Wymiar biznesowy ataków na systemy uczące się. Cz. 1

Mariusz Rafało

1. Wprowadzenie

W dzisiejszych czasach większość procesów biznesowych jest zinformatyzowana, a wiele z nich jest realizowanych wyłącznie w świecie cyfrowym. Informatyzacja sprawia, że poszczególne kroki procesu, generowane dane czy inne informacje, są rejestrowane w bazach danych. Dane rejestrowane w ten sposób określa się mianem cyfrowego śladu. Służą one monitorowaniu procesu i analizowaniu aktywności uczestników procesu (Surma, 2017). Przykładowo: w bazach danych dostępne są informacje o złożonych zamówieniach, wystawionych fakturach, dostępnych produktach, zalogowanych klientach itd. Rejestrowane są wszelkie działania i aktywności klientów wykonywane na stronie WWW, w aplikacjach mobilnych, a także coraz częściej w sklepach i punktach obsługi klienta. Dane mogą być również zbierane bez aktywnego udziału klienta – np. za pomocą czujników RFID, geolokalizacji GSM, GPS czy lokalizacji wi-fi. Niezależnie od branży cyfrowe ślady interakcji firmy z klientami są rejestrowane w systemach informatycznych. Przy czym coraz częściej dzieje się to w czasie niemal rzeczywistym, co oznacza przykładowo, że jeśli klient złożył reklamację w placówce firmy, to niemal w tej samej chwili informacja o tym jest dostępna w systemie CRM, zatem gdy klient zadzwoni na infolinię tej firmy, zostanie obsłużony z wykorzystaniem najbardziej aktualnej wiedzy.

Dane pochodzące z cyfrowego śladu mogą być wykorzystane dwojako. Po pierwsze, mogą służyć bieżącemu wsparciu procesów biznesowych. Po drugie, dane można wykorzystywać w celach analitycznych. Służą temu głównie dane historyczne, które można wykorzystać jako zbiór uczący dla systemów uczących się czy, ogólnie mówiąc, systemów sztucznej inteligencji (AI). Systemy AI, mające „wiedzę" na temat historycznych transakcji, pozwalają nie tylko na automatyzację działań operacyjnych, lecz także na wspieranie w podejmowaniu decyzji. Algorytmy sztucznej inteligencji opierają się na dobrze ugruntowanych zasadach matematyki, statystki i ekonometrii. W wielu przypadkach złożoność algorytmów i systemów AI jest jednak tak duża, że są one raczej postrzegane jako „czarne skrzynki", które realizują konkretne działania, w sposób nie zawsze zrozumiały dla użytkownika. Ta złożoność, niedostępna dla ludzkiej percepcji, może być wykorzystana do potencjalnego „oszukania" systemu AI. Celem takich działań może być znalezienie luki w „czarnych skrzynkach" modeli sztucznej inteligencji. Można tego dokonać, stosując zaawansowane systemy analizy danych lub przez wprowadzanie do modelu specjalnie spreparowanych danych. Zaatakowany w ten sposób proces może zadziałać niepoprawnie (np. przyznać kredyt osobie, która nie powinna go otrzymać, zignorować transakcję, która jest oszustwem itp.) lub może całkowicie przestać działać.

Celem autora tego rozdziału jest przedstawienie zagrożeń wynikających z zastosowania autonomicznych systemów sztucznej inteligencji (robotów programowych), które wspierają lub całkowicie realizują procesy biznesowe. Niezawodność i ciągłość procesów biznesowych stanowią nie tylko o efektywności działania firmy, ale niejednokrotnie o jej działaniu w ogóle. Przykładami mogą być tu sklep internetowy, który traci możliwość finalizacji składanych zamówień, lub bank, który utracił zdolność rozpoznawania ryzykownych kredytobiorców. Tego typu zagrożenia dla systemów uczących się wymagają zastosowania odpowiednich metod zarządzania ryzykiem.

2. Robotyzacja i automatyzacja procesów biznesowych

Proces biznesowy to uporządkowany zestaw działań, które są określone, mierzalne i prowadzą do uzyskania konkretnego rezultatu. Może on obejmować krótkie sekwencje działań (np. wystawienie faktury, przyjęcie reklamacji) lub bardziej złożone aktywności (np. realizacja zamówienia online, z możliwością odbioru produktu w wybranej lokalizacji). Ujęcie procesowe pozwala zarządzającym na kontrolowanie i monitorowanie działań firmy, zgodnie z prowadzonymi działaniami sprzedażowymi, marketingowymi czy obsługą klienta.

Automatyzacja procesu biznesowego polega na realizacji tego procesu (lub jego fragmentu) za pomocą technologii, bez udziału pracownika lub przy jego minimalnym udziale (zazwyczaj ograniczającym się do nadzoru). Obecnie większość procesów biznesowych jest wspierana przez informatyczne systemy zarządzania lub jest realizowana w całości w świecie cyfrowym. Takie procesy, jak elektroniczny obieg dokumentów, sprzedaż w kanałach elektronicznych czy śledzenie przesyłek kurierskich online, to przykłady procesów zautomatyzowanych.

Jeśli natomiast automatyzacja jest realizowana za pomocą oprogramowania, które wykonuje określone, często powtarzalne zadania, to można mówić o robotyzacji tego procesu. Robotyzacja jest rodzajem automatyzacji, który polega na tym, że proces (lub jego fragment) jest realizowany przez system informatyczny, tj. robota.

2.1. Robotyzacja procesów

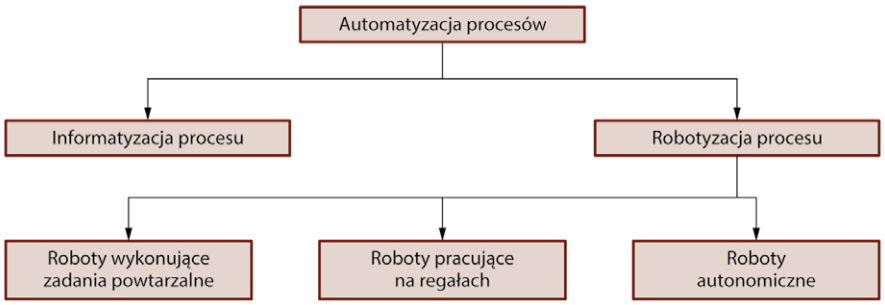
Roboty kojarzą się powszechnie z automatami, które wykonują określone zadania w fabrykach, na liniach produkcyjnych lub halach montażowych (pomijając stereotypowy wizerunek robota w literaturze i kinie). Roboty, które wspierają lub realizują procesy biznesowe, są najczęściej specjalizowanymi systemami komputerowymi. Określanie tych programów mianem robotów stanowi metaforę; roboty programowe realizują zadania w świecie cyfrowym, analizują dane i wykonują określone akcje.

Robotyzacja procesów biznesowych (ang. *Robotic Process Automation* – RPA) to jeden z rodzajów automatyzacji procesów, który jest realizowany za pomocą robota programowego (Sobczak, 2020 a). Założeniem robotyzacji procesu jest zastąpienie powtarzalnych zadań wykonywane przez pracowników specjalistycznym oprogramowaniem (Sobczak, 2018). Robotyzacja zwiększa efektywność realizowanego procesu biznesowego w kilku wymiarach. Po pierwsze, w ujęciu wydajności pracy: robot może przetworzyć i zweryfikować setki dokumentów, wprowadzić tysiące informacji do systemu lub zweryfikować jakość danych pomiędzy systemami. Drugi wymiar dotyczy jakości wykonywanych zadań: przy powtarzalnych zadaniach o określonej strukturze robot raczej nie popełnia błędów. Trzeci aspekt to ciągłość pracy: robot może pracować bez przerwy, a jeśli zajdzie taka potrzeba, może zostać zeskalowany, aby pracować równolegle.

Popularnym przykładem robotów programowych, opartych na regułach, są chatboty i voiceboty. Pierwsza kategoria służy do obsługi klienta za pomocą czatu, najczęściej na stronie internetowej, druga służy obsłudze klienta na linii telefonicznej. W obu przypadkach robot obsługuje komunikację z klientem i realizuje wybrane zadania, wynikające z informacji otrzymanych od klienta.

Podstawowe ujęcie robotyzacji dotyczy zatem realizacji większej liczby operacji oraz wykonywania ich w sposób nieprzerwany i bezbłędnie. Inne, nieco mniej oczywiste zastosowania robotów obejmują nie tyle automatyzację działań, ile ich samodzielne inicjowanie. Przykładowo robot programowy, mający autonomię, może ocenić zdolność kredytową klienta banku, może zweryfikować online poprawność transakcji internetowych w ciągu dziesiętnych części sekundy czy też zaproponować maklerowi najbardziej optymalną decyzję zakupową. Roboty autonomiczne działają na nieco innych zasadach niż roboty „klasyczne", oparte na powtarzalnych regułach.

Systemy klasy RPA służą do budowania robotów programowych, uruchamiania ich w środowisku procesu biznesowego oraz do sterowania robotami. Możliwości automatycznej realizacji zadań przez roboty programowe



Rys. 1. Klasyfikacja automatyzacji procesów biznesowych Źródło: opracowanie własne

pochodzą ze zdefiniowanych reguł postępowania w konkretnych przypadkach. System działa według określonych sekwencji i wykonuje określone akcje. Przykładowo mogą to być akcje ekranowe (wprowadzenie informacji, nawigowanie po ekranie itp.), akcje związane z komunikacją z innymi systemami (np. połączenie do bazy danych czy uruchomienie innego systemu) itp. Wraz ze wzrostem poziomu autonomii robota programowego komponenty, które odpowiadają za podejmowanie decyzji, wykorzystują coraz bardziej zaawansowane techniki analiz danych. W zależności od posiadanej autonomii systemy programowe można podzielić na (rys. 1):

- **roboty, które wspierają działania rutynowe:** weryfikacja dokumentów, automatyczne skanowanie dokumentów PDF, odczytywanie danych z jednego systemu i wpisywanie ich do innego itp.;
- **roboty oparte na regułach:** systemy działają, wykonując akcje, które są wyzwalane przez określone zdarzenia czy dane, np. obsługa klienta za pomocą chatbota;
- **roboty autonomiczne:** mogą podejmować samodzielne decyzje, np. w zakresie rekomendacji produktu klientowi lub w zakresie oceny zdolności kredytowej klienta.

Warto także wspomnieć, że istnieją zrobotyzowane procesy, które w ogóle nie mogłyby być realizowane, gdyby nie wykonywał ich robot. Przykładowo proces spersonalizowanej rekomendacji produktów w kanale internetowym byłby niemożliwy do realizacji bez udziału specjalizowanego systemu[1]. Sam proces prezentowania określonej treści klientom online oczywiście istnieje, ale ma on niewiele wspólnego z rekomendacją personalizowaną, ponieważ każdy (lub niemal każdy) klient otrzymuje ten sam komunikat. Dopiero wprowadzenie algorytmów eksploracji danych pozwala na to, aby każdego (lub niemal każdego) klienta traktować indywidualnie i prezentować mu treści, które mogą być dla niego interesujące.

2.2. Sztuczna inteligencja w robotyzacji procesów

Zaawansowane techniki i technologie analizowania danych stały się w ostatnich latach bardziej dostępne dla organizacji. W sposób szczególny przyczynia się do tego rozwój oprogramowania otwartego (ang. *open source*), głównie w domenie *big data*, oraz rozwój systemów oferowanych w chmurze (ang. *cloud computing*). Pokusa jest duża: zastosowanie AI pozwala bowiem na istotne optymalizacje i korzyści we wspieranych przez nie procesach biznesowych. Dla niektórych firm stanowi to o optymalizacji realizowanych procesów, dla innych stanowi podstawowy czynnik przewagi konkurencyjnej (Davenport i Harris, 2007; Surma, 2009). Nawet relatywnie proste systemy AI mogą zostać wykorzystane do poprawy procesów decyzyjnych czy do wsparcia procesów biznesowych. Przykładowo sieć neuronowa może wspierać decydentów w instytucji finansowej w doborze produktów finansowych w portfelu inwestycyjnym (Culkin i Das, 2017). Inny moduł, oparty na uczeniu maszynowym, może implementować chatbota, aby obsługiwał wybrane zlecenia klientów. Bot tekstowy (lub głosowy) może istotnie obniżyć koszty obsługi klienta, zwłaszcza przy większej liczbie klientów i pewnej strukturyzacji zadań.

Obecnie takie procesy biznesowe, jak rekomendacje produktów, analiza ryzyka kredytowego, wycena szkód ubezpieczeniowych czy identyfikacja nadużyć są z powodzeniem realizowane przez roboty programowe, bez udziału lub przy minimalnym udziale człowieka (Sobczak, 2020 b). Procesy te, ze względu na specyfikę (brak powtarzalności i stałych reguł), są realizowane przez roboty programowe wykorzystujące techniki sztucznej inteligencji.

Wykorzystanie przez roboty programowe zaawansowanej analityki nie świadczy jeszcze o autonomii systemu RPA. Istnieją bowiem systemy, które wyłącznie wskazują rozwiązania problemów decyzyjnych, zaś podjęcie tej decyzji (podjęcie działania) pozostawiają operatorom. Takie rozwiązanie spotykane jest np. w systemach służących ustaleniu prawdopodobieństwa odejścia klienta lub rezygnacji klienta z usług firmy (tzw. analizy *churn*) w banku lub firmie telekomunikacyjnej. Rolą systemu jest analiza danych w celu zidentyfikowania i oznaczenia klientów, którzy w najbliższym czasie zrezygnują z usług firmy. Dalsze decyzje dotyczące tych klientów (zaproponowanie klientom atrakcyjnych promocji, obniżenie ceny świadczonych usług itp.) są już realizowane poza systemem RPA. Podobną sytuację można zauważyć na rynkach finansowych, gdzie przepływ danych następuje online, a systemy AI analizują te dane w czasie rzeczywistym i rekomendują decyzje.

W obu przytoczonych przykładach spotkać można jednak zastosowania, gdzie robot, poza rekomendacją decyzji, także ją podejmuje. W przypadku analizy *churn* robot może sam dokonać wysłania kampanii marketingowej do wytypowanych klientów, proponując im określone produkty czy promocje. W przypadku systemów finansowych samodzielne decyzje o zakupie czy sprzedaży mogą przynieść konkretne korzyści, bowiem system ma możliwość analizy pełnej dynamiki rynku (ceny akcji, kursy walut, stopy procentowe itp.), a także zachodzących pomiędzy nimi interakcji.

Jest to szczególnie przydatne w tych zastosowaniach, gdzie nie ma czasu na weryfikację proponowanej klasyfikacji przez osobę nadzorującą. Niekiedy decyzja musi być podjęta natychmiast, ponieważ już za kilka minut może ona być już nieadekwatna do sytuacji w otoczeniu. Takie scenariusze, jak blokowanie podejrzanych transakcji finansowych, oferowanie pożyczek gotówkowych w bankomatach czy wysyłanie powiadomień do klientów, którzy znajdują się w pobliżu sklepu, wymagają działań automatycznych. Wynika to z faktu, że wartość każdej informacji eroduje w czasie (Koźmiński, 2004).

Takie przekazanie decyzyjności systemom RPA jest efektywne ekonomicznie. Z jednej strony firma jest w stanie obsłużyć większą liczbę klientów, realizować personalizowane rekomendacje czy nadzorować pracę linii produkcyjnej. Z drugiej strony otwiera to furtkę dla potencjalnych ataków, których celem może być zatrzymanie procesu lub jego niepoprawne działanie. Atak może się odbywać przez dostarczenie do systemu określonych „wadliwych" danych. Takie dane, rozpoznane przez system AI, mogą spowodować określone działania systemu: zatrzymanie linii produkcyjnej, błędne decyzje zakupowe czy niepoprawne decyzje dotyczące oceny zdolności kredytowej.

Proces budowy modeli opartych na uczeniu maszynowym opiera się na poszukiwaniu powiązań i regularności w danych. Nauka modelu odbywa się na bazie dostarczonych danych, tzw. danych uczących. Algorytm uczący jest trenowany na zbiorze uczącym, który zawiera informacje o wyniku predykcji. Zakłada się, że po nauczeniu modelu można go wykorzystać do predykcji wyników także dla innych przypadków. Model, nie mając dostępu do innych danych, posługuje się zatem uogólnieniami i regułami wyuczonymi z danych uczących. To ważna cecha, która powoduje, że model jest możliwy do zastosowania na danych, które nie są znane wcześniej. Zachowanie poziomu ogólności wynika z zagrożenia tzw. przeuczeniem modelu. Przeuczony model charakteryzuje się wysoką szczegółowością odnalezionych reguł. Szczegółowe reguły doskonale odzwierciedlają stan zbioru danych uczących, jednak w przypadku jakichkolwiek innych danych okazują się one zbyt wyspecjalizowane. Model nie ma zdolności do klasyfikowania przypadków nieco innych, bo jego reguły są zbyt precyzyjne (Provost i Fawcett, 2013).

Podatność systemów AI na ataki (lub działania niezamierzone) wynika z faktu, że systemy te relatywnie słabo radzą sobie z adaptacją do nowych warunków (do nowych danych) oraz z sytuacjami wyjątkowymi. Jeśli dane wejściowe dla robota programowego będą istotnie różne od tych, które robot już zna (od danych uczących), jego zachowanie może nie być deterministyczne. Dodatkowym ograniczeniem jest fakt, że systemy odbierają informacje za pomocą innych niż człowiek zmysłów. Dlatego możliwe jest, że klasyfikowany obiekt (zdjęcie, dźwięk czy cechy klienta) jest błędnie oceniany przez model, podczas gdy człowiek nie miałby problemu z poprawną oceną. Te cechy robotów programowych (a także algorytmów i systemów AI w ogóle) stanowią o ich podatności na ataki spowodowane wygenerowanymi sztucznie danymi.

3. Ryzyko operacyjne w procesach biznesowych

Automatyzacja podejmowania decyzji generuje ryzyko, że w przypadku awarii systemu podejmie on błędną decyzję lub w ogóle zatrzyma się. Jest to szczególnie obserwowalne w przypadku pojawienia się sytuacji (danych), które odbiegają od normy. Jeśli robot programowy napotka sytuację nieprzewidzianą, która nie została uwzględniona przy jego projektowaniu, to może on zachować się dwojako. Po pierwsze, może zgłosić anomalię do administratora lub innego systemu – jest to możliwe tylko wówczas, gdy projektant robota zaimplementował taką funkcję. W przeciwnym wypadku robot będzie działał nadal, ale jego zachowanie będzie nieadekwatne do sytuacji (Sobczak, 2020 a). Owo niedeterministyczne zachowanie może prowadzić do zatrzymania procesu biznesowego (np. jeśli robot chatbot ulegnie awarii, obsługa klientów tym kanałem staje się niemożliwa) lub do jego wadliwego funkcjonowania, np. dane wprowadzane przez

robota programowego do systemu są niepoprawne. W przypadku robotów wspieranych sztuczną inteligencją awaria może także prowadzić do zatrzymania działania systemu lub do jego błędnego działania, jednak w przypadku tych systemów skutki tej awarii mogą być dalece szersze, przykładowo: błędnie przydzielane kredyty, błędne decyzje zakupowe, błędne rekomendacje produktów klientom.

3.1. Problematyka ryzyka

Istotne jest zatem, aby robotyzowane i automatyzowane procesy biznesowe monitorować oraz by zarządzać ryzykiem utraty ciągłości ich funkcjonowania. Intuicyjnie wydaje się, że ryzyko awarii systemu opartego na RPA jest niższe niż ryzyko pomyłki w przypadku, gdy te działania byłyby realizowane ręcznie. Biorąc pod uwagę możliwość popełnienia błędu, można przyjąć, że tak jest w rzeczywistości: maszyny oczywiście mogą się „mylić" (zadziałać niepoprawnie), ale prawdopodobieństwo tego jest znikome – zwłaszcza dla czynności o dobrze znanej strukturze. Każde ryzyko sklasyfikować można nie tylko względem prawdopodobieństwa wystąpienia, lecz także wpływu, jaki zmaterializowane ryzyko będzie miało. Ten drugi wymiar klasyfikacji ryzyka wypada już nieco mniej optymistycznie dla robotów RPA: ze względu na wysoką automatyzację roboty programowe wykonują zadania o wysokim poziomie istotności lub wykonują ich tak dużo, że sama ilość sprawia, iż są istotne. Z tej perspektywy awaria robota, choć mało prawdopodobna, może mieć wagę krytyczną dla procesu biznesowego lub całej organizacji.

Istotne jest także rozróżnienie pomiędzy ryzykiem a niepewnością. Ryzyko określa zdarzenia mające określoną strukturę oraz określone prawdopodobieństwo wystąpienia. Znając skutki ryzyka, prawdopodobieństwo jego wystąpienia oraz jego strukturę, można podjąć działania, które będą przeciwdziałać temu ryzyku. Możliwości są trzy i wynikają ze struktury ryzyka:

- można ograniczać prawdopodobieństwo wystąpienia zdarzenia;
- można ograniczać skutki wystąpienia zdarzenia, kiedy już wystąpi;
- można wpływać na zakres zdarzenia, aby go modyfikować.

W przypadku niepewności wskazane wcześniej możliwości mitygacji nie są dostępne. Niepewność jest zdarzeniem kompletnie nieznanym, dla którego nie jest możliwe określenie stanów ani prawdopodobieństw zajścia. Nie istnieje zatem strukturalny sposób na redukcję niepewności (Bielecki, 2001). Wskazuje się na elementy, takie jak informacja, która redukuje niepewność (Koźmiński, 2004) czy kapitał intelektualny (Kwiatkowski, 2000), który pozwala lepiej radzić sobie ze skutkami ryzyka.

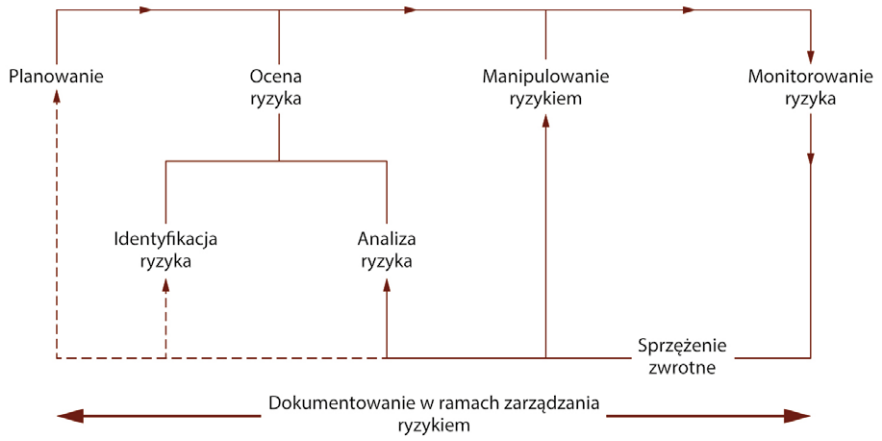
3.2. Zarządzanie ryzykiem

Zarządzanie ryzkiem to systematyczny proces służący do zidentyfikowania, oceny i kontrolowania ryzyka. Dodatkowo w wielu obszarach funkcjonowania organizacji wymóg posiadania strukturalnych metod kontroli ryzyka nie tylko wynika już z decyzji kierownictwa, lecz także stanowi standard (Tupa i in., 2017). Przykładowo zgodnie z międzynarodową normą ISO 31000:2009 zarządzanie ryzykiem można zdefiniować jako skoordynowane działania dotyczące kierowania i nadzorowania organizacją w odniesieniu do ryzyka (Niesen i in., 2016). Warunkiem skutecznego zarządzania ryzykiem jest zastosowanie ujęcia procesowego, czyli określenie zasad postępowania z sytuacjami ryzykownymi, a także zakresu integracji zarządzania ryzykiem z procesami biznesowymi (Conrow, 2000; Zawiła-Niedźwiecki, 2013). Jest to konieczne, ponieważ ryzyka materializują się właśnie w procesach biznesowych. Także z perspektywy procesu biznesowego możliwa jest ocena wpływu ryzyka.

Strukturalne podejście do ryzyka wymaga uporządkowanych, nazwanych i mierzalnych działań – modeli zarządzania ryzykiem. Za podstawowe elementy zarządzania ryzykiem uznaje się (Sadgrove, 2015): identyfikację ryzyka, jego ocenę, monitorowanie, ustalanie zasad działania, wdrożenie tych zasad oraz testowanie ich skuteczności. Z kolei organizacja COSO (*Committee of Sponsoring Organizations of the Treadway Commmission*) definiuje następujące obszary związane z zarządzaniem ryzykiem (Zawiła-Niedźwiecki, 2013):

1. Identyfikacja środowiska organizacji.
2. Określenie celów zarządzania ryzykiem.
3. Określenie wewnętrznych i zewnętrznych ryzyk.
4. Ocena i analiza ryzyka.
5. Działania w odpowiedzi na ryzyko.
6. Polityka kontroli i procedury weryfikacji.
7. Zakres i forma komunikacji ryzyk.
8. Monitorowanie i ocena działań związanych z zarządzaniem ryzykiem.

Natomiast Conrow (2000) określa model zarządzania ryzykiem obejmujący podobne elementy i aktywności, nadając im formę procesu, w którym mogą zachodzić pętle sprzężeń zwrotnych (rys. 2).



Rys. 2. Funkcjonalny model zarządzania ryzykiem

Źródło: Conrow, E. (2000). Effective Risk Management: Some Keys to Success. American Institute of Aeronautics and Astronautics

Tabela 1. Ryzyka związane z uczeniem maszynowym – względem autonomii

Autonomia	Przykład robota	Poufność	Integralność	Dostępność
Niska	Chatbot obsługujący klienta na stronie WWW	Zasady uczenia modelu dostępne dla osób niepowołanych	Chatbot rekomenduje niewłaściwe rozwiązania problemów	Chatbot nie reaguje poprawnie na zapytania klientów
Średnia	Analiza zdjęć pojazdów (likwidacja szkód ubezpieczeniowych)	Dostęp do zdjęć, które posłużyły do uczenia modelu	Model klasyfikuje błędnie niektóre zdjęcia	Model nie klasyfikuje zdjęć
Wysoka	System weryfikujący poprawność transakcji online (identyfikacja nadużyć)	Dostęp do zmiennych, które model bierze pod uwagę	Model błędnie klasyfikuje wybrane transakcje (nadużycia klasyfikowane są jako poprawne transakcje)	Model nie klasyfikuje transakcji

Klasyczny model zarządzania ryzykiem określa zależność między zarządzaniem ryzykiem a zarządzaniem ciągłością. Jest on często określany jako triada: ryzyko – bezpieczeństwo – ciągłość działania. Model ten obejmuje realizację trzech kluczowych funkcji (Zawiła-Niedźwiecki, 2013):

- analiza – aby ryzyko poprawnie zidentyfikować i nazwać;
- prewencja – aby minimalizować prawdopodobieństwo jego wystąpienia;
- terapia – aby redukować skutki ryzyka (gdy już się zmaterializuje).

Funkcje te wskazują na trzy możliwości zarządzania ryzykiem. Każda funkcja modelu obejmuje określone działania, które powinny być realizowane w procesach biznesowych.

3.3. Ryzyko w RPA działających z wykorzystaniem systemów uczących się

Rozpatrując proces zarządzania ryzykiem dla zrobotyzowanych procesów biznesowych, zwłaszcza tych, które zrobotyzowane są za pomocą AI, można stwierdzić, że im więcej autonomii udzielone jest systemom RPA w procesie biznesowym, tym większy jest wpływ ryzyka na funkcjonowanie tego procesu i/lub całej firmy. Na jeszcze większy wpływ ryzyka narażone są firmy, które opierają na automatyzacji całe modele biznesowe (a nie tylko wybrane procesy). Szczególnie takie innowacje, jak Internet Rzeczy (ang. *Internet of Th ings*) czy *big data* otworzyły możliwości dla nowych modeli biznesowych, opartych na danych (Minelli, 2013). Modele te są szczególnie narażone na niepoprawne działanie spowodowane faktem, że do systemu decyzyjnego trafiły „złe" dane.

Przez ostatnie lata nie było potrzeby adresowania tego zagadnienia, ponieważ istotność i krytyczność AI w procesach była relatywnie niska (a co za tym idzie, ryzyko awarii takiego systemu miało mały wpływ). Dopiero relatywnie niedawno to zagadnienie zyskało na znaczeniu, gdyż systemy AI są wykorzystywane coraz szerszej i mają coraz większą autonomię działania. W tych uwarunkowaniach pojawia się konieczność zaadresowania nowych kategorii ryzyk, zagrażających organizacjom w takich obszarach, jak sztuczna inteligencja, *big data* czy robotyzacja procesów (Niesen i in., 2016).

W tej sytuacji pojawiają się zagrożenia związane z zatrzymaniem pracy modeli, ich niepoprawnym działaniem lub ich niedeterministycznym działaniem (gdy niemożliwe jest określenie powodów wskazania przez model danego wyniku). Te zagrożenia wpisują się w model zarządzania bezpieczeństwem informacji (tabela 1). Model triady bezpieczeństwa obejmuje (Andress, 2011):

- Poufność – dotyczy zapewnienia bezpieczeństwa poszczególnych elementów modelu (zbiór uczący i testujący, zmienne, parametry modelu itp.). Dostęp do każdego z elementów modelu powoduje ryzyko ich wykorzystania, aby model oszukać.
- Integralność – dotyczy głównie monitorowania i zapewnienia powtarzalności wyników dla określonego modelu lub określonych reguł. Jednym z narzędzi zapewnienia integralności są miary jakości modelu (macierz pomyłek, pole pod krzywą ROC, dokładność itp.).
- Dostępność – dotyczy zapewnienia ciągłości funkcjonowania modelu. Niedostępność systemu może skutkować zablokowaniem procesu biznesowego.

Triada Poufność – Integralność – Dostępność jest szczególnie istotna, ponieważ klasyczne podejścia do zarządzania ryzykiem koncentrują się na ciągłości procesów, dzięki którym konkretny model biznesowy jest realizowany; rzadziej obejmują cały model biznesowy. Kompletny model podejścia do zarządzania ryzykiem powinien być zatem złożony z dwóch części: (1) utrzymania ciągłości modelu biznesowego firmy oraz (2) oceny i modyfikacji modelu biznesowego (Niemimaa i in., 2019). ■

1. Przykładowy system rekomendacyjny firmy Amazon oferowany jako usługa: https://aws.amazon.com/personalize/

Bibliografia dostępna pod linkiem:
nis.com.pl/bibliografia.html

Fragment pochodzi z książki:
Hakowanie sztucznej inteligencji,
Jerzy Surma (redakcja naukowa)
Wydawnictwo Naukowe PWN,
Warszawa 2020