyliczenia kosztów napraw. Jednak wykorzystanie AI do wykrywania nadużyć wiąże się z pewnymi zagrożeniami. Przede wszystkim, istnieje ryzyko łączenia danych, które nie powinny być ze sobą powiązane. AI opiera się na analizie dużych ilości danych, a czasem te dane mogą pochodzić z różnych, niepowiązanych ze sobą źródeł. Jeśli dane te są łączone bez odpowiedniego kontekstu lub zrozumienia, wyniki mogą być mylące. Na przykład AI może połączyć dane dotyczące częstotliwości wypadków samochodowych z danymi na temat koloru pojazdu, prowadząc do niewłaściwych wniosków, takich jak stwierdzenie, że czerwone samochody są częściej zaangażowane w wypadki. Choć takie powiązanie może wydawać się logiczne dla algorytmu, nie ma rzeczywistego powiązania przyczynowego między tymi dwoma zestawami danych.



Korzyści i zagrożenia wynikające ze stosowania AI

Korzyści

Wykorzystanie AI pozwala na automatyzację i optymalizację wielu procesów biznesowych w instytucjach finansowych, pozwalając na szybsze i zazwyczaj dokładniejsze analizowanie określonych zbiorów danych. W wielu przypadkach, zwłaszcza w bardziej dojrzałym etapie wykorzystania AI, wiąże się to również ze spadkiem kosztów obsługi danego procesu (system AI jest co do zasady tańszy niż odpowiednia grupa pracowników wykonujących jego zadanie). Przykładem może być tu np. element systemu przeciwdziałania praniu pieniędzy i finansowaniu terroryzmu odpowiedzialny za skanowanie i weryfikację dokumentów klientów.

Bardzo często wykorzystanie AI przekłada się na dodatkową wartość dla klientów instytucji finansowych. Optymalizacja procesu może z jednej strony wiązać się z jego większą przyjaznością dla klienta (np. skróceniem czasu oczekiwania na decyzję banku w sprawie kredytu), lepszym dopasowaniem usługi do jego potrzeb (system AI może przeanalizować więcej danych na jego temat), ale również powstaniem nowych kanałów obsługi klienta (założenie konta w banku przez Internet w kilka minut nie byłoby możliwe bez zastosowania technologii AI).

Stosowanie systemu AI, zwłaszcza bez zachowania odpowiednich standardów może się jednak wiązać z istotnymi zagrożeniami dla klienta – zarówno jego prywatności, jak i interesów majątkowych.

Zagrożenia – rozważania ogólne

Jak zauważa prof. Krzysztof Jajuga, system AI powinien służyć rozwiązaniu konkretnego problemu, zadanemu przez użytkownika. Jest to ważne, gdyż końcowym odbiorcą systemu AI w naszych rozważaniach jest konsument. Nie ma systemu AI, który byłby uniwersalny i mógłby być zastosowany dla dowolnego problemu.

Narzędzie AI jest to z reguły tzw. czarna skrzynka, czyli narzędzie, którego sposób działania nie jest zrozumiały dla końcowego użytkownika (w szczególności konsumenta). Co więcej, może się zdarzyć, że twórca narzędzia nie rozumie problemu, któremu ma

służyć dany system AI, jak również nie jest ekspertem w zakresie metod, które są wykorzystywane w systemie AI, a jedynie zna się na stronie czysto technicznej (informatycznej).

Główne zagrożenia związane ze stosowaniem narzędzi AI można podzielić na te związane z:

zakresem stosowania narzędzia:

- niedostosowanie narzędzia AI do celu konsumenta (końcowego użytkownika);
- brak transparentności narzędzia AI („czarna skrzynka”);
- brak możliwości wyjaśnienia i interpretacji stosowanego rozwiązania;

oraz zakresem jakości narzędzia:

- wykorzystanie niewiarygodnych danych w tworzeniu systemu;
- wykorzystanie niewłaściwych metod w tworzeniu systemu;
- brak walidacji systemu.

W ekstremalnym przypadku system nieodpowiedni do potrzeb konsumenta jest dodatkowo systemem niskiej jakości. Największe wyzwanie dla systemów AI wiąże się z tymi sytuacjami, w których system powinien być:

- kreatywny;
- w maksymalnym stopniu dostosowany do potrzeb użytkownika (spersonalizowany).

Te dwa warunki oznaczają brak skalowalności tworzonych systemów, czyli konieczność poniesienia bardzo wysokich jednostkowych nakładów. Spośród wszystkich rozwiązań realizowanych przez AI, z których potencjalnie może korzystać konsument usług finansowych, najbardziej zaawansowane powinny być narzędzia robodoradztwa, wykorzystywane w planowaniu finansowym gospodarstw domowych. Jednak konieczność spełnienia obu powyższych warunków oznacza, że bardzo trudne (jeśli w ogóle możliwe na obecnym poziomie rozwoju) jest stworzenie profesjonalnego systemu „pełnego” robodoradztwa (bez udziału człowieka pełniącego funkcję doradcy finansowego).

Jak się wydaje, zdecydowanie większą użyteczność mogą mieć systemy „hybrydowego” robo doradztwa, w których:

- analiza dużej ilości danych z zastosowaniem zaawansowanych metod ilościowych (pod kątem profilu gospodarstwa domowego) jest przeprowadzana przez AI (w tym generowane są możliwe rozwiązania);
- wybór najlepszego rozwiązania (lub zbioru rozwiązań) jest w gestii doradcy finansowego (z udziałem gospodarstwa domowego).

Oczywiście jest konieczność weryfikacji profesjonalności narzędzia AI oraz doradcy finansowego.

Można założyć, że są trzy niezbędne warunki, których spełnienie może ograniczyć zagrożenia dla konsumenta związane z AI:

- weryfikacja narzędzi AI pod kątem wiarygodności danych, odpowiedniości stosowanych metod oraz transparentności narzędzia dla końcowego użytkownika;
- wprowadzenie standardów etycznych w zakresie tworzenia narzędzi AI;
- wprowadzenie regulacji, które jednak nie mogą blokować innowacyjności w zakresie AI.

Bias i halucynacje – metody przeciwdziałania

Szczególnie istotnymi zagrożeniami przy obecnym stosowaniu systemów AI są bias i halucynacje.

Bias to uprzedzenie lub skłonność, które mają wpływ na proces uczenia maszynowego. Często wynika ono z niedoskonałości w danych wejściowych, które są używane do trenowania modeli sztucznej inteligencji. Zdarza się np., gdy dane treningowe są nierównomiernie rozłożone między różnymi klasami w uczeniu nienadzorowanym – model może nauczyć się faworyzować jedną klasę kosztem innych. Przykładem uprzedzenia w bankach może być sytuacja, gdy algorytmy stosowane do oceny zdolności kredytowej faworyzują osoby o określonej płci, rasie, wieku czy pochodzeniu społecznym. Trzeba pamiętać, że często wykorzystuje się historyczne dane treningowe, które zostały stworzone przez człowieka i jeśli zawierają treści dyskryminujące, model może nauczyć się i naśladować te uprzedzenia, co prowadzi do niesprawiedliwych decyzji

kredytowych. Innym przykładem może być sytuacja, gdy algorytm sztucznej inteligencji faworyzuje osoby o wyższych dochodach, ignorując inne czynniki, które mogą wskazywać na zdolność do spłaty kredytu, takie jak stabilność zatrudnienia czy historia kredytowa.

Halucynacje to cecha modeli generujących tekst. Odwołuje się do sytuacji, gdy model generuje informacje, które nie były zawarte w danych wejściowych, nie mają podstaw w rzeczywistości lub są nieadekwatne do danego kontekstu. Na przykład jeśli poprosimy model o napisanie opisu osoby, która nigdy nie istniała, lub o przewidzenie wydarzeń, które jeszcze nie miały miejsca, model może "halucynować" szczegóły, które nie są oparte na faktach. W niektórych sytuacjach halucynacje mogą prowadzić do generowania treści niezwiązanych, niekonsekwentnych, nieodpowiednich czy nawet fałszywych.

Aby przeciwdziałać takim sytuacjom, stosuje się:

- interpretowalność modelu (model explainability), czyli zdolność modelu do zrozumienia i wyjaśnienia, dlaczego i jak podejmuje określone decyzje. Jest to miara, jak łatwo ludzie mogą zrozumieć proces podejmowania decyzji przez system AI.
- strategie i praktyki zapewniające, że dane używane przez organizację są dokładne, niezawodne, spójne i bezpieczne. W kontekście modeli AI i uczenia maszynowego jakość danych i ich zarządzanie są kluczowe dla skuteczności modeli, ponieważ modele te są "nauczane" na podstawie dostarczonych im danych.
- bariery ochronne (guiderails), pomagają zapobiegać potencjalnym problemom, takim jak niewłaściwe użycie modeli, naruszenie prywatności lub niesprawiedliwe i uprzedzone wyniki. Jednym z ich przykładów może być system, który monitoruje i kontroluje generowanie treści przez model, zapobiegając generowaniu takich, które są obraźliwe, uprzedzone, nieodpowiednie lub które naruszają zasady ochrony prywatności.

W celu lepszego zrozumienia korzyści i zagrożeń płynących dla klienta ze stosowania przez instytucje finansowe technologii AI przywołaliśmy poniżej kilka przykładów powszechnie wykorzystywanych technologii.

Korzyści i zagrożenia na przykładzie robodoradztwa

Dobrym przykładem ułatwiającym zrozumienie z korzyści oraz zagrożeń płynących ze stosowania technologii AI jest robodoradztwo. [Prof. Krzysztof Waliszewski wskazuje na następujące zalety tej usługi:](#)

- ułatwienie dostępu do usługi doradztwa inwestycyjnego i zapewnienie możliwości jej świadczenia dla szerokiego kręgu klientów, przez co dokonuje się upowszechnienie tej usługi wśród osób, które nie skorzystałyby nigdy z porady finansowej,
- obniżenie minimalnej kwoty będącej przedmiotem tej usługi np.: do 100 zł, czyli zniesienie bariery wejścia, która występuje przy klasycznym doradztwie inwestycyjnym,
- obniżenie kosztów świadczenia usługi robodoradztwa w porównaniu z tradycyjnym doradztwem inwestycyjnym
- dobór takich instrumentów finansowych, które będą uznane za odpowiednie dla szerokiej grupy klientów, głównie ETF,
- możliwość inwestowania w instrumenty o charakterze globalnym np.: pośrednio ETF, akcje i obligacje globalne, a przez to uniezależnienie wyników od bieżącej koniunktury rynkowej w kraju,
- przewalutowanie po korzystniejszych kursach,
- możliwość zakupu ułamka ETF, co nie jest możliwe przy inwestowaniu aktywnym przez rachunek maklerski,
- brak kosztów prowizji giełdowych za zakup i sprzedaż ETF na giełdach,
- uniezależnienie decyzji podejmowanej dla klienta od czynników psychologicznych, którymi mógłby kierować się sam klient lub tradycyjny doradca finansowy.

Wady i zagrożenia robodoradztwa dla użytkowników są następujące:

- dostęp do ograniczonej ilości portfeli inwestycyjnych, które nie do końca mogą pasować do indywidualnego profilu inwestora,

- ograniczenie bezpośredniego kontaktu klienta z pracownikami firmy inwestycyjnej na etapie badania odpowiedniości (model hybrydowy) i w całym procesie (model czysty), co może utrudniać lub uniemożliwiać wychwycenie indywidualnej specyfiki konkretnego klienta,
- ryzyko modelu, czyli błędnego zaprogramowania algorytmu na wstępie,
- niska elastyczność, ponieważ nie można wypłacić środków z części obligacyjnej lub akcyjnej.

Zagrożenia na przykładzie chatbotów

Podstawową wadą chatbotów jest ryzyko wykorzystywania przez nich nieaktualnych danych. Najczęściej wiedza chatbota, jest uzupełniana nie na bieżąco – jak strony internetowe czy regulaminy, ale etapami. Tym samym podstawa generowanych treści może być nieaktualna. Aby się upewnić, że chatbot ma aktualną wiedzę, zawsze warto zadać mu pytanie: „z jakiego okresu pochodzi Twoja wiedza?” Ponadto chatbot, który posiada niewystarczającą wiedzę na dany temat może zacząć halucynować (zjawisko to opisaliśmy powyżej).

Odrębną, ale nie mniej istotną kwestią jest bezpieczeństwo danych przekazywanych przez klientów chatbotom w trakcie „rozmowy”. Wdrożenie modeli języka maszynowego (LLM), które szybko i skutecznie dostarczą oczekiwane odpowiedzi, będzie wymagało korzystania z przeznaczonych do tego serwerów. Zazwyczaj banki nie są właścicielami tych serwerów, ale korzystają z nich na zasadzie chmury obliczeniowej lub jako kompleksowe usługi zewnętrzne. W rezultacie dane, które przesyłamy do banku, nawet w zaszyfrowanej formie, są uzupełniane, filtrowane i przekazywane do dostawców zewnętrznych, aby umożliwić zrozumienie i generowanie odpowiedzi za pomocą modeli LLM.

Nielogiczne lub niespójne odpowiedzi to problem, który może wystąpić podczas korzystania z chatbotów opartych na modelach LLM do generowania odpowiedzi. „Halucynacje” chatbotów polegają na tym, że generują one odpowiedzi na podstawie niepowiązanych ze sobą informacji. To oznacza, że zamiast tworzyć logiczne i spójne odpowiedzi na podstawie wprowadzonych danych lub pytań, chatboty te mogą tworzyć odpowiedzi, łącząc przypadkowo wybrane fragmenty informacji. Odpowiedź będzie

poprawna językowo, ale nie bazująca na faktach. Przykładowo, dwa regulaminy podobnie brzmiącej promocji „Otwórz konto Junior” i „Konto dla Seniora”, mogą być użyte do wygenerowania odpowiedzi, gdy zapytamy: „jakie są warunki wiekowe do otwarcia konta?”

Zagrożenia związane z hiperpersonalizacją

Jak wskazano powyżej, technologia AI pozwala na stworzenie oferty danej usługi (np. ubezpieczenia) niemal „idealnie skrojonej pod klienta”. Jednak i to zjawisko nie jest pozbawione zagrożeń.

Źle zastosowane algorytmy sztucznej inteligencji mogą zaklasyfikować klienta do niewłaściwej grupy konsumentów, przez co klienci są ekspozowani tylko na produkty, usługi i informacje, które są dostosowane do ich preferencji, co ogranicza ich eksplorację i odkrywanie nowych rzeczy. Przykład: klient zaklasyfikowany do grupy posiadających mieszkanie może nie otrzymać nigdy reklamy z korzystną ofertą hipoteczną, ponieważ statystycznie ci klienci nie kupują więcej niż jedną nieruchomość.

Hiperpersonalizacja może być również wykorzystywana do manipulowania użytkownikami, np. poprzez wyświetlanie im reklam, które wykorzystują ich ulubiony kolor, markę samochodu, ulubionego influencera i ulubiony cytat. Może to prowadzić do podejmowania przez klientów nieracjonalnych decyzji.

Wreszcie stosowanie hiperpersonalizacji może prowadzić do dalszego ograniczenia prywatności klienta, i tak już znacznie okrojonej w dobie cyfrowej gospodarki.

AI w finansach a regulacje prawne – co musi przedsiębiorca, a co może konsument

Jak wskazaliśmy w rozdziale I, większość przepisów AI Act będzie obowiązywać dopiero od 2 sierpnia 2026 r. Nie oznacza to jednak, że wykorzystywanie systemów AI nie podlega już teraz regulacjom.

Zakazane systemy AI w ramach AI Act

Od 2 lutego 2025 r. obowiązują rozdziały I i II AI Act. Zgodnie z rozdziałem II na terenie UE zakazane jest wprowadzanie do obrotu, oddawanie do użytku lub wykorzystywanie systemów AI, które mogą stanowić istotne naruszenie praw i wolności obywateli oraz interesów konsumentów poprzez m.in.³:

- stosowanie technik podprogowych lub celowych technik manipulacyjnych prowadzących do podejmowania decyzji wyrządzających poważną szkodę,
- wykorzystywanie słabości fizycznej ze względu na wiek, niepełnosprawność lub szczególną sytuację społeczną, lub ekonomiczną,
- stosowanie tzw. scoringu społecznego,
- ocenę osób fizycznych pod kątem ryzyka popełnienia przez nie przestępstwa wyłącznie na podstawie profilowania tej osoby, jej cech osobowości i cech charakterystycznych,
- wykorzystanie systemów zdalnej identyfikacji biometrycznej w czasie rzeczywistym w przestrzeni publicznej.

Każdy, kto nie przestrzega powyższych przepisów, podlegać będzie administracyjnej karze pieniężnej w wysokości do 35 mln euro lub – jeżeli sprawcą naruszenia jest przedsiębiorstwo – w wysokości do 7% jego całkowitego rocznego światowego obrotu z poprzedniego roku obrotowego (w zależności od tego, która z tych kwot jest wyższa)⁴.

³ Art. 5 AI Act

⁴ Art. 99 ust. 3 AI Act

Informacje dotyczące egzekwowania powyższych zakazów zostaną uzupełnione po przyjęciu ustawy stosującej AI Act.

Ochrona danych osobowych (RODO) a wykorzystanie AI

RODO⁵, obowiązujące od 25 maja 2018 r. unijne przepisy regulujące bezpośrednio zasady przetwarzania danych osobowych, nie zawierają pojęcia „sztucznej inteligencji”. Art. 22 RODO określa natomiast zasady zautomatyzowanego podejmowania decyzji w indywidualnych przypadkach, w tym profilowania. Wykorzystywanie systemu AI do podjęcia danej decyzji przez przedsiębiorcę (np. ocena zdolności kredytowej) będzie w większości przypadków tożsame ze zautomatyzowanym podejmowaniem decyzji, tzn. stosowanie AI będzie podlegało pod art. 22 RODO. Przepis ten stanowi, że podmiot podejmujący zautomatyzowaną decyzję musi wdrożyć *właściwe środki ochrony praw, wolności i prawnie uzasadnionych interesów osoby, której dane dotyczą, a co najmniej prawa do uzyskania interwencji ludzkiej ze strony administratora, do wyrażenia własnego stanowiska i do zakwestionowania tej decyzji.*

Oznacza to, że osoba, której dane są przetwarzane przy wykorzystaniu AI do podjęcia decyzji, ma prawo zwrócić się do podmiotu podejmującego taką decyzję i ją zakwestionować, żądając interwencji ludzkiej w procesie decyzyjnym czy wyjaśnień dotyczących przesłanek podjęcia takiej decyzji.

Ponadto osoba, której dane są przetwarzane w sposób zautomatyzowany (a więc która np. przechodzi proces oceny kredytowej), powinna otrzymać następujące informacje⁶:

- że dana decyzja podejmowana jest w sposób zautomatyzowany,
- informacje o zasadach podejmowania tej decyzji,
- informacje o znaczeniu i konsekwencjach takiego przetwarzania dla tej osoby.

Co więcej, jeżeli przetwarzanie odbywa się w sposób zautomatyzowany, osoba, której dane dotyczą, *ma prawo otrzymać w ustrukturyzowanym, powszechnie używanym*

⁵ Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Dz. U. UE. L. z 2016 r. Nr 119, str. 1 z późn. zm.).

⁶ Art. 13 ust. 2 lit. f RODO, również 14 ust. 2 lit. g i 15 ust. 1 lit. h

formacie nadającym się do odczytu maszynowego dane osobowe jej dotyczące, które dostarczyła administratorowi, oraz ma prawo przesłać te dane osobowe innemu administratorowi bez przeszkód ze strony administratora⁷.

Prawo bankowe a wykorzystanie AI

Prawo bankowe⁸ wprowadza po stronie banków dodatkowe obowiązki w przypadku podejmowania zautomatyzowanego podejmowania decyzji w przedmiocie oceny zdolności kredytowej. Po pierwsze, dokonanie oceny w sposób zautomatyzowany wymaga *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*. Co więcej, określa zamkniętą listę danych, które mogą być wykorzystane do zautomatyzowanej (a więc m.in. wykorzystującej AI) oceny kredytowej, wykluczając m.in. tzw. szczególne kategorie danych osobowych (dotyczące pochodzenia rasowego lub etnicznego, poglądów politycznych, zdrowia, seksualności itd.)⁹.

W przypadku żądania podjęcia interwencji przez człowieka wniosek kredytowy podlega analizie dokonanej przez analityka, który może także korzystać z narzędzi wspierających analizę, ale to on ostatecznie podejmuje decyzję. Decyzja ta może, ale nie musi być odmienna od tej wskazanej przez system zautomatyzowanego przetwarzania. Uprawnienie do zakwestionowania decyzji przez wnioskodawcę i wskazania uzasadnienia może zostać wykorzystane do dalszej analizy dokonywanej przez człowieka. Ponadto każda osoba, wobec której dokonano oceny zdolności kredytowej, w tym w sposób zautomatyzowany, ma prawo uzyskać bezpłatnie wyjaśnienie dotyczące tej oceny¹⁰. Powinna ona zawierać informacje na temat czynników, które miały wpływ na dokonaną ocenę zdolności kredytowej. Informację o tym uprawnieniu instytucja finansowa powinna przekazać razem z informacją o dokonanej ocenie kredytowej.

⁷ Art. 20 RODO

⁸ Ustawa z dnia 29 sierpnia 1997 r. Prawo bankowe (t.j. Dz. U. z 2024 r. poz. 1646 z późn. zm.).

^{9 9} Art. 105a ust. 1b i 1c Prawa bankowego

¹⁰ Art. 70a Prawa Bankowego

Przepis będący podstawą do realizacji tego uprawnienia milczy w zakresie tego, jak powinno być ono realizowane przez obowiązującą do tego instytucję, pozostawiając duże pole do interpretacji, co ma szczególne znaczenie w przypadku stosowania bardziej zaawansowanych rozwiązań z zakresu sztucznej inteligencji.

Z pomocą może przyjść tutaj jednak komunikat Urzędu Komisji Nadzoru Finansowego z 2020 r.¹¹ dotyczący realizacji uprawnienia do uzyskania wyjaśnień na temat dokonanej oceny zdolności kredytowej. Dokument ten nie odnosi się wprawdzie konkretnie do zautomatyzowanej oceny, ale bez wątpliwości może stanowić istotne wyjaśnienie wątpliwości związanych z zakresem udzielanych informacji. UKNF wskazał w swoim piśmie, że „optymalnym i pożądanym przez UKNF, a także czyniącym zadość celowi art. 70a ustawy rozwiązaniem w tym [wyjaśnienie podstaw decyzji kredytowej] zakresie jest uwzględnianie w odpowiedziach na wnioski klientów: zindywidualizowanej i szczegółowej informacji, w tym informacji na temat środków, które powinien przedsięwziąć wnioskujący, aby usunąć negatywne skutki determinujące decyzję kredytodawcy o nieprzyznaniu kredytu”. Urząd wskazał także, że takie wyjaśnienia powinny zawierać informacje dotyczące czynników, w tym także danych osobowych wnioskującego, które wpłynęły na dokonaną ocenę zdolności kredytowej.

Prawo ubezpieczeniowe a wykorzystanie AI

Przepisy prawa ubezpieczeniowego¹² wprowadzają dość podobne wymogi dotyczące podejmowania zautomatyzowanych decyzji w sprawach dokonania oceny ryzyka ubezpieczeniowego oraz likwidacji szkód – ustalania przyczyn i okoliczności zdarzeń losowych, jak również ustalania wysokości szkód oraz rozmiaru odszkodowań, oraz innych świadczeń należnych uprawnionych z umów ubezpieczenia lub umów gwarancji ubezpieczeniowych¹³.

Prawo ubezpieczeniowe uzależnia podejmowanie takiej decyzji od *zapewnienia osobie, której dotyczy zautomatyzowana decyzja, prawa do otrzymania stosownych wyjaśnień co do podstaw podjętej decyzji, zakwestionowania tej decyzji, wyrażenia własnego*

¹¹ https://www.knf.gov.pl/knf/pl/komponenty/img/Komunikat_UKNF_ws_prawa_do_uzyskania_wyjasnien_nt_oceny_zdolnosci_kredytowej_wersja_szczegolowa_70332.pdf

¹² Ustawa z dnia 11 września 2015 r. o działalności ubezpieczeniowej i reasekuracyjnej (t.j. Dz. U. z 2024 r. poz. 838 z późn. zm.).

¹³ Art. 41 tejże ustawy.

stanowiska oraz do uzyskania interwencji ludzkiej. Ustala również zamknięty katalog danych, z którego mogą być wykorzystywane w celu podejmowania takich decyzji.

Opisane wyżej systemy zautomatyzowanej oceny zdolności kredytowej czy ryzyka ubezpieczeniowego, czy systemy likwidacji szkód są szczególną formą rozwiązania, o którym mowa w art. 22 RODO. Tak więc i pozostałe systemy, które działają na podobnej zasadzie, będą wymagały realizacji szczególnych praw podmiotów danych, a więc klientów instytucji finansowych.

Robodoradztwo

Doradztwo inwestycyjne zostało zdefiniowane w ustawie z dnia 29 lipca 2005 r. o obrocie instrumentami finansowymi w sposób następujący¹⁴: „doradztwo inwestycyjne polega na przygotowywaniu (z inicjatywy firmy inwestycyjnej albo na wniosek klienta) oraz przekazywaniu klientowi, 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 rekomendacje dotyczące powstrzymania się od wykonania takiej czynności”.

Zgodnie z wytycznymi Europejskiego Urzędu Nadzoru Rynku Finansowego (ESMA)¹⁵ podmiot świadczący usługi robodoradztwa powinien spełnić względem klienta określone obowiązki informacyjne, aby zniwelować ewentualne niedostatki w jego wiedzy ekonomiczno-finansowej i istotnie ograniczyć dla niego ryzyko:

- czytelnie i precyzyjnie wyjaśnić stopień i zakres zaangażowania człowieka oraz poinformować, czy i w jaki sposób klient może zwrócić się do firmy z prośbą o kontakt z pracownikiem;
- wyjaśnić, że udzielone przez klientów odpowiedzi w ramach przeprowadzenia oceny odpowiedniości będą miały bezpośredni wpływ na rekomendowane instrumenty finansowe;

¹⁴ Art. 76 ustawy z dnia 29 lipca 2005 r. o obrocie instrumentami finansowymi (t.j. Dz. U. z 2024 r. poz. 722 z późn. zm.).

¹⁵ https://www.esma.europa.eu/sites/default/files/2023-04/ESMA35-43-3172_Guidelines_on_certain_aspects_of_the_MiFID_II_suitability_requirements_PL.pdf, wytyczna 17

- opisać źródła informacji wykorzystywanych do generowania rekomendacji inwestycyjnych (np. w przypadku wykorzystania formularza elektronicznego firma inwestycyjna powinna wyjaśnić, że odpowiedzi udzielone w kwestionariuszu mogą stanowić jedyną podstawę do sporządzenia spersonalizowanej rekomendacji inwestycyjnej ew. wskazać, czy firma ma dostęp do innych informacji lub kont/rachunków klienta);
- wyjaśnić, w jaki sposób i kiedy informacje dotyczące klienta będą aktualizowane z uwzględnieniem jego sytuacji majątkowej, okoliczności osobistych itp.

W 2020 r. Urząd Komisji Nadzoru Finansowego wydał stanowisko w sprawie świadczenia usługi robodoradztwa¹⁶, które określa wymagania stawiane podmiotom świadczącym te usługi, zarówno w aspekcie wewnętrznym (jakość danych, nadzór człowieka), jak i zewnętrznym – w kontakcie z klientem.

Obowiązki te można podzielić na obszar:

1. informacyjny związany z przekazywaniem klientowi stosownych informacji zarówno przed, jak i w trakcie świadczenia usługi oraz
2. związany z kontaktem z pracownikiem instytucji, czyli umożliwieniem klientowi wyjaśnienia wątpliwości z udziałem człowieka.

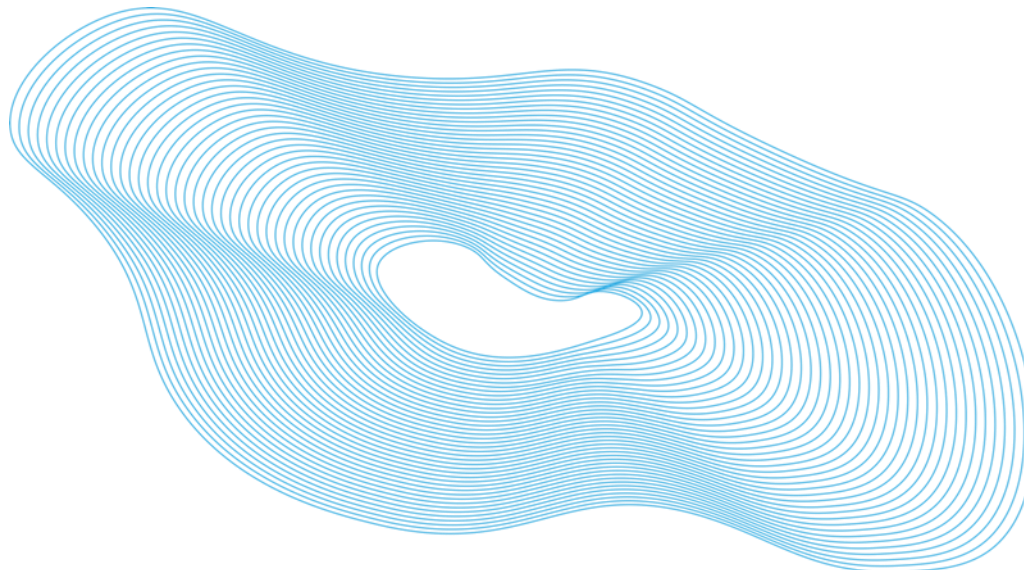
W przypadku obowiązków informacyjnych można posłużyć się podsumowaniem, które zaproponował UKNF w swoim stanowisku wskazując, że „firma inwestycyjna powinna przekazać w sposób przystępny i rzetelny informacje na temat sposobu świadczenia usługi robodoradztwa. W tym celu może wykorzystywać funkcjonalności wspomagające wyświetlanie treści strony internetowej, w tym przyciągające uwagę użytkownika strony”. Oznacza to konieczność przekazywania nie tylko informacji związanych stricte z usługą doradztwa inwestycyjnego, ale także podstawowych założeń działania systemu, który stoi za automatyzacją tej usługi czy funkcjonalności, z których może skorzystać klient.

Rozciąga się to także na konieczność zapewnienia odpowiednich informacji w regulaminie świadczenia usługi, gdzie konieczne jest wskazanie, chociażby jaki jest zakres

¹⁶ https://www.knf.gov.pl/knf/pl/komponenty/img/Stanowisko_UKNF_ws_swiadczenia_uslugi_robo_doradztwa_71303.pdf

automatyzacji oraz udziału czynnika ludzkiego, a także podstaw działania systemu (algorytmu), w kontekście generowanych rekomendacji. Nie oznacza to, że firma inwestycyjna będzie zobowiązana do przekazania szczegółowych informacji dotyczących jej algorytmu (te mogą stanowić tajemnicę przedsiębiorstwa), jednak jej klient powinien być w stanie zrozumieć, jakie są założenia samej usługi i z jakim ryzykiem może się ona wiązać. Niewypełnienie tych obowiązków może wiązać się z odpowiedzialnością samej firmy inwestycyjnej, jeżeli jest związane to z naruszeniem przepisów prawa odnoszących się do usługi doradztwa inwestycyjnego, np. w zakresie przejrzystości.

Wspomniane już stanowisko nie nakazuje wprowadzić firmie inwestycyjnej, aby ta zapewniała zawsze kontakt klienta z człowiekiem, jednak zdaniem urzędu jest to dobra praktyka, która może zapewnić wyjaśnienie tych kwestii, które mogą budzić wątpliwości. Sposób realizacji tej rekomendacji pozostaje po stronie instytucji, której powinno jednak zależeć na zapewnieniu klientowi dostępności informacji.



AI Act – czyli jak będzie już za nieco mniej niż rok

Termin wejścia w życie

Jak wskazaliśmy w rozdziale IV, AI Act został przyjęty 13 czerwca 2024 r. i część jego przepisów, w tym te dotyczące zakazu stosowania pewnych rodzajów systemów AI, już obowiązuje.

Większość przepisów jednak będzie obowiązywać dopiero od 2 sierpnia 2026 r. Dla pełnego i skutecznego stosowania AI Act wymagane jest również przyjęcie przez polski parlament ustawy stosującej, czyli przepisów określających krajowe zasady stosowania (i egzekwowania) zakazów i nakazów określonych w rozporządzeniu.

Systemy wysokiego ryzyka

Systemami wysokiego ryzyka będą w szczególności następujące systemy AI:

- związane z dostępem do podstawowych usług prywatnych oraz podstawowych usług i świadczeń publicznych, a wśród nich:
 - systemy AI przeznaczone do wykorzystywania do celów oceny zdolności kredytowej osób fizycznych lub ustalenia ich scoringu kredytowego, z wyjątkiem systemów AI wykorzystywanych w celu wykrywania oszustw finansowych;
 - systemy AI przeznaczone do wykorzystywania przy ocenie ryzyka i ustalaniu cen w odniesieniu do osób fizycznych w przypadku ubezpieczenia na życie i ubezpieczenia zdrowotnego;
- związane ze ściganiem przestępstw,
- związane z biometrią,
- przeznaczone do wykorzystywania jako związane z bezpieczeństwem elementy procesów zarządzania krytyczną infrastrukturą cyfrową, ruchem drogowym i procesów ich działania lub zaopatrzenia w wodę, gaz, ciepło lub energię elektryczną.
- związane z kształceniem i szkoleniem zawodowym,

- związane z zatrudnieniem, zarządzaniem pracownikami i dostępem do samozatrudnienia,
- związane z zarządzaniem migracją, azylem i kontrolą graniczną,
- sprawowaniem wymiaru sprawiedliwości i procesami demokratycznymi.

Pełną listę systemów wysokiego ryzyka określają art. 6 i załącznik III do AI Act. Nas oczywiście najbardziej interesować będą dwa pierwsze systemy wysokiego ryzyka wymienione powyżej – związane ze scoringiem kredytowym i działalnością ubezpieczeniową.

AI Act wprowadza bowiem wobec systemów AI wysokiego ryzyka szereg wymogów:

- ustanowienie i wdrożenie systemu zarządzania ryzykiem,
- rozwijanie systemów AI na podstawie odpowiednich danych treningowych, walidacyjnych i testowych,
- stworzenie odpowiedniej dokumentacji technicznej dla systemu AI,
- stworzenie zautomatyzowanych rejestrów zdarzeń w ramach systemu AI,
- zapewnienie przejrzystości działania i dołączanie instrukcji,
- zapewnienie nadzoru człowieka na etapie projektowania i rozwoju,
- zapewnienie odpowiedniego poziomu dokładności, solidności i cyberbezpieczeństwa.

Wymogi wobec podmiotów wykorzystujących systemy AI wysokiego ryzyka

Podmioty wykorzystujące systemy AI wysokiego ryzyka zostaną zobowiązane do spełnienia szeregu dodatkowych wymogów¹⁷, w szczególności:

- zapewnienia nadzoru nad systemem AI ze strony eksperta (człowieka) posiadającego niezbędne kompetencje, przeszkolenia i uprawnienia, a także niezbędne wsparcie w ramach organizacji,
- zapewnienie adekwatności i wystarczającej reprezentatywności danych wejściowych (jeżeli sprawują nad nimi kontrolę),

¹⁷ Art. 26 AI Act

- monitorowania systemu AI pod kątem generowanego przez niego ryzyka,
- przechowywania informacji o zdarzeniach związanych z funkcjonowaniem systemu AI przez okres nie krótszy niż 6 miesięcy lub inny opisany przepisami,
- informowania osób fizycznych, wobec których wykorzystują systemy AI, o tym fakcie.

Podmioty nieprzestrzegające ww. wymogów podlegać będą administracyjnej karze pieniężnej w wysokości do 15 mln euro lub – jeżeli sprawcą naruszenia jest przedsiębiorstwo – w wysokości do 3% jego całkowitego rocznego światowego obrotu z poprzedniego roku obrotowego, w zależności od tego, która z tych kwot jest wyższa¹⁸.

Prawa osób, na które systemy AI mają wpływ

Osoby, na które AI ma wpływ, powinny mieć prawo do uzyskania wyjaśnienia¹⁹, jeżeli:

- decyzja opiera się głównie na wynikach określonych systemów AI wysokiego ryzyka i
- decyzja ta wywołuje skutki prawne lub podobnie znacząco oddziałuje na te osoby w sposób, który ich zdaniem ma niepożądany wpływ na ich zdrowie, bezpieczeństwo lub prawa podstawowe.

Wyjaśnienie powinno być jasne i merytoryczne oraz powinno dawać osobom, na które AI ma wpływ, podstawę do korzystania z ich praw.

Prawo do wniesienia skargi

Każda osoba fizyczna lub prawna mająca podstawy, by uznać, że zostały naruszone przepisy AI Act, będzie mogła wnieść skargę do odpowiedniego organu nadzoru rynku²⁰. Krajowe przepisy określające tę ścieżkę nie zostały jeszcze przyjęte.

Informacje dotyczące egzekwowania powyższych uprawnień zostaną uzupełnione po przyjęciu ustawy stosującej AI Act.

¹⁸ Art. 99 ust. 4 lit. e

¹⁹ Art. 86 AI Act

²⁰ Art. 85 AI Act

Kodeks postępowania w zakresie AI ogólnego przeznaczenia

Kodeks postępowania w zakresie AI ogólnego przeznaczenia to dokument stworzony przez niezależnych ekspertów na zlecenie Komisji Europejskiej, który ma wspierać wdrażanie AI Act. Kodeks został opublikowany 10 lipca 2025 r., aktualnie państwa członkowskie i Komisja oceniają jego adekwatność. Kodeks zostanie dodatkowo uzupełniony wytycznymi Komisji dotyczącymi kluczowych pojęć związanych z modelami AI ogólnego przeznaczenia, które zostaną opublikowane jeszcze w lipcu 2025 roku.

Stosowanie Kodeksu ma charakter dobrowolny, a zatem obowiązek przestrzegania jego postanowień mają tylko jego sygnatariusze, w przeciwieństwie do AI Act, który musi być przestrzegany przez wszystkich dostawców modeli ogólnego przeznaczenia. Kodeks podpisała już firma Open AI (Chat GPT), a zamiar podpisania Kodeksu zgłosiły takie podmioty jak Google (Gemini), Microsoft (Microsoft Copilot), czy Anthropic (AI. Claude). Pomimo dobrowolnego charakteru Kodeksu, jego przyjęcie stwarza domniemanie, że dostawca modeli ogólnego przeznaczenia wypełnił obowiązki wynikające z AI Act, natomiast dostawcy, którzy nie podpisali Kodeksu, muszą przedstawić Komisji alternatywne środki inne Kodeks potwierdzające, że wypełnia zobowiązania wynikające z AI Act.

Kodeks składa się z trzech rozdziałów, pierwszy dotyczy przejrzystości rozumianej jako obowiązek wyjaśnienia przez dostawcę modelu ogólnego przeznaczenia architektury modelu, czyli w jaki sposób dany model podejmuje decyzje, skąd pobiera dane testowe, jakie ma ograniczenia. Dostawca ma obowiązek dokumentowania wspomnianych danych oraz przechowywania ich przez okres 10 lat od chwili wprowadzenia modelu na rynek, aby dane były wystandaryzowane. Kodeks zawiera gotowy formularz do wypełnienia przez dostawcę. Użytkownik ma prawo zwrócić się do dostawcy o udostępnienie wspomnianych danych, a dostawca ma obowiązek zrealizować prośbę w terminie 14 dni. Dzięki temu użytkownik może zorientować się, kto odpowiada za błędy popełnione przez AI oraz ocenić przydatność modelu do danego celu uwzględniając jego ograniczenia.

Drugi rozdział Kodeksu dotyczy praw autorskich. Dostawcy modeli ogólnego przeznaczenia są zobowiązani do przestrzegania prawa autorskiego Unii Europejskiej, czyli m.in. zapewnienia, że dane do modeli są pozyskiwane z poszanowaniem praw

autorskich, a także stosowanie technologii uniemożliwiającej obchodzenie zabezpieczeń typu paywall, czy technologii DRM (Digital Rights Management – technologii cyfrowego zarządzania prawami autorskimi np. zapewnienie streamingu bez możliwości kopiowania, ochrona e-booków przed kopiowaniem, umożliwienie dostępu do treści tylko subskrybentom itp.), a także wykluczenia pobierania danych z serwisów uznanych prawomocnie za uporczywie naruszające prawa autorskie.

W trzecim rozdziale Kodeksu sformułowano zasady dotyczące zarządzania ryzykiem systemowym, czyli ryzykiem wynikającym z najbardziej zaawansowanych modeli sztucznej inteligencji. Dostawcy mają obowiązek przeprowadzania testów bezpieczeństwa identyfikujących ryzyka dla zdrowia, bezpieczeństwa, praw podstawowych, środowiska i demokracji, a także obejmujące scenariusze manipulacji, nadużyć, halucynacji, toksycznych treści oraz błędnych odpowiedzi. Dzięki temu możliwe będzie mitygowanie ryzyka np. nierównego dostępu do usług bankowych wynikającego z uprzedzeń modelu na tle płci, rasy, czy poglądów politycznych.

