

# Współczesne narzędzia cyfryzacji organizacji

Witold Bartkiewicz  
Piotr Czerwonka  
Anna Pamuła

# **Współczesne narzędzia cyfryzacji organizacji**



WYDAWNICTWO  
UNIwersytetu  
Łódzkiego

**CYFRYZACJA**

# **Współczesne narzędzia cyfryzacji organizacji**

**Witold Bartkiewicz**

**Piotr Czerwonka**

**Anna Pamuła**

Witold Bartkiewicz, Piotr Czerwonka, Anna Pamuła – Uniwersytet Łódzki, Wydział Zarządzania  
Katedra Informatyki, 90-237 Łódź, ul. Matejki 22/26

RECENZENT  
*Gabriela Idzikowska*

REDAKTOR INICJUJĄCY  
*Monika Borowczyk*

OPRACOWANIE REDAKCYJNE  
*Anna Dziadzio*

SKŁAD I ŁAMANIE  
*Munda – Maciej Torz*

KOREKTA TECHNICZNA  
*Leonora Gralka*

KONCEPCJA GRAFICZNA OKŁADKI I STRON TYTUŁOWYCH  
*Katarzyna Turkowska*

PROJEKT OKŁADKI  
*Agencja Reklamowa efectoro.pl*

Zdjęcie wykorzystane na okładce: © Depositphotos.com/grandfailure

© Copyright by Authors, Łódź 2020  
© Copyright for this edition by Uniwersytet Łódzki, Łódź 2020

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego  
Wydanie I. W.09041.18.0.K

Ark. wyd. 11,5; ark. druk. 15,125

ISBN 978-83-8220-211-3  
e-ISBN 978-83-8220-212-0

<https://doi.org/10.18778/8220-211-3>

Wydawnictwo Uniwersytetu Łódzkiego  
90-131 Łódź, ul. Lindleya 8  
[www.wydawnictwo.uni.lodz.pl](http://www.wydawnictwo.uni.lodz.pl)  
e-mail: [ksiegarnia@uni.lodz.pl](mailto:ksiegarnia@uni.lodz.pl)  
tel. 42 665 58 63

# Spis treści

<b>Wstęp</b>	<b>9</b>
--------------	----------

## **Rozdział 1**

Zastosowanie nowych technologii w biznesie	13
1.1. Wprowadzenie	13
1.2. Chmura obliczeniowa i rozwój technologii mobilnych	14
1.3. Rynek usług w chmurze obliczeniowej i rola użytkowników mobilnych	21
1.4. Blockchain	26
1.4.1. Założenia projektowe łańcuchów bloków	28
1.4.2. Budowa łańcucha bloków	31
1.4.3. Struktura bloku	31
1.4.4. Zastosowania technologii Blockchain	34
1.5. Przemysłowy Internet Rzeczy	34
1.6. Rzeczywistość rozszerzona	35
1.7. Podsumowanie	37
Literatura	37

## **Rozdział 2**

Systemy informatyczne i aplikacje biznesowe w zarządzaniu organizacją	<b>39</b>
2.1. Wprowadzenie	39
2.2. Systemy ERP jako efekt rozwoju systemów informatycznych dla biznesu	40
2.2.1. Systemy informatyczne zarządzania i ich typologia	40
2.2.2. Geneza i rozwój zintegrowanych systemów zarządzania	44
2.2.3. Moduły systemów ERP	51
2.2.4. Platformy współpracy między dostawcami a klientami	56
2.2.5. Technologie systemów ERP	59
2.2.6. Wdrożenie i zmiany w implementacji systemów klasy ERP	64
2.2.7. Analiza informacji gromadzonych w systemach informatycznych	70

2.3. Zastosowanie nowych technologii w systemach informatycznych	77
2.3.1. Systemy informatyczne w transformacji cyfrowej przedsiębiorstw	77
2.3.2. Inteligentne systemy ERP i produkcja w chmurze obliczeniowej	84
2.3.3. Automatyzacja procesów biznesowych a aplikacje biznesowe	90
2.4. Podsumowanie	95
Literatura	96

### **Rozdział 3**

Środowisko analityczno-decyzyjne organizacji	99
3.1. Wprowadzenie	99
3.2. Decyzje w organizacji	99
3.2.1. Decyzje menedżerskie i współczesna organizacja	99
3.2.2. Typy decyzji podejmowanych w organizacji	102
3.2.3. Ogólna charakterystyka wspomagania decyzji w organizacji	109
3.3. Systemy Wspomagania Decyzji	115
3.3.1. Pojęcie Systemu Wspomagania Decyzji (DSS)	115
3.3.2. Modele w Systemach Wspomagania Decyzji	121
3.3.3. Podstawowe metody wspomagania decyzji stosowane w DSS	125
3.3.3.1. Modele dziedzinowe	125
3.3.3.2. Modele optymalizacyjne	127
3.3.3.3. Modele wielokryterialne	132
3.3.3.4. Decyzje klasyfikacyjne	136
3.3.3.5. Badania symulacyjne i decyzje w warunkach niepewności	139
3.3.3.6. Modele analizy decyzyjnej i decyzje w warunkach ryzyka	144
3.4. Środowisko analityczne organizacji	147
3.4.1. Struktura środowiska analitycznego organizacji	147
3.4.2. Dane jako zasób organizacji – centralne zarządzanie danymi	152
3.4.3. Hurtownie danych	162
3.4.4. Eksploracja danych	165
3.4.4.1. Charakterystyka procesu eksploracji danych	165
3.4.4.2. Analiza regresji	169

3.4.4.3. Klasyfikatory	172
3.4.4.4. Analiza skupień (grupowanie danych)	180
3.4.4.5. Analiza reguł asocjacyjnych	186
3.4.5. Eksploracja tekstu i sieci Internet	190
3.4.5.1. Eksploracja tekstu (informacji)	190
3.4.5.2. Eksploracja sieci Internet	196
3.4.6. Analityka Big Data	199
3.5. Narzędzia sztucznej inteligencji	200
3.5.1. Pojęcie sztucznej inteligencji	200
3.5.2. Systemy ekspertowe	204
3.5.2.1. Ogólna charakterystyka systemów ekspertowych	204
3.5.2.2. Baza wiedzy systemu ekspertowego	205
3.5.2.3. Podsystem wnioskujący systemu ekspertowego	206
3.5.2.4. Mechanizmy objaśniające	208
3.5.2.5. Podsystem gromadzenia wiedzy	209
3.5.3. Sztuczne sieci neuronowe	210
3.5.3.1. Pojęcie i rodzaje sztucznych sieci neuronowych	210
3.5.3.2. Charakterystyka zastosowań sieci neuronowych	212
3.5.3.3. Tworzenie modeli sieci neuronowych	215
3.5.4. Systemy z logiką rozmytą	216
3.5.4.1. Nieprecyzja lingwistyczna i zbiory rozmyte	216
3.5.4.2. Systemy z logiką rozmytą	218
3.5.5. Algorytmy genetyczne	220
3.6. Podsumowanie	222
Literatura	222

<b>Analiza przypadku 1</b>	<b>225</b>
----------------------------	------------

<b>Analiza przypadku 2</b>	<b>227</b>
----------------------------	------------

<b>Analiza przypadku 3. Ryzyko wykorzystania niepewnej informacji</b>	<b>233</b>
---	------------

<b>Zakończenie</b>	<b>239</b>
--------------------	------------





# Wstęp

Menedżerowie każdej firmy koncentrują się na innowacjach i nowych technologiach, aby usprawnić działalność operacyjną, utrzymać się na rynku lub zdobyć na nim przewagę konkurencyjną. Nowe rozwiązania pozwalają zarówno zmodernizować działania i obniżyć koszty, jak również dostosować się do nowych regulacji. Obecnie, gdy cyfrowa transformacja jest nie tylko określonym celem, ale ciągłym procesem, organizacje muszą budować strategie i podejmować decyzje, które sprawią, iż sprostać one wyzwaniom coraz bardziej konkurencyjnego rynku.

Poprawa efektywności działania organizacji przez zastosowanie technologii informatycznych ma już długą historię, zapoczątkowaną przez duże jednostki komputerowe i dziedzinowe systemy wspierające wybrane obszary i określone procesy biznesowe. Kolejne generacje systemów obejmują swoim zasięgiem coraz więcej procesów biznesowych oraz zasobów organizacji.

Sprawne funkcjonowanie podmiotów gospodarczych opiera się na dobrze zintegrowanych aplikacjach pracujących na podstawie najnowocześniejszych technologii. System ERP to naturalny fundament transformacji cyfrowej – dotyka ona każdego obszaru firmy i zapewnia sprawny kontakt z partnerami i klientami. Trudno wyobrazić sobie dzisiejszy świat biznesu bez systemów klasy ERP. Wspomagane ręczne procesy biznesowe ograniczyłyby rozwój, a globalna produkcja i inne formy internacjonalizacji, w tym handel elektroniczny, byłyby niemożliwe. W przypadku wielu firm spełniły one oczekiwania i zapewniły zwrot z inwestycji – choć należy pamiętać, że znane są też sytuacje porażek prób ich wdrożenia.

Cyfrowa transformacja biznesu prowadzi do kolejnych znacznych usprawnień: łączenia źródeł danych, optymalizacji procesów biznesowych, przetwarzania informacji w czasie rzeczywistym nie tylko w ramach organizacji, ale także ekosystemu partnerów – co daje szansę rozwoju wielu stronom. Trzeba brać pod uwagę fakt, że starsze systemy funkcjonujące w organizacjach nie zostaną nagle zastąpione przez nowe. Budowana przez lata infrastruktura i dobre praktyki związane z wdrażaniem technologii mogą w pewnym sensie stać się podstawą wprowadzania zmian. Menedżerowie „od zawsze” oczekiwali technologii, które

przysporzą ich firmom nowych klientów, dadzą się łatwo zaimplementować i zintegrować z istniejącą już infrastrukturą, a przy tym pozwolą zminimalizować koszty i zoptymalizować wykorzystanie zasobów. Obecne systemy ERP wychodzą poza dotychczasowe ramy. To już nie tyle systemy zarządzania zasobami organizacji, co bardzo często systemy zarządzania całościowym potencjałem biznesowym przedsiębiorstwa.

Współczesna organizacja zмага się z wieloma wyzwaniami, musi m.in. redukować koszty i spełniać wymagania interesariuszy, często nie dysponując odpowiednią i aktualną informacją niezbędną w procesach decyzyjnych. Cyfrowa transformacja biznesu nie jest możliwa bez danych, których analizę usprawniają nowe technologie i rozwiązania. To one stanowią fundament pozwalający danej organizacji na czynienie kluczowych postanowień, umożliwiających sprawne działania.

Równie istotnym elementem pracy współczesnej organizacji jest nastawienie na klienta. Aplikacje, dzięki którym uzyskuje się analizy opisowe, pomagają zrozumieć historię danych; analizy predykcyjne wskazują natomiast zachowania i sygnalizują wzorce i możliwości. Z kolei analizy preskryptywne dostarczają organizacjom rekomendacji dotyczących optymalnych działań potrzebnych do osiągnięcia celów biznesowych związanych z obsługą klienta, zyskami i efektywnością operacyjną.

Rozwój usług środowiska chmury obliczeniowej i aplikacji mobilnych zmienił sposób nawiązywania i prowadzenia interakcji między klientami i firmami, pozwolił na stworzenie nowych modeli biznesowych i strumieni przychodów. Potencjał, jaki niesie ze sobą zastosowanie takich narzędzi, jak sztuczna inteligencja, robotyka i Internet Rzeczy wpływa zarówno na firmy i użytkowników, jak również narzędzia umożliwiające nawiązywanie i podtrzymywanie relacji między biznesowymi partnerami.

Do tej pory to firmy usługowe chętniej i szybciej sięgały po nowe rozwiązania. Przedsiębiorstwa produkcyjne, posiadające stabilne rozwiązania w zakresie systemów zarządzania, reagowały wolniej na modernizację. Niemniej firmy te generują ogromną ilość danych w porównaniu do innych sektorów gospodarki – cyfrowa transformacja biznesu skutkuje zatem innowacjami, które zmieniają samą naturę tego obszaru poprzez modyfikacje i rozwój łańcucha dostaw oraz aplikacji biznesowych niezbędnych dla wdrażania w gospodarce idei Przemysłu 4.0.

Niniejsza monografia jest drugą z serii *Cyfryzacja* wydawanej przez Wydawnictwo Uniwersytetu Łódzkiego. Pierwsza, zatytułowana *Zarzą-*

*dzanie danymi w organizacji*, pod red. Beaty Gontar, stanowi kompendium wiedzy z zakresu architektury danych, hurtowni, bezpieczeństwa, wizualizacji i nowych technologii, takich jak Internet Rzeczy i chmura obliczeniowa. Celem niniejszej pracy jest przedstawienie kierunków rozwoju aplikacji biznesowych i technologii mających wpływ na rozwój organizacji. W rozdziale pierwszym uzupełnione zostały informacje o technologiach opisanych w pierwszej pozycji cyklu, scharakteryzowano także nowe, jak np. Blockchain czy rozszerzona rzeczywistość – z zaznaczeniem ich oddziaływania na rozwój aplikacji oraz możliwości, jakie daje zastosowanie nowych koncepcji w biznesie. W drugim rozdziale zaprezentowano historię rozwoju aplikacji wykorzystywanych w organizacjach, wskazując na system klasy ERP jako podstawowy system wspierający zarządzanie organizacją. Poszerzenie funkcjonalności systemów biznesowych o bieżącą dostępność wieloaspektowych analiz i włączenie opcji wsparcia w procesie decyzyjnym to powody, dla których w rozdziale trzecim przedstawiono typy decyzji, jakie podejmowane są w organizacjach oraz narzędzia, które je wspomagają – w tym rozwiązania z grupy sztucznej inteligencji.



# Rozdział 1

## Zastosowanie nowych technologii w biznesie

### 1.1. Wprowadzenie

Szybki rozwój nowych technologii wymusza zmiany w organizacjach i ich relacjach z partnerami. Ma on ogromny wpływ na to, jak pracują oraz jakie rozwiązania stosują współcześnie organizacje. Wśród technologii tych można wyróżnić: chmurę obliczeniową (*Cloud Computing* – CC), Internet Rzeczy (*Internet of Things* – IoT), przemysłowy Internet Rzeczy (*Industrial Internet of Things* – IIoT), analitykę danych, sztuczną inteligencję (*Artificial Intelligence* – AI), druk przestrzenny (*3D Printing*), cyfrowe odpowiedniki (*Digital Twin* – DT), wirtualną i rozszerzoną rzeczywistość (*Virtual Reality/Augmented Reality* – VR/AR), Big Data, automatyzację i robotykę, roboty współpracujące (*Collaborative Robots* – CR), Blockchain, a także specyficzne dedykowane oprogramowanie obniżające koszty prototypowania produktów i urynkowienie ich.

Rozwiązania takie, jak chmura obliczeniowa, Internet Rzeczy czy Blockchain pojawiły się stosunkowo niedawno i już zmieniły paradygmaty dotyczące sposobu korzystania z aplikacji i bezpieczeństwa w sieci komputerowej. W niniejszym rozdziale omówiona zostanie głównie technologia chmury obliczeniowej oraz Blockchain.

## 1.2. Chmura obliczeniowa i rozwój technologii mobilnych

Pojęcie chmury obliczeniowej (*Cloud Computing* – CC) zaczęło funkcjonować w przestrzeni społecznej pod koniec XX wieku. Wraz ze wzrostem możliwości korzystania z Internetu i popularyzacją usług mobilnych składowanie i przetwarzanie danych w chmurze stało się dla większości ludzi czymś zupełnie naturalnym, a często wręcz niezauważalnym. Na popularyzację rozwoju tego modelu współpracy z Internetem miała również wpływ konieczność sprostania problemowi obsłużenia i zagwarantowania ciągłości pracy kolejnym masowym usługom, oferowanym np. przez portale społecznościowe i serwisy udostępniające media. Olbrzymi wzrost transferu danych w Internecie w XXI wieku i problem ich składowania przyczyniły się do dynamicznego rozwoju branż usług informatycznych wokół chmury obliczeniowej.

Istnieje wiele definicji *Cloud Computing*. Według NIST (National Institute of Standards and Technology) można ją określić jako „sposób dostępu poprzez sieć komputerową do współdzielonych i łatwo konfigurowalnych zasobów obliczeniowych (sieci, serwerów, magazynów danych, aplikacji i usług), które na żądanie, dynamicznie mogą być przydzielane i zwalniane, przy równoczesnym minimalnym zaangażowaniu serwisów technicznych” [Mell i Grance, 2011]. Definicja proponowana przez Uniwersytet Kalifornijski w Berkeley kładzie nacisk na skalę i dostępność zasobów: „chmura obliczeniowa powinna dawać poczucie nieskończoności zasobów obliczeniowych, zdejmować z użytkowników końcowych konieczność angażowania w ich konfigurację i umożliwiać płacenie tylko za wykorzystane zasoby” [Armbrust, Fox, Griffith, Joseph i Katz, 2009]. Definicje proponowane przez profesora Rajkumara Buyyę podkreślają aspekty wirtualizacji i rozproszenia infrastruktury: „chmura to równoległy i rozproszony system komputerowy składający się z zespołu połączonych systemów wirtualnych, które są dynamicznie tworzone i udostępniane jako jeden lub kilka ujednoliconych zasobów komputerowych, bazując na umowie SLA (*Service-level agreement*) zawartej pomiędzy usługodawcą a klientem” [Buyya i inni, 2009]. Z kolei McKinsey & Co. kładzie nacisk na cechy istotne dla przedsiębiorstw – definiuje chmurę jako „usługi oparte o sprzęt komputerowy, oferujące możliwości oblicze-

niowe, sieciowe i magazynowe, gdzie zarządzanie platformą sprzętową jest ukryte przed usługobiorcą, który może traktować koszty infrastruktury jako zmienne koszty operacyjne, a same zasoby infrastruktury charakteryzują się wysoką elastycznością” [McKinsey & Co., 2009].

Wymagania stawiane przed chmurą obliczeniową mają być realizowane dzięki funkcjonalnościom zdefiniowanym przez NIST. Główne wymagania dla chmury obliczeniowej to:

- wykorzystanie pul zasobów – oferowane przez usługodawcę zasoby komputerowe są skonfigurowane w pulach dla wspierania współdzielonego środowiska zasobów przez wielu klientów (*multi-tenant*<sup>1</sup>), składającego się z zasobów fizycznych i wirtualnych konfigurowanych dynamicznie w zależności od potrzeb; konfiguracja zasobów przez klienta odbywa się na wysokim poziomie abstrakcji i usługobiorca właściwie nie może określić, gdzie fizycznie znajdują się jego zasoby (mamy do czynienia z poczuciem niezależności lokalizacji danych); na pulę zasobów mogą się składać takie elementy, jak pamięć, przestrzeń magazynowa, czas obliczeniowy procesorów, przepustowość sieci;
- automatyzacja większości procesów odpowiedzialnych za alokację zasobów i wdrażanie nowych aplikacji – klient powinien być w stanie jednostronnie przydzielić wymagane możliwości obliczeniowe, takie jak czas serwera czy wielkość magazynów danych, bez konieczności interwencji pomocy technicznej;
- duża elastyczność i szybka reakcja na zmieniające się wymagania systemów – zasoby i usługi mogą być przydzielane w sposób elastyczny, często również na zasadzie automatycznej, dla zagwarantowania szybkiego skalowania usług w zależności od potrzeb klientów; usługobiorca powinien mieć możliwość przydzielenia zasobów w każdej ilości i w każdym momencie;
- szeroki dostęp sieciowy – zasoby chmury są dostępne przez sieć za pomocą standardowych mechanizmów, które promują wykorzystanie heterogenicznych platform cienkich (*thin*) i grubych (*thick*) klientów, np. telefonów, tabletów, laptopów, stacji roboczych;

---

<sup>1</sup> Wykorzystanie tych samych zasobów lub aplikacji przez wielu klientów należących do tej samej firmy lub różnych organizacji. Możliwość zastosowania rozwiązań *multi-tenant* w przypadku chmury obliczeniowej oznacza konieczność wdrożenia takich automatycznych mechanizmów gwarantujących segmentację, izolację, zarządzanie, odpowiednie poziomy usług i model zliczania oraz pobierania opłat, aby zapewnić bezpieczną i stosownie odseparowaną pracę podmiotów korzystających z zasobów chmury obliczeniowej.



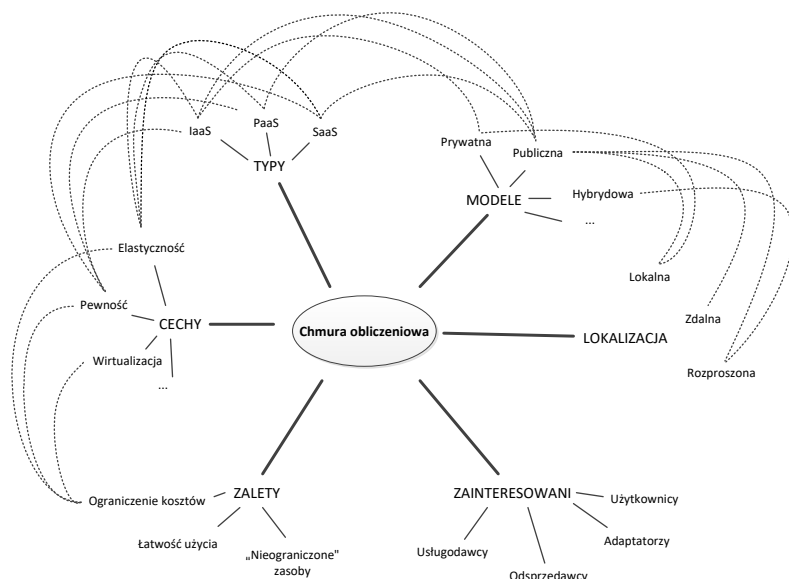
- dostosowanie parametrów pracujących systemów w zależności od ich potrzeb w ramach możliwości dostępnych komponentów sprzętowych;
- monitorowanie – system zarządzający chmurą obliczeniową kontroluje i optymalizuje wykorzystanie zasobów przez udostępnienie miar wydajnościowych dla różnego rodzaju usług i zasobów wykorzystywanych w chmurze; praca systemu może być monitorowana, a raportowanie na różnych poziomach szczegółowości udostępnione zarówno usługodawcy, jak i klientowi.

Z powyższych definicji wynikają konkretne korzyści – podstawowymi cechami chmury obliczeniowej, które mogą być atrakcyjne dla klienta, są bowiem:

- 1) wysoka elastyczność zasobów,
- 2) bezobsługowość infrastruktury,
- 3) wysoka dostępność wynikająca ze skali,
- 4) model płacenia tylko za wykorzystane zasoby.

Na rysunku 1.1 przedstawiono główne wymiary formujące systemy chmurowe, które pozwalają na zdefiniowanie podstawowych cech chmury i jej właściwości – wynikających z definicji oraz sposobu wykorzystania. Każde wdrożenie chmury obliczeniowej powinno charakteryzować się elementami zawartymi w jej eksplikacji oraz przynosić co najmniej jedną zaletę w stosunku do tradycyjnych modeli przetwarzania danych. Typ, model oraz lokalizacja pozwalają na dokładniejsze zdefiniowanie sposobu i celu implementacji, a zainteresowani określają interesariuszy rozwiązania.

Z punktu widzenia podmiotu korzystającego z chmury obliczeniowej jej zasoby mogą wydawać się nieograniczone. Zwiększenie lub zmniejszenie zapotrzebowania na moc obliczeniową, przepustowość sieci, wielkość magazynów danych mogą być – w zależności od umowy SLA – automatycznie dostosowywane do potrzeby wykonywanych zadań. Oznacza to, że można dopasować parametry pracy aplikacji w zależności od wymagań – w okresie, kiedy nie jest to konieczne, nie płać za wolne zasoby, a w sytuacji zwiększenia zapotrzebowania na moc obliczeniową wykonać niezbędne operacje z pełną wydajnością określonych w SLA zasobów. Uzyskuje się również możliwość szczegółowego raportowania wykorzystanych zasobów, co pozwala na lepsze dobranie kontraktowanych parametrów i optymalizację pracy systemu informatycznego.



**Rysunek 1.1.** Widok głównych wymiarów formujących systemy chmurowe

**Źródło:** opracowanie własne na podstawie [Kostantos, Kapsalis, Kyriazis i Cunha, 2013].

Skala zasobów sprzętowych odpowiedzialnych za pracę chmury obliczeniowej, automatyzacja procesów związanych z zarządzaniem i zaawansowane mechanizmy monitorowania jej funkcjonowania również powinny zagwarantować wysoką dostępność usług w chmurze obliczeniowej. Na ich pracę nie mogą mieć wpływu lokalne awarie zasilania czy uszkodzenia poszczególnych serwerów lub łącz internetowych od jednego z dostawców. Oczywiście w historii firm oferujących usługi w chmurze można odnotować większe lub mniejsze awarie, które potrafiły doprowadzić nawet do długotrwałych utrudnień dla klientów w dostępie do ich danych. Tego typu sytuacje spotkały gigantów rynku – zarówno Microsoft [Whittaker, 2013], jak i Amazon [Perez, 2014].

Chmura obliczeniowa jest zazwyczaj kojarzona z kilkoma dużymi podmiotami i popularnymi usługami internetowymi. Za teoretycznie nieskończone możliwości dużych chmur obliczeniowych są odpowiedzialne wielkie centra danych i setki tysięcy serwerów. W 2013 roku Google posiadał trzynaście centrów danych utrzymujących 900 tysięcy serwerów, które używały 260 milionów watów (taka moc pozwala zaopatrzyć w energię elektryczną 200 tysięcy domów). Amazon wraz ze swoimi

siedmioma centrami i 450 tysiącami serwerów jest na drugim miejscu. Wiele kolejnych firm, jak Microsoft, Facebook, IBM, posiada własne centra i setki tysięcy serwerów.

Wraz z rozwojem sprzętu i oprogramowania chmura obliczeniowa nie musi być jednak rozwiązaniem kojarzonym tylko z taką skalą usług. W zależności od modelu samej chmury i modelu jej realizacji potencjalny klient otrzymuje szerokie możliwości swojego zaistnienia w chmurze. Można wyodrębnić cztery podstawowe modele chmury obliczeniowej:

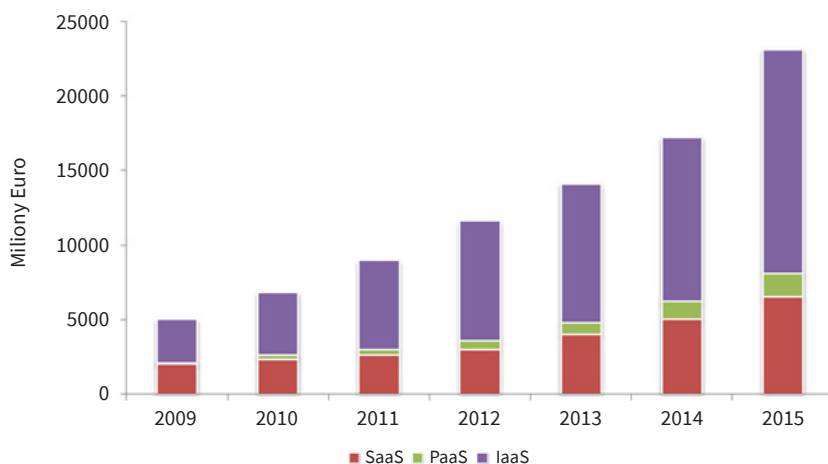
- 1) chmurę prywatną – jej infrastruktura jest skonfigurowana dla wyłącznego wykorzystania przez pojedynczą organizację, mogącą jednak zawierać wielu konsumentów chmury (np. jednostki biznesowe); chmura prywatna może być zarządzana, utrzymywana i być własnością zarówno organizacji, jak i podmiotu trzeciego lub ich kombinacji, a także zlokalizowana w siedzibie organizacji lub/i poza nią;
- 2) chmurę publiczną – jej infrastruktura znajduje się w otwartym publicznym dostępie; może być zarządzana i utrzymywana, a także stanowić własność organizacji biznesowej, uczelni, organizacji publicznej lub ich kombinacji; infrastruktura chmury jest ulokowana w siedzibie usługodawcy;
- 3) chmurę społecznościową – jej zasoby są przeznaczone do wyłącznego użytku przez konkretną społeczność klientów organizacji, które współdzielą jakiś cel (np. misję, wymagania bezpieczeństwa, względy jakościowe); może być zarządzana i utrzymywana, a także być własnością jednej lub kilku organizacji w społeczności, organizacji zewnętrznej lub ich kombinacji;
- 4) chmurę hybrydową – jej infrastruktura jest połączeniem dwóch odmiennych infrastruktur (prywatnej, publicznej lub społecznościowej), które zachowują unikalną osobowość, ale są ze sobą połączone przez standardową lub zastrzeżoną technologię umożliwiającą przenoszenie danych i aplikacji.

Sposób i rodzaj udostępniania zasobów przez chmurę zależy od jej modelu realizacji. NIST definiuje trzy podstawowe odmiany:

- 1) Oprogramowanie jako usługa (*Software as a Service* – SaaS) – w modelu tym dostawca oferuje konkretne oprogramowanie gotowe do wykorzystania; klient nie posiada informacji o platformie (infrastrukturze sprzętowej), interesuje go jedynie funkcjonalność oferowana przez poszczególne aplikacje; model ten jest najczęściej wyko-

- rzystywany przez osoby prywatne, może być również atrakcyjny dla małych firm, ponieważ pozwala uniknąć zakupu i utrzymania serwerów, oprogramowania systemowego i użytkowego, a usługi konfiguracyjne mogą ograniczyć się do jednorazowej, wstępnej konfiguracji;
- 2) Platforma jako usługa (*Platform as a Service* – PaaS) – w modelu tym dostawca oferuje pakiet aplikacji powiązanych ze sobą i wyposażonych w jednolity interfejs użytkownika; najczęściej dostęp realizowany jest przez przeglądarkę internetową, co oznacza możliwość korzystania z oprogramowania na wielu różnych platformach sprzętowych i systemowych; usługodawcy PaaS udostępniają oprogramowanie, które pozwala łatwiej programować i tworzyć skalowalne rozwiązania na warstwie infrastruktury;
  - 3) Infrastruktura jako usługa (*Infrastructure as a Service* – IaaS) – w tym modelu oferta obejmuje infrastrukturę sprzętową, oprogramowanie systemowe i użytkowe oraz usługi serwisowe; możliwe są jednak różne warianty, np. instalacja oprogramowania dostarczonego przez klienta czy też rozdzielenie usług administracyjnych między dostawcę a klienta; mimo że technologia pozwala na budowanie prywatnych infrastruktur chmury obliczeniowej, podstawą większości usług w niej dostępnych są publiczni dostawcy infrastruktury; „posiadany przez nich kapitał i zaplecze eksperckie pozwala budować i utrzymywać duże centra danych z globalnym dostępem oraz bronić własne zasoby przed nieustannym fizycznym, a zwłaszcza elektronicznym atakiem” [Armbrust, Fox, Griffith, Joseph i Katz, 2009].

W związku z dużymi nakładami finansowymi niezbędnymi do uruchomienia i rozwoju usług chmury obliczeniowej opartej na IaaS zajmuje ona pierwsze miejsce pod względem wartości rynkowej spośród wszystkich modeli chmury. Na drugim miejscu plasuje się warstwa SaaS – te usługi są bowiem najczęściej oferowane klientom końcowym. Obecny i prognozowany udział w rynku poszczególnych modeli został przedstawiony na rysunku 1.2. Największy wzrost w rynku chmury obliczeniowej według modelu wdrożenia należy do warstw IaaS i SaaS. Jest to związane m.in. z rosnącym ruchem sieciowym, zwiększaniem się liczby urządzeń mobilnych korzystających z Internetu i możliwością redukcji kosztów infrastruktury w chmurze obliczeniowej przez podmioty biznesowe. Warstwa PaaS to model najbardziej hermetyczny i zaawansowany – w naturalny sposób posiada więc najmniejszy udział w rynku.

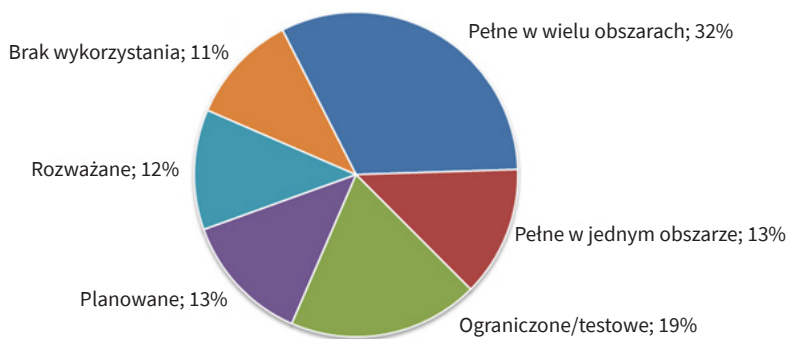


**Rysunek 1.2.** Udział poszczególnych modeli chmury obliczeniowej w rynku

**Źródło:** opracowanie własne na podstawie [Pierre Audoin Consultants].

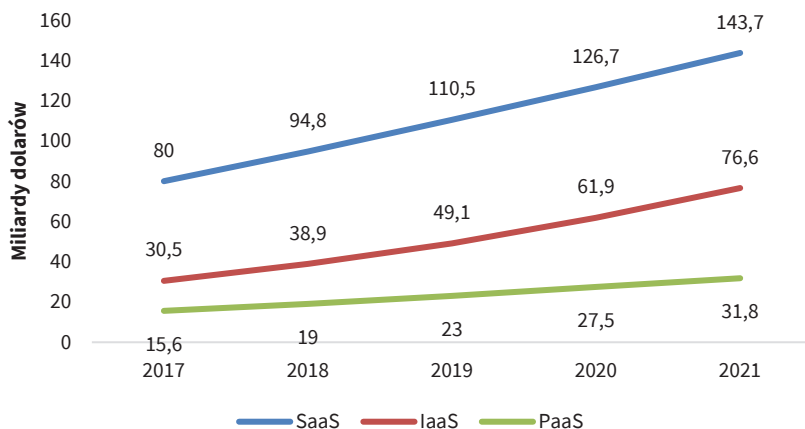
Wśród zainteresowanych chmurą obliczeniową, poza usługodawcami i użytkownikami, można również wymienić m.in. adaptatorów i odsprzedawców. Efektywne wykorzystanie możliwości technicznych chmury otwiera wiele nowych możliwości biznesowych przed podmiotami integrującymi i przenoszącymi do niej istniejące systemy informacyjne. Konkurencyjne cenowo i jakościowo rozwiązania bazujące na zasobach publicznych lub hybrydowych mogą stanowić ciekawą opcję dla organizacji poszukujących narzędzi do przetwarzania informacji. Na rysunku 1.3 zaprezentowano sposoby wykorzystania zasobów chmury obliczeniowej w krajach Unii Europejskiej w 2011 roku. Perspektywa pracy w chmurze obliczeniowej jest na tyle ciekawa, że wiele organizacji zastanawia się nad korzystaniem z jej zasobów lub używa ich w jednym lub w kilku obszarach swojej działalności.

Na przestrzeni lat dynamika rozwoju usług może się jednak zmieniać, a decyzje klientów dostosowywać do ewolucji samych technologii. Widać to na rysunku 1.4, gdzie przedstawiono aktualne prognozy przychodów z rynku chmury obliczeniowej. Atrakcyjna oferta zintegrowanych usług SaaS i ich dojrzałość skłoniły dużą część klientów do korzystania właśnie z tego modelu przetwarzania danych w chmurze obliczeniowej – co stoi w sprzeczności z analizami wykonywanymi jeszcze kilka lat temu.



**Rysunek 1.3.** Sposoby wykorzystania chmury obliczeniowej w krajach Unii Europejskiej w 2011 roku

**Źródło:** opracowanie własne na podstawie [European Commission, 2012].



**Rysunek 1.4.** Przychody z rynku chmury obliczeniowej

**Źródło:** opracowanie własne na podstawie [Vargas, 2019].

### 1.3. Rynek usług w chmurze obliczeniowej i rola użytkowników mobilnych

Usługi w chmurze obliczeniowej są szeroko dostępne już od kilku lat. W wielu przypadkach wiedza o tym, że aplikacja w jakikolwiek sposób bazuje na rozwiązaniach CC nie jest powszechna i klient często pośrednio lub bezpośrednio korzysta z funkcjonalności chmury, nie zdając sobie z tego

sprawy. Ze względu na zainteresowanie różnymi aspektami oferowanych usług (np. czy jest to usługa CC i charakteryzuje się w związku z tym pewnymi atrakcyjnymi cechami, ale również potencjalnym ryzykiem w niektórych obszarach), sposobem pracy z nimi oraz poszukiwaniem konkretnych funkcjonalności użytkowników można podzielić na trzy grupy:

- 1) klientów indywidualnych,
- 2) odbiorców biznesowych,
- 3) odbiorców rządowych.

Z jednej strony każda z tych grup przyczynia się do zwiększenia zapotrzebowania na różne modele i usługi CC, z drugiej zaś – sam rozwój i szeroka dostępność korzyści płynących z istnienia chmur obliczeniowych w naturalny sposób wpływa na ich rosnącą popularność. Szerokopasmowy dostęp do Internetu nie jest już czymś nadzwyczajnym – liczba wymagających go urządzeń (również mobilnych) ciągle rośnie.

Tabele 1.1 i 1.2 zawierają przykłady usług dostępnych w chmurze obliczeniowej, których adresatem może być zarówno klient indywidualny, jak i biznesowy. Funkcjonalności przez nie oferowane w dużej mierze odpowiadają za organizację pracy i zwiększenie funkcjonalności przedsiębiorstwa przez wykorzystanie nowoczesnych środków komunikacyjnych.

**Tabela 1.1.** Przykłady usług w chmurze obliczeniowej oferowanych klientowi indywidualnemu

Nazwa usługi	Opis usługi według strony usługodawcy
Amazon Cloud Drive	„Amazon Cloud Drive to osobisty dysk twardy umiejscowiony w chmurze. Można na nim przechowywać muzykę, filmy, zdjęcia i dokumenty. Do zarządzania dyskiem wystarczy przeglądarka internetowa. Oprogramowanie klienckie dostępne na wszystkie popularne systemy operacyjne pozwala automatycznie archiwizować zasoby lub odciażać pamięć telefonów poprzez przenoszenie ich do chmury” ( <a href="http://www.amazon.com/clouddrive/learn-more">www.amazon.com/clouddrive/learn-more</a> ).
Apple iCloud	„Zestaw usług <i>online</i> oraz oprogramowania firmy Apple służące do synchronizowania usług sieciowych dla telefonów iPhone, tabletów iPad, odtwarzacza iPod touch oraz komputerów Mac. Można go obsługiwać poprzez dwa interfejsy: przeglądarkowy oraz oprogramowanie zainstalowane na komputerze. Urządzenia Apple z systemem iOS w wersji minimum 5.0 mają wbudowany dostęp do usługi iCloud i mogą korzystać z synchronizowania usług poczty, kalendarzy, kontaktów, przypomnień, zakładek, notatek, dokumentów i danych, strumieni zdjęć” ( <a href="http://www.icloud.com">www.icloud.com</a> ).
Dropbox	Usługa świadczona przez Dropbox Inc., polegająca na udostępnieniu przestrzeni dyskowej na serwerach tej firmy. Wysyłanie, przeglądanie i pobieranie danych jest możliwe poprzez zwykłą przeglądarkę lub dedykowaną aplikację zainstalowaną na komputerze ( <a href="http://www.dropbox.com">www.dropbox.com</a> ).

Nazwa usługi	Opis usługi według strony usługodawcy
Evernote	Evernote pozwala udostępniać notatki i zdjęcia oraz organizować i wyszukiwać dane za pomocą wielu platform. Użytkownik może tworzyć notatki, zapisywać strony internetowe, a także zdjęcia z telefonu komórkowego oraz tworzyć listy zadań i nagrywać wiadomości głosowe. Wszystkie dane są zsynchronizowane z serwisem Evernote i dostępne dla użytkowników Windows, Mac, Web oraz urządzeń przenośnych ( <a href="http://www.evernote.com">www.evernote.com</a> ).
Google Apps	Google Apps pozwala użytkownikom na utworzenie skrzynki pocztowej (Gmail), strony internetowej i stron wiki (Google Sites), organizowanie pracy dzięki kalendarzom i udostępnianie wydarzeń znajomym (Google Calendar), współdzielenie dokumentów, prezentacji i arkuszy kalkulacyjnych (Google Docs) ( <a href="http://www.google.com/apps/intl/en/business/index.html">www.google.com/apps/intl/en/business/index.html</a> ).
OneDrive	Wirtualny dysk autorstwa Microsoft, część serwisu Windows Live ( <a href="http://explore.live.com/skydrive">http://explore.live.com/skydrive</a> ).
SpiderOak	„SpiderOak udostępnia proste, bezpieczne, skonsolidowane i darmowe rozwiązanie <i>backupu</i> sieciowego, synchronizacji, dostępu i składowania. [...] Dzięki elastycznemu zaprojektowaniu usługi dajemy ci możliwość pracy z danymi z dowolnego systemu operacyjnego (Mac, Windows, Linux) i urządzenia (zewnętrzne dyski, dyski sieciowe, pendrive) z wykorzystaniem tego samego konta” ( <a href="http://www.spideroak.com">www.spideroak.com</a> ).
Spotify	Powstały w październiku 2008 roku w Szwecji muzyczny serwis oferujący swoim użytkownikom dostęp do milionów utworów w strumieniu za pomocą telefonów, tabletów i komputerów. Pozwala on na słuchanie na żądanie muzyki z wielu wytwórni płytowych, zarówno tych wielkich, jak i tych bardziej niezależnych. Serwis prezentuje społecznościowe podejście do słuchania i odkrywania nowej muzyki np. listy znajomych, rekomendacje itp. ( <a href="http://www.spotify.com">www.spotify.com</a> ).
Waze	Waze to darmowa, społecznościowa aplikacja nawigacyjna i informująca o ruchu. Kiedy instalujesz Waze, nie tylko otrzymujesz darmową nawigację, ale stajesz się częścią lokalnej społeczności kierowców łączącej siły w ominięciu korków, oszczędzaniu czasu i poprawieniu warunków dojazdu ( <a href="http://www.waze.com">www.waze.com</a> ).

**Źródło:** opracowanie własne.

**Tabela 1.2.** Przykłady usług w chmurze obliczeniowej oferowanych klientowi biznesowemu

Nazwa usługi	Opis usługi według strony usługodawcy
Amazon EC2/S3/EBS/Glacier	Amazon Elastic Compute Cloud (Amazon EC2) to usługa webowa, udostępniająca w chmurze skalowalne zasoby obliczeniowe. Programiści i administratorzy poprzez interfejs w przeglądarce internetowej mogą tworzyć i zarządzać instancjami maszyn wirtualnych, dostępnymi w różnych konfiguracjach systemów operacyjnych i dostępnych zasobów. Przy zakupie usługi można wybrać wiele scenariuszy rozliczania wykorzystania usługi i dobrać model dzierżawy maszyn wirtualnych najlepiej do swoich potrzeb ( <a href="http://aws.amazon.com/ec2">aws.amazon.com/ec2</a> ).



Tabela 1.2 (cd.)

Nazwa usługi	Opis usługi według strony usługodawcy
Amazon EC2/ S3/EBS/Glacier	<p>Amazon S3 to infrastruktura magazynowa dla Internetu, oparta na obiekowym (dokumentów, filmów, muzyki, plików graficznych) modelu przechowywania danych (<a href="http://aws.amazon.com/s3">aws.amazon.com/s3</a>). Zasoby są przechowywane i zarządzane jako obiekty, a nie partycje czy wolumeny dyskowe. Pozwala to na zoptymalizowanie pod kątem dostępu sieciowego procesów replikacji, wersjonowania i zlokalizowanego pobierania danych.</p> <p>Amazon EBS pozwala na tworzenie wolumenów blokowych (tradycyjne dyski, w których można zainicjować systemy plików) w chmurze i podłączenie ich do maszyn wirtualnych utworzonych w usłudze EC2. Klient ma do dyspozycji trzy rodzaje wolumenów, charakteryzujące się różnymi parametrami wydajnościowymi i ceną, dzięki czemu może wybrać rozwiązanie najbardziej zoptymalizowane do swoich potrzeb i zadań obliczeniowych wykonywanych w EC2.</p> <p>Amazon Glacier to dedykowane rozwiązanie służące do archiwizacji danych i wykonywania kopii zapasowych. Usługodawca oferuje rozliczanie na podstawie wolumenu składowanych danych, wielkości transferu i liczby odwołań do zasobów. Dzięki takiemu rozliczaniu i niskiej cenie za przechowywanie danych (0,01 USD za 1 GB/miesiąc), klient może otrzymać nieograniczone i bezpieczne miejsce do przechowywania elektronicznych archiwów.</p>
Concur	<p>Zestaw oprogramowania służącego do optymalizacji pracy firmy i wspierania procesów zarządzania organizacją. W skład pakietu wchodzi m.in.:</p> <ul style="list-style-type: none"> <li>• <i>Expense</i> – zarządzanie wydatkami;</li> <li>• <i>Travel Booking</i> – pomoc przy planowaniu podróży, automatyzacji rezerwacji biletów, pobytu, restauracji;</li> <li>• <i>Invoice</i> – rejestracja faktur i zarządzanie płatnościami w firmie; oprogramowanie posiada funkcje automatyzacji płatności i możliwość projektowania własnych przepływów zdarzeń, które pozwalają na tworzenie zaawansowanych procesów obiegu dokumentów finansowych w organizacji;</li> <li>• <i>Insight</i> – komponent analityczny, którego celem jest przedstawianie w przystępnej formie (np. wykresy) danych zebranych przez pozostałe moduły;</li> <li>• <i>Tripit</i> – oprogramowanie służące organizowaniu informacji dotyczących podróży.</li> </ul> <p>Poszczególne moduły są bardzo wyspecjalizowane, ale mogą ze sobą współpracować, a przez to zwiększać funkcjonalność całości, np. łącząc <i>Travel</i> i <i>Expense</i>, otrzymujemy wgląd w zestawienie kosztów podróży pracowników. W zależności od wybranego planu, organizacja może dobrać odpowiednie poziomy funkcjonalności i komponenty do swoich potrzeb (<a href="http://www.concur.com">www.concur.com</a>).</p>
Basecamp	<p>„Basecamp to najbardziej popularna na świecie aplikacja web dla przechowywania, koordynowania i zarządzania projektami, zadaniami, dyskusjami i decyzjami przedsiębiorstwa” (<a href="http://basecamp.com">basecamp.com</a>). Producent oprogramowania oferuje swoje rozwiązanie w modelu <i>pas-as-you-go</i>, co w tym przypadku oznacza uzależnienie kosztów usługi od liczby aktywnych projektów.</p>

Nazwa usługi	Opis usługi według strony usługodawcy
SAP JAM	<p>Usługa, której celem jest zorganizowanie i poprawienie mechanizmów komunikacji w organizacji przez budowanie sieci społecznościowej wewnątrz firmy. Oprogramowanie implementuje kilka podstawowych funkcji:</p> <ul style="list-style-type: none"> <li>• sieć społecznościowa (komunikacja, współdzielenie plików i zasobów medialnych, wyszukiwanie danych i pracowników, raportowanie aktywności zainteresowań, dostęp z urządzeń mobilnych);</li> <li>• zorganizowana współpraca i rozwiązywanie problemów (analiza za i przeciw, dynamiczny konstruktor agendy, rankingi, osie czasu, zatwierdzanie decyzji, pomysły);</li> <li>• współpraca z zewnętrznymi partnerami (zorganizowane grupy do pracy z partnerami, klientami, dostawcami);</li> <li>• integracja procesów biznesowych i sieci społecznościowych (integracja z innymi komponentami SAP, np. BizX dla <i>social-learningu</i>, integracja z innymi portalami społecznościowymi i usługami zewnętrznymi, jak Gmail, Twitter, YouTube).</li> </ul>
Box	<p>„Box oferuje bezpieczne i skalowalne współdzielenie zasobów. Box pozwala na składowanie wszystkich zasobów <i>online</i>, dzięki czemu możesz mieć do nich dostęp, zarządzać i udostępniać z dowolnego miejsca. Box zapewnia wszystko, co jest potrzebne do pracy na zasobach umieszczonych w sieci [...]: upewnienie się, że każdy jest na tej samej stronie i tej samej wersji dokumentu [...], wymiana informacji zwrotnych w jednym miejscu, zarówno w przypadku szybkiego komentarza, jak i głębokiej dyskusji [...], zagwarantowanie ciągłości projektów poprzez zarządzanie zadaniami na poziomie pliku. Uzyskiwanie szczegółowego widoku w czasie rzeczywistym co się dzieje z zasobami <i>online</i>” (www.box.com).</p>
Huddle	<p>„Bezpieczna przestrzeń robocza do pracy w chmurze” (www.huddle.com). Zestaw funkcjonalności dostępny przez przeglądarki internetowe, udostępniający możliwość zarządzania plikami, pracę grupową na dokumentach, zarządzanie projektami i organizację czasu pracy.</p>
Microsoft Office 365	<p>Zestaw usługi i oprogramowania służącego usprawnieniu mechanizmów komunikacji w organizacji i zarządzaniu dokumentami i danymi w chmurze (www.office365.com). Komplet usług obejmuje obsługę zaawansowanej poczty elektronicznej, platformę do współdzielenia dokumentów i zarządzania projektami Sharepoint, internetowy dysk pozwalający na synchronizację danych z komputerami oraz komunikator umożliwiający obsługę telekonferencji. Office 365 jest silnie powiązany z pakietem biurowym MS Office i oferuje duży poziom integracji usług sieciowych w tym z pakietem oprogramowania biurowego, ale pozwala też na pracę nad dokumentami i arkuszami kalkulacyjnymi tylko z wykorzystaniem przeglądarki internetowej.</p>
Salesforce.com	<p>Jedno z najbardziej popularnych rozwiązań typu <i>Customer Relationship Management</i> (CRM) na świecie (Salesforce.com), które jest oferowane w wielu wersjach (Sprzedaż, Usługi, Marketing, Społeczność, Analiza) i pakietach funkcjonalnych dla automatyzacji i wspierania procesów zarządzania relacjami z klientem. Dodatkowo Salesforce oferuje własne platformy PaaS do tworzenia aplikacji biznesowych w chmurze i umieszczania ich w sklepie AppExchange.</p>

**Źródło:** opracowanie własne.

Praktycznie każda z opisywanych ofert stawia na użytkownika mobilnego lub korzystającego z wielu urządzeń. Taka charakterystyka pracy zwiększa potrzebę niezawodności usług i jest odpowiedzialna za znaczny przyrost ruchu sieciowego, który należy obsłużyć. W zestawieniu nie ujęto ofert portali społecznościowych – one również w dużym stopniu ją pogłębiają. Z badań przeprowadzonych przez Pew Research Internet Project [Pew Research Center, 2019] wynika, że ponad 60% użytkowników telefonów korzysta na nich z aplikacji obsługujących media społecznościowe, a 28% robi to codziennie.

## 1.4. Blockchain

Pierwszy pomysł utrzymania łańcucha bloków w celu skonstruowania odpornego na manipulacje znacznika czasu dokumentów cyfrowych powstał w 1991 roku. Koncepcja ta nie została jednak spopularyzowana, dopóki autor o pseudonimie Satoshi Nakamoto – którego tożsamość wciąż nie jest znana – zaprezentował łańcuchy bloków i ich prawdziwe zastosowanie w zdecentralizowanych sieciach w 2008 roku w artykule zatytułowanym *Bitcoin: Peer-to-Peer Electronic Cash System* [Nakamoto, 2019].

Później, w 2009 roku, Satoshi Nakamoto stworzył referencyjną implementację technologii Blockchain o nazwie Bitcoin. Było to pierwsze – i wciąż pozostaje najpopularniejsze – wdrożenie elektronicznego systemu gotówki opartego na Blockchain.

Chociaż Blockchain ma wiele definicji, najprościej można go opisać jako strukturę danych w postaci łańcucha bloków, które zostają połączone w celu utworzenia zbioru rekordów, zwanego rejestrem głównym, z kryptografią będącą kluczowym składnikiem całego procesu. Technologia ta nie ma zadeklarowanego mechanizmu przechowywania – zamiast tego posiada zestaw protokołów regulujących sposób powiększania łańcucha danych. Blockchain może być więc przechowywany w plikach płaskich lub w bazie danych.

Chociaż eksperci Blockchain podsuwają różne definicje tej technologii, wszystkie one mają wspólny człon – rejestr główny [Raj, 2019]. Twórcy popularnej implementacji Blockchain 2.0 – Hyperledger – proponują następującą eksplikację: „Blockchain to rozproszony rejestr *peer-to-peer* stworzony w drodze konsensusu, w połączeniu z systemem »inteligent-

nych umów« i innymi technologiami wspomagającymi». Antony Lewis – dyrektor w firmie R3, która wymyśliła oparte na łańcuchu bloków rozwiązanie Coda dla biznesu – charakteryzuje Blockchain w bardziej obrazowy sposób: „Blockchain to specyficzna forma lub podzbiór technologii rozproszonego rejestru, która tworzy chronologiczny łańcuch bloków, stąd nazwa »łańcuch bloków«”. Jeszcze prostszą definicję, odwołującą się tylko do charakteru danych w łańcuchu, podaje Andreas Antonopoulos – ewangelista Bitcoin: „Struktura danych Blockchain to uporządkowana, wstecznie połączona lista bloków” [tamże].

Aby zrozumieć tę technologię, warto sięgnąć do genezy Bitcoina i przypomnieć sobie główny powód powstania kryptowaluty. Dla twórcy Bitcoina wielkim ograniczeniem handlu w Internecie było jego uzależnienie od instytucji finansowych, pełniących funkcję zaufanej trzeciej strony w procesie przetwarzania płatności elektronicznych. Podczas gdy system działał wystarczająco dobrze w przypadku większości transakcji, nadal cierpiał z powodu słabości modelu opartego na zaufaniu. Ze względu na potencjalne spory mediacyjne i niemożność ich unikania przez instytucje finansowe zagwarantowanie całkowicie nieodwracalnych transakcji okazuje się praktycznie niemożliwe. Koszty mediacji wpływają na wzrost kosztów transakcji oraz podniesienie progu ich wielkości minimalnej, co ogranicza opłacalność dokonywania drobnych operacji. Możliwość cofnięcia transakcji oznacza również, że wzrasta zapotrzebowanie za zaufanie i pewność. Podmioty transakcji – sprzedający i kupujący – muszą na siebie wzajemnie uważać, próbując zdobyć jak najwięcej informacji, a jednocześnie minimalizując ryzyko niepowodzenia transakcji. Nigdy jednak nie da się zagwarantować 100% pewności operacji, gdyż pewien odsetek oszustw w Internecie jest nieunikniony. Można wyeliminować te koszty i niepewności, stosując fizyczną walutę, ale w przypadku dokonywania płatności w Internecie zawsze występuje trzecia strona zaufana (np. portal aukcyjny).

Satoshi Nakamoto wskazał, że:

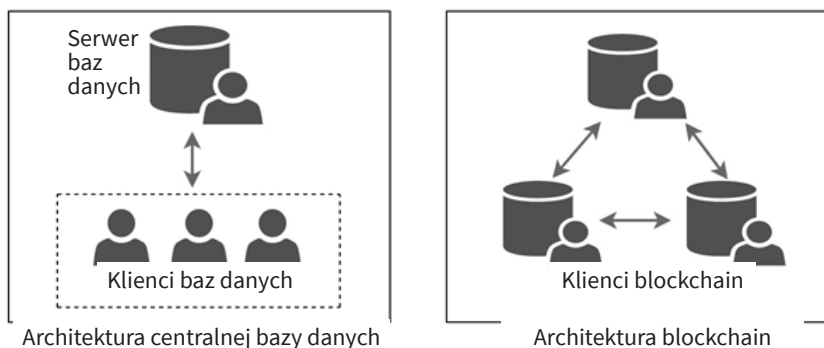
[...] potrzebny jest elektroniczny system płatności oparty na dowodzie kryptograficznym zamiast zaufania, pozwalający dwóm chętnym stronom na bezpośrednią transakcję bez potrzeby korzystania z zaufanej trzeciej strony. Transakcje, które są obliczeniowo niepraktyczne do odwrócenia, chroniłyby sprzedawców przed oszustwem, a rutynowe mechanizmy depozytowe można łatwo wdrożyć, aby chronić kupujących... Proponujemy rozwiązanie problemu podwójnych wydatków za pomocą rozproszonego serwera czasu opartego na sieci *peer-to-peer* do generowania

obliczeniowego dowodu chronologicznej kolejności transakcji. System jest bezpieczny, dopóki uczciwe węzły wspólnie kontrolują większą moc procesora niż jakkolwiek inna współpracująca grupa węzłów atakujących [Nakamoto, 2019].

### 1.4.1. Założenia projektowe łańcuchów bloków

Blockchain to metodologia przechowywania danych tylko do odczytu, z możliwością dołączania elementów na końcu łańcucha. Bloki można tworzyć i czytać tylko w rejestrze głównym – nie jest dozwolone ich aktualizowanie i usuwanie. W publicznym łańcuchu bloków nie ma kontroli dostępu, ponieważ jest on otwarty zarówno dla operacji odczytu, jak i dopisywania.

Relacyjne bazy danych działają zgodnie z modelem operacyjnym tworzenia, odczytu, aktualizacji i usuwania (*Create, Read, Update, Delete* – CRUD). W przeciwieństwie do Blockchain, każda baza danych ma administratora, który przypisuje kontrolę dostępu innym użytkownikom. Relacyjne bazy danych są w większości obsługiwane przez jeden kontrolujący wszystkie dane aplikacji podmiot, podczas gdy technologia Blockchain została zaprojektowana dla aplikacji zdecentralizowanych.



**Rysunek 1.5.** Różnica pomiędzy rozproszoną architekturą Blockchain a centralną architekturą baz danych

**Źródło:** opracowanie własne.

Jednym z głównych założeń Blockchain jest niezmienność (*immutability*) łańcucha bloków. Transakcja, która została dopisana do zbioru danych w łańcuchu i rozpowszechniona pomiędzy węzłami sieci, powinna być nieodwracalna i na zawsze pozostać w danych łańcucha. Block-

chain, najprościej mówiąc, to łańcuch połączonych ze sobą za pomocą kryptografii bloków. Wskaźniki skrótów kryptografii (*hash pointer*) są używane jako odniesienia do łączenia każdego bloku w łańcuchu bloków publicznego rejestru. Teoretycznie to rozwiązanie bezpieczne, ponieważ wykorzystane techniki kryptograficzne uniemożliwiają dowolne modyfikowanie łańcucha bloków i podmianę zapisanych w nim treści. Rejestr publiczny nie jest jednak nadzorowany przez żaden podmiot centralny i każdy użytkownik łańcucha bloków posiada jego całkowicie transparentną kopię – każdy węzeł *peer-to-peer* może w związku z tym wstawić nowy blok i stworzyć własną wersję łańcucha (w jednym punkcie w czasie wiele węzłów ma niekiedy wiele różnych wersji łańcucha). W końcu podmioty modyfikujące łańcuch mogą propagować blok do każdego węzła w sieci i próbować udowodnić, że ich łańcuch bloków jest prawidłowy. To pokazuje, że niezmienności nie da się osiągnąć tylko poprzez połączenie wszystkich bloków w bezpieczny sposób – musi jej towarzyszyć jakiś zdecentralizowany mechanizm ekonomiczny, dający każdemu węzłowi sieci uczciwe szanse głosowania w sprawie tworzenia bloków, a także utrudniający odtworzenie bloków po ich dodaniu.

Proponowane przez Satoshi Nakamoto rozwiązanie tego problemu jest jedynym powodem, dla którego Bitcoin można wdrożyć w zdecentralizowanym środowisku. Zastosowany w nim algorytm konsensusowy PoW (*Proof of Work*) był pierwszym – i nadal jest najbardziej znany – rozwiązaniem gwarantującym niezmienności łańcucha. Algorytm ten zabezpiecza rejestr publiczny nawet w sieci bez zaufania i gwarantuje wysoki stopień jego niezmienności. Z perspektywy kryptowaluty węzły wykonujące PoW nazywane są węzłami wydobywczymi (*mining nodes*) lub potocznie „koparkami”. Jak sama nazwa wskazuje, wydobywanie jest czynnością polegającą na wykuwaniu nowych bloków, które mają być dołączane do łańcucha bloków. Ilość pracy, jaką zajmuje kopanie, zapewnia, że Blockchain staje się niezmienny, a manipulowanie przy jakichkolwiek wcześniejszych transakcjach jest prawie niemożliwe.

Wynika to z faktu, że każdy węzeł – chcący manipulować danymi z przeszłości – powinien być w stanie zrekonstruować wszystkie bloki, korzystając z mechanizmu PoW i konkurując z wszystkimi innymi węzłami wydobywczymi. Okazuje się to niemalże niewykonalne, chyba że dany węzeł posiada większość mocy obliczeniowej sieci – w takiej sytuacji atakujący miałby szansę pokonać pozostałe węzły. Właśnie dlatego

algorytm konsensusu Bitcoin stosuje się szeroko w publicznych aplikacjach Blockchain w celu osiągnięcia wyższej niezmienności zapisów.

Jednak niezmienność to nie jedyna cecha wykazana przez technologię Blockchain. Ze względu na zdecentralizowany charakter łańcucha bloków każda transakcja w łańcuchu jest replikowana we wszystkich węzłach sieci. Replikacja informacji zapewnia większą jego niezawodność. Replikowane transakcje muszą być zatwierdzone przez każdy węzeł, aby osiągnąć konsensus. Dzięki temu operacje są publicznie widoczne, a wszystkie dane z łańcucha bloków stają się przejrzyste dla sieci. Przejrzystość zapewniana przez Blockchain może być dobrodziejstwem dla niektórych przypadków użycia i przekleństwem dla innych. Właśnie dlatego stworzono warianty Blockchain.

Pierwszy rodzaj tej technologii to publiczne łańcuchy bloków, będące projektami typu *open source*. Umożliwiają one bezpłatne dołączenie do sieci węzłów, rozwijanie projektu oraz wsparcie społeczności. Można również bez żadnych opłat wykorzystać taki łańcuch w prowadzeniu własnej działalności biznesowej.

Drugim rodzajem projektów Blockchain są prywatne łańcuchy bloków – czyli takie, w których jakaś organizacja sprawuje nadzór nad siecią węzłów i ogranicza do niej dostęp: mają go tylko użytkownicy i podmioty zatwierdzone przez organizację kontrolującą.

W każdej z tych kategorii znajdują się projekty z łańcuchami blokowymi ogólnego przeznaczenia, takimi jak Ethereum – <https://ethereum.org/pl/> [Tormen, 2019] – można zatem potraktować technologię Blockchain jako rozwiązanie ogólnego zastosowania i użyć konkretnej implementacji do zapewnienia podstawowej infrastruktury i elementów konstrukcyjnych do tworzenia własnej aplikacji. Korzystając z odpowiedniego oprogramowania klienckiego, możliwe okazuje się podłączenie do każdego rozwiązania Blockchain opartego na Ethereum.

Występuje też wiele sytuacji, w których technologia Blockchain jest wykorzystana w bardziej celowy sposób, a sam projekt kierowany do konkretnego sektora lub branży. Mamy wtedy do czynienia z łańcuchami bloków specyficznymi dla aplikacji wspierającej innowacyjne modele biznesowe i dostarczającymi rozwiązanie wykorzystujące zalety technologii Blockchain.

Zdarzają się też przypadki użycia łańcuchów bloków charakterystycznych dla konkretnej branży, czyli dopasowanych do aplikacji. Koncentru-

ją się one tylko na jednym sektorze lub rynku, gdzie budują innowacyjne modele biznesowe i dostarczają rozwiązań wykorzystujących zalety technologii Blockchain.

### **1.4.2. Budowa łańcucha bloków**

Jeśli potraktujemy łańcuch bloków jako strukturę danych, wówczas bloki będą zagregowanymi zestawami danych, których używa się do tworzenia łańcucha bloków. Tworzenie łańcucha bloków jest podobne do kreacji listy połączonych elementów, gdzie każdy składnik odwołuje się do następnego składnika w sekwencji. W przypadku łańcucha bloków każdy blok ma odniesienie do poprzedniego węzła, co tworzy połączenie aż do początkowego bloku (zwanego blokiem genezy) łańcucha. Jak wspomniano, łańcuch bloków można przechowywać w formacie pliku płaskiego lub bazy danych. Bitcoin używa LevelDB do przetrzymywania metadanych dotyczących wszystkich pobieranych na dysk bloków (rozmiar łańcucha bloków Bitcoin na początek roku 2020 to około 15 GB danych).

Każdy blok posiada wskaźnik będący identyfikatorem bloku. Są to zaszyfrowane wartości danych nagłówka bloku, które można uznać za unikalny identyfikator o stałym rozmiarze, reprezentujący każdy blok. Ponieważ wszystkie bloki są połączone tą wartością skrótu, każdy z nich będzie posiadał identyfikator poprzedniego bloku. Poprzedni blok jest określany z kolei jako blok macierzysty, a pozostałe mogą mieć tylko jednego rodzica.

Każdy blok można również zidentyfikować, określając jego wysokość w łańcuchu bloków. Wysokość łańcucha jest niczym innym, jak odległością bloku lub liczbą bloków od bloku genezy. Wysokość okazuje się ważnym atrybutem łańcucha bloków, ponieważ łatwiej odnieść się do bloku z wykorzystaniem prostej wartości, niż szukać bloku za pomocą skrótu. W przypadku łańcucha bloków wykorzystywanego przez Bitcoin algorytmem mieszającym odpowiedzialnym za generowanie unikalnego identyfikatora bloku jest SHA256.

### **1.4.3. Struktura bloku**

Wszystkie wersje Blockchain wykorzystują mechanizm łączenia ze sobą bloków w celu utworzenia niezmiennego rejestru – w zależności jednak od rozwiązania bloki mogą różnić się strukturą i rodzajem przechowywanych



danych. Przykładowo: łańcuchy bloków wspierające kontrolę dostępu będą inne niż łańcuchy pozbawione tej funkcjonalności. W przypadku najbardziej popularnej wersji Blockchain dla Bitcoin (niewspierającej kontroli dostępu) struktura bloku wygląda następująco: nagłówek bloku i część odpowiedzialna za przechowywanie transakcji są najważniejszymi częściami bloku, ponieważ odpowiadają za wartość skrótu, czyli unikalną tożsamość bloku. Rozmiar bloku to jednocześnie ogólny rozmiar całości. Nagłówek bloku zawiera wszystkie metadane bloku, a licznik transakcji numer transakcji dla danego bloku. Wreszcie wszystkie transakcje są przechowywane w bloku (rysunek 1.6).

Rozmiar bloku
Nagłówek bloku
Licznik transakcji
Transakcje

**Rysunek 1.6.** Struktura bloku łańcucha bloków Bitcoin

**Źródło:** opracowanie własne.

Jak wspomniano, Blockchain zaczyna się od początkowego bloku zwanego blokiem genezy. Jeśli cofniemy się wzdłuż łańcucha z dowolnego bloku, zawsze dotrzemy do bloku genezy, co potwierdzi, że cały łańcuch jest legalny i ważny. Blok genezy okazuje się często kodowany statycznie w łańcuchu publicznym lub łańcuchu niewspierającym kontroli dostępu. W przypadku łańcuchów z kontrolą dostępu blok genezy kreuje pierwszy uczestnik łańcucha – najczęściej twórca sieci węzłów.

Nagłówek bloku zawiera jego metadane i przechowuje informacje potrzebne do połączenia bloków w łańcuchu. Są to minimalne pola wymagane w łańcuchu bloków bez kontroli dostępu, czyli w takim jak Bitcoin, aby skutecznie utworzyć blok, który można dołączyć do niezmiennego łańcucha. Pole Hash poprzedniego bloku jest odniesieniem do ostatniego utworzonego. Korzeń Merkle to wartość drzewa skrótu Merkle – podsumowuje wszystkie transakcje w bloku. Algorytm konsensusowy PoW wykorzystuje znacznik czasu, cel trudności i wartość jednorazową do rozwiązania zagadki mieszającej (rysunek 1.7).

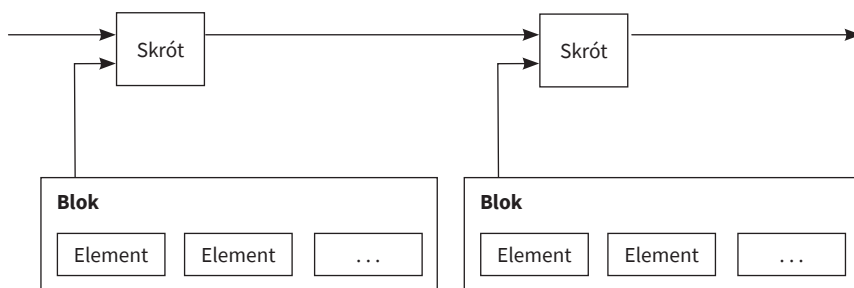
Wersja
Skrót poprzedniego bloku
Korzeń drzewa Merkle
Znacznik czasu
Cel trudności
Wartość jednorazowa

**Rysunek 1.7.** Struktura nagłówka

**Źródło:** opracowanie własne.

Ogólnie bloki są połączone w łańcuch bloków za pomocą referencji – tutaj jednak bloki zostały zestawione poprzez odwołanie się do wartości skrótu (identyfikatora) poprzedniego bloku. Każdy pełny węzeł w sieci Blockchain przechowuje kopię pełnego łańcucha bloków i dołącza nowy blok, jeżeli takowy jest anonsowany przez węzeł sieci *peer-to-peer*. Ze względu na zdecentralizowany charakter łańcucha bloków każdy węzeł musi zweryfikować blok przed podłączeniem go do lokalnego łańcucha.

Obliczona wartość skrótu każdego bloku to kombinacja skrótu poprzedniego bloku i danych bloku własnego. Powoduje to utworzenie praktycznie nierozdzielnej więzi pomiędzy sąsiednimi blokami (rysunek 1.8).



**Rysunek 1.8.** Łączenie bloków Bitcoin przy pomocy skrótów

**Źródło:** opracowanie własne.

Wszystkie elementy są mieszane, a blok jest oznaczony znacznikiem czasu, co sprawia, że następny blok będzie zawierał właśnie ten znacznik czasu, tworząc uporządkowany łańcuch bloków.

Każdy węzeł w sieci Blockchain wykonuje prosty proces dodawania nowych bloków do istniejącego lokalnego łańcucha. Ilekroć węzeł odbiera blok z sieci, sprawdza poprzedni skrót blokowy. Jeśli wartość skrótu jest zgodna z wartością skrótu ostatniego bloku w lokalnym łańcuchu bloków węzła, wówczas węzeł akceptuje ten blok i dołącza go do bieżącego łańcucha bloków. Dopóki będzie to najdłuższy znany łańcuch bloków, dopóty bloki są uważane za ważne przez wszystkie elementy równorzędne w sieci w łańcuchu bloków opartym na PoW.

#### **1.4.4. Zastosowania technologii Blockchain**

Technologia Blockchain w naturalny sposób jest najchętniej wykorzystywana w sytuacjach związanych z transakcjami bezgotówkowymi, bez potrzeby angażowania zaufanej trzeciej strony. Po cechy łańcuchów bloków coraz częściej mogą sięgać organizacje chcące korzystać z dobrodziejstw technologii, a nie bazować tylko na Bitcoinie. Staje się to możliwe dzięki powstaniu projektów Blockchain, które można dostosowywać do potrzeb organizacji m.in. poprzez modyfikację struktury bloków. To zachęta np. dla organizacji z branży bankowej czy ubezpieczeniowej [Badr, 2018], ale również aplikacji niezwiązanych w żaden sposób z potwierdzaniem operacji finansowych. Przykładem będzie tu zastosowanie technologii łańcuchów bloków do dystrybucji danych w sieciach P2P urządzeń IoT [Deters i Samaniego, 2016].

### **1.5. Przemysłowy Internet rzeczy**

Internet Rzeczy (*Internet of Things* – IoT) i Przemysłowy Internet Rzeczy (*Industrial Internet of Things* – IIoT) to technologie o ogromnym wpływie na proces wdrażania idei społeczeństwa informacyjnego i Przemysłu 4.0. Od czasu, gdy pojęcie po raz pierwszy zostało zaproponowane przez Ashтона Kevina w 1999 roku, rozwinęło się szereg koncepcji związanych z komunikacją między rzeczywistymi lub wirtualnymi przedmiotami, stworzeniami, urządzeniami.

Zarówno IoT, jak i IIoT oparte są na tej samej podstawie – dostępności inteligentnych i połączonych urządzeń. Jediną różnicą między nimi jest ich ogólne użycie. Podczas gdy Internet Rzeczy stosuje się najczęściej do użytku konsumenckiego, IIoT jest wykorzystywany do celów przemysłowych, takich jak produkcja, monitorowanie łańcucha dostaw i system zarządzania nimi.

Firma IDC zidentyfikowała w swojej ankiecie cztery typowe etapy wdrożenia Internetu Rzeczy [Abas Business Solutions Poland, 2019]:

- etap nr 1: obiekty są łączone w sieć,
- etap nr 2: wprowadzany jest monitoring obiektów,
- etap nr 3: przedsiębiorstwa wykorzystują zgromadzone dane do optymalizacji procesów,
- etap nr 4: przedsiębiorstwa przenoszą zdobyte informacje na swój asortyment produktów i opracowują nowe oferty i usługi.

Według wyników tych samych badań większość projektów związanych z Internetem Rzeczy realizowanych przez przedsiębiorstwa średniej wielkości znajduje się między pierwszym a trzecim etapem. Oznacza to, że zastosowanie IoT ogranicza się do doskonalenia procesów wewnętrznych. Firmy nie wykorzystują jeszcze ani potencjału zwiększenia korzyści dla klientów przez podłączone do sieci produkty i usługi, ani szans na zdobycie w ten sposób nowych źródeł przychodów [tamże].

Zagadnienie Internetu Rzeczy i obszarów jego zastosowań została opisana w monografii *Zarządzanie danymi w organizacji* pod red. Beaty Gontar [Gontar, 2019]. W kolejnych rozdziałach niniejszej pracy znajdują się objaśnienia dotyczące wykorzystania wybranych technologii, m.in. automatyzacji i robotyki, analityki Big Data oraz sztucznej inteligencji.

## 1.6. Rzeczywistość rozszerzona

Rzeczywistość rozszerzona (*Augmented Reality* – AR), której nie należy mylić z rzeczywistością wirtualną (*Virtual Reality* – VR), prezentuje treści cyfrowe (tekst, obrazy, animacje itp.) w widoku dostrzegalnym przez rzeczywistego użytkownika.

Encyklopedia Britannica podaje następującą definicję rzeczywistości rozszerzonej: „Rzeczywistość rozszerzona w programowaniu komputerowym to proces łączenia lub rozszerzania obrazów wideo lub

fotografii poprzez nakładanie na obraz przydatnych i generowanych komputerowo informacji” [Hosch].

Tworzenie rzeczywistości rozszerzonej polega na generowaniu w czasie rzeczywistym widoku z przydatnymi informacjami i nakładaniu go na widok realnego świata. Informacje są generowane przez lokalny procesor i źródło danych (a także zdalne źródło danych/bazę danych) i uzupełniane przez dane sensoryczne, takie jak dźwięk, wideo lub informacje o położeniu i lokalizacji. Często używane zamiennie pojęcia „rzeczywistość wirtualna” i „rzeczywistość rozszerzona” nie są tym samym. Pierwsza przenosi użytkownika do całkowicie odizolowanego świata generowanego komputerowo, zwykle z zaledwie trzema stopniami swobody (*Degrees of Freedom* – DoF) umożliwiającymi śledzenie ruchu obrotowego. Druga zapewnia dodatkowe informacje wizualne nałożone na otaczający użytkownika świat oraz sześć stopni swobody (6DoF), co pozwala zarówno na śledzenie obrotu głowy/ciała użytkownika AR, jak i zmiany jego pozycji.

Wymagania technologiczne rzeczywistości rozszerzonej są znacznie większe niż w przypadku rzeczywistości wirtualnej, dlatego rozwój tej pierwszej trwał dość długo. Jednak kluczowe elementy potrzebne do zbudowania systemu AR pozostały takie same aż od czasu pionierskich prac Ivana Sutherlanda z lat sześćdziesiątych.

Do stworzenia wrażenia rzeczywistości rozszerzonej niezbędne są wyświetlacze, moduły śledzące oraz komputery graficzne i odpowiednie oprogramowanie. W procesie generowania rzeczywistości rozszerzonej stosowane są różne technologie, w tym systemy projekcji optycznych, wyświetlacze, urządzenia mobilne (jak tablety i smartfony) oraz systemy wyświetlania noszone przez osobę w postaci okularów lub kasku. Ze względu na sposób użycia urządzenia AR charakteryzuje się często jako urządzenia do noszenia (*wearables*). Wykorzystują one kilka następujących technologii:

- 1) Procesor graficzny GPU (*Graphics Processing Unit*) do sterowania wyświetlaczami.
- 2) Urządzenia wyświetlające/projekcyjne do tworzenia obrazów:
  - optyka do kierowania obrazów w pole widzenia użytkownika.
- 3) Czujniki:
  - skierowane do przodu kamery do analizy,
  - czujniki do pozycjonowania położenia użytkownika w przestrzeni (np. GPS),

- czujniki ruchu,
  - czujnik wysokości,
  - czujniki ruchu gałki ocznej do śledzenia miejsca, w które patrzy użytkownik.
- 4) Systemy audio (mikrofony, przetwarzanie i głośniki) do komunikacji i rozszerzania realnego świata.
  - 5) Systemy identyfikacji i kategoryzacji obiektów, które rozpoznają, na co patrzy użytkownik AR (stół, krzesło, podłoga, ściany, okna, szklanki itp.), aby móc umieścić na nich lub nad nimi wirtualne obrazy. Systemy rozpoznawania obrazu mogą być również wykorzystane do identyfikacji osób, rozpoznawania twarzy czy widzianego tekstu.
  - 6) System operacyjny do sterowania wirtualnymi obrazami za pomocą głosu, oczu, dłoni i ruchów ciała.
  - 7) Bezprzewodowa komunikacja z urządzeniem – może być to zarówno smartfon, jak i dedykowana usługa w chmurze obliczeniowej.

## 1.7. Podsumowanie

Nowe technologie warunkują rozwój Przemysłu 4.0 i transformację cyfrową organizacji. Korzystanie z usług chmury obliczeniowej staje się powszechne. Blockchain stwarza nowe możliwości nie tylko w obszarach finansowych. Rzeczywistość rozszerzona jest stosowana m.in. w nauce, edukacji, przemyśle, logistyce, medycynie [Peddie, 2017]. W wielu przypadkach są to rozwiązania pionierskie, głównie ze względu na wysokie wymagania stawiane przed technologią: jednocześnie zapotrzebowanie na dużą moc obliczeniową i ograniczenia związane z koniecznością oszczędnego gospodarowania energią zasilającą wszystkie urządzenia wchodzące w skład systemu AR.

## Literatura

Abas Business Solutions Poland (2019), *Produkcja przyszłości: jak Internet Rzeczy (IoT) zmienia procesy biznesowe*, <https://www.erp24.pl/systemy-erp-artykuly/produkcja-przyszlosci-jak-internet-rzeczy-iot-zmienia-procesy-biznesowe.html> [dostęp: 01.03.2020].

- Armbrust M., Fox A., Griffith R., Joseph A.D., Katz R. (2009), *Above the Clouds: A Berkeley View of Cloud Computing*, <http://www-inst.cs.berkeley.edu/~cs10/sp11/lec/20/2010Fa/2010-11-10-CS10-L20-AF-Cloud-Computing.pdf> [dostęp: 10.10.2014].
- Badr B. (2018), *Blockchain by Example*, Pact Publishing, Birmingham.
- Buyya R., Yeo C.S., Venugopal S., Broberg J., Brandic I. (2009), *Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5<sup>th</sup> utility*, „Future Generation Computer Systems”, t. 25, nr 6.
- Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021.
- Deters R., Samaniego M. (2016), *Using Blockchain to push Software-Defined IoT Components onto Edge Hosts*, <https://dl.acm.org/doi/abs/10.1145/3010089.3016027> [dostęp: 07.12.2019].
- European Commission (2012), *Quantitative Estimates of the Demand for Cloud Computing in Europe and the Likely Barriers to Up-take – Final report*, <https://ec.europa.eu/digital-single-market/en/news/quantitative-estimates-demand-cloud-computing-europe-and-likely-barriers-take-final-report> [dostęp: 10.05.2014].
- Hosch W.L., *Augmented Reality*, <https://www.britannica.com/technology/augmented-reality> [dostęp: 01.03.2020].
- Kostas K., Kapsalis A., Kyriazis D., Cunha P.R. (2013), *Open-Source IaaS Fit for Purpose: A Comparison between OpenNebula and OpenStack*, [https://www.researchgate.net/figure/Non-exhaustive-view-on-the-main-aspects-forming-a-cloud-system-12\\_fig5\\_236109006](https://www.researchgate.net/figure/Non-exhaustive-view-on-the-main-aspects-forming-a-cloud-system-12_fig5_236109006) [dostęp: 09.09.2019].
- Lipovyanov P. (2019), *Blockchain for Business 2019*, Pact Publishing, Birmingham.
- McKinsey & Co. (2009), *Clearing the Air on Cloud Computing. Technical Report*, [http://www.isaca.org/Groups/Professional-English/cloud-computing/GroupDocuments/McKinsey\\_Cloud%20matters.pdf](http://www.isaca.org/Groups/Professional-English/cloud-computing/GroupDocuments/McKinsey_Cloud%20matters.pdf) [dostęp: 10.10.2017].
- Mell P., Grance T. (2011), *The NIST Definition of Cloud Computing, Recommendations of the National Institute of Standards and Technology*, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf> [dostęp: 07.12.2019].
- Nakamoto S. (2008), *Bitcoin: A Peer-to-Peer Electronic Cash System*, <https://bitcoin.org/bitcoin.pdf> [dostęp: 07.12.2019].
- Peddie J. (2017), *Augmented Reality – Where We Will All Live*, Springer, New York.
- Perez J.C. (2014), *Microsoft fixes Exchange Online outage after almost 9 hours*, <http://www.computerworld.com/article/2491194/desktop-apps/microsoft-fixes-exchange-online-outage-after-almost-9-hours.html> [dostęp: 25.06.2014].
- Pew Research Center (2019), *Social Media Fact Sheet*, <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/> [dostęp: 01.10.2019].
- Pierre Audoin Consultants, *PAC's Cloud Computing Worldwide by countries datamart*, <https://www.pac-online.com> [dostęp: 07.11.2014].
- Raj K. (2019), *Foundations of Blockchain*, Pact Publishing, Birmingham.
- Tormen R. (2019), *Blockchain for Decision Makers*, Pact Publishing, Birmingham.
- Vargas C. (2019), *Cloud Market Share 2019: AWS vs Azure vs Google – Who's Winning?*, <https://www.skyhighnetworks.com/cloud-security-blog/microsoft-azure-closes-iaas-adoption-gap-with-amazon-aws/> [dostęp: 07.12.2019].
- Whittaker Z. (2013), *Amazon Web Services suffers outage, takes down Vine, Instagram, others with it*, <http://www.zdnet.com/amazon-web-services-suffers-outage-takes-down-vine-instagram-flipboard-with-it-7000019842/> [dostęp: 04.06.2017].

# Rozdział 2

## Systemy informatyczne i aplikacje biznesowe w zarządzaniu organizacją

### 2.1. Wprowadzenie

Systemy informatyczne są wykorzystywane w organizacjach od czasu pojawienia się pierwszych komputerów używanych do realizacji celów komercyjnych. Zakres ich funkcjonalności i obszarów zastosowania ewoluował znacznie – od pierwszych prostych aplikacji wspomagających wybrane obszary aż do złożonych zintegrowanych systemów umożliwiających zarządzanie wszelkimi materialnymi i niematerialnymi zasobami organizacji oraz wspieranie niemal wszystkich procesów biznesowych.

Wdrożenie systemu informatycznego ma znaczny wpływ na funkcjonowanie firmy i zwykle oddziałuje na zmianę organizacji pracy wszystkich jej pracowników – stąd wiedza o tym, jak technologie mogą wspierać realizację celów przedsiębiorstwa, okazuje się dla nich niezbędna. Celem niniejszego rozdziału jest przedstawienie funkcjonalności i rozwoju systemów wspomagających zarządzanie organizacją, ich roli we wspieraniu i optymalizowaniu działania i procesów w niej zachodzących oraz wskazanie na aktualne trendy rozwoju oraz wyzwania związane z ich implementacją.



## **2.2. Systemy ERP jako efekt rozwoju systemów informatycznych dla biznesu**

### **2.2.1. Systemy informatyczne zarządzania i ich typologia**

System informatyczny to część systemu informacyjnego organizacji. W literaturze spotyka się wiele jego definicji, ale każda z nich podkreśla, iż jest to część systemu informacyjnego, realizowana za pomocą systemów komputerowych. System ten złożony jest z pięciu podstawowych współpracujących ze sobą elementów: jednostek komputerowych (w tym różnych urządzeń pozwalających na dostęp do danych), oprogramowania, użytkowników, procesów i danych.

Definicja informatycznego systemu zarządzania podkreśla z kolei rolę, jaką odgrywa on w procesach wspomagania procesów kierowania i zarządzania gospodarką jednostek organizacyjnych różnych szczebli [Niedzielska, 1983]. Aplikacje biznesowe pozwalają na gromadzenie, przechowywanie i zarządzanie danymi. Istotną ich cechą jest to, że umożliwiają udostępnianie, interpretowanie i prezentowanie danych w różnych formach.

Technologie informacyjne od czasu swego powstania są narzędziem, dzięki któremu firmy mogą wprowadzać innowacje. Zwrócono na to uwagę już wiele lat temu, wskazując główne obszary, w jakich zmiany dokonywane są najczęściej [Davenport, 1993]. Rozwój technologii potwierdził trafność tych założeń, wśród których za najważniejsze należy uznać [tamże]:

- automatyzację rozumianą jako usprawnienie wynikające z zastosowania technologii informatycznej (zamiast pracy ludzkiej) w ustrukturyzowanych i powtarzalnych procesach;
- zwiększanie dostępu do informacji, a tym samym poprawę jej jakości i skrócenie czasu dostępu, co wpływa na podejmowane decyzje w obszarach zarządzania, wykorzystania zasobów, kolejności działań w procesach itp.;
- zmiany w sekwencji wykonywanych procesów albo łączenie wykonywania kilku do tej pory sekwencyjnych procesów tak, aby realizowane były w jednym czasie;

- śledzenie położenia w przestrzeni wielu obiektów, co jest szczególnie istotne w planowaniu przedsięwzięć logistycznych;
- szybkie przetwarzanie i analizę zbiorów danych w dowolny sposób;
- ograniczanie barier geograficznych poprzez szerokie wykorzystanie Internetu i wirtualizację;
- możliwość integracji przestrzennej rozproszonych zasobów w organizacji, procesów w ramach jednego przedsiębiorstwa czy łańcucha dostaw oraz usuwania granic podziałów funkcjonalnych;
- wspieranie tworzenia wartości intelektualnej oraz gromadzenia wiedzy w procesie budowania łańcucha wartości;
- dostępność informacji w czasie rzeczywistym – dostawcy mają możliwość wielokanałowego komunikowania się bezpośrednio z klientami; klienci mogą wybierać i oceniać oferty.

Rozwój systemów informatycznych biznesu przyczynił się do powstania wielu kategorii systemów. Istnieje szereg kryteriów ich podziału, np. według generacji, skali funkcjonowania, zakresu merytorycznego, poziomu kompleksowości, stopnia integracji, stopnia uniwersalności [Adamczewski, 2001].

Ze względu na wsparcie procesów zarządzania na jego poszczególnych szczeblach (operacyjnym, taktycznym i strategicznym) oraz zaspokojenie potrzeb informacyjnych określonych grup użytkowników w ujęciu klasycznym wyróżnia się następujące typy systemów [Kisielnicki, 2009; Laudon i Laudon, 2013]:

- transakcyjne – rejestrujące podstawowe zdarzenia gospodarcze (*Transactions Systems*),
- automatyzacji pracy biurowej (*Office Automation Systems*),
- systemy informowania kierownictwa (*Management Information Systems*),
- systemy informacyjne kierownictwa wyższych szczebli zarządzania (*Executive Information System*),
- systemy wspomagania decyzji (*Decision Support System*),
- systemy ekspertowe (*Expert Systems*).

Dynamiczny rozwój systemów, większe wymagania użytkowników oraz przenikanie się poszczególnych generacji stały się przyczynami stworzenia kompleksowych zintegrowanych systemów informatycznych, wykorzystywanych przez pracowników różnych szczebli zarządzania [Kisielnicki, 2009].

W literaturze wyróżniane są także typy systemów związane z zadaniami i działaniami prowadzonymi przez organizacje, takie jak:

- systemy planowania zasobów produkcyjnych (*Manufacturing Resources Planning*);
- systemy informacji sprawozdawczej;
- systemy wspomagające zarządzanie zasobami przedsiębiorstwa (*Enterprise Resources Planning*);
- systemy informatyczne logistyki (*Logistic Information Systems*):
  - systemy wspomagające zarządzanie dystrybucją (*Distribution Resources Planning*),
  - systemy wspomagające zarządzanie magazynami (*Warehouse Management Systems*);
- systemy analityki biznesowej (*Business Intelligence*);
- systemy zarządzania wiedzą;
- systemy wczesnego ostrzegania (*Early Warning Systems*);
- systemy biznesu elektronicznego.

Zakres funkcjonalny aplikacji wspomagających zarządzanie przedsiębiorstwem można pogrupować w cztery główne obszary: sprzedaż i marketing, planowanie i produkcja, finanse i controlling oraz zarządzanie zasobami ludzkimi. W praktyce gospodarczej dla potrzeb analiz dotyczących wspierania procesów przez systemy informatyczne można zastosować bardziej szczegółową klasyfikację zaproponowaną przez Iwonę Obłąk i Ewę Ziembę [Obłąk i Ziembę, 2012], obejmującą następujące obszary:

- strategia i zarządzanie,
- zarządzanie zasobami informacyjnymi,
- zarządzanie zasobami ludzkimi,
- zarządzanie zasobami finansowymi i materialnymi,
- zarządzanie relacjami zewnętrznymi,
- zarządzanie doskonaleniem i zmianą,
- planowanie oraz rozwój produktów i usług,
- proces produkcji,
- zarządzanie zaopatrzeniem,
- marketing,
- sprzedaż.

We współczesnych organizacjach dominuje podejście procesowe, a do wspierania obsługi procesów biznesowych, optymalizacji działań i podejmowania decyzji stosują one wiele systemów informatycznych. Szczególne

znaczenie mają procesy określane mianem procesów głównych, służące bezpośrednio realizacji celów organizacji tworzących wartość dodaną dla klienta i przynoszących zyski. Procesy główne i pomocnicze są określane indywidualnie przez każdą organizację – stąd konieczność sporządzania map procesów, zwłaszcza przed decyzją o implementacji systemów informatycznych. Relacje z klientami oraz udział w łańcuchu dostaw powodują, że procesy biznesowe obejmują działania wychodzące poza obszar organizacji, co pozwala na wydzielenie systemów związanych z:

- zarządzaniem zasobami przedsiębiorstwa (*Enterprise Resource Planning* – ERP),
- zarządzaniem relacjami z klientami (*Customer Relationship Management* – CRM),
- zarządzaniem relacjami z dostawcami (*Supplier Relationship Management* – SRM),
- zarządzaniem wiedzą – pozwalające pozyskiwać i przekazywać wiedzę w organizacji.

W ciągu ostatnich kilkudziesięciu lat wiele organizacji dokonało inwestycji w systemy, które miały zapewnić im wsparcie w przepływie informacji, integracji procesów biznesowych i zarządzaniu podstawowymi zasobami (materiałami, urządzeniami, zasobami ludzkimi i finansowymi). Z powodu coraz większego uzależnienia gospodarki od technologii IT i Internetu obecną gospodarkę określa się mianem „gospodarki elektronicznej”; natomiast ze względu na znaczenie tworzenia i współdzielenia wiedzy – jako „gospodarkę opartą na wiedzy”.

Wdrażane w organizacjach biznesowych inwestycje IT służą przede wszystkim:

- osiągnięciu doskonałości operacyjnej procesów,
- wspomaganiu procesu wytwarzania nowych dóbr i usług,
- nawiązywaniu bliskich relacji z klientami i dostawcami w łańcuchu sprzedaży i obsługi posprzedażowej,
- poprawieniu procesu podejmowania decyzji,
- zdobywaniu przewagi konkurencyjnej lub utrzymaniu się na rynku.

W wielu przypadkach inwestycje te wymagają znacznych nakładów finansowych – nie zawsze przynosząc oczekiwane korzyści i zyski. Tradycyjne rozwiązania systemu informatycznego instalowanego lokalnie wiążą się z kosztami związanymi z zakupem infrastruktury sprzętowej i oprogramowania, szkoleniami pracowników, bieżącym utrzymaniem i doskonaleniem

systemów, a także zapewnieniem kadry o odpowiednich kompetencjach. Prowadzone w tym zakresie badania wskazują, iż ponad 67% wdrożeń nie osiąga nawet połowy zakładanych rezultatów, co wynika nie tylko z samego procesu implementacji, ale też utrzymywania i rozszerzania funkcjonalności systemu w fazie jego użytkowania [Pei-Fang i inni, 2015].

### **2.2.2. Geneza i rozwój zintegrowanych systemów zarządzania**

Pojęcie „zarządzanie zasobami przedsiębiorstwa” (*Enterprise Resource Planning* – ERP) oznacza zarządzanie organizacją poprzez planowanie jej zasobów. Określa ono również klasę systemów informatycznych wykorzystywanych do wspierania zarządzania procesami wykonywanymi w organizacji poprzez gromadzenie oraz umożliwienie wykonywania operacji na zebranych danych [Woźniak, 2002].

Systemy ERP są kontynuacją, rozwinięciem funkcjonalnym systemów MRP (*Manufacturing Resource Planning*), przeznaczonych do planowania zasobów wytwórczych przedsiębiorstwa, których geneza wywodzi się od projektowanych w latach pięćdziesiątych XX wieku systemów ewidencyjnych wspomagających sterowanie zapasami (*Material Requirements Planning*) [Durlik, 1998].

Ewolucja systemu MRP do obecnego obszaru funkcjonalności systemu zarządzania przedsiębiorstwem trwała wiele lat. Pierwszym znaczącym etapem było w latach osiemdziesiątych XX wieku wyposażenie go w funkcje dotyczące planowania zakupów materiałów na podstawie harmonogramu produkcji (*Manufacturing Resource Planning* – MRP II); kolejnym – rozszerzenie o funkcje pozwalające na kontrolowanie pozostałych zasobów przedsiębiorstwa, w tym środków trwałych oraz zasobów ludzkich. Zintegrowane Systemy Informatyczne klasy ERP ukształtowały się w latach dziewięćdziesiątych. Przeznaczone początkowo dla firm o charakterze produkcyjnym, wkrótce znalazły zastosowanie w przedsiębiorstwach usługowych oraz podmiotach administracji publicznej [Hadaś i Cyplik, 2012]. Systemy klasy ERP przez niektórych autorów nazywane są również MRP III (*Money Resource Planning*), ponieważ pozwalają na zaawansowane planowanie i śledzenie zasobów finansowych, obejmując takie obszary, jak: rachunkowość zarządcza, *cash flow*, rachunek kosztów [Bojarski, 2003; Lech, 2012].

Trzonem systemu ERP jest MRP II, stworzony przez Amerykańskie Stowarzyszenie Sterowania Produkcją i Zapasami (*American Production and Inventory Control Society* – APICS) [Woźniak, 2002] – i zgodnie ze standardem obejmuje on następujące obszary [Landarter i Gray, 1989; Parys, 1998]:

- planowanie produkcji i sprzedaży (*Sales and Operation Planning* – SOP),
- zarządzanie popytem (*Demand Management* – DEM),
- harmonogramowanie planu produkcji (*Master Production Scheduling* – MPS),
- planowanie zapotrzebowania materiałowego (*Material Requirements Planning* – MRP),
- podsystem struktur wyrobów (*Bill of Material Subsystem*),
- podsystem transakcji materiałowych (*Inventory Transaction Subsystem* – INV),
- podsystem harmonogramów spływu (*Scheduled Receipts Subsystem* – SRS),
- sterowanie produkcją (*Shop Floor Control* – SFC),
- planowanie zdolności produkcyjnych (*Capacity Requirements Planning* – CRP),
- zarządzanie środowiskiem roboczym (*Input/Output Control*),
- zaopatrzenie (*Purchasing* – PUR),
- planowanie dystrybucji (*Distribution Resource Planning* – DRP),
- pomoce warsztatowe (*Tooling*),
- interfejs do planowania strategicznego (*Financial Planning Interface*),
- symulacja (*Simulation*),
- pomiar działania systemu (*Performance Measurement*).

Zaimplementowanie w systemie schematu sprzężenia zwrotnego między procesem planowania a procesem produkcji MRP II pozwoliło na wspomaganie procesów zarządzania na wszystkich jego poziomach: strategicznym, taktycznym i operacyjnym [Durlík, 1998]. Kolejne rozszerzenia funkcjonalności systemu umożliwiły szeroką integrację działań przedsiębiorstwa na każdym ze szczebli zarządzania oraz procesów dotyczących zaopatrzenia i produkcji – w celu reagowania na zmiany popytu [Klonowski, 2004]. Systemy ERP zapewniają integrację danych i procesów nie tylko wewnątrz organizacji, ale i w procesach wiążących ją z otoczeniem administracyjnym i partnerami w łańcuchu dostaw, poprzez implementację gotowych narzędzi pozwalających na automatyzację

procesów wymiany danych. „Proces integracji zasobów informacyjnych i usług informatycznych w organizacji jest zjawiskiem permanentnym, widocznym przez pryzmat wdrażania całej rodziny systemów zintegrowanych” [Zaskórski, 2012].

Rozwój zintegrowanych systemów informatycznych przebiegał w kilku płaszczyznach [Lech, 2003]:

- podstawowej – zapewniającej sprawne i wydajne funkcjonowanie systemu baz danych w systemie partnerów łańcucha dostaw,
- procesowej – realizującej zasady planowania zasobów przedsiębiorstwa i wspomagającej wykonywanie poszczególnych procesów biznesowych oraz zarządzanie nimi;
- analitycznej – pozwalającej na przetwarzanie zgromadzonych w bazie danych według dowolnych kryteriów,
- portalowej – dającej możliwość wykorzystania usług i zasobów Internetu.

Przyjęcie za podstawę integracji IT ze strategią biznesową oraz wyjście poza obszar organizacji – poprzez połączenie i współdzielenie danych z wieloma kooperantami w zakresie tego samego łańcucha dostaw z użyciem technologii internetowych i rozwiązań mobilnych – utworzyło model systemu określany mianem ERP II [Bojarski, 2003; Kasprzak, 2006].

ERP to główny system organizacji przeznaczony do zbierania danych o zdarzeniach gospodarczych i użytkownikach procesów biznesowych. Przez wiele lat jego rozwój dopasowany był do ekonomiki produktu i transakcji. Rozszerzony o dane z innych źródeł pozwala na optymalizację procesów podejmowania decyzji, a odpowiednia architektura systemowa umożliwia bieżącą aktualizację procesów biznesowych, usług IT, modeli danych i systemu raportowania. Zastosowanie elementów gamifikacji<sup>1</sup> wpływa na zaangażowanie i identyfikację pracowników oraz partnerów z celami organizacji.

Systemy informatyczne firmy (określane też mianem *Enterprise Information Systems* – EIS) nowej generacji, skoncentrowanej na połączeniu „wszystkiego”: danych, usług, sensorów, urządzeń i ludzi, będą cechowały się [Panetto i inni, 2016]:

- wszechobecnością – wcześniejsze generacje były implementowane w sieciach komputerowych organizacji i Internecie z uwzględnieniem potrzeby konkretnych aplikacji; rozszerzenie usług na urzą-

---

<sup>1</sup> Określana też mianem „grywalizacji” – obejmuje działania mające na celu zmianę zachowania pracowników poprzez wykorzystanie mechanizmów stosowanych w grach.



dzenia mobilne oraz ich wzajemna komunikacja zapewnią obecność i dostęp do systemu w każdym miejscu;

- architekturą sterowaną rozproszonymi modelami, które kontrolują i modyfikują funkcjonalność oraz modele biznesowe współpracujących organizacji, co wpłynie na zmiany cyklu życia produktu;
- otwartością – technologiczną i prawną dostępnością do oprogramowania i zasobów czy współdzielenia wiedzy; w szerszym kontekście oznacza to transparentność struktur organizacji, tj. możliwości współtworzenia środowisk i narzędzi wspierających kreowanie nowych produktów i usług;
- dynamiczną rekonfiguracją – pozwalającą na różne sposoby realizacji zadań i procesów z wykorzystaniem tego samego lub ulepszanego zbioru zasobów poprzez stworzenie interfejsów do modeli procesów i działań;
- zwielokrotnieniem użytkowników funkcjonalnych mających uprawnienia do wykonywania określonych działań; w tym przypadku mogą to być także agenty programowe albo system nadzorujący procesy współpracy;
- świadomością semantyczną na poziomie sensorów – pozwalającą zbierać i przetwarzać dane na poziomie rozproszonych urządzeń, a na wyższych poziomach zarządzać procesami współpracy;
- zastosowaniem zaawansowanych modeli wnioskowania.

Krótką charakterystykę kolejnych generacji systemów informatycznych zarządzania oraz determinanty ich rozwoju przedstawiono w tabeli 2.1.

Systemy informatyczne generacji od pierwszej do piątej były skoncentrowane na wsparciu formy sprzedaży, w której każde zamówienie traktowano jako nową transakcję sprzedaży, a wprowadzanie zmian w produkcji/usłudze wiązało się z tworzeniem nowej oferty produktowej; stawiały one na operacyjną doskonałość, niezawodność, gromadzenie danych i bezpieczeństwo. Generacje szósta i siódma systemów koncentrują się m.in. na modelu dystrybucji, w tym modelu subskrypcji. Model ten, określany także mianem abonamentowego, jest znacznie bardziej skomplikowany, a sam proces sprzedaży powiązany z wieloma etapami relacji z klientem i jego zmiennymi decyzjami. System abonamentowy to nie tylko zakup i obsługa posprzedażowa produktu/usługi. Abonament podlega ciągłym modyfikacjom, takim jak: zmiana zakresu, poszerzenie, obniżenie lub anulowanie umowy – wszystkie te zjawiska muszą znaleźć swoje odzwierciedlenie w systemie, stanowiąc główną oś działania. Różnice pomiędzy systemami zaprezentowano na rysunku 2.1.



**Tabela 2.1.** Determinanty rozwoju generacji systemów ERP

<b>Determinanty rozwoju</b> <i>t</i> – technologiczne <i>b</i> – biznesowe <i>o</i> – organizacyjne		<b>Generacja/ technologia</b>	<b>Zakres</b>  <b>Wyzwania</b>
<i>t</i>	bazy danych	<b>1. MRP</b> <b>lata osiemdziesiąte XX wieku</b> <b>(sterowanie zapasami)</b>	systemy lokalne, pojedyncze aplikacje  integracja, strukturalizacja danych
<i>b</i>	skuteczność		
<i>o</i>	wsparcie wybranych działów		
<i>t</i>	bazy danych	<b>2. MRP II</b> <b>standard MRP</b>	planowanie produkcji, zarządzanie popytem  reguły biznesowe
<i>b</i>	sprzężenie zwrotne: planowanie–produkcja		
<i>o</i>	wsparcie dla wszystkich poziomów zarządzania		
<i>t</i>	klient–serwer system zarządzania bazą danych	<b>3. ERP</b> <b>lata dziewięćdziesiąte XX wieku</b>	całe przedsiębiorstwo  integracja danych, procesy biznesowe
<i>b</i>	wydajność		
<i>o</i>	wsparcie wszystkich działów i relacji z klientami		
<i>t</i>	komunikacja, Internet, SOA	<b>4. ERP II</b> <b>2000</b>	cyfryzacja, zarządzanie procesami (BPM)  interoperacyjność, dynamiczne dane
<i>b</i>	skuteczność biznesu		
<i>o</i>	wsparcie zintegrowanego łańcucha dostaw		
<i>t</i>	media społecznościowe, sieci semantyczne, chmura obliczeniowa, mobilność	<b>5. ERP zaangażowania i współpracy</b> <b>Human Centric</b> <b>2010</b>	złożone wirtualne środowiska  prostota obsługi, bezpieczeństwo, standardy, dane w czasie rzeczywistym
<i>b</i>	elastyczność		
<i>o</i>	media społecznościowe		
<i>t</i>	analityka, chmura obliczeniowa, Internet Rzeczy (IoT) – sensory, model subskrypcji	<b>6. ERP (EIS) przedmiotów</b> <b>2015</b>	cyberprzestrzeń  platformy integracyjne
<i>b</i>	wrażliwość		
<i>o</i>	wsparcie dla sieci sensorów		
<i>t</i>	model subskrypcji usług, urządzenia	<b>7. ERP (EIS) wszystkiego</b> <b>2020</b>	sieci wzajemnych powiązań  otwartość, współpraca, bezpieczeństwo, środowisko <i>plug&amp;play</i> , inteligentne wnioskowanie
<i>b</i>	proaktywność		
<i>o</i>	wsparcie sieci powiązań		

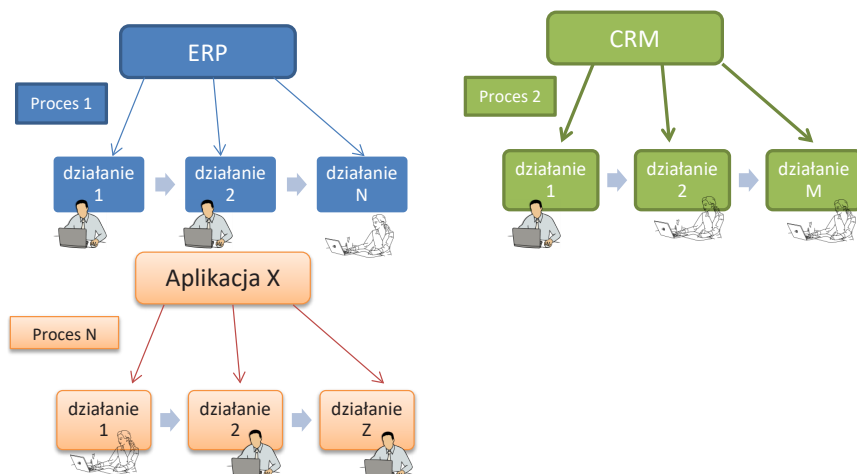
**Źródło:** opracowanie własne na podstawie [Panetto i inni, 2016; Essex, 2019].



**Rysunek 2.1.** Elementy procesu sprzedaży w ekonomice produktu i ekonomice subskrypcji

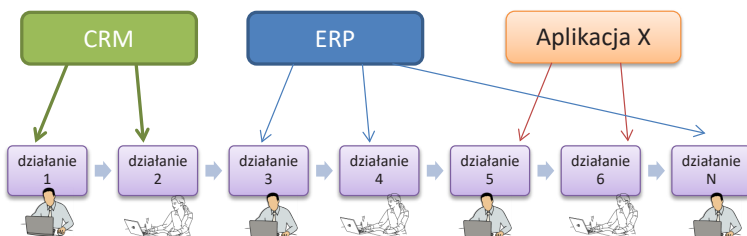
**Źródło:** opracowanie własne na podstawie [Somer, 2017].

Aplikacje biznesowe są wykorzystywane w organizacji do wspierania procesów biznesowych, które mogą być wspomagane przez pojedynczą aplikację (rysunek 2.2) lub przez szereg aplikacji (rysunek 2.3). Za wykonanie działań odpowiedzialnych może być z kolei wielu pracowników.



**Rysunek. 2.2.** Obsługa działań w procesach wspomaganych przez niezależne aplikacje

**Źródło:** opracowanie własne.



**Rysunek 2.3.** Obsługa działań w procesie wspomaganym przez różne aplikacje

**Źródło:** opracowanie własne.

Rozwój systemów informatycznych zarządzania wcześniejszych generacji dążył m.in. do osiągnięcia kompleksowego zarządzania produktem/usługą, obejmując wszystkie działania: od tworzenia wymagań dla nowego produktu, poprzez projektowanie, produkcję testową, wprowadzanie na rynek, dystrybucję, aż po wycofywanie produktu z rynku i zastępowanie go całkowicie nowym produktem lub jego nowszą wersją. Kolejne obszary, o które wzbogacane były systemy, obejmowały [Zalewski, 2011]:

- MES (*Manufacturing Execution System*) – system do śledzenia oraz nadzorowania produkcji i przepływu materiałów,
- SCADA (*Supervisory Control and Data Acquisition*) – oprogramowanie do zbierania danych dotyczących sterowanego procesu i sterowania w czasie rzeczywistym,
- HMI (*Human-Machine Interface*) – interfejs zapewniający bezpośredni dialog człowieka i maszyny,
- PLM (*Product Lifecycle Management*) – zarządzanie cyklem życia produktu,
- PDM (*Product Data Management*) – zarządzanie danymi o produkcie,
- *workflow* – mechanizmy sterowania przepływem prac,
- APS (*Advanced Planning and Scheduling*) – zaawansowane planowanie i harmonogramowanie produkcji,
- EAM/CMMS (*Enterprise Asset Management/Computerised Maintenance Management Systems*) – zarządzanie majątkiem przedsiębiorstwa oraz wspomaganie utrzymania ruchu.

W modelu subskrypcji system informatyczny musi posiadać funkcjonalność pozwalającą na śledzenie zmian każdego z elementów umowy abonamentowej oraz szybką integrację z pozostałymi systemami, co wymaga koncentracji na procesach systemu rejestracji i obsługi umowy abonamentowej, obsługi płatności, modelowania umowy i analiz zakupowych decyzji i zachowań klientów. Oferenci systemów ERP, tworzonych i optymalizowanych dla ekonomiki produktu, stoją przed wyzwaniem konieczności zapewnienia znacznej skalowalności systemu w następujących obszarach [Somer, 2017]:

- całościowej i jednolitej informacji o kliencie – systemy ERP były tworzone z myślą o śledzeniu i analizowaniu produktów, nie klientów, nawet jeśli w kolejnych etapach poszerzano je o moduły CRM; obecnie systemy muszą oferować nowe funkcjonalności pozwalające na ocenę i wycenę relacji z klientem;

- całego cyklu życia abonamentu od momentu zainteresowania klienta i jego rejestracji za pomocą dowolnego medium (telefonu, Internetu), zwłaszcza odnawiania abonamentu i oferty usług dodatkowych;
- rejestrowania, obsługi i analizy złożonych procesów i modeli płatności (wystawiania rachunków, rejestracji płatności, egzekwowania zadłużeń, analizy zysków – przy zastosowaniu modeli odzwierciedlających stopień i częstotliwość korzystania z usługi oraz modeli *pay-as-you-go*);
- szybkiego dodawania nowych usług i wpisywania ich do abonamentów klientów w postaci stałego lub tymczasowego (wprowadzonego na okres próbny) elementu w modelach B2B (*Business to Business*) oraz B2C (*Business to Customer*);
- prowadzenia zaawansowanych analiz finansowych dotyczących pomiaru i określania zależności pomiędzy rezerwacjami (abonentami), płatnościami, przepływem gotówki i osiąganym zyskiem.

Przed współczesnym przedsiębiorstwem staje więc wyzwanie sprawnej obsługi nie tylko systemów dla transakcji produktu, ale także coraz powszechniejszego modelu subskrypcji.

### 2.2.3. Moduły systemów ERP

Systemy ERP są zintegrowanymi aplikacjami o strukturze modułowej, złożonej z niezależnych części systemu, które pozwalają na bezpieczny obieg potrzebnych informacji i wsparcie procesów biznesowych firmy. Integracja modułów odbywa się bez powielania danych niezbędnych do realizacji procesów przez nie obsługiwanych. Logika biznesowa i szeroki zakres wspomaganych procesów sprawiają, iż zwykle są one podzielone na oddzielne aplikacje nazywane modułami tematycznymi, co pozwala organizacji na wybór strategii wdrażania kolejności implementowanych modułów, odpowiednio do potrzeb i posiadanych zasobów finansowych [Grudzewski, 2004]. Zaawansowanie technologiczne – wraz z zastosowaniem standardów dotyczących planowania produkcji MRP II, procedur finansowych (jak *cash flow*, metoda *Activity Based Costing* – ABC), kompleksowego zarządzania jakością (*Total Quality Management* – TQM) oraz standardami norm ISO 9000 – gwarantuje zgodność z aktualnymi wymogami sprzętowymi oraz programowymi,

gdzie istnieje możliwość przemieszczenia na nowe płaszczyzny sprzętu komputerowego i systemów operacyjnych [Kisielnicki, 2013].

Każdy moduł systemu ERP specjalizuje się we wspieraniu określonych procesów biznesowych. Nazwy modułów i ich szczegółowe funkcjonalności różnią się w zależności od dostawcy, niemniej jednak do głównych modułów należą:

- 1) Zarządzanie finansami – to kluczowa funkcja biznesowa w przedsiębiorstwach. Moduły obejmują funkcjonalności księgi głównej, która jest podstawą zarządzania finansami. W przypadku dużych systemów pozwalają one na wielofirmowość, tj. obsługę zgodną z regulacjami każdego kraju zaimplementowanego w systemie. Moduły te zawierają m.in. funkcje związane z obsługą kont płatniczych, zobowiązań, należności, sprawozdawczości finansowej.
- 2) Kontroling – dla potrzeb rachunkowości zarządczej oferowane są moduły dotyczące planowania i realizowania strategicznych celów organizacji oraz bieżącego (operacyjnego) planowania i monitorowania bieżącej rentowności, płynności i efektywności przedsiębiorstw.
- 3) Zarządzanie aktywami – pozwala na zarządzanie pozycjami majątku przedsiębiorstwa: wartościami niematerialnymi i prawnymi oraz wyposażeniem. Funkcjonalność ta umożliwia rejestrację wartości zakupu, obliczanie amortyzacji i wartości bieżącej według różnych modeli, tworzenie harmonogramów konserwacji urządzeń, rejestracji czasu pracy oraz przestojów urządzeń. Dla przedsiębiorstw sieciowych posiadających bogatą infrastrukturę moduł jest często wiązany z systemami GIS.
- 4) Logistyka: Zarządzanie łańcuchem dostaw (ZŁD) – moduły tego obszaru wspierają obsługę procesów całego łańcucha dostaw. Zawierają funkcjonalności dotyczące procesów dystrybucji, zaopatrzenia, produkcji oraz magazynowania. W obszarze logistyki dystrybucji funkcjonalność pozwala na rejestrację zamówień, zarządzanie nimi, m.in. poprzez udzielenie bonifikat, planowanie dostaw, obsługę zwrotów (i opakowań). Moduły te umożliwiają wielokanałowe zautomatyzowanie składania zamówienia, wprowadzanie do niego zmian lub anulowanie zgodnie z przyjętymi procesami biznesowymi danej organizacji. Przydzielone uprawnienia pozwalają użytkownikowi na przeglądanie oraz śledzenie statusu zamówienia. W logistyce zaopatrzenia funkcjonalność wspiera organizację

w optymalizowaniu kosztów i czasu dostaw materiałów na podstawie planu sprzedaży i planu produkcji, co umożliwia łączenie zapotrzebowania spływającego od wielu użytkowników i przeprowadzenie wyboru dostawcy zgodnie z przyjętymi w organizacji zasadami. Moduły produkcji pozwalają na planowanie i harmonogramowanie obciążenia zdolności produkcyjnych na podstawie marszrut i specyfikacji materiałowych (*Bill of Material* – BOM). Zarządzanie magazynem usprawnia śledzenie procesów dystrybucji związanych z surowcami, półproduktami i produktami gotowymi oraz pozwala na pełną synchronizację i kontrolę nad zapasami w magazynach. Funkcjonalność obejmuje szereg raportów dotyczących analiz pozycji magazynowych i ich rotacji.

- 5) Zarządzanie kapitałem ludzkim (*Human Capital Management* – HCM) – wspomaga zarówno zadania operacyjne, jak i strategiczne. W pierwszym zakresie umożliwia szczegółowy opis działań kadrowo-placowych, administracji oraz zgodności z regulacjami rządowymi. W drugim pozwala na zarządzanie rozwojem kadr, kompetencjami i talentami.
- 6) Zarządzanie projektami – umożliwia budowanie harmonogramu przedsięwzięć, analizę kosztów oraz czasu wykonywanego projektu, zarządzanie portfelem przedsięwzięć, zasobami i rozliczaniem ich.
- 7) Zarządzanie relacjami z klientami (*Customer Relationship Management* – CRM) – pozwala na gromadzenie wszystkich danych na temat klientów, realizacji zamówień i projektów, poziomu satysfakcji, tworzenia ofert i kampanii marketingowych czy segmentacji klientów. Funkcjonalność wzbogacana jest o szeroki zakres analiz.
- 8) Cykl życia produktu (*Product Lifecycle Management* – PLM) – umożliwia śledzenie produktu podczas całego jego cyklu życia: od powstania koncepcji, poprzez wprowadzane zmiany, do końca okresu eksploatacji. Funkcjonalność pozwala na zarządzanie portfelem oraz projektowanie produktów.

Klasyczne systemy informatyczne do planowania zasobów przedsiębiorstwa znacznie usprawniły zarządzanie nimi. ERP zapewnia dzisiaj zintegrowany widok podstawowych procesów biznesowych w różnych działach: od zaopatrzenia, produkcji i logistyki, przez sprzedaż, aż do księgowości i płac. Systemy te gromadzą wiele danych na temat kosztów,

sprzedaży, zaopatrzenia, dostawców czy klientów, umożliwiając sprawną realizację procesu i określenie wyniku ekonomicznego oraz tworzenie sprawozdań i raportów. Nie wspierają w wystarczającym stopniu firm, które chcą znać całkowity wpływ swojej aktywności w ramach rozszerzonego łańcucha wartości – oraz pragną nim zarządzać. Obecnie konsumenci oczekują nie tylko jakości i ceny. Wraz ze wzrostem świadomości społecznej bardziej dbają oni o społeczne i środowiskowe oddziaływanie produktów i usług, które nabywają. Klienci są zainteresowani np. składem produktu, chcą znać pochodzenie surowców, otrzymywać informacje, czy produkt zawiera substancje toksyczne, czy procesy produkcyjne są przyjazne dla środowiska, jakie są praktyki bezpieczeństwa pracy w całym łańcuchu dostaw. Podobne żądania związane z pozyskaniem informacji środowiskowych mogą kierować inni interesariusze, np. organy rządowe, władze lokalne czy stowarzyszenia [Odenwald, 2014].

Współczesne systemy informatyczne i metody sprawozdawczości w większości przypadków traktują czynniki, takie jak emisje, odpady i ochrona pracowników jako efekty zewnętrzne. Jednak wiele z nich, w tym niepewność gospodarcza, wzrost liczby ludności, zmiany klimatyczne i rosnące zapotrzebowanie na zasoby naturalne, wywiera nową presję na przedsiębiorstwa, aby przy podejmowaniu decyzji organizacje brały pod uwagę szerszy zestaw elementów. Zwiększa się nacisk społeczny na wykorzystanie surowców wtórnych i logistykę zwrotną, co pociąga za sobą zmianę cyklu życia produktu. Realistyczny model ekonomiczny organizacji powinien odzwierciedlać pełen zakres jej działalności: od pierwotnego zaopatrzenia w materiały do samego końca cyklu życia produktu. Systemy informacyjne powinny zatem pomagać przedsiębiorstwom w ocenie i ujawnianiu wpływu, jaki ich rozszerzone łańcuchy wartości mają nie tylko na wyniki finansowe, ale również na środowisko, społeczeństwo i gospodarkę [tamże]. Organizacje mają możliwość i są w stanie analizować szczegóły dotyczące emisji dwutlenku węgla lub zużycia wody w łańcuchu dostaw, ale będą musiały zintegrować te informacje z wewnętrznymi danymi, pochodzącymi z ich własnych procesów.

Podążając za oczekiwaniami, dostawcy systemów informatycznych wspierających zarządzanie organizacją muszą rozszerzyć model obsługi klienta o model subskrypcyjny, oferować bieżącą analitykę danych, a także dostarczać holistycznego rozwiązania zarządzania zasobami – systemy te należy więc rozpatrywać szerzej: nie tylko jako narzędzia do mierzenia zy-

sków, ale także do uwzględniania potrzeb ludzi i środowiska naturalnego. Takie podejście uwzględnia m.in. gromadzenie danych o tym, skąd pochodzą substancje i surowce składające się na produkt; ile energii potrzeba do wytworzenia produktu; ile odpadów i emisji generują poszczególne etapy oraz w jakim stopniu praktyki biznesowe dostawców są zgodne z podstawowymi prawami człowieka i standardami pracy (rysunek 2.4).



**Rys 2.4.** Obszary zarządzania zasobami ERP dla potrzeb organizacji i interesariuszy

**Źródło:** opracowanie własne.

Obecnie wielu producentów wymaga od swoich dostawców zapisywania na platformie technologicznej odpowiednich danych o produktach – zgodnie z zasadami zrównoważonego rozwoju. W organizacji prowadzi to często do korzystania z szeregu platform, które dostawcy muszą utrzymywać dla różnych klientów. Przejście na sieci współdzielonych danych w chmurze obliczeniowej to sposób na usprawnienie tego procesu, pozwalający firmom na łatwiejsze i tańsze śledzenie informacji. Umożliwia to także skuteczne gromadzenie, zarządzanie i ponowne wykorzystywanie danych dotyczących zrównoważonego rozwoju produktów, przy jednoczesnej oszczędności czasu i pieniędzy dla wszystkich uczestników sieci [tamże].



Rozwijanie zdolności do tworzenia zagregowanych raportów z istniejących systemów ERP to klucz do optymalizacji wydajności zrównoważonego rozwoju. Przykładem zastosowania takiego rozwiązania jest Danone, międzynarodowa korporacja spożywcza z siedzibą w Paryżu, mierząca i analizująca wpływ produktów na zrównoważony rozwój na poziomie zakładów – poprzez wykorzystanie informacji z istniejących systemów przedsiębiorstwa. Kiedy poszczególne zakłady otrzymują zlecenia produkcyjne, zarówno Danone, jak i podległe firmy są w stanie rozpocząć obliczanie związanej z tym emisji dwutlenku węgla i poziomu zużycia wody. Pracownicy mają również możliwość określenia tego wpływu w czasie rzeczywistym – i jednocześnie zastosowania algorytmów symulacyjnych wspomagania decyzji, np. efektu wykorzystania strategii łągodzących, takich jak wymiana materiałów, alternatywny transport lub wprowadzenie nowych rodzajów odnawialnych źródeł energii. Po zoptymalizowaniu produkcji, harmonogramów logistycznych i zaplanowaniu ich przy niższych emisjach dwutlenku węgla, mniejszym zużyciu energii i redukcji ilości odpadów przedsiębiorstwa określają możliwości generowania długoterminowej zrównoważonej wartości. Stosując to podejście, Danone analizuje 35 000 produktów w skali miesiąca i osiąga spadek emisji na poziomie 35% [tamże].

#### **2.2.4. Platformy współpracy między dostawcami a klientami**

Od czasu, gdy rozwinął się rynek *e-commerce* systemy ERP przestały być wystarczającym narzędziem do wspierania wszystkich procesów organizacji. Rozpoczął się wtedy rozwój platform zintegrowanego łańcucha dostaw dostarczających większej funkcjonalności w obszarze kontaktów z klientami i dostawcami. Od końca XX wieku organizacje interesowały się rozwojem rynku B2B, co podyktowane było tradycyjnym modelem biznesowym opartym na ekonomii produktu. Ekonomia subskrypcji skłania obecnie organizacje do szerszego zastosowania modelu B2C.

Jako jedne z pierwszych platform wspomagających procesy biznesowe organizacji w obszarze kontaktu z klientem pojawiły się platformy CRM. Podstawą ich budowy były aplikacje monitorujące i usprawniające pracę przedstawicieli handlowych. Do tej pory część przedsiębiorstw wdraża ich funkcjonalność tylko w zakresie automatyzacji sprzedaży. Systemy CRM posiadają wiele funkcji wspierających operacje handlowe i pozwalają na

skuteczne zarządzanie sprzedażą. Stanowiąc centrum systemu, gromadząca szereg danych baza firm zapewnia handlowcom informacje, które mogą okazać się niezbędne nawet przy pierwszym kontakcie z potencjalnym klientem. Dostęp do niej pozwala pracownikom bieżąco sprawdzić: historię transakcji, terminowość płatności, realizowane projekty, wartość zakupów klienta itp. Taka wiedza umożliwia odpowiednie podejście do negocjacji oraz przygotowanie oferty odpowiadającej wymaganiom i możliwościom klienta. Z punktu widzenia dystrybucji istotną funkcją w systemach CRM jest zarządzanie możliwościami/okazjami poprzez zarejestrowanie w systemie nawet najmniejszej wzmianki o zainteresowaniu kupnem. Na podstawie zdobytych informacji handlowiec może ocenić szansę powodzenia transakcji, określić priorytet zamówienia oraz nakreślić plan działania. W ten sposób budowana baza wiedzy – poprzez rejestrowanie każdego kontaktu z klientem lub wiadomości z otoczenia – zapewnia większą szansę możliwości skutecznej sprzedaży. Zintegrowane z innymi platformami systemy CRP mogą automatycznie przekształcać pozyskane zamówienia w zlecenia. Niezbędna jest bieżąca aktualizacja bazy także o działania i efekty posprzedażowe. Obecne systemy posiadają proste, intuicyjne narzędzia, dzięki czemu tworzenie analiz nie wymaga specjalistycznej wiedzy oraz umiejętności. Systemy wyposażone w CRM analityczny umożliwiają ponadto skomplikowane analizy i prognozy rentowności, realizacji planów czy dynamiki sprzedaży.

Systemy CRM wspomagają również automatyzację w procesie obsługi klienta. Ten obszar funkcjonalności najlepiej sprawdza się w kontaktach ze stałymi klientami, dokonującymi powtarzalnych zakupów. Platformą, która wyróżniała się innowacyjnym rozwiązaniem, był Salesforce – doskonale wpasował się on w trendy zachodzące w biznesie. Innowacyjność systemu polegała na zaoferowaniu dostępu do systemu w modelu chmurowy obliczeniowej. Ponadto Salesforce wyróżniał się otwartym interfejsem programowania aplikacji (*Application Programming Interface* – API), co pozwoliło innym firmom na tworzenie dodatkowych narzędzi bazujących na Salesforce. Obecnie to już nie tylko platforma CRM, ale platforma współpracy łącząca partnerów biznesu, umożliwiająca tworzenie nowych produktów, wspierająca działania prosumpcji i współtworzenia wartości.

Podobna sytuacja miała miejsce w przypadku procesów zaopatrzeniowych. Dzięki funkcjom systemu ERP możliwe jest usprawnianie procesów zakupowych i relacji z dostawcami w ograniczonym zakresie.

Systemy ERP przestały być na tyle wystarczające, by w pełni zoptymalizować proces zaopatrzenia, zarówno pod względem obsługi samego procesu, jak i kosztów. ERP nie reaguje w pełni na coraz większą potrzebę efektywnego zarządzania bazą dostawców i kontraktów, analizowania wydatków pod kątem grup materiałowych i źródeł zaopatrzenia. Aby sprostać tym wymaganiom, pojawiły się na rynku narzędzia SRM.

Platformy SRM łączą dostawców i odbiorców dzięki integracji rozproszonych wcześniej w różnych systemach procesów zakupowych. Największe korzyści z ich implementacji mogą odnieść organizacje, które korzystają już z szerokiego zakresu narzędzi ERP. Konsolidacja ERP z SRM daje możliwość udostępnienia dostawcy danych o stanach magazynowych lub planach produkcyjnych – dzięki czemu można powierzyć mu uzupełnianie materiałów zgodnie z ustalonymi potrzebami. Połączenie modułu SRM z systemem dostawcy, czyli integracja zewnętrzna – o ile korzysta on z narzędzi umożliwiających elektroniczną wymianę danych – to kolejne uproszczenie procesów zakupowych pozwalające na szybką wymianę danych i minimalizację błędów wprowadzanych przez użytkownika.

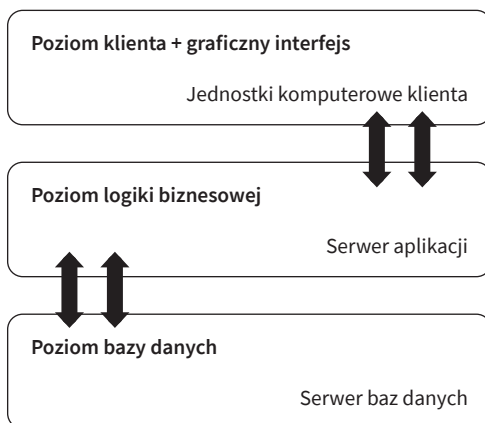
Ważną funkcją systemu SRM, optymalizującą proces zaopatrzenia operacyjnego, są tzw. zamówienia samoobsługowe, często wykorzystywane przy zamówieniach pośrednich – pozwalają one wykluczyć operacyjną rolę działu zakupów z procesu. Uprawniony do zakupów pracownik może samodzielnie dokonać zakupu poprzez łatwo dostępny katalog produktów z cenami wynegocjowanymi wcześniej. W zamówieniach bezpośrednich, które wymagają planu, SRM wspomaga system ERP, gdyż zapotrzebowanie zarejestrowane w ERP, na podstawie planów produkcyjnych, kierowane jest do SRM jako zamówienie zewnętrzne, gdzie możliwe jest natychmiastowe wybranie źródła dostawy i złożenie zamówienia. Do głównych funkcji wspieranych przez SRM należą [Capgemini, 2017]:

- analizy zaopatrzeniowe,
- zarządzanie cyklem życia produktu,
- zaopatrzenie strategiczne,
- zarządzanie danymi dostawcy,
- zarządzanie wydajnością dostawcy,
- zarządzanie kontraktami,
- zarządzanie katalogami,
- zaopatrzenie operacyjne,
- wydatki na zasoby zewnętrzne.

Rozwój mediów społecznościowych spowodował rozszerzenie funkcjonalności systemów SRM o mechanizmy, takie jak: rekomendacje, oceny społeczności branżowej czy tzw. *crowdsourcing*. Istotnym trendem jest też dostęp do aplikacji za pomocą urządzeń mobilnych, co pozwala na szybkie realizowanie procesów zatwierdzania zamówienia, negocjacji i komunikacji.

## 2.2.5. Technologie systemów ERP

Tradycyjne systemy ERP najczęściej pracują w technologii klient–serwer, opierając się na warstwowej architekturze, której centrum stanowi serwer bazodanowy. Typowa architektura przedstawiona została na rysunku 2.5.



**Rysunek 2.5.** Architektura klient–serwer systemu ERP

**Źródło:** opracowanie własne.

W takiej architekturze za dostępność poszczególnych funkcjonalności, według przyjętej logiki biznesowej, odpowiada serwer aplikacji. Integralność i spójność danych jest zapewniona na poziomie baz danych (odpowiada za nią konkretny serwer), a dający możliwość wykonywania podstawowych operacji graficzny interfejs użytkownika wyświetla na jednostkach komputerowych żądane informacje.

Dostępne w systemach informatycznych wersje architektury klient–serwer mogą posiadać różny zakres funkcjonalny [ERP-VIEW.pl, 2019]:

- prosty – umożliwiający zdalną prezentację danych,
- wzbogacony – pozwalający na wstępne przetwarzanie danych w zdalnej stacji roboczej,
- rozbudowany – wspierający formułowanie zapytań i przygotowanie odpowiedzi w stacji roboczej (w wersji zaawansowanej to rozproszona baza danych).

Zintegrowany system informatyczny ERP, który umożliwia użytkownikowi uruchomienie wybranej funkcji systemu z dowolnej stacji roboczej, posiada określone cechy [Lech, 2003; Bytniewski, 2005]:

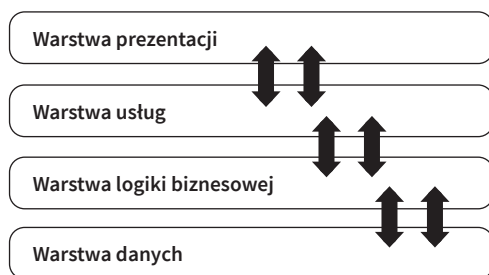
- użytkownik stacji roboczej może uruchomić dowolną funkcję systemu zgodnie z przyznanymi uprawnieniami w realizacji procesu biznesowego;
- interfejs w obszarze całego systemu jest jednolity;
- dane są wprowadzane tylko raz oraz automatycznie aktualizują stan systemu; okazują się także widoczne i dostępne dla wszystkich użytkowników tego systemu zgodnie z przydzielonymi uprawnieniami;
- zapewnia kompleksowość funkcjonalną – system swoim zakresem obejmuje wszystkie obszary działalności techniczno-ekonomicznej organizacji;
- zapewnia integrację danych i procesów – wymiana danych wewnątrz systemu (pomiędzy poszczególnymi modułami) oraz z jego otoczeniem;
- pozwala na elastyczność strukturalną i funkcjonalną – maksymalne dostosowanie rozwiązań sprzętowo-programowych do potrzeb przedsiębiorstwa w chwili instalowania i uruchamiania systemu; umożliwia dynamiczne dopasowanie systemu przy zmiennych wymaganiach i potrzebach generowanych przez otoczenie;
- jest otwarty – gwarantuje możliwość rozszerzania systemu o nowe moduły;
- posiada skalowalną architekturę oraz daje możliwość tworzenia połączeń z systemami zewnętrznymi;
- jest zaawansowany merytorycznie i oferuje pełne informatyczne wspomaganie procesów informacyjno-decyzyjnych;
- jest zaawansowany technologicznie i gwarantuje zgodność z aktualnymi standardami sprzętowo-programowymi oraz daje możliwość migracji na nowe platformy sprzętu komputerowego, systemów operacyjnych, mediów i protokołów komunikacyjnych;

- jest zgodny z przepisami, np. z ustawą o rachunkowości, a w szczególności z zasadami prowadzenia ksiąg rachunkowych z wykorzystaniem technologii informatycznej, ustalania i raportowania wyników finansowych przedsiębiorstwa oraz sporządzania sprawozdań finansowych itp.

W pierwszej dekadzie XXI tradycyjna architektura systemów ERP zaczęła być zastępowana architekturą opartą na usługach (*Service Oriented Architecture* – SOA) – co zmieniło ideę sposobu wspierania procesów biznesowych przez aplikacje. SOA nie dotyczy konkretnej technologii, czyli okazuje się raczej filozofią niż metodą projektowania. To pojęcie związane z infrastrukturą niezbędną do obsługi usługi (w tym przypadku usługa rozumiana jest jako odrębny fragment rozwiązania informatycznego wykonującego konkretne zadanie – pakiet ściśle powiązanych standardowych funkcji). Zgodnie z podejściem SOA moduł konkretnej usługi może być umieszczony w dowolnym miejscu architektury i wywoływany przez inne moduły jej oprogramowania, co pozwala na stworzenie systemu informatycznego z powiązanych ze sobą modułów funkcjonalnych różnego pochodzenia [ERP-VIEW.pl, 2019]. Główną ideą leżącą u podstaw stosowania architektury SOA okazuje się więc wielokrotne korzystanie z dobrze zdefiniowanych usług wywoływanych przez różne, luźno połączone komponenty IT. Takie rozwiązanie zapewnia niezależność od wykorzystywanej platformy, technologii i języków programowania. Przykładem takich usług mogą być: wykonanie fragmentu logiki biznesowej, transformacja danych, zapytanie do baz danych, wysłanie komunikatu. Dodatkowo usługa internetowa rozumiana jest jako usługa, do której można uzyskać dostęp za pomocą standardowych protokołów internetowych i opartych na XML.

W przypadku systemów tradycyjnych procesy biznesowe są obsługiwane w sposób wymuszony działaniem aplikacji, np. poprzez narzucenie kolejności wykonywanych działań. Mogą one nie być niezbędne dla obsługi procesu biznesowego, ale wynikają z konieczności zapewnienia prawidłowego działania programu.

Architektura SOA pozwala organizacji określić, w jaki sposób aplikacja ma działać, jak powinien przebiegać proces, a nawet jak zaprojektować poszczególne widoki ekranu. Wspomagająca proces biznesowy usługa IT może zostać wywołana za pomocą dowolnego urządzenia: interfejsu użytkownika, portalu internetowego, aplikacji mobilnej itp. Przykład podstawowej architektury SOA dla systemu ERP został pokazany na rysunku 2.6.



**Rysunek 2.6.** Architektura SOA ERP

**Źródło:** opracowanie własne.

Rozważając jako przykład sklep internetowy, można zauważyć, że w zakresie wsparcia procesów sprzedaży SOA pozwala klientowi wybrać z zaprezentowanego na stronie katalogu produktów/usług interesujący towar, dodać go do koszyka i dokonać akceptacji zamówienia. Przedstawienie oferty może być często zmieniane, zależnie od marketingowych koncepcji i prowadzonych kampanii. Modyfikacjom mogą podlegać drobne elementy, takie jak np. kolorystyka i czcionka, umieszczenie opisu nad lub pod produktem czy wykorzystanie efektów 3D dla wizualizacji produktu. Przeobrażenia te nie wpływają jednak na główne procesy biznesowe. Usługi informatyczne, które pozyskują dane do katalogów np. z systemów ERP i pozwalają przeprowadzić proces zamówienia, zostają takie same, dzięki oddzieleniu ich od sposobu prezentacji.

Zmiany w procesach biznesowych są nieuniknione. W miarę rozwoju firma dodaje poziomy złożoności i alternatywne przebiegi. Wspierające je usługi informatyczne nie ulegają silnym modyfikacjom – zmienia się jedynie sposób ich wykorzystania w ścieżkach procesów biznesowych. Jeśli firma wprowadza do sprzedaży produkt, który wymaga nieco innej ścieżki postępowania w procesie sprzedaży (np. produkty alkoholowe lub leki na receptę), to w procesie sprzedaży potrzebne będzie wywołanie usługi weryfikacji wieku klienta lub numeru recepty. W sytuacji, kiedy takie usługi weryfikacyjne już istnieją w systemie procesu zamówienia, będą one wywoływane w odpowiednim kroku procesu. W przeciwnym razie należy je zaprojektować i wprowadzić do katalogu usług. Decyzja o wykorzystaniu usługi weryfikacji nie jest powiązana z warstwą prezentacji, a z obsługującym konkretne zamówienie procesem biznesowym.

Odseparowanie procesów biznesowych od usług i warstwy prezentacji, czyli zastosowanie architektury SOA, ma wpływ na uproszczenie procesów integracji – co w przypadku systemów ERP pozwala na zwiększanie funkcjonalności poprzez dołączanie aplikacji uzupełniających od innych producentów. Wykorzystanie SOA nie ogranicza się tylko do systemów klasy ERP – używa się jej również w innych obszarach i procesach biznesowych, takich jak np. CRM lub SCM.

Ogromna popularność urządzeń mobilnych oraz wygoda korzystania z nich powodują, że zarówno pracownicy, jak i klienci używają ich nie tylko w celach prywatnych, ale i biznesowych – w wielu lokalizacjach, o dowolnych porach. Funkcjonalności udostępniane na urządzenia mobilne w obszarze sprzedaży i CRM poprawiają efektywność pracy i skracają czas komunikacji. Łatwość dostępu do obszarów HCM pozwala pracownikom na samodzielne, sprawne załatwianie spraw związanych z zatrudnieniem, zwiększaniem kompetencji czy rozliczaniem delegacji. W obszarze produkcji i dystrybucji mogą oni być informowani na bieżąco o zaistniałych problemach czy sytuacjach awaryjnych – oraz zdalnie je rozwiązywać.

Aby firma mogła czerpać kompleksowe korzyści z mobilności, aplikacje na urządzenia przenośne powinny być zintegrowane z systemami transakcyjnymi (np. ERP, CRM) lub analitycznymi (*Business Intelligence*). Do tego należy umożliwiać szybkie, przejrzyste i intuicyjne zarządzanie danymi.

W niemal każdej organizacji można wyróżnić procesy i działania, które nie zmieniają się zbyt często: firmy produkcyjne sprzedają przez dłuższy okres określone produkty, usługowe koncentrują się na sprzedaży oferty usług. Niektóre z elementów działania i oferty firm podlegają jednak częstszym przeobrażeniom, np. prezentacja katalogu aktualnych produktów, cenniki, kampanie reklamowe.

Oczekiwaniem biznesu w stosunku do IT jest niemal natychmiastowe skalowanie możliwości i odpowiedzi na zmiany wymagań rynkowych. Modyfikacje dotyczą zwykle trzech wymiarów: liczby przetwarzanych danych, liczby interakcji między użytkownikami (aplikacjami/klientami) oraz zmiany w funkcjonalnościach aplikacji [Bluesoft, 2018]. Biznes oczekuje szybkiego i sprawnego zaimplementowania przeobrażeń, aby osiągać zakładane cele. Zmiana zakresu przetwarzania danych wiąże się z masowym gromadzeniem danych (w tym danych z mediów



społecznościowych) oraz rozwojem IoT i technologiami Big Data. Liczba interakcji dotyczy natomiast użytkowników i częstotliwości korzystania z aplikacji i usług, z kolei zmiany w funkcjonalnościach są wymaganiem pojawiającym się wraz z rozwojem firmy, ofertą nowych produktów i tworzeniem nowych modeli biznesowych [tamże].

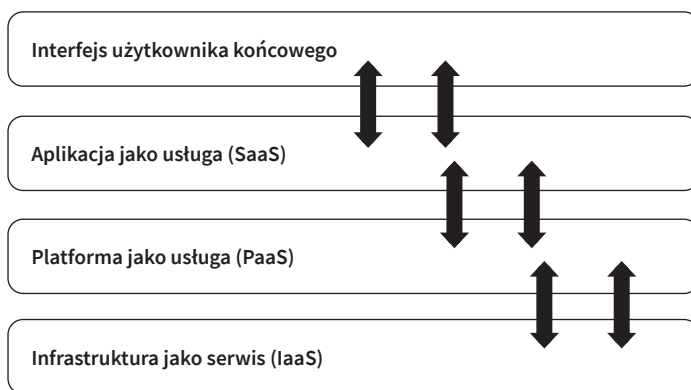
W celu szybkiej reakcji na zmiany i zapotrzebowania rynku (np. wprowadzenie krótkiej promocji) prostsze staje się tworzenie małych, łatwiejszych w zarządzaniu aplikacji. Zaletą podzielenia aplikacji na mniejsze usługi jest poprawa jej modularności i ułatwienie zrozumienia każdej z nich, co przekłada się na tempo ich rozwoju. Taki sposób projektowania został nazwany „architekturą mikrousług” (*Microservices Architecture* – MSA) – zakłada ona, że struktura systemu bazuje na zbiorze luźno powiązanych usług, rozumianych jako oddzielnie uruchamiane programy komputerowe.

System złożony z małych usług, z których każdą można uruchomić na niemal każdym rodzaju infrastruktury, pozwala natychmiastowo reagować na zwiększone zapotrzebowanie równolegle obsługiwanych interakcji. Istotną częścią tego rozwiązania jest możliwość zwiększania i zmniejszania liczby serwerów w zależności od potrzeb, co w praktyce wymaga korzystania z infrastruktury w chmurze obliczeniowej [tamże].

## **2.2.6. Wdrożenie i zmiany w implementacji systemów klasy ERP**

Implementacja systemu może przebiegać w modelu tradycyjnym (*on-premise*), całkowicie w środowisku chmury lub jako model hybrydowy, gdzie część funkcjonalności dostępna jest *on-premise*, a część w modelu chmury.

Model świadczenia usług w chmurze ma strukturę warstwową (rysunek 2.7). Na poziomie infrastruktury zdefiniowane są zestandaryzowane podstawowe usługi dotyczące przetwarzania, przechowywania czy udostępniania sieciowego – tu dostawcy mogą oferować usługi dotyczące systemów operacyjnych. Warstwa platformy dostarcza usługi związane z wytwarzaniem oprogramowania, testowaniem i utrzymaniem. Z kolei poziom aplikacji to odpowiednie funkcjonalności biznesowe; warstwa interfejsu użytkownika pozwala na komunikację.



**Rysunek 2.7.** Warstwowa architektura świadczenia usług w chmurze obliczeniowej

**Źródło:** opracowanie własne.

Za ERP w chmurze uważa się model, w którym w pełni licencjonowane oprogramowanie jest dostępne poprzez usługę hostingu – również dostępną w środowisku chmury. Obecnie określenie to odnosi się także do oprogramowania oferowanego jako usługa (SaaS). SaaS ERP to model, w którym dostawca jest właścicielem oprogramowania (*host off-site*), a klient płaci za wykorzystywane usługi. SaaS ERP jest zatem typem modelu środowiska chmury, przy czym nie każdy system ERP w chmurze musi być oferowany jako model SaaS. Cele stosowania systemu ERP są takie same, bez względu na typ wykorzystanego hostingu; wybór modelu musi zostać przez organizację starannie rozważony przed implementacją, z wzięciem pod uwagę szeregu aspektów – każdy z modeli niesie za sobą bowiem jednocześnie szereg korzyści i barier [Czerwotka i inni, 2018].

Rozwój chmury obliczeniowej zmienia sposób, w jaki organizacje korzystają z systemów informatycznych. Dotychczasowi wiodący dostawcy systemów ERP oferują wersje systemów w postaci SaaS z propozycją migracji dotychczasowych rozwiązań, zmuszając niejako klientów do przejścia na ten model wykorzystywania systemu. Istnieje możliwość rozwiązań hybrydowych. Pojawia się wielu mniejszych dostawców proponujących produkty tylko w wersji chmury. Każdy model implementacji daje zarówno korzyści, jak i przynosi ograniczenia. W tabeli 2.2 zaprezentowano wady i zalety różnych aspektów implementacji systemów ERP.

**Tabela 2.2.** Zalety i wady modeli implementacji systemów ERP

Aspekt	SaaS ERP/ model chmury ERP	on-premise
zakres, nadzór nad systemem i danymi	+ duża skalowalność  – ograniczony	+ wyższy stopień kontroli nad systemem i danymi  – wymagany znaczny udział firm zewnętrznych przy implementacji, co zwiększa koszt rozwiązania – wymagane wysokie kompetencje zespołu IT
implementacja, dostosowanie (kustomizacja)	+ szybka implementacja  – dostosowanie ograniczone, co może mieć wpływ na implementację unikalnych funkcjonalności, a tym samym na przewagę konkurencyjną	+ szeroki zakres kustomizacji (dopasowania systemu dla potrzeb klienta)  – wysokie koszty
integracja	+ łatwiejsza: wykorzystanie interfejsu API lub innych standardów ma wpływ na obniżenie kosztów  – integracja wymaga ustalenia wielu umów SLA; w przypadku, gdy integrowane systemy są dostarczane przez różnych dostawców, należy rozważyć problem przechowywania danych w sytuacji awarii jednego z systemów	+ na poziomie architektury klient-serwer
aktualizacja i dostęp do nowych wersji	+ częsta, w niektórych przypadkach nawet cotygodniowa; pozwala szybko nadążać za zmiennymi regulacjami + podnosi skalowalność + szybki dostęp do nowych technologii	– dłuższy i kosztowny dostęp do nowych rozwiązań – częste aktualizacje mogą generować błędy dla rozwiązań z wysokim stopniem kustomizacji
dostęp poprzez urządzenia mobilne	+ ze względu na naturę internetową są natywnie mobilne i wyposażone w standardowe aplikacje mobilne	– oferowane jako rozszerzenia do systemu muszą być starannie sprawdzone, zwłaszcza w przypadku, gdy są proponowane jako usługi innych dostawców
bezpieczeństwo i niezawodność	+ wzmocnienie bezpieczeństwa poprzez szyfrowanie i używanie dodatkowych zabezpieczeń zbudowanych na platformach, takich jak usługa Amazon Web Services	– zapewniane wewnętrznie przez organizację, generuje znaczny koszt, wymaga kompetencji i przygotowania procedur na wypadek ataków

Aspekt	SaaS ERP/ model chmury ERP	on-premise
koszt licencji	<ul style="list-style-type: none"> <li>+ model wyceny oparty na subskrypcji miesięcznej lub rocznej</li> <li>– złożony model wyceny opierający się na wyborze wielu czynników, takich jak liczba użytkowników, transakcje, ilość danych lub inne jednostki miary</li> </ul>	<ul style="list-style-type: none"> <li>– koszt licencji plus koszty bieżącego wsparcia; ceny negocjowane indywidualnie</li> </ul>
koszt wdrożenia	<ul style="list-style-type: none"> <li>+ niższe koszty początkowe i niższe koszty operacyjne</li> <li>+ wykorzystanie ekonomiki skali pozwala na redukcję cen usług przez dostawcę; niemniej ceny całego wdrożenia i subskrypcji muszą być starannie przeanalizowane</li> <li>+ możliwość wprowadzenia oszczędności poprzez rezygnację z określonych aktualizacji</li> <li>+ korzystanie z usług wsparcia dla systemu, także w modelu chmury</li> <li>+ krótszy czas i koszt wdrożenia (głównie z tytułu redukcji zapotrzebowania na specjalistyczne usługi i personel)</li> <li>– koszty przygotowania i dodatkowe koszty bieżące</li> </ul>	<ul style="list-style-type: none"> <li>– wdrożenie wymaga zwykle korzystania ze specjalistycznych usług firm zewnętrznych</li> <li>– większy koszt początkowy inwestycji</li> </ul>

**Źródło:** [Czerwotka i inni, 2018].

Nowe technologie i aplikacje ICT pozwalają na coraz efektywniejsze wspomaganie procesów biznesowych i osiąganie nowych korzyści biznesowych, co jest jednym z głównych czynników decydujących o wprowadzeniu lub wymianie posiadanego przez organizację systemu informatycznego. Wśród najważniejszych powodów podejmowania tej decyzji menedżerowie wymieniają większą funkcjonalność oraz zmianę strategii IT organizacji. Wiele firm podkreśla również takie kwestie, jak: potrzeba integracji, rozwój organizacji i strategia działań producenta aktualnie posiadanego oprogramowania, potrzeba implementacji szerokiego zakresu funkcji analitycznych, restrukturyzacja przedsiębiorstwa, wymuszający migrację posiadanego systemu na inny (wybrany przez korporację) *rollout* korporacyjny, specyfika branżowa lub zmiany kapitałowe w przedsiębiorstwie.

Wdrożenie systemu informatycznego, bez względu na przyjęty model dostępu i metodykę implementacji, to projekt, w który powinni

być zaangażowani użytkownicy końcowi oraz menedżerowie wyższych szczebli zarządzania, kontrolujący na bieżąco przebieg procesu. Już sam wybór systemu okazuje się równie skomplikowany i powinien przede wszystkim koncentrować się na spełnianiu wymagań biznesowych i na długoterminowych korzyściach z inwestycji. Do czynników, które firma musi rozważyć, należą: cena, aktualna technologia systemu, wykorzystanie i przystosowanie do branży, koszt całego wdrożenia – w tym szkoleń pracowników.

Podstawą do osiągnięcia sukcesu przy wdrażaniu strategii IT jest zgodność ze strategią biznesową organizacji i racjonalne podejmowanie decyzji o informatyzacji. Należy jasno określić cele, jakie firma chce osiągnąć przy pomocy systemu informatycznego i poprzeć wybór różnymi analizami oraz ocenami. Wprowadzanie zmiany w organizacji w postaci wdrożenia rozwiązań ICT umożliwia firmom nie tylko wspomaganie i optymalizację aktualnie określonych procesów, ale także pozwala na znajdowanie usprawnień w postaci zmiany wykonywanej pracy i dotychczasowych procesów [Lech, 2003].

Wdrożenie systemu klasy ERP wiąże się z ryzykiem niepowodzenia, dlatego dla procesu wyboru i implementacji ustanawia się dedykowane projekty. Sama struktura systemu nie zawsze pozwala na łatwe przystosowanie go do indywidualnych potrzeb organizacji. Złożoność oferowanych systemów ERP zwykle powoduje, iż klienci mają tendencję do wdrażania większej funkcjonalności niż wymagana, co ma wpływ na zwiększenie kosztów i kompleksowość procesów. O kosztach implementacji decyduje także przyjęty model *on-premise* lub usługi w chmurze. W obu przypadkach należy przeprowadzić szczegółową analizę samego systemu, zakresu jego funkcjonalności, modelu i dostawcy. Szacowanie zwrotów z inwestycji i wyliczanie konkretnych korzyści z implementacji systemu są niezwykle trudne. Organizacje korzystają z różnych metod. Jedna z nich polega na policzeniu, w jakim czasie zwraca się inwestycja w system – dłuższy okres i znaczny wkład finansowy wiążą się z podniesieniem stopnia ryzyka i zmniejszeniem oczekiwanych rezultatów. Druga oparta jest na rozważaniach alternatywnych metod inwestycji, gdzie przychody generowane po wdrożeniu systemu ERP porównuje się z innymi rozwiązaniami, które w tym czasie firma mogła zastosować (np. depozyt bankowy, inwestycje w park maszynowy). W praktyce szacunki wyrażone w walucie mają sens, gdy tempo rozwoju firmy nie jest

wysokie. Innym sposobem oceny inwestycji jest analiza efektów wdrożenia systemu w postaci identyfikacji obszarów usprawnionych dzięki wdrożeniu systemu ERP i sprawdzeniu, w jakim stopniu zostały zrealizowane cele – sprawdza się tu analiza scenariuszowa funkcjonowania firmy bez rozważanej implementacji [CEO.com.pl, 2017]. Nie można zapominać także o istotnym czynniku ryzyka wdrożenia systemu, czyli wyborze dostawcy.

Podstawowym zadaniem firmy przed wyborem rozwiązania IT jest analiza potrzeb organizacji, która powinna być przeprowadzona zgodnie z określonymi zasadami, wypracowanymi zarówno przez teoretyków, jak i praktyków, oraz obejmować następujące kroki [BPSC, 2017]:

1. Określenie celu – wdrożenie rozwiązań z obszaru IT musi być poprzedzone szczegółową analizą celów ich wykorzystywania w najbliższym okresie i przewidywanej rozbudowy, w tym integracji z innymi rozwiązaniami w przyszłości.
2. Przygotowanie zespołu – w skład zespołu powinni wchodzić, oprócz osób z zarządu organizacji, także potencjalni użytkownicy systemu.
3. Mapowanie procesów – analizie należy poddać nie tyle procesy obszaru, którego dotyczy wdrażane rozwiązanie, co wszystkie procesy organizacji. Etap ten pozwala na optymalizację i zmianę procesów; jest skomplikowany i wymaga wykorzystania narzędzi informatycznych.
4. Określenie harmonogramu i okresu – wdrożenie systemów ERP może przebiegać według przyjętego modelu, o czym decydują takie czynniki, jak: rozmiar organizacji, jej typ, liczba procesów. Okres implementacji waha się od kilku miesięcy do kilku lat. W tym czasie należy przewidzieć momenty na przygotowanie procesów, migracje danych, testowanie i szkolenie pracowników.
5. Możliwość łatwej modyfikacji i elastyczność systemu – co pozwala na szybkie przystosowanie systemu do zmian otoczenia.

Nabywcy i użytkownicy systemów ERP dążą do pozyskania systemu cechującego się wysoką wygodą użytkowania i prostym interfejsem użytkownika. Świadczą o tym m.in. badania przeprowadzone w Polsce wśród małych i średnich przedsiębiorstw, które wskazały, że to najbardziej pożądane właściwości systemu. Ponadto użytkownicy chcą, aby system charakteryzował się elastycznością i stabilnością pracy, a dostawca oferował wsparcie techniczne.

## **2.2.7. Analiza informacji gromadzonych w systemach informatycznych**

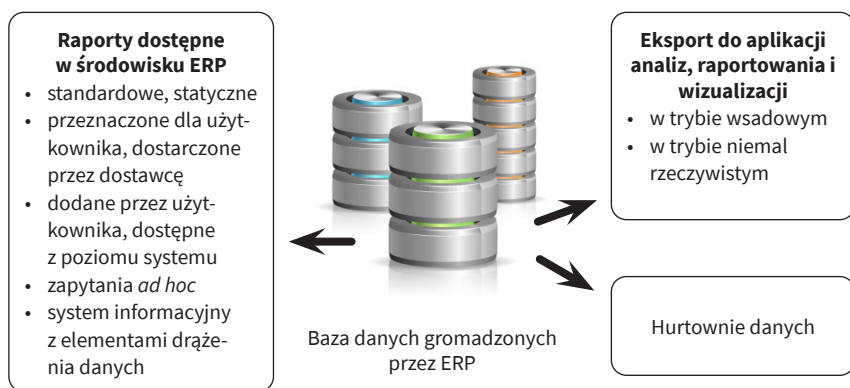
W nowoczesnej gospodarce informacja jest postrzegana jako zasób niezwykle istotny. Integracja systemów informatycznych powoduje, że wszyscy użytkownicy posiadają aktualny obraz obsługiwanych procesów w czasie niemal rzeczywistym, a jednocześnie minimalizowane są nakłady pracy dotyczące wielokrotnego wprowadzaniem tych samych danych w różnych miejscach systemu. Możliwości innowacyjne IT polegają na udostępnianiu, przesyłaniu i analizie danych, co tworzy nową jakość informacji [Davenport, 1993; Lech, 2003]. Dzięki ICT informacja jest aktualna, dostępna, niezależna od lokalizacji oraz elastyczna. Przedsiębiorstwo powinno umiejętnie wykorzystywać i interpretować dane w celu lepszego dopasowania swojej działalności i jej efektów do potrzeb klientów – oraz dokładnego ich poznania.

Obecnie w wielu organizacjach zintegrowane systemy ERP stanowią główne źródło danych i pozyskiwania informacji o zasobach i stanie realizacji procesów. Jednym z podstawowych zadań, jakie pełnią, jest dostarczanie aktualnych i dopasowanych do potrzeb użytkowników raportów i wskaźników. Rozwój technologii sieciowych i internetowych spowodował, iż dane mogą być przesyłane w miejsca oddalone znacznie od faktycznej lokalizacji serwerów aplikacji i serwerów baz danych. Jednocześnie rozwój mediów społecznościowych spowodował, iż informacje z tych źródeł są dla organizacji równie istotne.

Systemy informatyczne ERP dają wiele możliwości w zakresie raportowania danych. Standardowy zestaw raportów jest zwykle dostarczany wraz systemem, natomiast dodatkowe zestawienia są tworzone zgodnie z zapotrzebowaniem klientów. Ograniczają się one do informacji zawartych w bazie danych systemu. Systemy posiadają funkcje pozwalające na tworzenie zapytań i dołączanie raportów sporządzanych za pomocą innych narzędzi, a wbudowany system informacyjny umożliwia drążenie danych. Z kolei wbudowane narzędzia są ściśle powiązane ze strukturą danych systemów. Raporty pozwalają analizować bieżące i okresowe wyniki, np. sprzedaż za dany okres; listę zamówień niezakończonych, niezapłaconych faktur, zamówień na dany produkt/usługę za dany okres; rotację towarów w magazynie itp. Każdy moduł wyposażony jest w szereg raportów przystosowanych do obsługiwanych procesów biznesowych.

Dodatkowo systemy posiadają funkcjonalność pozwalającą na pozyskiwanie danych z baz SQL za pomocą języka SQL lub prostych kreatorów. Wygodną – dla określenia stanu procesu biznesowego – okazuje się opcja umożliwiająca śledzenie dokumentów powstałych w trakcie jego obsługi, np. wszystkich związanych z przetwarzaniem zamówienia klienta.

Oprócz narzędzi dostarczanych przez dostawców systemów ERP, dostępne są także narzędzia dostawców zewnętrznych, pozwalające na tworzenie raportów i analizowanie zgromadzonych danych. Zwykle wymagają implementacji struktury tabel z systemu ERP. Zaletą tego typu systemów jest możliwość łączenia danych z różnych aplikacji (rysunek 2.8).



**Rysunek 2.8.** Raportowanie z danych gromadzonych przez systemy ERP

**Źródło:** opracowanie własne.

Obejmowanie coraz większych obszarów i gromadzenie znacznej ilości danych spowodowało wzrastające zapotrzebowanie na ich analizę. Do analizy danych w organizacji wykorzystywane są narzędzia analityczne oparte na idei hurtowni danych oraz eksploracji danych (zob. podrozdział 3.4 niniejszej monografii). W celu zapewnienia szybkiego i prostego transferu danych w systemach ERP zaczęły pojawiać się interfejsy do najczęściej wykorzystywanych systemów analitycznych.

Rozwój aplikacji obsługujących różne obszary biznesowe pokazuje, jak były one integrowane, co zapewniało szersze wsparcie działań i dostęp do gromadzonych danych. Poziomą integrację wewnętrzną odzwierciedla konsolidację procesów biznesowych z systemami do tej pory funkcjonującymi niezależnie – przyczynia się to do większej spójności gromadzonych



informacji. Z kolei integracja zewnętrzna umożliwia połączenie systemów organizacji z systemami innych podmiotów gospodarczych i administracji publicznej poprzez technologie wymiany danych.

Wyzwaniem dla raportowania danych z systemu będzie w tym przypadku zestawienie w jednym sprawozdaniu finansowych i niefinansowych danych oraz analiz. Ani publikowane sprawozdania finansowe przedsiębiorstwa, ani coraz bardziej popularne sprawozdania o zrównoważonym rozwoju nie dadzą pełnego obrazu wartości i wyników przedsiębiorstwa – dopóki oba te aspekty nie zostaną zintegrowane w kompletnym i spójnym raporcie [Odenwald, 2014].

W organizacjach – oprócz systemów ERP – do wspierania procesów biznesowych często wykorzystywane są inne systemy informatyczne. Gromadzenie danych oraz zapewnienie wykonywania na nich działań może być połączone z systemami wspomagania pracy grupowej i systemami obsługującymi przepływy pracy [Januszewski, 2008].

Głównym problemem integracji systemów informatycznych w ramach danej organizacji lub partnerów w łańcuchu dostaw jest stosowanie różnego rodzaju formatów danych i nośników informacji. Sama różnorodność systemów informatycznych zaimplementowanych u partnerów biznesowych okazuje się wyzwaniem dla integracji, ponieważ wymusza wprowadzanie zmian organizacyjnych, zasad przepływu informacji i ich dokumentacji, a zwykle również dokonanie modyfikacji w procesie funkcjonowania używanego systemu informatycznego [Lech, 2003]. W celu optymalizacji procesów biznesowych na etapie analizy i wdrożenia systemu ERP wykorzystywane są często aplikacje służące do modelowania procesów biznesowych. Modele zaimplementowanych procesów udostępniane są w systemach ERP, np. w systemie pomocy i *e-learningu*, przez co umożliwiają ciągły dostęp do szczegółowych informacji o procesach, ich wzajemnym powiązaniu oraz przyjętych miarach efektywności.

Obszarem, który ściśle wiąże się z systemami ERP, okazuje się sfera relacji z klientami i analiz sprzedaży. Niektóre z systemów ERP zawierają elementy systemów CRM, niemniej jednak organizacje zwykle implementują systemy CRM o szerokim zakresie funkcjonalności, pozwalającym na prowadzenie segmentacji projektów dla klienta i kampanii reklamowych. Równie często wdrażane są programy służące do badania sprzedaży, umożliwiające dokonywanie zaawansowanych analiz z wykorzystaniem szerokiej gamy narzędzi statystycznych i metod wizualizacji [zob. Gontar, 2019].

Podstawowym elementem gwarantującym uzyskanie odpowiednich informacji z systemu ERP jest baza danych, do której dostęp mają użytkownicy systemu za pośrednictwem odpowiednich funkcjonalności jego modułów. Zaletą systemu ERP okazuje się centralny proces przetwarzania danych, zapewniający użytkownikom z odpowiednimi prawami dostępu aktualną i rzetelną informację, niezbędną do obsługi procesu.

Tradycyjnie systemy klasy ERP udostępniane były w postaci aplikacji kompilowanych za pomocą niskopoziomowych języków programowania ogólnego zastosowania. Obecnie systemy muszą oferować szybką możliwość implementacji zmian i wprowadzania nowych funkcjonalności – odpowiednio do zmieniających się wymagań biznesu. Ze względu na wysoki poziom złożoności duże systemy ERP posiadają własne specjalizowane platformy rozwoju, pozwalające na [Wojtkowski, 2010]:

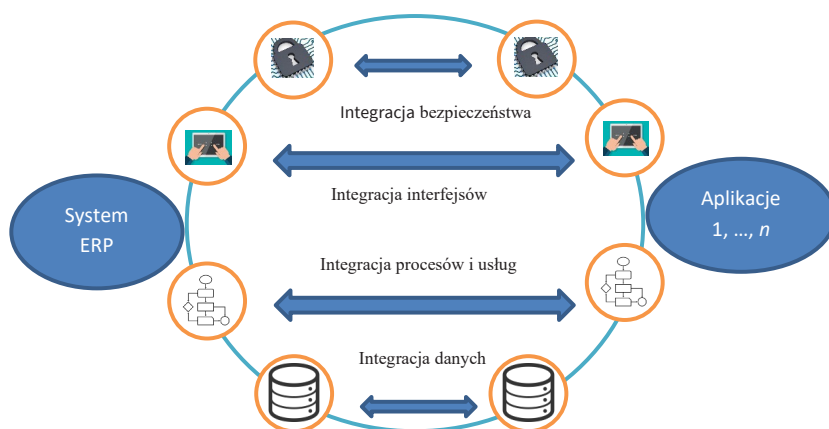
- zarządzanie parametryzacją i złożonością wymaganej logiki funkcjonalności biznesowej oraz kontrolę i utrzymanie poszczególnych obszarów funkcjonalnych;
- elastyczne i szybkie wprowadzanie zmian bez konieczności kompilacji całej aplikacji;
- integrację wielu aplikacji, źródeł danych, plików z danymi – zgodnie ze standardami dostępnych interfejsów danych, komunikacji czy technologii;
- wsparcie samego procesu wdrożenia poprzez odpowiednią metodologię obejmującą takie etapy, jak: określenie notacji wymagań, analizy przedwdrożeniowe, wersjonowanie obiektów, zarządzanie zmianami, dokumentacja, zarządzanie zgłoszeniami i czynnościami utrzymania i rozwoju systemu.

Integracja systemu ERP z innymi systemami, np. z CRM, wymusza analizę procesów biznesowych i podjęcie decyzji, jakiego rodzaju wspomaganie informatyczne wymaga ich obsługa. Planowanie integracji wiąże się z rozważeniem następujących zagadnień:

- jakie aplikacje, oprócz CRM i modułów ERP, będą włączone w proces integracji;
- gdzie przechowywane są dane;
- jaki jest oczekiwany czas reakcji na zapytanie – czy możliwa jest aktualizacja raz dziennie z rozwiązaniem wsadowym, czy też wymagany jest dostęp w trybie niemal rzeczywistym do niezależnych źródeł danych.

Model integracji z innymi aplikacjami (rysunek 2.9) musi zatem rozważać szereg jej aspektów:

- integrację danych poprzez wykorzystanie różnych metod i protokołów (np. import, eksport, transfer plików CSV, API, integrację z *master data* itp.);
- integrację na poziomie logiki biznesowej, wraz z jej procesami – pozwalającą na całościową obsługę procesów biznesowych niezależnie od modelu, w jakim poszczególne działania są wspierane przez system (SaaS, *cloud*, *on-premise*) organizacji lub systemy jej partnerów;
- integrację usług wspierających procesy biznesowe;
- interfejs użytkownika (UI, UX dla aplikacji mobilnej) – umożliwiający kontekstowe wykorzystanie funkcjonalności podczas pracy na różnych urządzeniach;
- integrację bezpieczeństwa, w ramach której można wykorzystać już istniejące w organizacji rozwiązania centralnego systemu logowania.



**Rysunek 2.9.** Obszary integracji aplikacji

**Źródło:** opracowanie własne.

Wiele systemów informatycznych zarządzania wykorzystuje relacyjne bazy danych [zob. Gontar, 2019]. Obecnie, poza bazami relacyjnymi, istnieją także inne możliwości przechowywania danych, np. bazy danych tekstowych, grafowych czy przechowywanie i przetwarzanie danych w pamięci (*in-memory*).

Rozwój chmury obliczeniowej umożliwił także udostępnianie platform jako usług, czego przykładem jest SAP HANA Platform. To otwarta platforma dostępna w trybie *on-premise* lub w module SaaS, przeznaczona do tworzenia aplikacji biznesowych w środowisku, które zawiera zestaw kompleksowych usług i narzędzi pozwalających zmieniać, tworzyć, rozszerzać i integrować aplikacje biznesowe w chmurze. Podstawą pracy okazuje się tu innowacyjny system zarządzania relacyjnymi bazami danych w pamięci (*in-memory processing*) [SAP.com, 2017]. SAP HANA przetwarza dane jednocześnie na wielordzeniowych procesorach – wszystkie dane są dostępne w pamięci głównej, co pozwala uniknąć obniżenia wydajności poprzez minimalizację operacji wejścia/wyjścia dysku. W tradycyjnych systemach baz danych tabela jest dwuwymiarową strukturą danych zorganizowaną w wierszach i kolumnach [por. Gontar, 2019], natomiast pamięć komputera to struktura liniowa. Tabela może być reprezentowana w kolejności wierszy lub kolumn. Organizacja zorientowana na wiersze przechowuje tabelę jako sekwencję rekordów; w kolumnowym przetwarzaniu to dane z całej kolumny są przetrzymywane w ciągłym obszarze pamięci. SAP HANA obsługuje oba przypadki, ale jest szczególnie zoptymalizowana pod kątem przetwarzania kolumnowego, jako że zapewnia wysoki poziom kompresji danych i brak konieczności tworzenia dodatkowych struktur indeksów [SAP.com, 2017].

Rozpatrując platformę SAP HANA jako przykład nowoczesnego, opartego na usługach rozwiązania, trzeba zauważyć, iż zawiera ona pewne kategorie usług charakterystyczne dla nowoczesnych rozwiązań informatycznych. Pierwsza z nich to usługi aplikacji przeznaczone do tworzenia aplikacji internetowych [tamże]. Dotyczą one takich elementów, jak zarządzanie cyklem życia aplikacji (od modelowania struktury produktu, przez projektowanie aplikacji, po tworzenie, transport danych, walidację i konfigurację), rozwój interfejsu internetowego i obsługa otwartych standardów programistycznych, np. SQL, HTML5, Java Script, JDBC, ODBC, JSON i OData [tamże].

Druga grupa usług jest związana z przetwarzaniem danych. SAP HANA to rozwiązanie, które znacznie rozszerza tradycyjną rolę serwera bazy danych – funkcjonuje jako kompleksowa platforma do przetwarzania i analizowania dużej ilości danych w pamięci oraz możliwości przetwarzania równoległego. Tak zaprojektowane rozwiązanie

dostarcza szeregu usług z obszaru zaawansowanej analityki danych (np. analizy szeregów czasowych, danych strumieniowych, analizy tekstu, analityk predykcyjnych), narzędzi integracji (w tym integracji z danymi z urządzeń typu *smart*) oraz usług związanych z samym zarządzaniem bazą danych (przetwarzanie kolumnowe, wieloprocessorowe i kompresja, zapewnienie jakości oraz bezpieczeństwa). Możliwość przetwarzania wielu typów danych w jednym systemie umożliwia tworzenie aplikacji nowej generacji. Z kolei zaawansowane możliwości przetwarzania, takie jak algorytmy predykcyjne, wizualizacja i eksploatacja danych *online* oraz biblioteki funkcji biznesowych, pomagają w przyspieszeniu działania aplikacji i ograniczają transmisję danych.

Trzecia grupa usług platformy to usługi integracji. Dane w organizacji są generowane i przechowywane w rozmaitych systemach i różnego rodzaju typach baz. Inteligentne usługi konsolidacyjne pozwalają przetwarzać je w jednym miejscu; dają dostęp do danych gromadzonych w systemach silosowych, replikują i przenoszą dowolny rodzaj danych na platformę SAP HANA w czasie rzeczywistym, zarówno w przypadku implementacji w modelu chmury, jak i *on-premise*. Platforma SAP HANA zawiera także wzorzec strukturalny do tworzenia przez użytkownika niestandardowych adapterów przeznaczonych do replikacji danych lub integracji zapytań z różnych źródeł. Wbudowany procesor strumieniowy może na bieżąco przechwytywać i analizować strumienie danych oraz kierować je do odpowiednich miejsc przechowywania, przetwarzania lub pulpitów menedżerskich. Remote Data Sync synchronizuje dane między SAP HANA a tysiącami zdalnych baz danych (SQL Anywhere, UltraLite). Platforma SAP HANA zapewnia także szereg integracji z Hadoop.

Integracja aplikacji i osiągnięcie interoperacyjności<sup>2</sup> w ramach integracji napotyka szereg trudności – natury nie tylko technologicznej, ale także organizacyjnej i konceptualnej. W tabeli 2.3 wskazano przykładowe obszary barier, przed jakimi staje w takim przypadku organizacja.

---

<sup>2</sup> Interoperacyjność oznacza zdolność dwóch lub większej liczby systemów informatycznych (albo ich komponentów) do wymiany informacji i ich użycia.

**Tabela 2.3.** Bariery integracji aplikacji

<b>Obszar</b> <b>Typ</b>	<b>Biznes</b>	<b>Procesy</b>	<b>Usługi</b>	<b>Dane</b>
konceptualne	wizja, strategia, kultura	semantyka, syntaktyka, zawartość	semantyka, syntaktyka, zawartość	semantyka, syntaktyka, zawartość
technologiczne	strategia IT, stopień digitalizacji	sposób implementacji działania	poziom szczegółowości i pakietów	formaty wymiany
organizacyjne	struktura, aspekty prawne, metody pracy	przebieg procesu	zarządzanie usługami	klasyfikacja, prawa własności i dostępu

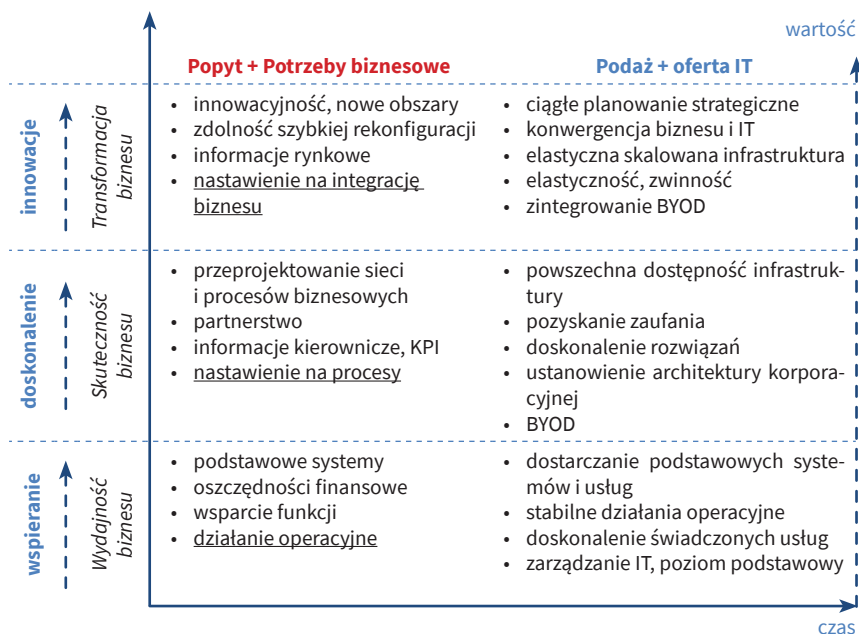
**Źródło:** [Weichhart i inni, 2016].

## 2.3. Zastosowanie nowych technologii w systemach informatycznych

### 2.3.1. Systemy informatyczne w transformacji cyfrowej przedsiębiorstw

Oczekiwania organizacji w zakresie wsparcia rozwiązaniami IT przebiegają na różnym poziomie (rysunek 2.10): od potrzeby wspomożenia podstawowych działań operacyjnych, poprzez usprawnienie procesów biznesowych i ich doskonalenie zapewniające skuteczność biznesu, po poszukiwanie partnerstwa dla nowych innowacyjnych pomysłów i ofert.

Dalsza ewolucja systemów informatycznych zarządzania jest nieunikniona, a jej kolejne generacje okażą się – w coraz większym stopniu – interoperacyjnym, elastycznym, ściśle dopasowanym do strategii IT organizacji rozwiązaniem, w którym procesy biznesowe obsługiwane będą za pomocą rozmaitych aplikacji i usług, dostępnych na różnego rodzaju platformach i chmurach obliczeniowych, przez użytkowników zlokalizowanych na całym świecie (rysunek 2.11). Stawia to przed organizacjami nowe wyzwania – także w zakresie zmiany dzisiejszego modelu współpracy w obszarze komunikacji systemów informatycznych i platform wymiany oraz przetwarzania danych.



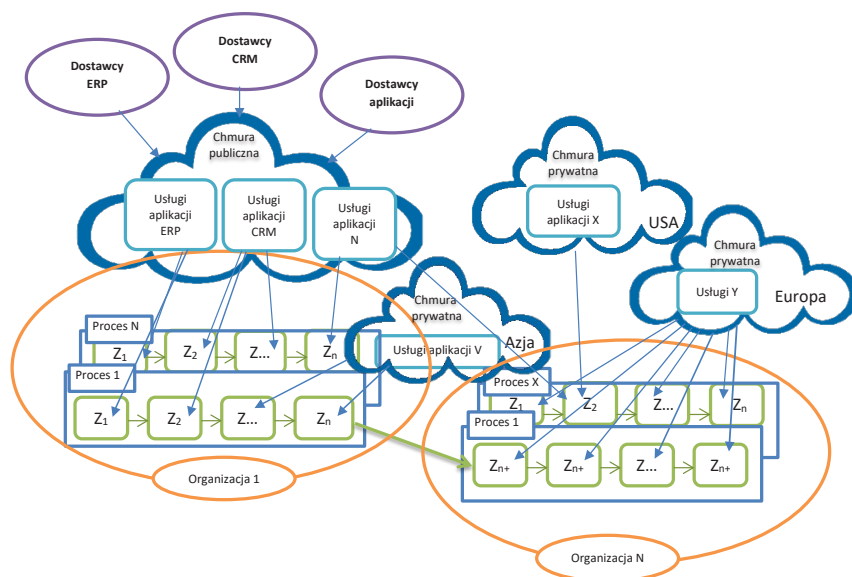
**Rysunek 2.10.** Potrzeby biznesowe i podaż IT

**Źródło:** opracowanie własne na podstawie [Sidenko, 2015].

W pierwszej dekadzie XXI wieku wielkość i złożoność wdrożeń ERP wzrosła do tego stopnia, że do implementacji dużych systemów organizacje powoływały osobne grupy projektowe. Konfiguracja i parametryzacja systemu okazały się kosztownym przedsięwzięciem, zajmującym wiele miesięcy. W przypadku nowych wersji systemów ich *upgrade* również wymagał kosztownych projektów.

Możliwość kastomizacji i przystosowania do potrzeb klienta systemów ERP sprawiła, że stały się one dostępne dla wielu organizacji. Obecnie dostosowanie to może stać się zarówno przeszkodą w przejściu na wersje SaaS ERP, jak i mieć znaczny wpływ na filozofię projektowania usług przez ich dostawców, którzy starają się zapewnić pewną elastyczność w doborze rozwiązań klientom bez eskalowania kosztów, z minimalizacją ryzyka, ale przy jednoczesnej odpowiedniej złożoności – zgodnie z zapotrzebowaniem biznesu. Chmura obliczeniowa jest rewolucyjnym rozwiązaniem dla systemów aplikacji biznesowych, takich jak ERP. Możliwość uruchamiania złożonego oprogramowania z centrów danych do-

stawców lub innych firm oznacza, że działy IT nie muszą utrzymywać infrastruktury, a kolejne wersje nie wymagają aż tak złożonych nakładów pracy i kosztów. Wdrożenie w chmurze, zwłaszcza w modelu SaaS, z krótszymi cyklami aktualizacji i rozszerzeniami niewymagającymi działań ze strony klienta, pomogło też uprościć systemy ERP – w innym modelu dalszego rozszerzania funkcjonalności stałyby się one bardziej złożone.



**Rysunek 2.11.** Schemat wsparcia procesów biznesowych aplikacjami z wykorzystaniem modeli chmury obliczeniowej

**Źródło:** opracowanie własne na podstawie [Schulte i inni, 2015].

Aby osiągnąć określone cele biznesowe, firmy były zorganizowane jako zamknięte, scentralizowane i hierarchiczne struktury charakteryzujące się biurokratyzowanym modelem – miało to sens, kiedy stawiano przede wszystkim na zminimalizowanie kosztów transakcji i asymetrii informacyjnych oraz dostarczenie (względnie) statycznego produktu lub usługi na (względnie) stabilny rynek krajowy [Fenwick i inni, 2019]. Implementowane systemy informatyczne wspierały te działania.

Systemy ERP ciągle pełnią kluczową funkcję w organizacjach, zwłaszcza w tradycyjnym obszarze *back-office*, związanym z zarządzaniem finansami, magazynem czy produkcją.



Rozważania na temat doboru aplikacji do potrzeb danej firmy rozpoczęły się na przełomie XIX i XX wieku. Nie zawsze moduły systemu ERP stawały się najlepszym rozwiązaniem – w niektórych przypadkach ich funkcjonalność okazywała się niewystarczająca i dla organizacji korzystniejsze było wykorzystanie oprogramowania wyspecjalizowanych dostawców. Chmura obliczeniowa zaczęła mieć największy wpływ na ewolucję systemów klasy ERP w 2010 roku – wraz z rozwojem oferty aplikacji w modelu SaaS, które obsługują jedną funkcję biznesową. Najpopularniejsze z nich to systemy klasy CRM i aplikacje do zarządzania kapitałem ludzkim (HCM).

Samo zastąpienie starego systemu ERP nie jest jednak wystarczające – firmy powinny spojrzeć całościowo na zakup aplikacji wspierających ich procesy i działania wychodzące poza organizację.

Wyzwaniem dla współczesnych systemów zarządzania, w tym systemów klasy ERP, staje się integracja, połączenie wszystkich procesów, danych i innych zasobów. Zadaniem inteligentnego systemu ERP jest już nie tylko wspomaganie zarządzania zasobami organizacji, ale także tworzenie wartości, zapewnienie większego stopnia kompleksowej automatyzacji, co pozwoli na eliminację rutynowych, powtarzalnych operacji. Nowe technologie, sztuczna inteligencja i Internet Rzeczy modernizują ERP i wkrótce spowodują powstanie autonomicznego oprogramowania odpowiedzialnego za wiele codziennych procesów biznesowych.

Kolejne technologie wprowadzały swego rodzaju „fale” implementacji w organizacjach, wpływając na ich funkcjonowanie. Począwszy od komputerów typu *mainframe*, poprzez technologie klient-serwer, monolityczne systemy ERP, Internet i rozwiązania mobilne, aż po IoT można obserwować rozwój cyfrowego łańcucha wartości przez poprawę wydajności obsługiwanych procesów, możliwości współpracy aplikacji i użytkowników, a także samego interfejsu i pracy z użytkownikiem. Wprowadzanie kolejnych technologii zazwyczaj wiązało się z większą złożonością architektury korporacyjnej, stawiało duże wymagania pracownikom i organizacji, aby osiągnąć poprawę sprawności biznesowej i wydajności całej firmy.

Elementem wyróżniającym implementację inteligentnych systemów ERP w modelu chmury obliczeniowej jest zmniejszenie stopnia złożoności. Chmura obliczeniowa to również model, w którym można uzyskać

wystarczającą moc obliczeniową i pamięć pozwalającą skrócić czas operacji w analizach Big Data i zastosować narzędzia sztucznej inteligencji. Jest jednocześnie bramą komunikacyjną między systemami ERP a ich użytkownikami. Te możliwości zmieniają cel i rozszerzają zasięg ERP daleko poza zarządzanie zasobami finansowymi, HR-em, produkcją czy łańcuchem dostaw. Obecnie systemy ERP obejmują również inne obszary, takie jak doświadczenie klienta czy satysfakcja interesariuszy.

Przykładem tego, w jaki sposób ERP w chmurze łączy pracowników na całym świecie jest system SAP ERP S/4HANA Cloud ERP, który wdraża Magna International Inc., producent części samochodowych na Florydzie. S/4HANA Cloud to wielodostępna wersja SaaS najnowszej generacji systemu ERP firmy SAP – zastosowanie i użytkowanie tej samej instancji systemu pozwoliło firmie skonsolidować pięć dotychczasowych systemów ERP, ujednolicić dane i procesy biznesowe w ramach globalnej działalności. Nie ma możliwości kastomizacji systemu S/4HANA, co jednak w pewnym zakresie sprzyja wymuszeniu standaryzacji procesów w organizacji.

SaaS to opcja wdrażania w chmurze sprawiająca, że inne aplikacje, z jakimi współpracują systemy ERP, np. platformy zakupów *online* lub narzędzia ankiet pracowniczych, są zintegrowane i dostępne niskim kosztem. To również mechanizm, dzięki któremu dostawcy ERP mogą dostarczać klientom nowe technologie znacznie szybciej. Stymulowany wszechobecnością Internetu SaaS jest obecnie centrum jednego z najważniejszych trendów ERP: demontażu tradycyjnej architektury ERP, procesu nazwanego przez Gartner Group dewolucją lub dekonstrukcją monolitycznego ERP, określanym także pojęciem „postmodernistyczne ERP”.

Trendem widocznym w systemach ERP, związanym z zarządzaniem doświadczeniami klientów (*Customer Experience Management*), okazuje się znaczne uproszczenie interakcji z systemami poprzez: wyposażenie w przyjazne dla użytkownika narzędzia, mniej złożone i łatwiejsze w nawigacji ekrany; głosowe interfejsy umożliwiające przekazywanie poleceń i wprowadzanie danych głosem oraz *chatboty*, które obsługują powtarzalne zadania. Sztuczna inteligencja, zwłaszcza uczenie maszynowe i przetwarzanie języka naturalnego (NLP), zmienia sposób interakcji ludzi z systemem ERP. Wszyscy główni dostawcy próbują znaleźć metody dostarczania aplikacji sztucznej inteligencji. Platformy ERP pozwalają na opracowywanie i dołączanie własnych „inteligentnych”

aplikacji. Z punktu widzenia dostawcy systemu nie ma użytkowników sztucznej inteligencji (*Artificial Intelligence* – AI) – są tylko przypadki jej użycia. Zastosowanie AI polega więc na wykorzystaniu w procesach dotyczących finansów, zamówień czy planowania działań uczenia maszynowego i automatyzacji do umożliwienia przepływu pracy, tworzenia dokumentów lub podjęcia decyzji. Przepływy pracy zarządzane z pomocą ERP z wbudowanymi elementami sztucznej inteligencji są bardziej skuteczne niż implementacja ogólnej platformy AI, z której większość użytkowników nie wie, jak korzystać. Aplikacje takie są osadzone w systemach ERP i służą do obsługi wąskich zadań w typowych przepływach pracy, takich jak przetwarzanie faktur, gdzie np. wykorzystując uczenie maszynowe, odpowiedni robot softwarowy wyodrębnia dane z obrazu zeskanowanej faktury, wprowadza na tej podstawie dane do bazy i przesyła informację do menedżera i klienta.

Przez wiele lat organizacje budowały własną architekturę IT, wprowadzając elementy nadzoru w obszarze zatasowania informatyki i dostosowując je do swoich modeli biznesowych.

Popularność modelu biznesowego platformy wzrosła w ostatnich latach wraz z rozwojem szeregu wzajemnie powiązanych technologii, takich jak Internet czy smartfony, dzięki czemu możliwa stała się szybka wymiana produktów, usług i informacji na dużą skalę. W ten sposób zaczęły tworzyć się globalne ekosystemy zachęcające zarejestrowanych użytkowników i konsumentów treści do dodawania większej wartości do platformy. Poprzez wielokrotne tworzenie treści następuje przyciąganie dodatkowych jej twórców i konsumentów – model platformy pozwala na czerpanie znacznych korzyści z „efektów sieciowych”. Stanowi on obecnie sposób na dostosowanie się do realiów szybko rozwijających się technologii i hiperkonkurencyjnych rynków światowych [Fenwick i inni, 2018].

Integracja osiągnięć technologicznych stanowi podstawę i infrastrukturę dla platform. Technologia cyfrowa oferuje rozwiązania typu *peer-to-peer*, a także pozwala użytkownikom dzielić się opiniami, doświadczeniami i wszelkimi innymi informacjami. Będąca szkieletem platformy technologia łączy też deweloperów i twórców z użytkownikami. W związku z tym wszystkie firmy, które chcą działać jako platformy, muszą myśleć i działać „tak, jakby” były one przedsiębiorstwami technologicznymi [tamże].

Kultura sukcesu platformy polega na zapewnieniu silnej i trwałej relacji pomiędzy platformą a jej uczestnikami. Powinna ona oferować elastyczne, unikalne, spersonalizowane doświadczenie, które „uruchamia” cykl platformowy konsumpcji, nauki, dzielenia się, tworzenia sieci i eksperymentowania. O efektywności i popularności platform decyduje kilka czynników [tamże]:

- technologia, na podstawie której działa platforma – musi być bezpieczna, pewna i niezawodna dla wszystkich użytkowników i uczestników platformy;
- doświadczenie użytkownika – przyjazny system pracy, proste metody korzystania z platformy i jej usług, stale aktualizowany i ulepszany interfejs;
- możliwość monetyzacji i prowadzenia działań analitycznych – technologia musi oferować użytkownikom i uczestnikom jednocześnie szybką i prostą interakcję, możliwość wyboru i wygodę użytkowania; platformy powinny ułatwiać kontakty ze społecznością użytkowników, która ma dla nich „znaczenie”, a także zachęcać użytkowników do kreatywności i zaangażowania (poprzez media społecznościowe, recenzje, blogi itp.);
- sposób dołączenia do platformy – szybki i prosty;
- odpowiednio zaprojektowany system zachęt, ocen i rekomendacji, mechanizmów informacji zwrotnej i modeli cenowych; platforma musi oferować uczestnikom dostępne, uczciwe i osobiste doświadczenie; uczestnicy powinni być w stanie zweryfikować reputację i zaufanie innych użytkowników;
- zarządzanie danymi – będące podstawą udanego kojarzenia interesariuszy platformy i odróżniające je od innych modeli biznesowych; skuteczny sposób kojarzenia zbiera dane od uczestników platformy i używa ich, aby ułatwić połączenia między grupami użytkowników.

Firmy działające w modelu platformy nie tylko wykorzystują nowe technologie w celu usprawnienia ekonomicznych lub społecznych interakcji pomiędzy zainteresowanymi stronami, ale zmieniają także swoją organizację pracy na bardziej płaski i integracyjny model, aby umożliwić szeroką współpracę wszystkich interesariuszy – w ten sposób zmaksymalizowana zostaje możliwość dostarczania ciągłych innowacji w zakresie usług i funkcjonalności platformy.

### 2.3.2. Inteligentne systemy ERP i produkcja w chmurze obliczeniowej

Podstawową ideą Przemysłu 4.0 jest zintegrowanie produkcji różnych inteligentnych fabryk wzdłuż łańcucha wartości w postaci cyber-fizycznych platform, dzięki czemu w czasie rzeczywistym dane przekładane są na informacje w całym łańcuchu wartości – co z kolei umożliwia dokładne podejmowanie decyzji w czasie rzeczywistym [Liu i Xu, 2017].

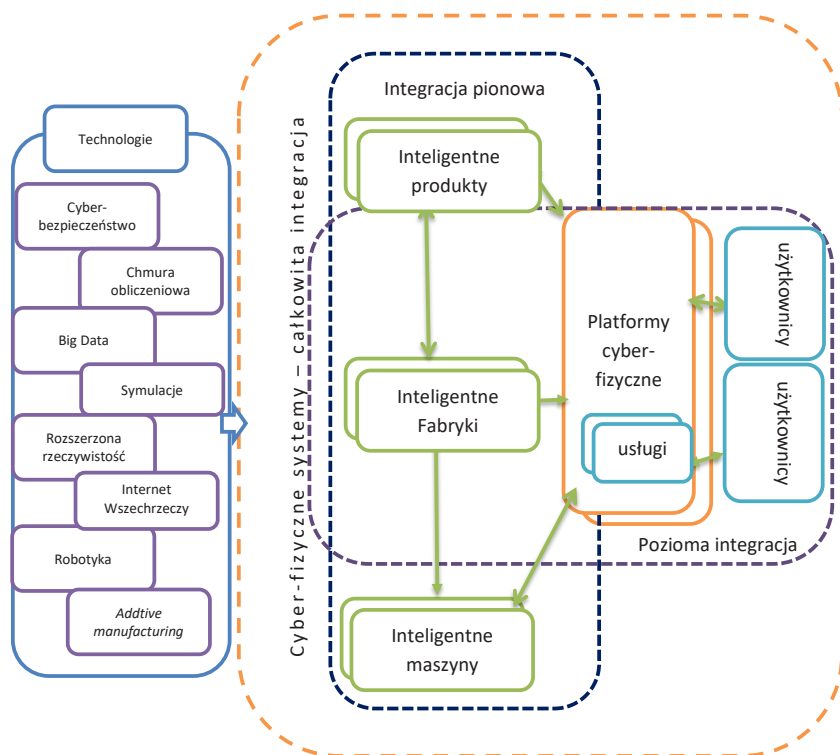
W przedsiębiorstwach produkcyjnych nowe technologie informatyczne pozwalają na kompleksowe modelowanie procesów, które mają być realizowane w zakładach na drodze ich wirtualizacji. Stanowią one element integrujący stosowane dotąd systemy wspomagania projektowania i oprogramowanie klasy ERP, tworzący tzw. cyfrowe fabryki (*digital factory*) [Gunia, 2019].

Powstanie cyber-fizycznych systemów odnoszących się do środowiska produkcyjnego jest wynikiem rozwoju wielu technologii, w tym takich jak: przemysłowy Internet Rzeczy, Big Data, chmura obliczeniowa, robotyka, rozszerzona rzeczywistość, *additive manufacturing* (proces wytwarzania trójwymiarowych obiektów na podstawie modeli komputerowych), symulacje czy technologie związane z bezpieczeństwem danych. To one przyczyniają się do powstania koncepcji cyber-fizycznych systemów, obejmujących m.in. koncepcje inteligentnych fabryk, inteligentnych produktów czy inteligentnych urządzeń (rysunek 2.12).

Koncepcja Przemysłu 4.0 wynika z idei i metod produkcji cyfrowej. Projekt wsparcia produkcji rozwiązaniami cyfrowymi istnieje i rozwija się w nauce oraz biznesie od kilkudziesięciu lat. Współpraca biznesowa pomiędzy partnerami i serwisyzacja z jednej strony, a rozwój Internetu i nowych technologii, w tym chmury obliczeniowej, z drugiej – nadały temu rozwojowi nową jakość poprzez ewolucję w kierunku wersji sieciowej o wysokim stopniu nasycenia technologiami cyfrowymi.

Chmura obliczeniowa stopniowo staje się jednym z głównych czynników umożliwiających przemysłowi produkcyjnemu osiągnięcie celu biznesowego w zakresie ścisłej i szybkiej współpracy z partnerami. *Cloud manufacturing* dotyczy określenia nie tylko technologii – jest rozumiane szerzej jako efekt synergii wynikający z połączenia najnowocześniejszych technologii, zapewniający bezpieczną i niezawodną platformę usługową przy stosunkowo niskich cenach i wspierający cały cykl produkcyjny

[Hao, 2016]. Przedsiębiorstwa produkcyjne mogą racjonalnie akceptować koncepcję przetwarzania w chmurze, aby stworzyć wersję produkcyjną chmury obliczeniowej określanej jako *cloud manufacturing*.



**Rysunek 2.12.** Technologie i koncepcje w systemach cyber-fizycznych

**Źródło:** opracowanie własne.

Pojęcie *cloud manufacturing* (produkcja lub wytwarzanie w chmurze) oznacza współpracę i komunikację w trybie *online* wszystkich partnerów we wszystkich aspektach operacji produkcyjnych ekosystemu, gdzie użytkownicy korzystają ze usług i wirtualnych zasobów dostępnych w chmurze, w celu wykonania przydzielonych zadań produkcyjnych na żądanie.

Nazwa Przemysłu 4.0 wywodzi się z czasów rewolucji przemysłowej, a wytwarzanie w chmurze wynika z zaawansowanych modeli i technologii. Koncepcja ta stara się stworzyć ramy funkcjonowania branży produkcyjnej w nadchodzącej epoce i przedstawić rozwiązania takich kwestii, jak: efektywność wykorzystania zasobów i energii, produkcja miejska czy

zmiany demograficzne – w postaci platform cyber-fizycznych. W warunkach rewolucji Przemysłu 4.0 każda firma może stać się aktywnym węzłem inteligentnej sieci, który podejmuje działania w zakresie wytwarzania, produkcji czy dystrybucji [Gontar, 2018], a także stanowi wysoce zdigitalizowany sieciowy paradygmat produkcji. Z kolei koncepcja produkcji w chmurze to zaawansowany model biznesowy produkcji, koncentrujący się na kwestiach bezpośrednio związanych z produkcją (np. współużytkowaniu zasobów i współpracy w chmurze obliczeniowej) – zwraca mniejszą uwagę na zagadnienia typu: problemy miasta, zmiany demograficzne, emisja dwutlenku węgla itp. [Liu i Xu, 2017].

Obie koncepcje odnoszą się do usług, przy czym produkcja w chmurze obejmuje także model „produkcja jako usługa” (zawierający komplet elementów niezbędnych do jej świadczenia, w tym zasoby i procesy produkcyjne) w pełnym cyklu życia produktu – co okazuje się szersze niż idee związane z produktami i procesami w Przemysle 4.0 [tamże].

Dwa główne pryncypia charakteryzujące wytwarzanie w chmurze to autonomia i agregacja. Pierwsze z nich oznacza, że organizacje lub jednostki współpracujące w tym środowisku są autonomiczne, drugie – że są one agregowane razem na podstawie zgłaszanej do tego środowiska potrzeby. Taki model umożliwia różnym producentom dzielenie się najlepszymi praktykami oraz unikalnymi lub zapasowymi zasobami/zdolnościami produkcyjnymi w puli zasobów specyficznych np. dla branży [Hao, 2016]. Zdolności (potencjał biznesowy) i zasoby mogą być oferowane i pobierane na żądanie zgodnie z paradygmatem zorientowanym na usługi (SOA), dzięki czemu wygodne udostępnianie różnorodnych rozproszonych zasobów produkcyjnych jest realizowane w sposób dynamiczny – wszystko opiera się na rzeczywistych potrzebach użytkowników. W tego typu proponowanej strukturze różni użytkownicy mogą wyszukiwać i wysyłać żądania z chmury produkcyjnej [tamże].

Rozróżnia się dwa poziomy systemów usług wytwarzania w chmurze. Poziom niższy to produkcyjna wersja sieci społecznościowej, która może wspierać udostępnianie usług organizacji. Działa jak portal internetowy i zapewnia komunikację w czasie rzeczywistym oraz wymianę informacji. Produkty końcowe lub dojrzałe usługi biznesowe klasyfikują i publikują w tym portalu dostawcy. Algorytmy wyszukiwania i dopasowywania są używane, aby pomóc klientom w przydzieleniu najlepiej pasujących produktów lub usług, a odpowiednie produkty lub usługi zostają dostarczane lokalnie



do klientów. Jest to produkcja z bezpośrednim wykorzystaniem technologii przetwarzania w chmurze, tj. produkcyjna wersja przetwarzania w chmurze bazująca na modelach typu PaaS, SaaS, IaaS. Dobrym przykładem tego poziomu systemu usług są: Alibaba, ThomasNet i GlobalSpec [tamże].

Wyższy poziom to platforma pozwalająca na tworzenie nowych innowacyjnych inicjatyw biznesowych, wymagająca specyficznej dla branży chmury wertykalnej opartej na ekonomii współdzielenia. W tym przypadku nie tylko zasoby ICT są świadczone jako usługi, ale cała działalność produkcyjna. W modelu tym synergia, połączenia najnowocześniejszych technologii, takich jak m.in. przetwarzanie w chmurze, Big Data i Internet Rzeczy, zapewniają bezpieczną i niezawodną platformę usługową po stosunkowo niskich cenach, wspierając cały cykl życia produkcji. Podczas gdy przetwarzanie w chmurze służy celowi obliczeniowemu i dzieli zasoby obliczeniowe, wytwarzanie w chmurze realizuje zamierzenia produkcyjne i udostępnia użytkownikom wiele niezbędnych do tego zasobów. Platforma zapewnia pełną integrację, tzw. *back-office*, z halą produkcyjną i oferuje współpracę w procesie produkcyjnym w różnych fabrykach. Model obejmuje również dostawców i odbiorców, ale nie skupia się na produktach końcowych – zamiast tego zasoby i możliwości produkcyjne są zwirtualizowane i zamknięte jako usługi [tamże]. Usługobiorcy mogą ich żądać w celu wykonania czynności produkcyjnych. Platforma wyznacza podstawowe ramy, które pomagają w zarządzaniu wszystkimi działaniami mającymi miejsce podczas wspólnego procesu produkcyjnego.

Pojęcie „cyfrowa fabryka” związane jest z cyfrowym obrazem realnej produkcji i bazuje na następujących elementach [Gunia, 2019]:

- cyfrowym opisie produktu z jego statycznymi i dynamicznymi właściwościami;
- cyfrowym planowaniu produkcji;
- cyfrowej symulacji przebiegu produkcji z możliwością zastosowania danych planowanych dla poprawy wskaźników efektywności procesów.

W literaturze można spotkać zastosowania technologii cyfrowej fabryki dotyczące różnych obszarów. Jako najistotniejsze wymieniane są [tamże]:

- działania wspomagające projektowanie nowych wyrobów (modelowanie i wirtualizacja wyrobów, technologia odwrotna, skanowanie 3D, druk 3D);
- działania na rzecz racjonalizacji systemów produkcyjnych z wykorzystaniem ich wirtualnych modeli;



- kształtowanie systemów pracy pod kątem ergonomii, bezpieczeństwa i higieny pracy, wydajności, jakości wykonania;
- rozwiązywanie problemów decyzyjnych w zakresie inwestycji, projektowania stanowisk pracy, systemów produkcyjnych itp.

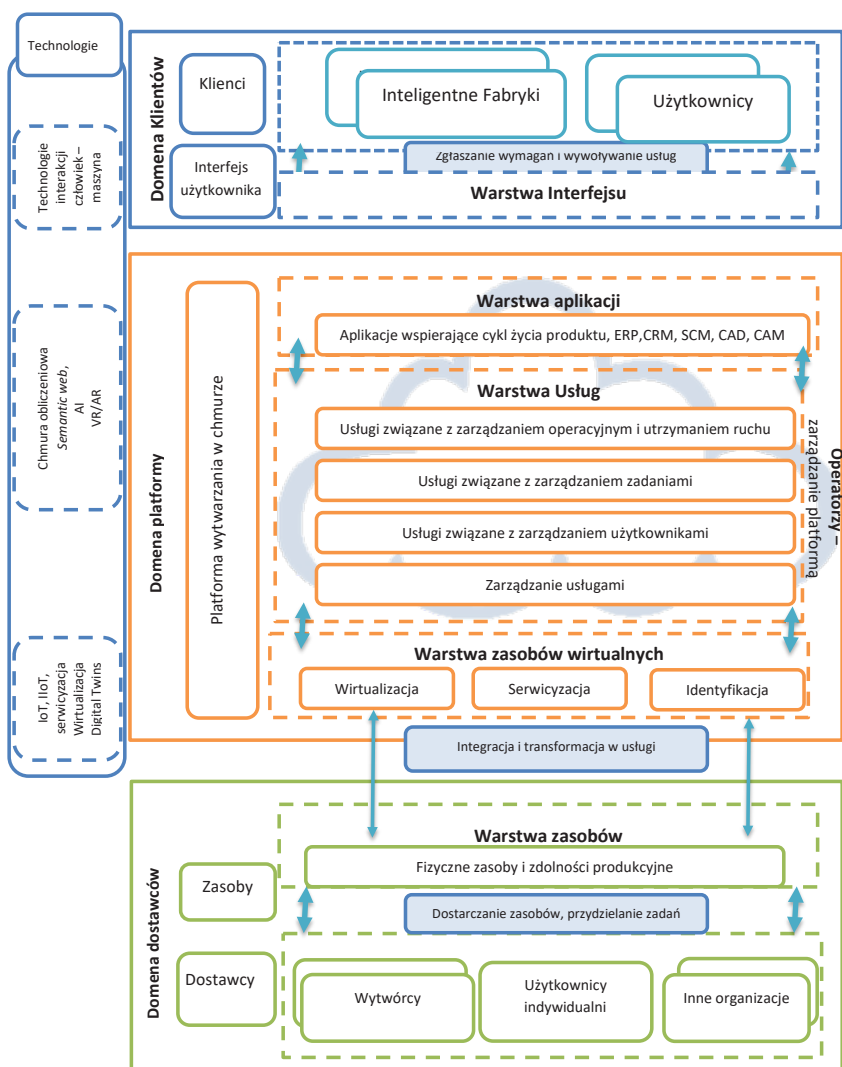
Koncepcja Przemysłu 4.0 wymusza współpracę i integrację, na którą można spojrzeć w trzech wymiarach. Inteligentna fabryka i inteligentny produkt to kluczowe elementy i główne cechy tej koncepcji – poprzez sieciowe systemy produkcji wymuszają one integrację pionową. Z kolei za pośrednictwem sieci wartości odbywa się integracja pozioma (zarówno w pojedynczej inteligentnej fabryce, jak i w różnych inteligentnych fabrykach). Dzięki integracji poziomej i integracji pionowej powstaje kompleksowa cyfrowa integracja w całym łańcuchu wartości i cyklu życia produktu.

Podstawą produkcji w chmurze jest utworzenie wielowarstwowej platformy opartej na wielu technologiach. Warstwy platformy dotyczą zasobów materialnych, niematerialnych i wirtualnych, a także usług, aplikacji i interfejsu [Liu i Xu, 2017]. Ogólny schemat modelu wraz z warstwami został przedstawiony na rysunku 2.13. Model produkcji w chmurze okazuje się ściśle związany z jej operacyjnym trybem pracy; określa on trzy domeny: dostawców, organizacji i użytkownika.

W pierwszej z nich następuje dostarczanie zasobów produkcyjnych, które są przekształcone w usługi, a następnie udostępnione na platformie wytwarzania w chmurze. Wprowadzenie pośredników/operatorów do zarządzania platformą pozwala zagwarantować i zapewnić odpowiednio wysoką jakość usług. Klienci mogą przysyłać do platformy swoje zapytania i wymagania dotyczące usług, poczynawszy od projektowania produktu, poprzez etap produkcji, aż po testowanie i zarządzanie na wszystkich etapach cyklu życia produktu [tamże].

W domenie dostawców zasoby produkcyjne można podzielić na fizyczne zasoby produkcyjne i zdolności produkcyjne (niematerialne i dynamiczne zasoby, które reprezentują zdolność organizacji do podejmowania działania). Zasoby fizyczne produkcji istnieją w postaci sprzętowej (np. komputery, serwery, surowce) lub programowej (np. możliwości symulacji, narzędzia analityczne, wiedza, dane, standardy itp.).

Zdolności obejmują z kolei możliwości projektowania produktu, symulacji, prowadzenia eksperymentów, produkcji, zarządzania i utrzymania [Xu, 2012].



**Rysunek 2.13.** Warstwowa architektura w modelu produkcji w chmurze

**Źródło:** opracowanie własne na podstawie [Liu i Xu, 2017].

Warstwa zasobów wirtualnych ma za zadanie zidentyfikowanie zasobów produkcyjnych, ich zwirtualizowanie (określenie zasobów logicznych i podpiętych pod nie zasobów fizycznych), a następnie pogrupowanie i przygotowanie pakietów dostępnych jako usługi w chmurze (serwicyzacja). Różne zasoby produkcyjne mogą być wirtualizowane

w rozmaity sposób. Linie i maszyny produkcyjne są zwykle zamapowane jako maszyny wirtualne niezależne od systemu, podczas gdy zasoby obliczeniowe i wiedza mogą być zwirtualizowane w podobny sposób jako określone, wykonujące zadania usługi. Wirtualni menedżerowie (np. *Virtual Machine Monitor*) ponoszą odpowiedzialność za komunikację z warstwą fizycznych zasobów oraz koordynowanie i przydzielanie maszyn wirtualnych do wykonywania zadań.

Warstwa usług odpowiada za udostępnianie ich i zarządzanie nimi. Zapewnia odpowiednią dostępność, przydzielanie, naliczanie opłat, monitorowanie zasobów produkcyjnych – przy czym nadal to dostawcy są odpowiedzialni za wykonanie zadań produkcyjnych i zapewnienie odpowiedniej jakości pracy [tamże].

Warstwa aplikacji pozwala na uruchomienie dowolnej funkcjonalności biznesowej (np. ERP, CRM, SCM), projektowej typu CAD, komputerowego wspomagania wytwarzania (CAM) czy sterowania urządzeniami CNC. Współpracuje ona z warstwą interfejsu użytkownika: zapewnia różnym urządzeniom pracę w złożonych systemach modelowania, symulacji i analiz; musi gwarantować współpracę dla opracowywania i współtworzenia nowych produktów. Model produkcji w chmurze przewiduje, że użytkownik może zdefiniować i zbudować aplikację produkcyjną za pośrednictwem zwirtualizowanych zasobów. Istotnym elementem tych warstw jest zapewnienie mobilności danych, odpowiednie naliczenie odpłatności za wykorzystane zasoby oraz zapewnienie bezpieczeństwa.

W warstwie aplikacji ważną funkcję pełnią też: moduł bazy danych, w którym gromadzone są dane wejściowe użytkownika; inteligentny moduł do przetwarzania danych, optymalizacji i generowania wykonalnych rozwiązań; moduł wspomagania decyzji do oceny i porównania rozwiązań [Caggiano i inni, 2019].

### **2.3.3. Automatyzacja procesów biznesowych a aplikacje biznesowe**

Dla większości przedsiębiorstw automatyzacja procesów ma wymiar strategiczny – jej fundamentem jest automatyzacja obciążenia pracą (*Workload Automation* – WLA). Przyczyną tego stanu są m.in. problemy z tradycyjnymi systemami informatycznymi. Automatyzacja obciążenia

pracą to praktyka wykorzystywania oprogramowania do planowania i zarządzania zadaniami związanymi z procesami i transakcjami biznesowymi. W tym kontekście obciążenie pracą może być postrzegane jako całkowite przetwarzanie, jakie komputer wykonuje w danym momencie. WLA pozwala na to, aby duża część tego przetwarzania odbywała się bez ingerencji człowieka. Duża liczba zadań jest zaplanowana, a nie wykonywana w czasie rzeczywistym.

Automatyzacja procesów przetwarzania danych i procesów biznesowych w organizacjach w znacznej części rozpoczyna się od zastosowania systemów wspomagających przepływ pracy (*workflow*) – one również mogą być wbudowane w systemy ERP: umożliwiają bowiem elektroniczny obieg dokumentów i minimalizują manualne procesowanie. Należy jednak zauważyć, że elektroniczny przepływ pracy razem z szybszym przepływem dokumentów może jednocześnie zwiększać tempo występowania kolejnych problemów. Automatyzacja jest możliwa dzięki różnym rozwiązaniom IT, takim jak tworzenie skryptów i aplikacji, integracja API, planowanie zadań, tworzenie makr czy RPA.

Automatyzacja procesów może być rozpatrywana w kontekście poziomów dojrzałości IT organizacji. Określa się tu następujące poziomy [Martinek-Jaguszewska, 2018]:

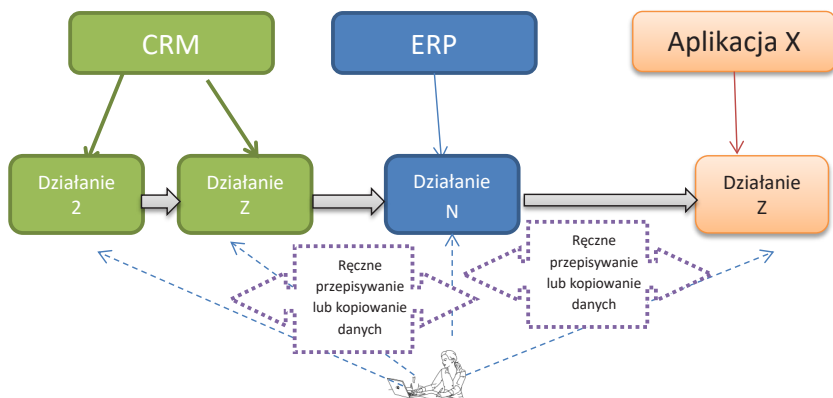
- 1) Komputeryzacja – implementacja i eksploatacja niepowiązanych narzędzi i aplikacji IT do wykonywania zadań przez pracowników;
- 2) Orkiestracja – standaryzacja narzędzi IT, zintegrowane zarządzanie procesami i zasobami organizacji, budowanie powiązań oraz integracji między narzędziami oraz systemami użytkowanymi w organizacji;
- 3) Automatyzacja procesów biznesowych – wykorzystanie dedykowanej technologii (w tym rozwiązań z zakresu RPA) do wykonywania powtarzalnych zadań bez aktywnej ingerencji pracownika (pełni on funkcję kontrolną lub wspierającą);
- 4) Automatyzacja kognitywna – technologia, która dzięki głębokiej analizie danych oraz uczeniu maszynowemu jest w stanie samodzielnie wybrać oraz zautomatyzować element działania procesu (dzięki wykrywaniu wzorców może przewidzieć warianty wykonanego scenariusza lub zapobiec jego spełnieniu się).

Aplikacje klasy ERP obejmują podstawowe obszary funkcjonowania firmy, ale nie zapewniają wsparcia jej specyficznych produktów i procesów. Część z nich jest obsługiwana przez inne aplikacje. Trzeba też

zauważyć, że coraz bardziej złożone i rozbudowane aplikacje są kosztowne w utrzymaniu i mało podatne na modyfikacje. Na stanowiskach pracy występuje zjawisko korzystania przez pracowników firmy z wielu aplikacji, w tym aplikacji kooperantów, udostępnianych poprzez Internet.

Z punktu widzenia integracji rozwiązaniem byłoby wdrożenie odpowiedniego procesu wprowadzania zmian do działających systemów, ale zwykle jest to projekt kosztowny o wysokim ryzyku niepowodzenia. Tego typu projekty są często odłożone w czasie – głównie ze względu na priorytetową konieczność implementacji innych aplikacji. Przy braku integracji systemów przenoszenie danych z systemu do systemu odbywa się zwykle metodą „kopiuj/wklej” [Kajrunajtys, 2017] – zob. rysunek 2.14. Aby ukończyć zadania w wyznaczonym procesie biznesowym, użytkownik w trakcie pracy musi wielokrotnie przełączać się między aplikacjami, wprowadzać te same dane do różnych aplikacji, co z kolei podnosi ryzyko popełnienia błędu, powoduje znudzenie i frustrację. W sytuacjach, gdy działania te są definiowalne, powtarzalne i oparte na procesach biznesowych oraz można dla nich wyznaczyć prosty algorytm działania, możliwe staje się zastosowanie robotów cyfrowych – oprogramowania do automatyzacji czynności człowieka wykonywanych na ekranie aplikacji (*Robotic Process Automation* – RPA). Roboty te działają poprzez istniejące interfejsy użytkownika w aplikacji i wykonują predefiniowaną pracę [tamże]. Realizują działania, do których miał uprawnienia pracownik na tym samym stanowisku, takie jak: dostęp do poczty elektronicznej i systemów, wykonywanie obliczeń, tworzenie dokumentów i raportów oraz sprawdzanie plików. Do podjęcia określonych działań robot wykorzystuje kontrolki klawiatury i myszy, przechwytuje i interpretuje elementy interfejsu aplikacji użytkownika pracującego na określonym stanowisku. Rozwiązanie to ma widoczne zalety: jest niskobudżetowe i neutralne wobec praw twórców aplikacji, może być wykorzystywane okresowo i łatwo zmodyfikowane, nie wymaga znajomości struktur baz danych, a jego wdrożenie okazuje się odwracalne [tamże].

Innym elementem, istotnym z punktu widzenia organizacji, które uzasadnia zastosowanie RPA, jest przystosowanie i przeszkolenie pracowników do nowego stanowiska pracy lub do pracy w zmienionym systemie. Wykorzystanie robota wymaga przygotowania i przetestowania scenariusza, według którego ma on pracować; po czym robot staje się gotowy do pracy w trybie ciągłym przez dwadzieścia cztery godziny na dobę.



**Rysunek 2.14.** Obsługa działań na stanowisku w procesie wspomagany przez różne aplikacje z ręcznym przenoszeniem danych

**Źródło:** opracowanie własne.

W rozwoju aplikacji biznesowych, takich jak ERP czy CRM, jednym z istotnych trendów okazuje się wykorzystanie automatyzacji do dalszego usprawnienia procesów biznesowych. Automatyzacja rozumiana jest tu podobnie jak w starym znaczeniu, zgodnie z którym maszyny przejmują ludzkie zadania – np. roboty na linii montażowej – tyle że w tym przypadku okazują się to roboty softwarowe. Aplikacje biznesowe (m.in. ERP, CRM czy SCM) są otwarte i przygotowane do wprowadzenia automatyzacji, ponieważ większość pracy użytkowników tych systemów polega na gromadzeniu danych, a następnie tworzeniu raportów i ich dystrybucji. W inteligentnym systemie ERP przetwarzanie danych odbywa się w tle, celem ich użytkowania jest zarządzanie i monitorowanie przy pomocy pulpitów menedżerskich – dostępnych poprzez strony WWW lub w aplikacjach mobilnych. Odpowiednie algorytmy informują menedżerów o problemie, pozwalając na przejęcie kontroli i podejmowanie decyzji wtedy, gdy będzie to absolutnie konieczne. Tego typu system może optymalizować nie tylko zasoby, którymi ERP obecnie zarządza, ale także wszelkie procesy, jak np. emisja dwutlenku węgla.

W wielu organizacjach istnieją możliwości wykorzystania i narzędzi automatyzacji przepływu pracy, i RPA w ramach efektywnej infrastruktury automatyzacji. Zrobotyzowana automatyzacja procesów jest idealna do wykonywania powtarzających się, pracochłonnych zadań. Najlepiej sprawdza się przy dużych nakładach pracy, ponieważ w rzeczywistości to seryjność

i mnogość operacji są warunkami koniecznymi, aby narzędzie to okazało się wartościowe. Nadaje się ono również do wykorzystania w środowiskach, w których brakuje specjalistycznej wiedzy informatycznej lub nakład zasobów IT został skierowany do innych obszarów. Narzędzia RPA są dobre w przenoszeniu danych ze stałych punktów – np. ze znanego punktu A do znanego punktu B. Jednak nie mogą one transportować danych ze znanego punktu A do zbioru punktów. Pomimo tych ograniczeń, narzędzia RPA mogą być nadal stosowane nawet w systemie wykorzystującym narzędzia WLA (nic nie stoi na przeszkodzie, by współpracowały). Oprogramowanie WLA może przekazywać niezbędne dane do oprogramowania RPA, a następnie pozwolić RPA na wykonywanie zadania [Watts, 2020].

Umożliwienie narzędziom WLA przekazywania niektórych zadań do narzędzi RPA pozostawia więcej zasobów dla narzędzi WLA do realizacji zadań złożonych – daje to jednocześnie efekt maksymalizacji efektywności narzędzi automatyzacji i wspiera rozwój efektywnej infrastruktury automatyzacji. Ponadto – przy prawidłowym zastosowaniu – WLA może być idealnym ułatwieniem dla RPA. Efektywne wykorzystanie ich razem owocuje bezpieczną, efektywną i wydajną automatyzacją w całym przedsiębiorstwie. Przy odpowiedniej strategii narzędzia te przedsiębiorstwa mogą stosować jako część skutecznego systemu automatyzacji [tamże].

Rozwój systemów RPA podąża w stronę inteligentnej automatyzacji procesów biznesowych (*Intelligent Process Automation* – IPA), u podstaw której leży wyłaniający się zestaw nowych technologii, łączący fundamentalne zmiany w projektowaniu procesów z automatyzacją procesów i uczeniem się maszyn [McKinsey & Co., 2016]. IPA naśladuje działania wykonywane przez ludzi i – dzięki postępowi w zakresie technologii głębokiego uczenia się oraz technologii poznawczych – z czasem uczy się realizować je jeszcze lepiej. Zastosowanie IPA może dla organizacji oznaczać radykalną poprawę wydajności, zwiększenie efektywności pracy pracowników, zmniejszenie ryzyka operacyjnego oraz usprawnienie czasu reakcji i doświadczeń klientów [tamże].

IPA w swoim pełnym zakresie obejmuje pięć podstawowych technologii [tamże]:

- 1) RPA – omawiane na początku roboty softwarowe, które automatyzują rutynowe zadania, takie jak ekstrakcja i czyszczenie danych, poprzez wykorzystanie istniejących interfejsów użytkownika;
- 2) Inteligentny przepływ pracy – narzędzie programowe do zarządzania procesem, które integruje zadania wykonywane przez grupy

ludzi i maszyny (np. nadzór nad grupą robotów softwarowych), co pozwala użytkownikom na inicjowanie i śledzenie stanu całego procesu w czasie rzeczywistym. Zakłada się, że docelowo to oprogramowanie będzie zarządzać relacjami pomiędzy różnymi grupami, w tym między robotami i ludźmi, oraz dostarczać analiz dotyczących wykonywania procesu i identyfikacji „wąskich gardeł”;

- 3) **Uczenie maszynowe i zaawansowana analityka** – to algorytmy, które identyfikują wzorce na podstawie analiz zbiorów danych strukturalnych, poprzez uczenie „nadzorowane” i „bez nadzoru”. Rozwiązania te zostały już szeroko wdrożone w wielu dziedzinach biznesu, w tym w zarządzaniu zasobami ludzkimi, w celu określenia i oceny kluczowych cech liderów i menedżerów, co pomaga rozwijać ścieżki kariery i planować sukcesję stanowisk (szerzej o uczeniu maszynowym – zob. podrozdział 3.5 niniejszej monografii);
- 4) **Generowanie języka naturalnego (NLG)** – które odnosi się do tworzenia tekstu lub mowy w języku naturalnym przez komputer, co z kolei umożliwia budowanie interakcji między ludźmi a technologią. Generowanie tekstu na podstawie ustrukturyzowanych danych jest wykorzystywane np. do automatycznego tworzenia raportów finansowych w korporacjach finansowych;
- 5) **Kognitywne agenty** – technologie, które łączą uczenie się maszynowe i generowanie tekstu w języku naturalnym, w celu zbudowania całkowicie wirtualnej siły roboczej (tzw. „agenta”), zdolnej do wykonywania zadań, komunikowania się, uczenia się na podstawie zestawów danych, a nawet podejmowania decyzji na podstawie „wykrywania emocji”; agenty te mogą być wykorzystywane np. do wspierania pracowników centrów obsługi klienta w rozmowach telefonicznych lub za pośrednictwem czatu. Są one dostępne przez całą dobę i dopasowywane do preferencji klienta pod względem cech osobowych, np. wyglądu, barwy głosu.

## 2.4. Podsumowanie

Przedsiębiorstwa ery cyfrowej i post-cyfrowej będą tworzyły ekosystemy oparte na platformach, w których uczestniczy wiele podmiotów. Organizacje zaczną działać „jako” platforma lub zostaną „zintegrowane”



w ramach platformy. Najbardziej wpływowymi firmami pozostaną właściciele platform – ze względu na kontrolę ich działania. Z tego powodu wszystkie organizacje (nie tylko przedsiębiorstwa technologiczne) starają się na nowo zdefiniować siebie jako platformy i wspomagać odpowiednimi technologiami [Fenwick i inni, 2018]

Widocznym trendem w rozwoju aplikacji biznesowych jest obecnie IPA. Integracja w systemach ERP, CRM, SCM, Internetu Rzeczy, technologii Blockchain i technologii inteligencji poznawczej pozwoli zrealizować zautomatyzowane procesy biznesowe na skalę masową. Zagadnienie dotyczące automatyzacji procesów biznesowych jest omawiane szerzej w kolejnej pozycji serii Cyfryzacja, zatytułowanej *Systemy inteligentne w sieci Internet*.

## Literatura

- Adamczewski P. (2001), *Informatyczne wspomaganie łańcucha logistycznego*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- BlueSoft (2018), *Jeden powód by stosować mikrousługi*, <https://bluesoft.com/pl/jeden-powod-by-stosowac-mikrousługi/> [dostęp: 01.03.2020].
- Bojarski R. (2003), *Systemy informatyczne w zarządzaniu przedsiębiorstwem*, Wydawnictwo Politechniki Śląskiej, Gliwice.
- BPSC (2017), *Zwrot z inwestycji w ERP? Tak, pod warunkiem, że szukamy odpowiedzi na pytanie „tak czy nie” a nie „ile”*, [https://www.erp-view.pl/erp/zwrot\\_z\\_inwestycji\\_w\\_erp\\_tak\\_pod\\_warunkiem\\_ze\\_szukamy\\_odpowiedzi\\_na\\_pytanie\\_tak\\_czy\\_nie\\_a\\_nie\\_ile.html](https://www.erp-view.pl/erp/zwrot_z_inwestycji_w_erp_tak_pod_warunkiem_ze_szukamy_odpowiedzi_na_pytanie_tak_czy_nie_a_nie_ile.html) [dostęp: 03.11.2017].
- Bytniewski A. (red.), (2005), *Architektura zintegrowanego systemu informatycznego zarządzania*, Wydawnictwo Akademii Ekonomicznej, Wrocław.
- Caggiano A., Lev Boun L., Deng B. (2019), *Intelligent cloud manufacturing platform for efficient resource sharing in smart manufacturing networks*, „Procedia CIRP”, nr 79.
- Capgemini (2017), *Supplier Relationship Management (SRM) Research 2016–2017, Solution Analysis and Business Insights*, Capgemini SE, Paris.
- CEO.com.pl (2017), *ERP a zwrot z inwestycji – czy w ogóle warto liczyć ROI?*, <https://ceo.com.pl/erp-a-zwrot-inwestycji-ogole-warto-liczyc-roi-26116/3> [dostęp: 03.11.2017].
- Czerwinka P., Gontar B., Pamuła A. (2018), *Korzyści i problemy implementacji systemów w środowisku chmury obliczeniowej – studia przypadków polskich przedsiębiorstw*, „Annales Universitatis Mariae Curie-Skłodowska. Sectio H Oeconomia”, t. 52, nr 2.
- Davenport T. (1993), *Process Innovation. Reengineering Work through Information Technology*, Harvard Business Press, Boston.
- Durlik I. (1998), *Inżynieria zarządzania. Strategia i projektowanie systemów produkcyjnych*, cz. 1, Wydawnictwo Placet, Warszawa.
- ERP-VIEW.PL (2019), *Raport ERP 2019*, <https://www.raport-erp.pl/> [dostęp: 01.03.2020].
- Essex D. (2019), *Mega-suites, cloud shape 20 years of ERP history*, <https://searcherp.techtarget.com/feature/Mega-suites-cloud-shape-20-years-of-ERP-history> [dostęp: 01.03.2020].

- Fenwick M., McCahery J.A., Vermeulen E.P.M. (2018), *The End of "Corporate" Governance (Hello "Platform" Governance)*, [https://ecgi.global/sites/default/files/working\\_papers/documents/finalfenwickmccaheryvermeulen.pdf](https://ecgi.global/sites/default/files/working_papers/documents/finalfenwickmccaheryvermeulen.pdf) [dostęp: 15.03.2020].
- Gontar B. (red.), (2019), *Zarządzanie danymi w organizacji*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2019.
- Gontar Z.H. (red.), (2018), *Smart Grid Analytics for Sustainability and Urbanization*, IGI Global, Hershey.
- Grudzewski W. (2004), *Metody projektowania systemów zarządzania*, Wydawnictwo Difin, Warszawa.
- Gunia G. (2019), *Zintegrowane systemy informatyczne przedsiębiorstw w kontekście Przemysłu 4.0*, „Zarządzanie Przedsiębiorstwem. Enterprise Management”, t. 22, nr 2.
- Hadaś Ł., Cyplik P. (red.), (2012), *Praktyczne aspekty wykorzystania systemów ERP w wybranych przedsiębiorstwach Wielkopolski*, L-Systems Robert Pawlak, Piła.
- Hao Y. (2016), *Cloud Manufacturing, Strategic Alignment between Manufacturing Industry and Cloud Computing*, [https://osuva.uwasa.fi/bitstream/handle/10024/7303/isbn\\_978-952-476-715-6.pdf?sequence=1&isAllowed=y](https://osuva.uwasa.fi/bitstream/handle/10024/7303/isbn_978-952-476-715-6.pdf?sequence=1&isAllowed=y) [dostęp: 15.03.2020].
- Januszewski A. (2008), *Funkcjonalność informatycznych systemów zarządzania. Zintegrowane systemy transakcyjne*, Wydawnictwo Naukowe PWN, Warszawa.
- Kajrunajtys D. (2017), *Robotic Process Automation jako czynnik poprawy jakości procesów testowania*, <https://www.saso.org.pl/static/IIkSASO/2017.06.30%20-%20DKajrunajtys.pdf> [dostęp: 28.02.2020].
- Kisielnicki J. (2009), *Typologia systemów informatycznych zarządzania* [https://wneiz.pl/nauka\\_wneiz/studia\\_inf/24-2009/si-24-157.pdf](https://wneiz.pl/nauka_wneiz/studia_inf/24-2009/si-24-157.pdf) [dostęp: 16.11.2017].
- Kisielnicki J. (2013), *Zastosowanie technologii informacyjnej, Systemy Informatyczne Zarządzania*, Wydawnictwo Placet, Warszawa.
- Klonowski Z. (2004), *Systemy informatyczne zarządzania przedsiębiorstwem. Modele rozwoju i właściwości funkcjonalne*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
- Landarter V.D., Gray C.D. (1989), *MRP II Standard System*, Oliver Wight Publication, Gloucester.
- Laudon K.C., Laudon J.P. (2013), *Essentials of Management Information Systems*, Pearson, London.
- Lech P. (2003), *Zintegrowane systemy zarządzania ERP/ERP II. Wykorzystanie w biznesie, wdrażanie*, Wydawnictwo Difin, Warszawa.
- Liu Y., Xu X. (2017), *Industry 4.0 and Cloud Manufacturing: A Comparative Analysis*, „Journal of Manufacturing Science and Engineering”, nr 139.
- Martinek-Jaguszewska K. (2018), *Znaczenie i rola automatyzacji procesów biznesowych – wyniki badań pilotażowych*, „Organizacja i Kierowanie”, nr 4.
- McKinsey & Co. (2016), *McKinsey on Digital Services Introducing the next-generation operating model*, <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Introducing%20the%20next-generation%20operating%20model/Introducing-the-next-gen-operating-model.ashx> [dostęp: 01.03.2020].
- Niedzielska E. (red.), (1983), *Wstęp do informatyki: technologia przetwarzania danych*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

- Oblak I., Ziemia E., (2012), *Systemy informatyczne w organizacjach zorientowanych procesowo*, „Problemy Zarządzania”, t. 10, nr 3.
- Odenwald T. (2014), *A New Perspective on Enterprise Resource Management*, <https://sloanreview.mit.edu/article/a-new-paradigm-for-managing-enterprise-resources/> [dostęp: 01.03.2020].
- Panetto H., Zdravkovic M., Jardim-Goncalves R., Romero D., Cecil J., Mezgar I. (2015), *New Perspectives for the Future Interoperable Enterprise Systems*, „Computers in Industry”, nr 79.
- Parys T. (1998), *MRP II przykładem systemu zintegrowanego*, „Informatyka”, nr 9.
- Pei-Fang H., HsiuJu R.Y., Jung-Ching Ch. (2015), *Assessing ERP Post-Implementation Success at the Individual Level: Revisiting the Role of Service Quality*, „Information and Management”, nr 52.
- Romanowicz W. (2014), *Nadchodzi era Internetu Przedmiotów*, „Biznes Benchmark Magazyn”, nr 9.
- SAP.com (2017), *The Definitive Guide to Value Creation with Intelligent ERP Cloud*, <https://www.sap.com/products/erp/erp-cloud.html#pdf-asset=826c6dbf-df7c-0010-82c7-edda71af511fa&page=3> [dostęp: 12.12.2019].
- Schulte S., Janiesch C., Venugopal S., Weber I., Hoenisch P. (2015), *Elastic Business Process Management: State of the art and open challenges for BPM in the cloud*, „Future Generation Computer Systems”, nr 46.
- Sidenko S. (2015), *Business Relationship Management as a Role, a Discipline, and an Organizational Capability*, *Symposium*, <https://www.slideshare.net/PMI-Montreal/symposium-2015-business-relationship-management-as-a-role-a-discipline-and-an-organizational-capability> [dostęp: 14.08.2020].
- Somer A. (2017), *5 way to reinvent your ERP systems*, <https://www.zuora.com/guides/five-ways-to-reinvent-your-erp/> [dostęp: 29.08.2017].
- Szymonik A. (2010), *Technologie informatyczne w logistyce*, Wydawnictwo Placet, Warszawa.
- Watts S. (2020), *Workload Automation (WLA) vs Robotic Process Automation (RPA): What's The Difference?*, <https://www.bmc.com/blogs/workload-automation-vs-robotic-process-automation/> [dostęp: 01.06.2020].
- Weichhart G., Molina A. Chen D., Whitman L., Vernadt F. (2016), *Challenges and Current Developments for Sensing, Smart and Sustainable Enterprise Systems*, „Computer in Industry”, nr 79.
- Wojtkowski R. (2010), *Po co aplikacjom ERP potrzebna jest platforma rozwoju (IDE)?*, <https://www.erp24.pl/systemy-erp-artykuly/po-co-aplikacjom-erp-potrzebna-jest-platforma-rozwoju-ide.html> [dostęp: 12.02.2020].
- Woźniak K. (2002), *American Production and Inventory Control Society, Charakterystyka zastosowań zintegrowanych systemów zarządzania przedsiębiorstwem klasy MRP, MRP II, ERP*, „Zeszyty Naukowe Akademii Ekonomicznej w Krakowie”, nr 574.
- Xu X., (2012), *From cloud computing to cloud manufacturing*, „Robotics and Computer-Integrated Manufacturing”, nr 28.
- Zalewski W. (2011), *Analiza systemów informatycznych wspomagających zarządzanie produkcją w wybranych przedsiębiorstwach*, „Ekonomia i Zarządzanie”, nr 4.
- Zaskórski P. (2012), *Integracja zasobów i usług informacyjnych w organizacji biznesowej*, „Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki”, nr 7.

# Rozdział 3

## Środowisko analityczno-decyzyjne organizacji

### 3.1. Wprowadzenie

Zarządzanie organizacją jest procesem, w którym określone cele organizacyjne osiągnąć się przy wykorzystaniu dostępnych zasobów. Realizacja poszczególnych funkcji zarządzania, takich jak planowanie, organizacja, koordynacja i kontrolowanie działalności, wymaga posiadania wartościowej informacji – dotyczącej zarówno stanu samego przedsiębiorstwa, jak i jego otoczenia – oraz możliwości sprawnego jej przetwarzania. Niniejszy rozdział został poświęcony dwóm niezwykle istotnym zagadnieniom: decyzjom menedżerskim i narzędziom wspomagającym ich podejmowanie, obecnie często wbudowanym w systemy informatyczne, a także zasobom informacyjnym jako podstawie procesów analitycznych.

### 3.2. Decyzje w organizacji

#### 3.2.1. Decyzje menedżerskie i współczesna organizacja

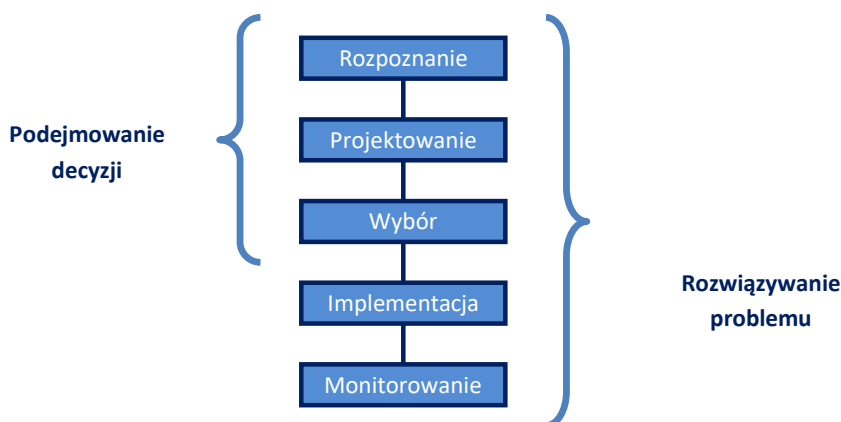
Informacje stanowią podstawę dla decyzji podejmowanych przez menedżerów zarządzających organizacją. Stąd posiadanie ich i właściwe wykorzystanie daje przewagę nad konkurencją, umożliwiając lepszą i szybszą reakcję na zachodzące zmiany.

Należy przy tym zwrócić uwagę, że ich znaczenie we współczesnym świecie zdecydowanie rośnie. Środowisko działania współczesnej organizacji staje się bowiem coraz bardziej złożone, co tworzy zarówno określone możliwości, jak i zagrożenia oraz problemy, które trudno wykorzystać i z którymi niełatwo sobie poradzić bez odpowiedniego wspomagania procesu decyzyjnego.

Wśród najistotniejszych elementów wywierających presję na organizację wymienić można [Sharda, Delen i Turban, 2017]:

- czynniki rynkowe – takie jak narastająca konkurencja, rozszerzanie się rynków globalnych, rozwój rynków elektronicznych w Internecie, innowacyjne metody marketingowe, rozwój możliwości *outsourcingu*, potrzeba realizacji transakcji w czasie rzeczywistym oraz na żądanie itp.;
- wymagania klientów – takie jak pragnienie indywidualnego traktowania, żądania jakości, różnorodności produktów i szybkości dostawy; klienci stają się przy tym coraz silniejsi oraz mniej lojalni itp.;
- czynniki technologiczne – takie jak większa innowacyjność, nowe produkty i usługi (produkty szybciej stają się przestarzałe), rosnące przeciążenie informacyjne, opiniotwórcze sieci społecznościowe, media społecznościowe itp.;
- czynniki społeczne – takie jak, z jednej strony, rozrastająca się sfera regulacji urzędowej, a z drugiej: deregulacja społeczeństwa (siła robocza jest bardziej zróżnicowana, starsza), zwiększenie społecznej odpowiedzialności przedsiębiorstw, większy nacisk na zrównoważony rozwój itp.

Presja środowiskowa coraz bardziej złożonego współczesnego świata wymaga od organizacji właściwych, pozwalających wykorzystywać stwarzane możliwości oraz rozwiązywać pojawiające się problemy reakcji, nowych strategii, zmian w modelach biznesowych, zwiększonej produktywności, reorganizacji kontaktów z partnerami. Działania te potrzebują odpowiednich decyzji, coraz bardziej złożonych, podejmowanych w coraz szybszym tempie w warunkach presji środowiskowej. Podejmowanie ich jest w zasadzie niemożliwe bez zintegrowanego informatycznego wspomagania oraz zasobów informacyjnych o charakterze operacyjnym i analitycznym. System informacyjny powinien asystować podczas całego tego zadania nie tylko poprzez integrację i dostarczanie niezbędnych danych, ale również przez aktywne ich przetwarzanie i wspomaganie procesu decyzyjnego.



**Rysunek 3.1.** Etapy procesów podejmowania decyzji i rozwiązywania problemu

**Źródło:** opracowanie własne.

Zadanie podejmowania decyzji musi być rozpatrywane na tle szerszego zagadnienia, nazywanego procesem rozwiązywania problemu [Stair, 1992; Laudon i Laudon, 1993; Jabłoński i Bartkiewicz, 2006; Sharda, Delen i Turban, 2017]. Jak widać na rysunku 3.1, proces rozwiązywania problemu obejmuje pięć etapów, z których początkowe trzy związane są z podjęciem decyzji, a kolejne dwa z jej implementacją oraz monitorowaniem skutków. Istotny jest również fakt, że nie ma on charakteru czysto sekwencyjnego. Podczas realizacji kolejnych faz może zachodzić potrzeba iteracyjnego powrotu do któregoś z wcześniejszych etapów i powtórnej realizacji wykonywanych podczas niego czynności.

W pierwszej fazie, na etapie rozpoznania, identyfikowane i definiowane są problemy (lub pojawiające się okazje), cele, jakie mają zostać w związku z tym osiągnięte i ogólne możliwości rozwiązania problemu (wykorzystania szansy). Zbierana jest informacja odnosząca się do zakresu zagadnienia: potencjalnych rozwiązań, ograniczeń i otoczenia.

W fazie projektowania buduje się możliwe sposoby poradzenia sobie z problemem, warianty decyzji (określane często jako alternatywy decyzyjne) niezbędnych do realizacji rozwiązań; szacowana jest również wstępnie wykonalność poszczególnych wariantów.

Etap wyboru to ostatnia faza podejmowania decyzji. Wymaga zdecydowania, która z rozważanych alternatyw decyzyjnych zostanie podjęta w celu rozwiązania problemu (wykorzystania okazji). Ten, na pierwszy

rzut oka łatwy, akt wyboru zwykle nie jest tak prosty, jak mogłoby się wydawać. Istnieje wiele trudnych do rozpoznania czynników, często o charakterze niejawnym, mogących wpływać na ostateczną decyzję.

Na etapie implementacji proces rozwiązywania problemu (wykorzystania okazji) wykracza już poza fazę podejmowania decyzji. Polega on na wykonaniu akcji wprowadzających w życie wybrane rozwiązanie.

Finalnym etapem procesu jest monitorowanie. W tej fazie decydent ocenia implementację rozwiązania, określając, czy zostały osiągnięte przewidywane wyniki i modyfikując ewentualnie proces w świetle nowej informacji zdobytej podczas wdrożenia. Może to wywołać sprzężenie zwrotne i powrót do któregoś z wcześniejszych etapów rozwiązywania problemu.

Z jednej strony procesy globalizacyjne, działalność na rynkach o zasięgu światowym, wymuszają na organizacji podporządkowanie się rozmaitym przepisom prawnym. Z drugiej zaś – mamy do czynienia z dążeniem do unifikacji, wskutek presji globalnej konkurencji i wymagań konsumentów. W powiązaniu z coraz bardziej dynamicznym charakterem gospodarki i koniecznością szybkiego reagowania na zachodzące zmiany, elementy te powodują wzrost niepewności, utrudniając przewidywanie efektów podejmowanych decyzji.

Rosnąca dynamika procesów gospodarczych skutkuje również częstymi i nieprzewidywalnymi zmianami, utrudniającymi uczenie się z doświadczenia oraz zwiększającymi potencjalne koszty popełniania błędów.

Wszystkie te zjawiska powodują, że organizacje coraz częściej i w coraz większym zakresie sięgają po informatyczne wsparcie swoich decyzji. Starają się w ten sposób zmniejszać lukę między bieżącym działaniem organizacji a pożądanym poziomem jej aktywności, wyrażonym w misji, celach oraz strategii ich osiągnięcia.

### **3.2.2. Typy decyzji podejmowanych w organizacji**

Jak wspomniano w poprzednim punkcie, podejmowanie decyzji polega na formułowaniu odpowiednich celów, a następnie na określeniu i wyborze sposobu ich realizacji. Oczywiście poszczególne sytuacje decyzyjne mogą się znacznie różnić między sobą pod względem stopnia złożoności. Niektóre decyzje są stosunkowo proste, przy podejmowaniu innych wybór najlepszej alternatywy przysparza wielu kłopotów. Niekiedy poważ-



nym wyzwaniem okazuje się nawet samo sformułowanie rozwiązywanego problemu, stanowiącego cel podejmowanej decyzji. W tym kontekście, w przypadku problemów decyzyjnych, mówić można o następujących kategoriach [Jabłoński i Bartkiewicz, 2006]:

- decyzje strukturalne (programowalne) – powtarzalne, rutynowe problemy decyzyjne, dla których istnieją standardowe rozwiązania, dające się często przedstawić w postaci określonej procedury postępowania, algorytmu; dla decyzji tego typu modele decyzyjne zostały zbudowane już wcześniej, podjęcie decyzji polega więc jedynie na zastosowaniu znanego rozwiązania; niejednokrotnie decyzje o charakterze strukturalnym podejmowane są automatycznie przez systemy informatyczne;
- decyzje niestukturalne (nieprogramowalne) – rozmyte, kompleksowe, nierutynowe problemy, dla których nie ma łatwych rozwiązań; nie ma dla nich wypracowanej metodologii ani modelu ich rozwiązania; aby je podjąć, niezbędne są subiektywne sądy i intuicja menedżera;
- decyzje częściowo strukturalne (częściowo programowalne) – w ich przypadku jedynie pewne fazy procesu podejmowania decyzji mają charakter strukturalny; nie mogą one zostać w pełni zautomatyzowane; wymagają subiektywnych ocen i osądów, w powiązaniu z formalną analizą danych i budową modeli.

Decyzje strukturalne należą do najłatwiejszych we wspomaganiu przez środowisko informatyczne. Ponieważ znane są metody ich podejmowania, a przy tym mają one charakter dobrze zdefiniowanych procedur, dominującą aktywność przy podejmowaniu tego rodzaju decyzji mogą przejmować systemy informatyczne.

Typowym problemem strukturalnym jest np. podjęcie decyzji o określeniu płacy pracownika. Procedury są tu na ogół jasne i oczywiste oraz w większości przypadków wykonywane całkowicie automatycznie przez systemy płacowe. Tu celem zastosowania wspomagania informatycznego będzie raczej polepszenie sprawności podejmowanych decyzji – kłopotliwy okazuje się bowiem fakt, że muszą zostać one podjęte wielokrotnie, czasem dla setek, a nawet tysięcy pracowników. Systemy informatyczne wykonają tę pracę znacznie szybciej niż ludzie.

Czy przyspieszenie procesu podejmowania decyzji jest jedynym celem stosowania systemów informatycznych w zagadnieniach strukturalnych? Oczywiście, że nie. Nawet w przypadku naszej przykładowej decyzji o określeniu płacy pracownika system informatyczny nie tylko obliczy



ją sprawniej niż człowiek, ale zazwyczaj popełni przy tym znacznie mniej błędów. A jeśli nawet jakieś nieścisłości wystąpią, to spowodowane one będą niemal zawsze pomyłkami ludzkimi na etapie wprowadzania danych lub budowy systemu informatycznego.

Zauważmy ponadto, że system informatyczny nie tylko podejmie decyzję niemal zawsze bezbłędną, i często dokładniejszą niż człowiek, ale również zrobi to niezawodnie. Program monitorujący stan zapasu magazynowego „nie zapomni” wygenerować ostrzeżenia w sytuacji jego spadku poniżej wyznaczonego poziomu (i ewentualnie podjąć dalszych akcji w celu jego uzupełnienia).

Należy również wziąć pod uwagę fakt, że znajomość metody podejmowania określonej decyzji wcale nie oznacza, że jest ona prosta i łatwa w realizacji. Podjęcie decyzji może wymagać wykonania bardzo skomplikowanych obliczeń i operacji. W wielu przypadkach człowiek nie byłby w stanie tego rodzaju czynności wykonać, a jeśli nawet, musiałby to zrobić przy użyciu znacznie gorszych, przybliżonych metod.

Biorąc pod uwagę wszystkie przytoczone argumenty, widać, że wspomaganie informatyczne może poprawiać nie tylko sprawność, ale również jakość decyzji strukturalnych. Większość zagadnień z tej grupy stanowią decyzje dosyć proste, ale wielokrotnie powtarzane i często o fundamentalnym znaczeniu dla organizacji, związane z operacyjną obsługą jej podstawowych procesów. Z tego powodu zazwyczaj nieco większy nacisk kładzie się tu na aspekt wspomagania ich sprawności.

Z nieco odmienną sytuacją mamy do czynienia w przypadku decyzji niestrukuralnych, jak również do pewnego stopnia częściowo strukturalnych. Przy tego rodzaju zagadnieniach nie ma prostych procedur tworzenia i wyboru wariantu decyzji; dosyć rzadko celem działania wspomagającego systemu informatycznego jest pełna ich automatyzacja i zastąpienie decydenta. Częściej zadaniem programów wspomagających okazuje się współpraca z decydem-człowiekiem i asystowanie mu podczas procesu podejmowania decyzji.

Zadaniem systemów informatycznych, również i w tym kontekście, może być oczywiście poprawa zarówno jakości, jak i sprawności podejmowania decyzji. Inaczej jednak niż w przypadku strukturalnym, częściej większe znaczenie ma tu aspekt jakościowy. W niektórych sytuacjach wykorzystanie wspomagających systemów informatycznych powoduje, że decydem znacznie dokładniej przeanalizuje określoną decyzję, poświę-

cając na to więcej zasobów i czasu, a więc pogarszając sprawność decyzji, poprawiając jednak jej jakość.

Celem wspomagania informatycznego decyzji niestukturalnych i częściowo strukturalnych okazuje się więc wspomaganie decydenta poprzez porządkowanie zagadnienia, jego wizualizację, wskazywanie pewnych rozwiązań, szacowanie zależności między zjawiskami i szereg innych działań, o których będzie mowa w dalszej części bieżącego rozdziału. Chodzi na ogół o poprawę jakości decyzji.

I znów – na ogół nie znaczy, że zawsze. Czasami zdarza się, że w pewnych zadaniach decyzyjnych, jakie człowiek wykonuje lepiej niż wspomagające systemy informatyczne (przynajmniej na obecnym poziomie ich rozwoju), zastępujemy go programem komputerowym. Powody mogą być rozmaite, np. dlatego, że realizuje swoje zadania zbyt wolno lub nie gwarantuje odpowiedniego poziomu bezpieczeństwa ich wykonania. Godząc się nawet z możliwym pogorszeniem poziomu jakości decyzji, automatyzacja może okazać się niezbędna dla samego zapewnienia realizacji wykonywanych zadań.

Z przytoczonych wcześniej przykładów czytelnik mógłby wysnuć wnioski, że decyzje strukturalne wiążą się z operacyjnym poziomem zarządzania organizacją, zaś niestukturalne i częściowo strukturalne – z wyższymi. Byłoby to jednak błędne rozumowanie. W tabeli 3.1 zebranych zostało kilka przykładów decyzji różnego typu, podejmowanych na poszczególnych poziomach zarządzania.

**Tabela 3.1.** Przykłady decyzji poszczególnych typów na różnych poziomach zarządzania

	<b>Zarządzanie operacyjne</b>	<b>Zarządzanie taktyczne</b>	<b>Zarządzanie strategiczne</b>
<b>Decyzje strukturalne</b>	rejestracja i przetwarzanie faktur, przyjmowanie zamówień, księgowanie należności i płatności	programowanie liniiowe dla optymalizacji wytwarzania, analiza budżetowa, raporty personalne	inwestycje portfelowe, lokalizacja magazynu
<b>Decyzje częściowo strukturalne</b>	handel akcjami i obligacjami, sterowanie zapasami	ustalenie budżetu harmonogramowanie projektu, projektowanie systemu wynagrodzeń	budowa nowej fabryki, analiza pozyskania kapitału, planowanie nowego produktu
<b>Decyzje niestukturalne</b>	wybranie zdjęcia ilustrującego artykuł w magazynie, akceptacja kredytu, wybór dostawcy surowca	zatrudnianie menedżerów, negocjacje	planowanie strategiczne, planowanie prac badawczo-rozwojowych

**Źródło:** opracowanie własne na podstawie [Stair, 1992; Sharda, Delen i Turban, 2017].

Oczywiście decyzje strukturalne częściej występują na poziomie operacyjnym, ale obecne są także na poziomie taktycznym i strategicznym. Przykładowo: typowa decyzja o lokalizacji magazynu ma charakter dosyć standardowy. Istnieje kilka znanych warunków, jakie należy określić, zaś ostateczna decyzja może być podjęta z wykorzystaniem standardowych podejść optymalizacyjnych. Procedura w większości wypadków jest więc znana, choć niekoniecznie prosta.

Z kolei decyzja o wyborze zdjęcia ilustrującego artykuł w magazynie okazuje się w miarę prosta do wykonania dla człowieka. Realizowana jest zwykle przez dziennikarza na poziomie operacyjnym, jeśli chodzi o zarządzanie magazynem. Natomiast bez wątpienia to decyzja niestrukturalna – trudno jest tu nawet określić reguły wyboru, zależą one od doświadczenia pracownika, jego osobistej intuicji (stanowi przykład decyzji bardzo trudnej dla systemów informatycznych).

Innym ważnym czynnikiem wpływającym na decydenta podczas analizy możliwych wariantów postępowania są rozmiary ryzyka związanego z daną decyzją. Pojęcie ryzyka ma oczywiście szeroki charakter i może być definiowane oraz interpretowane w rozmaity sposób. Zazwyczaj rozumie się je w sensie probabilistycznym – w tym znaczeniu ryzykiem będziemy określać ocenę prawdopodobieństwa faktu, że wybrany przez decydenta sposób postępowania może zaowocować nieoczekiwanymi i niepożądanymi skutkami. Z tego punktu widzenia można mówić o trzech rodzajach decyzji:

- decyzje podejmowane w warunkach pewności,
- decyzje podejmowane w warunkach niepewności,
- decyzje podejmowane w warunkach ryzyka.

Pierwsze występują w sytuacjach, gdy rozwiązywany problem jest na tyle prosty, a decydent w fazie rozpoznania zebrał na jego temat informacje na tyle wyczerpujące i dokładne, że można dogłębnie go zrozumieć i zdefiniować zależności pomiędzy występującymi w nim zmiennymi. Wszystkie istotne fakty mogące wpłynąć na wynik decyzji znane są więc decydentowi – jest on w stanie z góry z całą pewnością określić skutek wyboru każdego wariantu decyzji. Nie ma więc ryzyka, że jego zastosowanie przyniesie nieoczekiwane efekty. Oceniając warianty decyzji, decydent może więc z całą pewnością wyznaczyć wartości kryterium (funkcji celu) dla każdego z nich i dokonać wyboru najlepszego wariantu spośród rozważanych.

Oczywiście w praktyce znajomość absolutnie wszystkich faktów i zależności jest w zasadzie niemożliwa; tym niemniej w pewnych sytuacjach decydent może wiedzieć na tyle dużo i rozumieć sytuację na tyle dobrze, aby czynnik ryzyka został pominięty.

Decyzje podejmowane w warunkach niepewności występują w sytuacjach, gdy dostępne zasoby informacyjne nie są w stanie dostarczyć decydentowi żadnej wiedzy o pewnych faktach lub zjawiskach albo dziedzinie rozwiązywanego problemu okazuje się zbyt złożona, aby można było precyzyjnie zdefiniować zależności między występującymi w niej zmiennymi. W związku z tym decydent nie jest w stanie określić skutków rozważanych wariantów decyzji – pojawia się więc ryzyko, że wybrany wariant zaowocuje innymi efektami niż spodziewane.

Rozwiązaniem w takim przypadku będzie zastosowanie technik symulacyjnych, pozwalających na analizę potencjalnych możliwych skutków decyzji dla różnych scenariuszy rozwoju sytuacji lub wartości nieznanymi zmiennymi. Należy jednak zwrócić uwagę, że wyniki takich analiz bardzo rzadko pozwalają jednoznacznie wybrać dany wariant (sposób postępowania), który z pewnością będzie najlepszy. Symulacje dostarczają jedynie wiedzy o rozmaitych skutkach decyzji przy różnych scenariuszach.

Skoro nie jesteśmy w stanie jednoznacznie określić efektów, nie możemy wyznaczyć wartości kryterium (funkcji celu) oceniającego warianty decyzji i dokonać wyboru najlepszej opcji. W przypadku decyzji w warunkach niepewności nie potrafimy więc podjąć decyzji obiektywnie, w dodatku tej najbardziej optymalnej dla rozwiązania rozważanego problemu. Dlatego przy tego rodzaju decyzjach do wyboru wariantu wykorzystuje się na ogół kryteria subiektywne, zależne od psychologicznego nastawienia decydenta.

Typowym kryterium subiektywnym jest nastawienie decydenta do ryzyka (równoważnie można mówić tu o optymizmie odnośnie skutków decyzji). Istnieją metody porządkowania sytuacji decyzyjnej, pozwalające wskazać decydentowi najlepszy sposób postępowania (wariant decyzji), jeśli np. chce zaryzykować bądź przeciwnie – podjąć decyzję bezpieczną, zachowawczą. Pamiętajmy jednak, że samo nastawienie decydenta do ryzyka ma charakter czysto subiektywny, intuicyjny, psychologiczny i skala tego nastawienia nie wynika z żadnych obiektywnych przesłanek (ponieważ takich nie potrafimy określić).

Ostatnie, czyli decyzje podejmowane w warunkach ryzyka, wiążą się z rozwiązywaniem problemów, dla których również nie możemy przewidzieć jednoznacznie i z całą pewnością skutków, ale – w przeciwieństwie do sytuacji całkowitej niepewności – potrafimy określić prawdopodobieństwo realizacji różnych możliwych scenariuszy. Zatem przy decyzjach z tej grupy nadal nie jesteśmy w stanie wskazać wariantu decyzji, który będzie z pewnością najlepszy. Jeśli zdecydujemy się na wybór jakiegoś sposobu postępowania, istnieje ryzyko, że w rzeczywistości zaowocuje on niespodziewanymi i niekorzystnymi skutkami. Ponieważ jednak w warunkach ryzyka możemy określić prawdopodobieństwa realizacji różnych opcji, potrafimy oszacować prawdopodobieństwo pojawienia się niekorzystnego rozwoju sytuacji, czyli wielkość samego ryzyka.

W tym przypadku bierzemy pod uwagę także subiektywne kryteria wyboru wariantu. Możliwość oszacowania wielkości ryzyka pozwala jednak na uwzględnienie przez decydenta danej informacji przy określeniu jego nastawienia do ryzyka, a ponadto umożliwia tworzenie różnorodnych strategii zabezpieczających przed niekorzystnymi skutkami decyzji.

Przy decyzjach w warunkach ryzyka decydent dysponuje więc obiektywnym kryterium wyboru najlepszego sposobu postępowania, opartym na zasadzie optymalnej wartości oczekiwanej skutków decyzji. Dla przykładu, wybiera się nie opcję, która daje najwyższe zyski (ponieważ nie potrafimy ich jednoznacznie określić), lecz wariant przynoszący najwyższą wartość oczekiwaną zysków. Należy jednak pamiętać, że reguła optymalnych spodziewanych skutków decyzji nie wskazuje najlepszego pewnego sposobu postępowania, a jedynie taki, który jest najprawdopodobniej najlepszy.

Na koniec zwróćmy uwagę na fakt, że podejmowanie decyzji nieodłącznie wiąże się z towarzyszącym temu procesowi ryzykiem. W zasadzie, poza stosunkowo prostymi zagadnieniami, decydent często stoi przed koniecznością wyboru dalszego sposobu postępowania w sytuacji braku pełnej wiedzy o ważnych zmiennych decyzyjnych i parametrach mających wpływ na ocenę rozważanych alternatyw i stopień realizacji założonych celów. Stąd też istotną funkcję w procesie decyzyjnym pełnią metody analityczne, pozwalające na dostarczenie informacji oraz oszacowań niezbędnych do redukcji czynników ryzyka występujących w danej sytuacji decyzyjnej [Bartkiewicz, 2013; Marshall i Oliver, 1995].

Prognoza, oszacowanie czy powstająca przy użyciu metod analitycznych ekspertyza nieodłącznie obarczone są jednak błędem – wykorzystu-

jąc je w problemach decyzyjnych decydent musi być świadomy niepewności, jaka wiąże się z zastosowaniem wyników analizy oraz istniejącym w związku z tym ryzykiem. Jak widać, niemal każda decyzja podejmowana z wykorzystaniem prognoz, oszacowań, ekspertyz na ogół staje się decyzją w warunkach ryzyka.

### **3.2.3. Ogólna charakterystyka wspomagania decyzji w organizacji**

Poszczególne etapy procesu podejmowania decyzji związane są zazwyczaj z wykorzystaniem kilku typowych elementów, używanych przez decydenta do uporządkowania analizowanego zagadnienia i wyboru decyzji właściwych dla rozwiązywanego problemu. Scharakteryzujemy je pokrótce, ponieważ pełnią one istotne funkcje w komputerowym wspomaganiu tego procesu.

Jak wspomnieliśmy w poprzednim punkcie, podjęcie decyzji wymaga określenia zbioru potencjalnych możliwych rozwiązań danego problemu, nazywanych wariantami decyzji lub alternatywami (opcjami) decyzyjnymi, a następnie wyboru jednej z nich.

Dla porównywania wariantów decyzji i wyboru najlepszego niezbędne jest zdefiniowanie celów podejmowanej decyzji, wynikających z problemu, jaki ma ona rozwiązywać. Cele te zazwyczaj określane są w postaci pewnych warunków, nazywanych kryteriami decyzji (lub funkcją celu), oceniających daną opcję pod kątem jej dopasowania do realizacji tychże celów.

Tak więc kryterium (funkcja celu) stanowi miarę dopasowania danego wariantu decyzji do założonych celów. W procesie podejmowania decyzji interesować nas będzie wybór wariantu najlepszego (optymalnego) pod kątem założonego kryterium (funkcji celu). Do typowych kryteriów w decyzjach biznesowych należą m.in. minimalizacja kosztów lub maksymalizacja zysku. Kryterium może mieć jednak bardziej złożony charakter – być pochodną zysku lub kosztów (np. NPV) lub wiązać się z nimi w bardziej niejawny i zawikłany sposób. Przykładowo: kryterium wyboru wariantu decyzji jest jak najkrótszy czas realizacji określonych operacji tak, by uniknąć kosztów strat związanych, powiedzmy, ze starzeniem się produktu. Kryterium ma niekiedy charakter nawet trudniej uchwytnej czy wręcz sztucznej (jakość rozwiązania, jego użyteczność itp.).

Czasami funkcja celu nie jest określona jawnie, lecz wbudowana w logikę zagadnienia, np. w formie reguł lub procedur postępowania pozwalających ocenić, czy dany wariant decyzji będzie odpowiedni dla założonych celów, czy też nie. Okazuje się to częste przy tzw. decyzjach klasyfikacyjnych. Przykładowo: podjęcie decyzji o przyznaniu kredytu w przypadku konkretnego wniosku może wiązać się z określeniem jawnego kryterium oceny (np. w skali od 1 do 10), ale częściej polega na sprawdzeniu szeregu warunków dotyczących zdolności kredytowej aplikującego, jakości zabezpieczenia itp. W końcu kryterium wyrażane jest niekiedy w postaci pośrednich porównań parami poszczególnych wariantów decyzji (tzw. preferencji), dokonywanych przez decydenta.

Sytuacja komplikuje się w przypadku, gdy określając cele podejmowanej decyzji, trzeba wziąć pod uwagę kilka kryteriów oceny wariantów decyzji (rozwiązań problemu). Wówczas mówimy o decyzjach wielokryterialnych. Poszczególne kryteria zazwyczaj nie są w pełni zgodne między sobą (gdyby były, to prawdopodobnie nie musielibyśmy rozważać kilku kryteriów), a czasami nawet wręcz sprzeczne. Podejmując np. decyzję o zakupie jakiegoś urządzenia, kierujemy się zazwyczaj wręcz przeciwstawnymi celami, jakimi mogą być: cena, parametry jakościowe czy też koszty eksploatacyjne konkretnych modeli danego urządzenia. W związku z tym niezbędne jest scalenie (agregacja) poszczególnych kryteriów w sztuczną miarę oceny alternatyw, nazywaną zwykle użytecznością. Wymaga to zwykle dostarczenia przez decydenta informacji o istotności (lub preferencji) poszczególnych kryteriów lub budowy modelu agregacji na drodze analizy zbiorów danych historycznych.

Kolejne elementy modelu decyzji związane są z opisem samych wariantów (opcji) decyzji. Musi on uwzględniać wszelkie dane niezbędne do określania dla danego wariantu wartości kryterium. Można mówić tutaj o dwóch podstawowych rodzajach informacji. Wartości kryterium dla danego wariantu decyzji wyznaczane są bezpośrednio na podstawie dostarczanych przez decydenta informacji dla danej opcji decyzyjnej, nazywanych zmiennymi decyzyjnymi. Charakteryzują one więc dany wariant decyzji, a ich wartości różnicują między sobą poszczególne opcje, spośród których decydent dokonuje wyboru. W procesie podejmowania decyzji występują również pomocnicze informacje, nazywane parametrami decyzji – te na ogół nie są związane bezpośrednio z poszczególnymi wariantami, lecz opisują środowisko, otoczenie rozwiązy-



wanego problemu. Parametry mogą więc mieć zasadniczo takie same wartości dla wszystkich wariantów decyzji bądź ich znacznych podzbiorów. Natomiast są one niezbędne do wyznaczenia kryterium, a ponadto przy różnych wartościach parametrów optymalne mogą być różne warianty decyzji.

Ostatnim z najważniejszych standardowych elementów modelu podejmowania decyzji są ograniczenia. W rzeczywistych warunkach bowiem dopuszczalne są jedynie te rozwiązania problemu (alternatywy decyzyjne), które spełniają określone warunki. Typowymi przykładami ograniczeń przy podejmowaniu decyzji są np. wielkości zasobów, z jakich może korzystać organizacja przy realizacji celu.

Z punktu widzenia procesu podejmowania decyzji istotnym czynnikiem okazuje się również liczba rozważanych alternatyw decyzyjnych. Jeżeli nie jest ona zbyt duża, to wysiłki koncentrują się przede wszystkim na procesie budowy i oceny poszczególnych wariantów. Proces samego wyboru optymalnego wariantu staje się stosunkowo prosty – ponieważ nie mamy zbyt wielu alternatyw, możemy enumeratywnie wyznaczyć kryterium oceny dla każdej z nich i, w większości przypadków, bez żadnych poważniejszych problemów wybrać najlepszą czy też utworzyć ich ranking. W przypadku nieskończonej (lub skończonej, ale bardzo dużej) ilości alternatyw to oczywiście niemożliwe – zazwyczaj niezbędne staje się wtedy zdefiniowanie kryterium i ograniczeń w postaci formuł (równań i nierówności) matematycznych dla całych zbiorów rozwiązań, określonych przez zmienne decyzyjne i parametry, a następnie rozwiązanie matematycznego problemu optymalizacyjnego, polegającego na znalezieniu maksimum (minimum) funkcji celu przy danych ograniczeniach.

Przypomnijmy, że w punkcie 3.2.1 scharakteryzowano zadanie podejmowania decyzji jako trzyetapowy podproces zagadnienia rozwiązania problemu (lub wykorzystania pojawiającej się możliwości, okazji), złożony z fazy rozpoznania, projektowania możliwych alternatyw decyzyjnych oraz wyboru tej, która zostanie zrealizowana. Na koniec bieżących rozważań omówmy, oczywiście bardzo ogólnie, zadania informatycznego wspomagania na wszystkich tych etapach.

Rozpocznijmy od etapu rozpoznania. Jak wspomnieliśmy, w jego trakcie identyfikowane i definiowane są problemy, cele, jakie mają zostać osiągnięte i możliwości rozwiązań. Problem występuje zazwyczaj z powodu braku satysfakcji związanej z obecnym stanem rzeczy. Decydent



próbuję określić, czy problem istnieje, zidentyfikować symptomy, określić wielkość i jawnie go zdefiniować. Zbierana jest także informacja odnosząca się do zakresu problemu: możliwych rozwiązań, ogólnych ograniczeń i otoczenia, w jakim się pojawił.

Wspomaganie podejmowania decyzji na tym etapie wiąże się przede wszystkim z możliwością dostarczania niezbędnych informacji (zarówno ze źródeł wewnętrznych, jak i zewnętrznych) na temat zakresu problemu. Najistotniejsze znaczenie na ogół mają tu dane z systemów informatycznych zarządzania, monitorujących bieżący stan działań organizacji, a także z systemów raportowania (zarówno okresowego, jak i *ad hoc*) oraz z bezpośrednich zapytań do baz danych. Podstawowym źródłem informacji dotyczącej otoczenia problemu jest oczywiście Internet – jako medium pozyskiwania informacji handlowej, uwarunkowań transportu, przepisów prawnych, tranzytowych, importowych itp. [Papińska-Kacperek, 2007]. Do ważnych źródeł wspierających (w przypadku konieczności podjęcia wielu decyzji) zaliczyć należy wyspecjalizowane systemy informatyczne, takie jak systemy GIS (dostarczające informacji o charakterze geograficznym i przestrzennym) oraz systemy CRM (pozwalające na analizowanie klientów, ich historii zakupów itp.).

Zwróćmy jednak uwagę, że problem pozyskiwania danych sam w sobie stwarzać może szereg trudności. Operowanie na surowych informacjach coraz częściej stwarza kłopoty związane z rozmiarami zbiorów danych oraz trudnością ich interpretacji. Stąd też na etapie rozpoznania coraz większe znaczenie ma również wsparcie procesu podejmowania decyzji z wykorzystaniem rozmaitych technik analitycznych.

Istotną funkcję pełnić mogą systemy wyszukiwania danych, ułatwiające przeszukiwanie zasobów informacji niestukturalnej, zwłaszcza w Internecie. Należy wspomnieć również o hurtowniach danych dostarczających analitycznej informacji zagregowanej. Kolejnym ważnym narzędziem wspomagającym są narzędzia eksploracji danych oraz sieci web pozwalające na wykrywanie zależności oraz związków danych. W przypadku konieczności analizy dużych zbiorów informacji, zwłaszcza o charakterze niestukturalnym, dobrze skorzystać w fazie rozpoznania z systemów analitycznych Big Data. Gdy mamy do czynienia z wieloma decyzjami, pomocne w interpretacji danych będą narzędzia ich wizualizacji.

Na etapie projektowania budowane są możliwe rozwiązania problemu, warianty decyzji, które je realizują, oraz szacowana jest wstępnie wykonalność poszczególnych opcji. Definiuje się tu również kryteria oceny wariantów decyzji, zmienne decyzyjne i parametry modelu decyzji – w tej fazie procesu istotną funkcję pełnią zatem wspomniane w poprzednim punkcie systemy informatyczne: dostarczające, porządkujące, analizujące i wizualizujące informacje.

We wsparciu etapu projektowania najważniejsze znaczenie mają jednak różnego rodzaju modele dziedzinowe z zakresu rozwiązywanego problemu, pozwalające na określanie wartości liczbowych czy też logiczno-symbolicznych oraz różnego rodzaju zmiennych, jak np. rozmaite modele finansowe i rachunkowości zarządczej: kalkulacji kosztów, zysków, wyceny rozmaitych działań itp. W zależności od dziedziny problemu istotną funkcję mogą pełnić np. modele logistyczne (optymalizacji kosztów dostawy, tras przejazdów, wielkości produkcji, harmonogramowania dostaw itp.) czy techniczne. Są one zazwyczaj niezbędne do budowy poszczególnych wariantów decyzji oraz kryteriów ich wyboru.

Jak wspomnieliśmy, nie zawsze kryterium dla problemu decyzyjnego określa się w sposób bezpośredni. W przypadku tzw. decyzji klasyfikacyjnych wbudowane jest ono w logikę zagadnienia np. w postaci modelu pozwalającego ocenić, czy dany wariant decyzji będzie odpowiedni dla założonych celów, ewentualnie wskazującego, która opcja okaże się odpowiednia dla danej sytuacji decyzyjnej. W takim przypadku w fazie projektowania niezbędne staje się stworzenie procedury klasyfikującej. Może mieć ona formę logiczno-symboliczną – reguł decyzyjnych lub drzew decyzyjno-klasyfikacyjnych – a w bardziej złożonych przypadkach np. systemu ekspertowego. Innym typowym przykładem są klasyfikatory liczbowe, tworzone na podstawie zbioru przykładów, przy pomocy metod eksploracji danych.

Przy wielu decyzjach ważnym elementem okazuje się czynnik ryzyka. W takim przypadku w fazie projektowania należy również uwzględnić kwestie metod pozwalających na obliczenie prawdopodobieństw skutków decyzji oraz wynikającego z nich ryzyka. Wykorzystuje się tu zazwyczaj rozmaite modele probabilistyczne oraz eksploracji i statystycznej analizy danych.

Samo porządkowanie sytuacji decyzyjnej w warunkach ryzyka jest zadaniem dosyć złożonym. Niełatwo określić możliwe warianty decyzji,

czynniki niepewności, które wpływają na ich skutki, wartości możliwych zmiennych losowych oraz ich prawdopodobieństwa. Do wspomagania fazy projektowania przy tego rodzaju decyzjach stworzone zostały specjalne metody porządkowania i wizualizacji. Do najważniejszych należą probabilistyczne drzewa decyzyjne analizy danych oraz diagramy wpływu.

Często mamy do czynienia z sytuacją, w której warianty decyzji oceniane są pod kątem wielu (zazwyczaj skonfliktowanych) kryteriów. W takiej sytuacji w fazie projektowania musi zostać znaleziony również sposób agregacji poszczególnych kryteriów w funkcję użyteczności. Wymaga to określenia względnej istotności (lub preferencji) poszczególnych kryteriów oraz budowy modelu. Na ogół stosuje się liniowe modele agregacji, polegające na określeniu stałej wagi każdego kryterium w łącznej funkcji użyteczności. Czasami wykorzystuje się także rozwinięcia tej metody przy wykorzystaniu systemów rozmaitych kart punktowych; niekiedy ma miejsce tworzenie modelu agregacji na podstawie danych, za pomocą metod ich eksploracji.

Etap wyboru to ostatnia faza podejmowania decyzji. Wymaga on określenia toku działań realizowanych w celu rozwiązania problemu. W przypadku skończonego zbioru alternatyw w fazie tej szacowane są wartości kryterium dla wszystkich wariantów decyzji, stworzony zostaje ich ranking, a następnie dokonuje się wyboru najlepszego sposobu postępowania. Przy nieskończonej liczbie wariantów decyzji korzysta się z odpowiedniego modelu optymalizacyjnego. Dla decyzji o charakterze klasyfikacyjnym określa się – przy pomocy przygotowanych procedur klasyfikacyjnych – właściwe decyzje dla badanej sytuacji decyzyjnej.

W fazie wyboru wykorzystywane są więc przygotowane wcześniej, na etapie projektowania, modele definiujące poszczególne warianty decyzji oraz obliczające wartości kryteriów lub stosujące inną formę wyboru najlepszej alternatywy decyzyjnej. Nie jest to jednak tak proste, jak mogłoby się wydawać. Jak zostało wspomniane, istnieje wiele trudnych do rozpoznania czynników, często o charakterze niejawnym, mogących wpływać na ostateczny wybór wariantu decyzji. Pojawia się konieczność uwzględnienia elementu niepewności i ryzyka; wyniki działania modeli muszą być uważnie przeanalizowane.

Na tym etapie swoje zastosowanie znajdują przede wszystkim modele decyzyjne, analityczne i komputerowe metody badań operacyjnych, pozwalające na znalezienie rozwiązania optymalnego pod kątem założonych celów lub przynajmniej określenie najlepszego rozwiązania spośród rozważanych.

W przypadku występowania warunków całkowitej niepewności w fazie wyboru wykorzystuje się przygotowane wcześniej modele do symulacji skutków decyzji (określanych przez wartości kryteriów) dla różnych scenariuszy rozwoju sytuacji. Oprócz tego rodzaju symulacji, typu „co – jeżeli” (określanych również analizami wrażliwości), stosowane są również inne podejścia do badań z użyciem modelu, oparte na analizach sterowanych celami. W takim przypadku nakłada się różnego rodzaju warunki (cele) na możliwe skutki decyzji, a następnie próbuje się określić, jakie muszą być spełnione zależności związane z wartościami zmiennych decyzyjnych i parametrów, aby cele te mogły zostać osiągnięte.

Podjęte decyzje wprowadzane są następnie w życie. Etapy implementacji i monitorowania wykraczają już poza fazę podejmowania decyzji. Polegają one na wykonaniu akcji związanych z wybranymi wariantami decyzji oraz monitorowaniu rzeczywistych skutków prowadzonych działań.

Wspomaganie procesu rozwiązywania problemu na tych etapach związane jest przede wszystkim z wykorzystywaniem systemów informatycznych służących do obsługi procesu realizacji decyzji, rejestracji transakcji gospodarczych będących skutkiem zrealizowanej decyzji oraz bieżącej operacyjnej analizy ich efektów finansowych lub innych założonych wcześniej celów.

Istotne znaczenie na tym etapie ma również wspomaganie funkcji komunikacyjnych, w celu zbierania informacji o procesie realizacji decyzji, związane np. z wykorzystaniem jako medium komunikacyjnego usług sieci Internet czy też z realizacją podjętej decyzji za pośrednictwem mechanizmów elektronicznej wymiany danych EDI.

### **3.3. Systemy Wspomagania Decyzji**

#### **3.3.1. Pojęcie Systemu Wspomagania Decyzji (DSS)**

Systemy Wspomagania Decyzji (*Decision Support Systems* – DSS) są zorganizowanym zbiorem ludzi, procedur, baz danych i urządzeń wykorzystywanych w celu wspomagania podejmowania decyzji na wszystkich etapach tego procesu, poczynając od rozpoznania, czyli zdefiniowania

problemu i zakwalifikowania go do określonej grupy standardowej, następnie poprzez wybór odpowiednich danych, stworzenie i analizę modelu informacyjnego opisującego rzeczywistość, aż do pomocy w generowaniu wariantów dopuszczalnych rozwiązań oraz w wyborze najlepszego rozwiązania [Stair, 1992; Radzikowski, 1990].

DSS stosowane są w sytuacji, gdy podjęcie decyzji jest zadaniem skomplikowanym, a także kiedy mamy do czynienia z problemami słabo ustrukturalizowanymi. Zgodnie więc z tym, o czym wspomnieliśmy wcześniej, systemy te skupiają uwagę raczej na wspomaganiu, a nie automatyzacji decyzji. Tym samym ich celem jest przede wszystkim podnoszenie skuteczności i jakości decyzji, a nie sprawności zarządzania [Sroka, 1994; Radzikowski, 1990]. DSS pozwalają na realizację następujących zadań [Zieliński 2000]:

- wyszukiwanie danych jednostkowych, czyli wyodrębnionych ze zbiorów danych;
- swobodny dostęp do danych oraz ich analiza przyczynowo-skutkowa;
- dostarczanie danych zbiorczych wcześniej zdefiniowanych;
- przygotowywanie projektów możliwych decyzji;
- przedstawienie konsekwencji (ocena) proponowanych decyzji przy wykorzystaniu modeli obliczeniowych i symulacyjnych: „co – jeżeli”;
- określanie danych problemu niezbędnych do realizacji konkretnego celu, czyli wykonywanie analiz sterowanych celami;
- wybranie wariantu decyzji na podstawie zadanych kryteriów.

Analizując zadania DSS, można powiedzieć, że są to systemy informatyczne, które wspomagają decydentów w słabo ustrukturalizowanym środowisku decyzyjnym przy wykorzystaniu analitycznych modeli decyzyjnych z dostępem do baz danych. Biorąc pod uwagę powyższe cele Systemów Wspomagania Decyzji, możliwe staje się sformułowanie kilku cech tej klasy systemów [Jabłoński i Bartkiewicz, 2006]:

- 1) Orientacja problemowa – DSS koncentrują się na wspomaganiu rozwiązywania rzeczywiście powstających problemów decyzyjnych; są wyspecjalizowane w kierunku podejmowania ściśle określonych decyzji. DSS koncentruje się więc na konkretnym problemie (bądź stosunkowo wąskiej ich grupie), w przeciwieństwie do klasycznych systemów informatycznych dostarczających danych wspomagających kompleksowo proces zarządzania daną dziedziną. W konsekwencji DSS wspiera przede wszystkim pojedynczych decydentów i niewielkie grupy, a dopiero w dalszej kolejności całą organizację.

- 2) Wspomaganie decydentów głównie przy decyzjach niestrukturalnych i częściowo strukturalnych poprzez łączenie osądu człowieka i systemu informatycznego – DSS współpracuje z decydem, dostosowany jest do jego indywidualnych wymagań. Podczas rozwiązywania problemu decydent musi mieć kontrolę nad każdym krokiem procesu podejmowania decyzji.
- 3) Wsparcie we wszystkich fazach procesu decyzyjnego.
- 4) Do analizy sytuacji decyzyjnych DSS zazwyczaj wykorzystuje różne modele i procedury analityczne.
- 5) Konieczność wykonywania skomplikowanych, wyrafinowanych analiz i porównań przy wykorzystaniu zaawansowanych pakietów oprogramowania – zadania stawiane DSS wymagają zwykle zastosowania znacznie bardziej skomplikowanych algorytmów działania niż w przypadku tradycyjnie rozumianego przetwarzania danych. W związku z tym Systemy Wspomaganie Decyzji tworzone są często w specjalistycznych środowiskach programistycznych, dostarczających odpowiednich procedur obliczeniowych (jak np. SAS, Matlab czy w prostszych przypadkach MS Excel). DSS może również stanowić platformę integrującą dla samodzielnych zewnętrznych programów analitycznych.
- 6) Możliwość przetwarzania dużych ilości danych – DSS nie jest systemem przetwarzania danych masowych w tradycyjnym tego słowa znaczeniu. Wiele z Systemów Wspomaganie Decyzji wyposażonych zostaje jednak w możliwości przeszukiwania obszernych baz danych, co pozwala na integrację tradycyjnego przetwarzania danych z metodami badań operacyjnych, ułatwiającą kadrze menedżerskiej stosowanie ilościowych technik zarządzania. Należy ponadto zwrócić uwagę, że możliwości gromadzenia obszernych zbiorów informacji, oferowane przez współczesną technologię informatyczną, nie do końca idą w parze z możliwościami ich interpretacji i wykorzystania. Fakt ten uznawany jest obecnie za jedno z najistotniejszych „wąskich gardeł” w procesie rozwoju cywilizacyjnego. DSS wyspecjalizowane w analizie dużych zbiorów danych (określanych jako systemy Big Data) oraz analizujące zależności między danymi (przy pomocy metod eksploracji danych) stanowią jedną z najbujniej rozwijających się dziedzin informatyki.
- 7) Pobieranie i przetwarzanie danych z różnych źródeł – niektóre informacje mogą rezydować w bazach danych na komputerach

osobistych, kolejne znajdować się w innych systemach operacyjnych lub sieciowych. Dane są też przechowywane w arkuszach kalkulacyjnych lub specyficznych formatach związanych np. z oprogramowaniem analitycznym. DSS powinien mieć możliwości pozwalające na integrację danych z jak największej liczby źródeł.

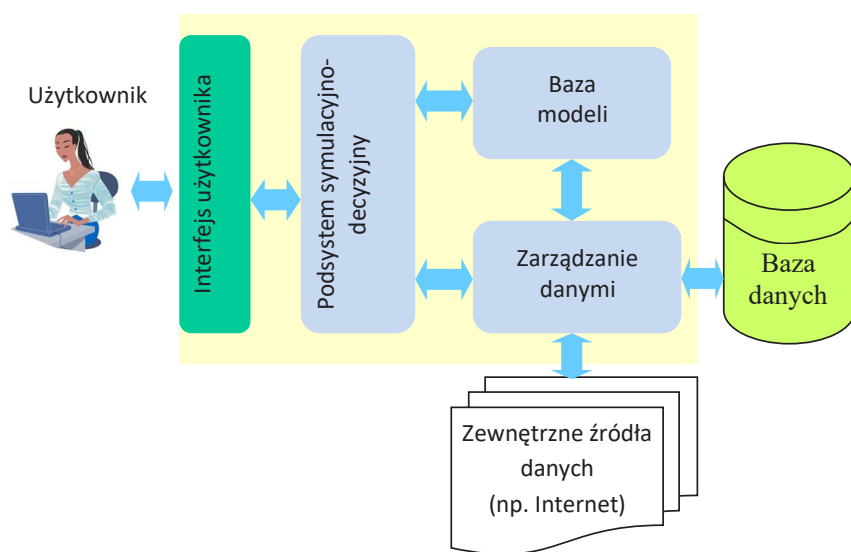
- 8) Tworzenie raportów i elastyczność prezentacyjna – jedną z przyczyn powstania DSS był fakt, że systemy transakcyjne i informatyczne zarządzania nie były dostatecznie elastyczne dla zaspokojenia wszystkich potrzeb informacyjnych i problemów decydentów. Wyjściami z tego typu systemów są zwykle drukowane raporty o ustalonej strukturze i formie. W przypadku Systemów Wspomagania Decyzji menedżerowie mogą otrzymać informacje, jakich potrzebują, w formie dostosowanym do ich indywidualnych wymagań. DSS opierają się przede wszystkim na interakcyjnych raportach na ekranie komputera, pozwalających na nawiązanie dialogu użytkownika z systemem i współpracę z nim w trybie *online*. W zależności od preferencji osoby rozwiązującej problem wyjście może być prezentowane również w formie drukowanej – zasadniczo jednak DSS są zorientowane ekranowo.
- 9) Orientacja tekstowa i graficzna – sposób prezentacji informacji wyjściowych w DSS jest zwykle bardzo elastyczny: obejmuje on zarówno format tekstowy, jak i graficzny (a coraz częściej również multimedialny). Współczesne DSS mogą tworzyć informacje tekstowe, tabelaryczne, wykresy liniowe, kołowe, linie trendu itp. Stosowanie elastycznej techniki prezentacyjnej pozwala decydentom na lepsze zrozumienie sytuacji i łatwiejszą interpretację wyników działania systemu. W powiązaniu z prostymi i wygodnymi technikami konwersacji z systemem DSS ułatwia przełamanie bariery psychologicznej decydentów wobec komputerów.
- 10) Elastyczność i adaptacyjność – Systemy Wspomagania Decyzji dostosowują się do zmian, jakie zachodzą w otoczeniu decydenta, umożliwiając indywidualne podejście do problemu decyzyjnego. Użytkownik często ma możliwość ingerencji w strukturę wewnętrzną DSS i konfigurowania jej – w zależności od swoich indywidualnych potrzeb i uwarunkowań sytuacji decyzyjnej.

Oczywiście nie wszystkie Systemy Wspomagania Decyzji posiadają pełny zestaw określonych wyżej charakterystyk. Jako że tworzone są



one na potrzeby konkretnych problemów decyzyjnych, stopień ich zaawansowania zależy w dużym stopniu od samego problemu, jak i potrzeb decyden-  
 ta. W wielu przypadkach Systemy Wspomagania Decyzji mają mniejszy zakres, posiadając jedynie wybrane z wymienionych elementów. Przy budowie lub wyborze gotowego DSS zyski z jego właściwości powinny być w równowadze z kosztami systemu, jego złożonością i stopniem kontroli nad nim. Z jednej strony możemy mieć więc do czynienia z wyrafinowanymi DSS o znacznym stopniu skomplikowania, a z drugiej – z prostymi rozwiązaniami, zaimplementowanymi np. w postaci modelu w arkuszu kalkulacyjnym.

Systemy Wspomagania Decyzji tworzone są więc z myślą o wspieraniu konkretnych indywidualnych decyzji. W związku z tym trudno jest mówić o jakimś uniwersalnym i jednolitym wzorcu ich budowy. DSS mogą mieć silnie zróżnicowaną strukturę wewnętrzną, w zależności od problemu, dla którego rozwiązania powstały, oraz preferencji użytkownika odnośnie działania systemu. Tym niemniej można wyróżnić pewien ramowy schemat ich budowy, obejmujący kilka standardowych podsystemów (rysunek 3.2):



**Rysunek 3.2.** Ogólna struktura Systemu Wspomagania Decyzji

**Źródło:** opracowanie własne.



Baza danych zawiera aktualne dane dotyczące działalności organizacji i jej otoczenia. Z tego powodu można czasami mówić o bazie danych wewnętrznych i zewnętrznych organizacji. Pierwsze pochodzą przede wszystkim z baz danych transakcyjnych oraz innych systemów informatycznych, drugie – od otoczenia gospodarczo-politycznego organizacji. Działanie DSS w sposób kluczowy zależy od jakości danych, tak więc powinny być one uważnie kontrolowane.

Baza modeli składa się z modułów, z których każdy zawiera opis odpowiednich zachowań związanych z daną sytuacją decyzyjną. Modele te wspierają podejmowanie decyzji na różnych poziomach zarządzania w zakresie rozmaitych funkcji kierowniczych i w odmiennych dziedzinach działalności obiektu. Z tego powodu ważne staje się zapewnienie – szczególnie dla tego elementu DSS – sposobności ciągłej modyfikacji i rozbudowy. Ta część systemu decyduje bowiem o rzeczywistych możliwościach całego systemu.

Podsystem symulacyjno-decyzyjny dokonuje – na podstawie żądań użytkownika oraz istniejących danych – wyboru kombinacji modeli niezbędnych do rozwiązania zadania oraz danych wejściowych dla tych modeli, a także wykonuje za ich pomocą konieczne obliczenia. Zauważmy, że przepływy informacyjne tego podsystemu z bazą danych mają charakter dwukierunkowy. Może on nie tylko pobierać dane wejściowe, ale również zapisywać w bazie danych informacje będące wynikiem działania DSS. Podobnie w przypadku bazy modeli – podsystem symulacyjno-decyzyjny może modyfikować je pod kątem konkretnego problemu oraz istniejących danych (np. poprzez reestymację ich parametrów), zapisując w bazie zmiany.

Interfejs użytkownika (menedżer dialogu) ma zapewnić wysoki komfort obsługi. Pozwala decydentowi na łatwy dostęp i manipulowanie DSS. Steruje pracą modułu symulacyjno-decyzyjnego oraz przekazuje użytkownikowi otrzymane od niego wyniki obliczeń. Użytkownik musi dostać to, czego zażąda, w możliwie różnorodnej formie, a system powinien być przygotowany na żądania niestandardowe. Nowoczesne interfejsy mają za zadanie jak największe uproszczenie sposobu komunikacji z systemem – poprzez komunikację graficzną, multimedialną, z wykorzystaniem powszechnie stosowanej terminologii biznesowej, a nawet przy użyciu języka naturalnego.

### 3.3.2. Modele w Systemach Wspomagania Decyzji

Jak mogliśmy zaobserwować, centralnymi elementami Systemu Wspomagania Decyzji, decydującymi o jego możliwościach, są znajdujące się w bazie DSS modele. Scharakteryzujemy je więc nieco bliżej.

Modelem nazywamy abstrakcję (czyli świadome i celowe uproszczone odwzorowanie) określonego fragmentu rzeczywistości. Stanowi on więc schematyczną reprezentację – rzeczywistość jest zbyt złożona, aby opisywać ją dokładnie, zaś większa część tej złożoności okazuje się nieistotna dla rozwiązania konkretnego problemu.

Przydatność modelu warunkują następujące czynniki [Krawczyk, 2001]:

- konstrukcji modelu musi towarzyszyć świadomość celu, w jakim został on stworzony;
- model powinien odzwierciedlać wszystkie istotne z punktu widzenia celu elementy rzeczywistości, ich własności oraz relacje między nimi;
- model powinien być wewnętrznie zgodny oraz zgodny z informacjami, które były podstawą jego konstrukcji;
- model powinien uwzględniać relacje między realnym fragmentem rzeczywistości, a jego otoczeniem.

Tworzenie modelu wymaga więc dokonania pewnej konceptualizacji problemu, wyboru reprezentowanych elementów rzeczywistości oraz stworzenia ich abstrakcyjnej postaci w formie składowych modelu, np. zmiennych i matematycznych lub symboliczno-logicznych zależności między nimi. Proces ten nazywamy modelowaniem.

Biorąc pod uwagę wybraną formę reprezentacji rzeczywistości, można mówić o modelach różnego rodzaju, m.in. fizycznych (zalicza się do nich przede wszystkim różnego rodzaju prototypy urządzeń mechanicznych, elektronicznych itp.), graficznych (reprezentujących modelowane aspekty problemu z wykorzystaniem różnego rodzaju diagramów i innych form obrazkowych) czy też matematycznych [Jabłoński i Bartkiewicz, 2006].

Z punktu widzenia możliwości wykorzystania modelu przez systemy informatyczne największe nasze zainteresowanie budzi ostatnia grupa – stąd dalsze rozważania ograniczamy do niej. Istnieje przy tym cały szereg metod reprezentacji świata rzeczywistego przez programy komputerowe, wśród których wyróżnić można dwa podstawowe podejścia:

- numeryczne – wykorzystujące do opisu stanu modelowanego wycinka rzeczywistości zmienne liczbowe, równania i formuły algebraiczne;
- symboliczne (logiczne) – opisujące rzeczywistość w postaci zestawów umownych symboli oraz zależności logicznych między nimi.

Podział ten nie ma oczywiście charakteru ostrego. W zasadzie można powiedzieć, że każdy program komputerowy stanowi model rzeczywistości wykorzystujący do jej opisu zarówno aparat matematyczny, jak i logiczny. W zależności od przewagi udziału każdego z nich można zakwalifikować go do którejś z grup. Jako przykłady krańcowych przypadków wymienia się, z jednej strony, skomplikowane modele matematyczne wymagające zastosowania złożonych algorytmów numerycznych, zaś z drugiej – systemy oparte na inżynierii wiedzy.

Dlaczego w ogóle prowadzi się proces modelowania i tworzy modele? Wymienić tutaj można wiele rozmaitych powodów. Do najważniejszych z nich zaliczamy [Krawczyk, 2001; Sharda, Delen i Turban, 2017]:

- manipulowanie modelem (zmiana zmiennych decyzyjnych lub środowiska) jest znacznie łatwiejsze niż manipulowanie prawdziwym systemem; eksperymentowanie okazuje się prostsze w realizacji i nie zakłóca codziennych działań organizacji;
- modele umożliwiają kompresję czasu, lata operacji można symulować w minutach lub sekundach systemowego czasu komputera;
- koszt analizy modelowania zwykle jest znacznie niższy niż koszt podobnego eksperymentu przeprowadzonego na prawdziwym systemie;
- koszt popełniania błędów podczas eksperymentu okazuje się znacznie niższy, gdy zamiast prawdziwych systemów używane są modele;
- otoczenie biznesowe wiąże się ze znaczną niepewnością; dzięki modelowaniu menedżer może oszacować ryzyko wynikające z określonych działań;
- modele matematyczne umożliwiają analizę bardzo dużej, czasem nieskończonej ilości możliwych rozwiązań; nawet w prostych problemach menedżerowie często mają do wyboru wiele alternatyw;
- modele ułatwiają i usprawniają proces uczenia się i szkolenia, a także wzmacniają jego efekty.

Ze względu na stopień naszej wiedzy o reprezentowanym fragmencie rzeczywistości wyróżniamy następujące rodzaje modeli [Jabłoński i Bartkiewicz, 2006]:

- modele algorytmiczne,
- modele dedukcyjne,
- modele indukcyjne.

Pierwsze wymagają dogłębnego zrozumienia natury problemu i istnienia wiedzy o sposobie jego rozwiązania, pozwalającej na wyspecyfikowanie równań lub algorytmów opisujących zachowanie modelowanego systemu w formie jawnie zdefiniowanych deterministycznych zależności. Są to tzw. „silne” modele, z silnymi założeniami i bez parametrów wolnych (szacowanych na podstawie danych). Stanowią one niewątpliwie najefektywniejszą metodologię rozwiązania problemu. Informacja uzyskana przy pomocy tego rodzaju modeli może być przyjmowana jako pewna. Zwykle jednak stosowane mogą być one jedynie w przypadku systemów stosunkowo prostych, dla których możliwe okazuje się precyzyjne zrozumienie i opisanie modelowanych fenomenów.

Modele dedukcyjne stosowane są w przypadku problemów, dla których nie jesteśmy w stanie zbudować precyzyjnej specyfikacji matematycznej lub algorytmicznej (logicznej). Niemniej jednak możemy na podstawie bezpośredniej obserwacji systemu wykryć pewne stałe wzorce jego zachowania. Pozwala to na określenie przez modelującego pewnych ogólnych zasad opisujących dynamikę systemu. Typowymi przykładami zastosowania rozumowania dedukcyjnego są systemy ekspertowe oraz wnioskowanie statystyczne. W pierwszym przypadku modelujący tworzy – wraz z „ekspertem” – bazę wiedzy określającą zachowanie systemu w postaci szeregu reguł logicznych o niewielkiej liczbie (lub nawet bez) parametrów wolnych. Systemy wnioskowania statystycznego, takie jak np. regresja liniowa, czynią silne założenia odnośnie natury związku między zmiennymi, umożliwiając oszacowanie parametrów wolnych na podstawie zaobserwowanych danych.

W miarę jak wzrasta złożoność systemu możliwość bezpośredniego, precyzyjnego określenia pewnych stałych wzorców jego zachowania zwykle maleje. Powiązania między zmiennymi stają się niejawne, nie można poczynić niemal żadnych założeń odnośnie ich natury. Cała wiedza na temat zachowania systemu musi zostać pozyskana wyłącznie na podstawie danych (obserwacji). Problemy tego typu rozwiązywane mogą być przy wykorzystaniu nieparametrycznych metod, takich jak sieci neuronowe, nieparametryczna regresja, adaptacyjne systemy rozmyte czy też algorytmy genetyczne. Zastosowanie modeli indukcyjnych w procesie

analizy danych pozwala na określenie nie tylko wartości parametrów, ale również kształtu odwzorowania między zmiennymi.

Jak widać, w miarę wzrostu złożoności modelowanego problemu stopień jego poznania *a priori* maleje i ciężar przesunął się w kierunku analizy wzorców zachowania na podstawie obserwacji. Systemy Wspomagania Decyzji wykorzystują w takiej sytuacji najczęściej nie tylko modele wspomagające sam proces decyzyjny, ale również modele analityczne – identyfikujące pewne istotne zmienne i fakty występujące w badanym problemie oraz porządkujące zebraną wiedzę o zależnościach między nimi. Modele te pełnią w systemach DSS rozmaite funkcje i wykonują wiele zadań, m.in. [Sharda, Delen i Turban, 2017]:

- mogą dostarczać informacji (lub ich szacunków), których nie można bezpośrednio zaobserwować ani zmierzyć;
- mogą definiować zależności między zmiennymi problemu decyzyjnego;
- mogą określać skutki analizowanych alternatyw decyzyjnych (wartości kryterium decyzyjnego);
- mogą generować alternatywne decyzje i wspierać wybór;
- mogą zarządzać grupami alternatywnych decyzji;
- mogą oceniać niepewność i ryzyko;
- mogą wyjaśnić problem decyzyjny, ułatwiając decydentowi analizę;
- mogą wyjaśnić decydentowi, dlaczego pewne zależności w problemie decyzyjnym mają taki charakter lub dlaczego pewne decyzje powinny być podejmowane, a inne nie.

Struktura modeli wspomagania decyzji może być oczywiście bardzo zróżnicowana, dostosowana do konkretnych realizowanych przez nie zadań. Zazwyczaj jednak odpowiada ona elementom modelu decyzji określonym w pierwszej części podrozdziału 3.2.3. Zjawiska i czynniki modelowanej sytuacji decyzyjnej reprezentowane są w formie zmiennych ilościowych lub symboliczno-logicznych. Jako typowe ich rodzaje wyróżniamy [tamże]:

- zmienne decyzyjne (zmienne kontrolowane) – zmienne wejściowe; zazwyczaj opisują poszczególne alternatywy decyzyjne, spośród których decydent dokonuje wyboru; ich wartości kontrolowane są więc przez decydenta;
- parametry (zmienne niekontrolowane) – opisują czynniki pozostające poza kontrolą decydenta, takie jak elementy środowiska, oto-

czenia rozwiązywanego problemu czy też charakterystyki modelu wspólne dla wszystkich wariantów decyzji (lub ich grup); są one jednak również niezbędne do wyznaczenia kryteriów decyzji, a także w sytuacji różnych wartości parametrów;

- zmienne wynikowe – zmienne wyjściowe, zależne; określają wartości kryterium decyzyjnego dla poszczególnych alternatyw albo inne zmienne niezbędne do podjęcia decyzji; zmienne wynikowe zależą od wartości zmiennych decyzyjnych i zmiennych niekontrolowanych;
- zmienne pośrednie – określają pośrednie wielkości konieczne do wyznaczenia zmiennych wynikowych modelu; przykładowo: model określający płacę pracownika może wykorzystywać zmienną pośrednią, przechowującą wartość obliczonej płacy zasadniczej.

Zależności między zmiennymi definiowane są przy pomocy formuł algebraicznych lub logicznych. Zasadniczo modele te określają pewne odwzorowania między zmiennymi decyzyjnymi a zmiennymi wynikowymi, wykorzystujące informacje przechowywane w zmiennych niekontrolowanych. Do magazynowania pośrednich wyników działania w modelu mogą zostać użyte pewne zmienne pośrednie.

### **3.3.3. Podstawowe metody wspomaganie decyzji stosowane w DSS**

#### **3.3.3.1. Modele dziedzinowe**

Wiedza dziedzinowa dostarcza narzędzi do ustalania istotnych informacji ekonomicznych, socjologicznych, technicznych, fizycznych czy też z innych, związanych z podejmowaną decyzją obszarów. Przykładowo: modele finansowe oferują instrumenty konieczne do obliczeń finansowych takich wielkości, jak np. strumień pieniądza czy stopy zwrotu. Modele techniczne pozwalają obliczyć m.in. wytrzymałość materiałów, zużycie energii, ilości wykorzystywanych surowców itp.

Kilka modeli dziedzinowych z różnych wybranych obszarów, wraz z typowymi przykładami wykorzystywanych przez nie zmiennych, zaprezentowano w tabeli 3.2.

**Tabela 3.2.** Przykładowe modele dziedzinowe z wybranych obszarów

Obszar działania	Zmienne decyzyjne	Zmienne niekontrolowane	Zmienne wynikowe
Finanse	alternatywy inwestycyjne i kwoty	stopa inflacji, stopa odsetkowa, dane o konkurencji	całkowity zysk, ryzyko, stopa zwrotu, zysk na akcję
Marketing	budżet reklamowy, warianty reklamy	dochody klienta, działania konkurencji	udział w rynku, satysfakcja klienta
Produkcja	co i ile produkować, poziomy zapasów	zdolność maszyny, technologia, ceny materiałów	koszt całkowity, poziom jakości
Księgowość	wykorzystanie środków trwałych, harmonogram audytu	dane techniczne środków trwałych, wysokość podatków, wymogi prawne	koszty środków trwałych, poziom błędu
Transport	harmonogram dostaw, dane pojazdów	odległości dostaw, regulacje prawne	całkowity koszt transportu, terminy płatności
Usługi	liczba pracowników	popyt na usługi	satysfakcja klienta

**Źródło:** opracowanie własne na podstawie [Sharda, Delen i Turban, 2017].

Modele te mogą być używane do oceny poszczególnych alternatyw decyzyjnych (w jakim stopniu odpowiadają one założonym celom decyzji) poprzez obliczenie wartości przyjętej funkcji celu (zysk, koszty itp.). W bardziej złożonych problemach decyzyjnych stosowane są również do wyznaczania wartości istotnych zmiennych, ograniczeń itp.

W zależności od sytuacji modele dziedzinowe mogą być bardzo proste – dobrym przykładem jest krótka formuła obliczająca zysk dla danego wolumenu produkcji:

$$\text{zysk(wolumen)} = \text{przychód(wolumen)} - \text{koszty zmienne(wolumen)} - \text{koszty stałe}$$

W innych przypadkach modele mogą mieć bardziej złożony charakter; niekiedy niezbędne okazuje się zastosowanie rozbudowanych algorytmów działania, łączenie wiedzy i informacji z różnych obszarów. Przykładowo: stworzony w arkuszu kalkulacyjnym model kosztów zakupu i eksploatacji maszyn wymaga znacznie większej liczby obliczeń – wykorzystuje szereg zmiennych pośrednich oraz informacje i wiedzę z rachunkowości i finansów (np. obliczenia stawki amortyzacyjnej, kosztów kre-

dytowania), logistyki produkcji (np. obliczenia zdolności produkcyjnych, kosztów materiałów produkcyjnych) czy z zakresu zarządzania zasobami ludzkimi (np. płace pracowników). Przykładowy model obliczeń przedstawia tabela 3.3.

**Tabela 3.3.** Przykładowy model obliczeń kosztów całkowitych zakupu i eksploatacji maszyn

Charakterystyka	Jednostka	WX Super-mon 2	Konstrix A12	
Koszt nabycia	zł	30000	62000	
Okres użytkowania	rok	6	6	
Wartość likwidacyjna	zł	0	8000	
Zdolność produkcyjna	szt./rok	8000	10000	=MIN(C12:D12)
Pensje obsługi	zł/rok	5000	5000	
Wynagrodzenia	zł/rok	25000	11000	
Inne koszty stałe	zł/rok	5000	6000	
Materiały	zł/rok	40000	48000	
Inne koszty zmienne	zł/rok	3000	3000	
Kalkulacyjna stopa procentowa	%	10%	10%	
Koszty zmienne		=C14+C16+C17	=D14+D16+D17	
Koszty zmienne znormalizowane		=C20*\$E\$12/C12	=D20*\$E\$12/D12	
Koszty stałe	Amortyzacja	=(C9-C11)/C10	=(D9-D11)/D10	
	Odsetki	=C18*(C9+C11)/2	=D18*(D9+D11)/2	
Koszty zmienne znormalizowane		=C21	=D21	
Koszty stałe razem		=SU-MA(C13;C15;	=SU-MA(D13;D15;	
		C22:C23)	D22:D23)	
Łączne koszty		=C25+C26	=D25+D26	

**Źródło:** opracowanie własne.

### 3.3.3.2. Modele optymalizacyjne

Jeżeli w danej sytuacji decyzyjnej mamy do wyboru nieskończenie wiele alternatyw decyzyjnych (możliwych rozwiązań problemu), to nie można dla każdej z nich określić jej wartości kryterium oceniającego alternaty-



wę pod kątem założonych celów. Podobne problemy mogą pojawić się w przypadku, gdy liczba możliwych alternatyw jest co prawda skończona, ale bardzo duża.

Pamiętajmy przy tym, że w rzeczywistych warunkach dopuszczalne są jedynie pewne rozwiązania (decyzje), spełniające określone warunki. Warunki te nazywamy ograniczeniami. Typowymi ich przykładami przy podejmowaniu decyzji z zakresu zarządzania są wielkości zasobów, z których może korzystać organizacja przy realizacji celu.

Wtedy niezbędne okazuje się zazwyczaj zdefiniowanie kryterium oceny decyzji (funkcji celu) i ograniczeń problemu w postaci formuł (wzorów) matematycznych, a następnie rozwiązanie matematycznego problemu optymalizacyjnego, polegającego na znalezieniu maksimum (minimum) funkcji celu przy danych ograniczeniach [Marshall i Oliver, 1995].

Jako ilustrację tego zagadnienia można rozważyć dosyć częstą sytuację, gdy decydent musi przeznaczyć określone ograniczone zasoby na różne konkurencyjne działania, aby zoptymalizować pewien mierzalny cel. Powiedzmy, że zysk jest funkcją wielu zmiennych nakładów różnego rodzaju: produkcyjnych, reklamowych itp. Interesuje nas znalezienie takiej ich kombinacji, dla której zysk będzie jak największy. Musi on być osiągnięty przy pewnych ograniczeniach – dla przykładu: łączna wartość wszystkich nakładów nie powinna przekraczać pewnej określonej z góry kwoty.

Liczba możliwych kombinacji poziomów nakładów okazuje się w tym przypadku nieskończona, a w najlepszym razie bardzo duża, tak więc przeanalizowanie ich wszystkich po kolei w celu ewaluacji każdego z nich (obliczenia zysku) jest po prostu niemożliwe. W celu znalezienia optymalnego zestawu nakładów dającego najwyższy zysk musimy więc zastosować model optymalizacyjny.

Do typowych obszarów wykorzystania modeli optymalizacyjnych należą:

- zadania określenia zestawów ilościowych wyrobów – *product-mix* (np. ile każdego wyrobu należy wyprodukować, aby uzyskać maksymalny zysk);
- transport (np. minimalizacja kosztów dostaw);
- zadania przydziału (najlepsze dopasowanie pewnych obiektów);
- planowanie inwestycji (maksymalizacja stopy zwrotu);
- proste modele zapasów (np. Ekonomiczna Wielkość Zamówienia);
- alokacja zasobów (np. budżetowanie kapitałowe).

Tworząc model optymalizacyjny, musimy wybrać zmienne decyzyjne  $x_1, x_2, \dots, x_n$ , (ich wartości chcemy określić) i zdefiniować w postaci pewnej funkcji celu  $f$  kryterium oceny naszej decyzji. Zadaniem modelu optymalizacyjnego będzie znalezienie takich wartości zmiennych  $x_1, x_2, \dots, x_n$ , dla których wartość funkcji celu jest największa (ewentualnie najmniejsza):

$$f(x_1, x_2, \dots, x_n) \rightarrow \max$$

Wartości wyznaczających najlepszą decyzję zmiennych  $x_1, x_2, \dots, x_n$  muszą przy tym spełniać ograniczenia – definiuje się je zwykle w formie zestawu równań i/lub nierówności:

$$g_i(x_1, x_2, \dots, x_n) \leq b_i$$

$$g_j(x_1, x_2, \dots, x_n) \geq b_j$$

$$g_k(x_1, x_2, \dots, x_n) = b_k$$

W praktyce często zdarza się, że wszystkie funkcje w modelu optymalizacyjnym, tzn. zarówno funkcja celu, jak i funkcje ograniczeń, mają charakter liniowy. W takim przypadku znalezienie maksymalnej wartości funkcji celu wiąże się z rozwiązaniem tzw. zadania programowania liniowego.

Rozważmy np. dane przedstawione w tabeli 3.4.

**Tabela 3.4.** Przykładowy model programowania liniowego

Batony	Nakłady jednostkowe			Zysk jednostkowy [zł/j.pr.]
	Praca urzędzeń [h/j.pr.]	Surowce [kg/j.pr.]	Robocizna [zł./j.pr.]	
$T_1$	3,0	35	200	250
$T_2$	1,8	70	100	110
<b>Zasoby</b>	10,0	425	550	

**Źródło:** opracowanie własne.

Zmienne decyzyjne:

- $x_1$  – rozmiar produkcji  $T_1$
- $x_2$  – rozmiar produkcji  $T_2$

Funkcja celu:

$$- \text{zysk}(x_1, x_2) = 250x_1 + 110x_2$$

Ograniczenia

- $3x_1 + 1,8x_2 \leq 10; x_1 \geq 0, x_2 \geq 0$
- $35x_1 + 70x_2 \leq 425$
- $200x_1 + 100x_2 \geq 550$

Firma produkuje dwa rodzaje wyrobów:  $T_1$  i  $T_2$ . Zysk jednostkowy z produkcji wyrobu  $T_1$  wynosi 250 złotych za jednostkę produkcyjną, dla wyrobu  $T_2$  – 110 złotych za jednostkę produkcyjną. Ponadto wiemy, że:

- wytworzenie jednostki produkcyjnej wyrobu  $T_1$  wymaga trzech godzin czasu pracy urządzeń, zaś jednostki produkcyjnej wyrobu  $T_2$  – 1,8 godziny; łącznie na produkcję przeznaczyć możemy 10 godzin czasu pracy urządzeń;
- wytworzenie jednostki produkcyjnej wyrobu  $T_1$  wymaga zużycia 35 kg surowca, zaś jednostki produkcyjnej wyrobu  $T_2$  – 70 kg; łącznie na produkcję możemy przeznaczyć 425 kg surowca;
- wytworzenie jednostki produkcyjnej wyrobu  $T_1$  wymaga poniesienia kosztów płacy pracowników w wysokości 200 złotych, wyrobu  $T_2$  – 100 złotych; przyjmijmy, że umówiliśmy się już z pracownikami, że zarobią na realizacji tego zadania co najmniej 550 złotych.

Teraz należy znaleźć optymalny plan produkcji, czyli wolumeny produkcji wyrobów  $T_1$  i  $T_2$ , dające najwyższy zysk, przy spełnieniu przedstawionych ograniczeń. Dla zadania programowania liniowego istnieje algorytm pozwalający efektywnie wyznaczyć dokładny punkt ekstremum funkcji celu w skończonej liczbie kroków. Algorytm ten nazywany jest algorytmem simpleks. Wartości zmiennych  $x_1$ ,  $x_2$ , maksymalizujące zysk w obrębie zdefiniowanych ograniczeń, mogą zostać znalezione bez trudu, choćby przy wykorzystaniu arkusza kalkulacyjnego.

W przypadku, gdy rozwiązanie zadania programowania liniowego znajdzie się w dziedzinie liczb całkowitych, mówimy o programowaniu całkowitoliczbowym.

Niestety w przypadku nieliniowych funkcji celu i funkcji definiujących ograniczenia nie ma już takich prostych algorytmów rozwiązania, jak algorytm simpleks (lub algorytmy pochodne). Szukanie ekstremum funkcji staje się wówczas procesem iteracyjnym, wymagającym często wielu kroków i pozwalającym na znalezienie jedynie wyniku przybliżonego.

Skomplikowana powierzchnia funkcji celu może mieć ponadto charakter wielomodalny, tzn. posiadać wiele lokalnych ekstremów. Wśród najważniejszych ogólnych kategorii metod znajdowania ekstremum funkcji można wymienić [Jabłoński i Bartkiewicz, 2006]:

- metody enumeratywne,
- metody losowe,
- metody gradientowe,
- metody heurystyczne.

Metody enumeratywne związane są z sytuacją, gdy zbiór wszystkich spełniających ograniczenia rozwiązań problemu jest skończony (i ponadto niezbyt duży). Jeśli potrafimy dla każdej alternatywy decyzyjnej obliczyć wartość funkcji celu, to wybierając rozwiązanie o wartości maksymalnej, rozwiązujemy zadanie. Rozważmy następujący przykład. Naszym celem jest znalezienie wariantu inwestowania o maksymalnej stopie zwrotu. Po analizie ofert inwestycyjnych okazało się, że ograniczenia (dotyczące np. wielkości zaangażowanego kapitału, zabezpieczeń, strategii inwestowania itp.) spełniają trzy z nich. Po obliczeniu stopy zwrotu widać natomiast, że przyjmują one odpowiednio wartości: 13,5, 12,9, 14,1%. Decyzją optymalną będzie oczywiście wybór wariantu trzeciego.

Możliwość zastosowania metod enumeratywnych jest oczywiście na ogół ograniczona. Inne sposoby rozwiązywania zadań optymalizacyjnych mają charakter iteracyjny: tu wychodzimy od pewnego rozwiązania początkowego, poprawiając je w każdym kroku algorytmu poprzez wybór rozwiązania lepszego pod względem wartości kryterium. W metodach losowych wybór kolejnego rozwiązania odbywa się na „chybił trafił” w określonym sąsiedztwie rozwiązania poprzedniego lub poprzez określone analizy probabilistyczne (metody pseudolosowe). Przykładem algorytmu tego typu może być wyżarzanie wykładnicze.

Metody gradientowe wykorzystują do poszukiwania rozwiązania informację o pochodnych. Wiadomym jest, że w punkcie ekstremum pochodna funkcji celu (dla funkcji wielu zmiennych nazywana gradientem) musi być równa zero. Jeśli wobec tego da się rozwiązać równanie (bądź w przypadku wielu zmiennych raczej układ równań) wyznaczające miejsce zerowe pochodnej, będzie to oznaczać znalezienie rozwiązania optymalnego. Niestety zastosowanie tego typu metod okazuje się ograniczone do pewnych niewielkich klas problemów (np. znajdowanie minimum kwadratowej funkcji błędu w zagadnieniu regresji liniowej dokonywane

jest przez rozwiązanie układu równań normalnych, wyznaczającego miejsce zerowe pochodnej funkcji błędu).

Znana jest również zależność zakładająca, że znak pochodnej wskazuje na to, czy funkcja maleje czy rośnie. W przypadku wielowymiarowym oznacza to, że poruszanie się w kierunku przeciwnym do gradientu funkcji celu to zbliżanie się do jej minimum. Większość metod gradientowych ma więc charakter iteracyjny. Kolejne przybliżenia rozwiązań wyznaczane są w kierunku przeciwnym do gradientu funkcji celu – pod warunkiem, że spełniają one ograniczenia. Do tej kategorii należą np. takie podstawowe algorytmy minimalizacji funkcji, jak np. metoda najszybszego spadku czy też sprzężonych gradientów.

Z kolei metody heurystyczne są metodami iteracyjnymi wyznaczającymi kolejne przybliżenia rozwiązań poprzez heurystyczne przekształcenia rozwiązań poprzednich. Przykładowo: algorytmy genetyczne bazują na heurystyce, że połączenie dwóch dobrych rozwiązań może zaowocować czymś znacznie lepszym.

Nieco innymi przykładami zadań optymalizacyjnych są zadania optymalizacji kombinatorycznej, łączące problematykę znajdowania ekstremum funkcji celu z teorią grafów. Wymienić tu można zagadnienia harmonogramowania czy też optymalizacji sieciowej, takie jak: optymalizacja sieci dystrybucji, tras przejazdów, analiza przedsięwzięcia itp.

### **3.3.3.3. Modele wielokryterialne**

Zdecydowana większość problemów menedżerskich zawiera wiele celów i oceniana musi być pod kątem wielu kryteriów. Jeśli np. rozważamy decyzję związaną z wyborem wariantów inwestowania, to bez wątpienia jednym z celów menedżera będzie maksymalizacja zysku, ale również należy wziąć tu pod uwagę m.in. minimalizację ryzyka. Modele decyzyjne, w których bierzemy pod uwagę wiele celów, nazywamy wielokryterialnymi. Doskonałe rozwiązanie polegałoby oczywiście na znalezieniu wariantu decyzji (rozwiązania problemu) najlepszego pod kątem wszystkich branych pod uwagę kryteriów. Rozwiązania dopuszczalne, optymalne pod względem wszystkich kryteriów, nazywamy idealnymi. Niestety świat nie jest idealny. Wielokryterialne zadanie decyzyjne rzadko posiada więc rozwiązanie idealne. Kryteria są ze sobą zwykle w znacznej mierze skonfliktowane, a nawet sprzeczne (jak w naszym przykładzie: wa-

rianty inwestowania najlepsze pod względem zyskowności będą na ogół najbardziej ryzykowne). Ponadto nawet jeśli rozwiązanie optymalne pod względem wszystkich kryteriów istnieje, to często znajduje się poza obszarem dopuszczalnym, wyznaczanym przez ograniczenia.

W większości przypadków w modelach wielokryterialnych trzeba zadowolić się wyznaczeniem rozwiązania niezdominowanego, tzn. takiego, dla którego nie ma rozwiązań dopuszczalnych o wartościach wszystkich kryteriów bliższych rozwiązaniu idealnemu – czyli znajduje się ono w obszarze dopuszczalnym (okazuje się najlepsze pod kątem choćby jednego kryterium).

Przyjrzyjmy się przykładowi przedstawionemu w tabeli 3.5. Zawiera ona dane kilku hipotetycznych samochodów, zestawione pod kątem ceny oraz mocy silnika. Gdybyśmy chcieli wybrać jak najlepszy samochód, kierując się oboma kryteriami, to idealnym rozwiązaniem okazuje się samochód jak najtańszy i o jak największej mocy silnika.

Widzimy, że spośród wszystkich rozważanych samochodów rozwiązaniem zdominowanym byłby jedynie wybór samochodu drugiego. Jest on droższy i ma mniejszą moc silnika od samochodu pierwszego. Pozostałe rozwiązania w analizowanym zbiorze dopuszczalnych alternatyw decyzyjnych są niezdominowane.

**Tabela 3.5.** Przykładowy model wielokryterialny wyboru samochodu

	Cena	Moc silnika	
Samochód 1	30 000	85	Zdominowane
Samochód 2	32 000	80	
Samochód 3	29 000	79	
Samochód 4	35 000	90	

**Źródło:** opracowanie własne.

Jak widać z powyższego przykładu, problem polega na tym, że rozwiązań niezdominowanych może być wiele. Powstaje więc pytanie, które z nich powinniśmy wybrać i w jaki sposób tego wyboru dokonać. Możliwych sposobów postępowania jest kilka. Jeden z nich polegać będzie na pozostawieniu tylko jednego kryterium, najważniejszego dla decydenta, oraz na zamianie wszystkich pozostałych na ograniczenia. W naszym zadaniu zakupu samochodu moglibyśmy więc określić, że interesują nas samochody w przedziale cenowym nie przekraczającym 30 000 złotych – wyboru dokonujemy

jedynie pośród nich, kierując się maksymalizacją kryterium mocy silnika. W ten sposób podjęlibyśmy decyzję o zakupie samochodu pierwszego.

Powyższa metoda pozwala nam uwzględnić w pewnym stopniu mniej istotne dla decydenta kryterium cenowe i podjąć decyzję na podstawie kryterium istotniejszego – mocy silnika. Natomiast bez wątpienia wiąże się ona ze znacznymi stratami informacyjnymi i w wielu przypadkach byłaby trudna do zaakceptowania przez decydenta. Dlatego częściej stosowanym podejściem jest scalenie wartości poszczególnych kryteriów w jedną, sztuczną miarę, nazywaną zwykle użytecznością decyzji, rozwiązania itp. [Marshall i Oliver, 1995].

Agregacja kryteriów cząstkowych wymaga również, podobnie jak poprzednie podejście, dostarczenia przez decydenta pewnych informacji (np. w formie wag) odnoszących się do istotności poszczególnych kryteriów. Zauważmy bowiem, że gdybyśmy przyjęli, iż wszystkie kryteria są dla decydenta równie ważne, to problem wyboru najlepszej decyzji byłby w tym przypadku w zasadzie nierozstrzygalny.

**Tabela 3.6.** Przykład liniowej funkcji użyteczności  $-0,003 \cdot \text{Cena} + 1 \cdot \text{Moc silnika}$

	<b>Cena -0,003</b>	<b>Moc silnika 1</b>	<b>Użyteczność</b>
Samochód 1	30 000	85	-5
Samochód 2	32 000	80	-16
Samochód 3	29 000	79	-8
Samochód 4	35 000	90	-15

**Źródło:** opracowanie własne.

Najczęściej stosowany jest liniowy model agregacji, wymagający określenia wag istotności poszczególnych kryteriów. W tabeli 3.6. przedstawiony został przykład agregacji kryteriów w rozważanym wcześniej przykładzie decyzji o wyborze samochodu. Waga dla kryterium ceny określona została przez decydenta jako  $-0,003$ , zaś dla kryterium mocy silnika w wysokości 1. Zwróćmy uwagę, że wagi obu kryteriów muszą być przeciwnych znaków, ponieważ działają one w przeciwnych kierunkach. Ponadto wielkość wagi powinna uwzględniać przedział wartości danego kryterium. Ponieważ ceny mają znacznie większe wartości niż moce silnika, waga dla kryterium cenowego musi być niższa, inaczej kryterium to zdominowałoby całkowicie łączną ocenę.

Zatem w naszym przykładzie funkcja liniowa obliczająca wartość użyteczności każdego z samochodów ma postać:

$$\text{użyteczność} = -0,003 \cdot \text{Cena} + 1 \cdot \text{Moc silnika}$$

Po scaleniu kryteriów cząstkowych można dokonać wyboru wariantu decyzji, kierując się już tylko jednym kryterium – użytecznością. Najlepszą opcją byłby więc, w warunkach określonych w naszym przykładzie, zakup samochodu numer 1.

Określenie przez decydenta konsekwentnego i zgodnego ze stanem rzeczywistym zestawu wag istotności poszczególnych kryteriów nie jest jednak zadaniem prostym, zwłaszcza w przypadku (wcale nierzadkim) istnienia wielu kryteriów cząstkowych oceny decyzji. Trudno bowiem w całościowy sposób konsekwentnie rozdysonować istotność nawet na kilka kryteriów składowych. Z tego powodu jedną z najpopularniejszych metod analizy wielokryterialnej jest AHP (*Analytic Hierarchy Process*), stworzona przez Thomasa Saaty'ego w latach siedemdziesiątych XX wieku [Krawczyk, 2001; Marshall i Oliver, 1995]. Polega ona na porównywaniu kryteriów parami i określaniu przez decydenta preferencji, które z dwóch porównywanych w parze kryteriów okazuje się istotniejsze i jaka jest siła tej przewagi. Dzięki takiemu lokalnemu podejściu do oceny decydentowi łatwiej określić swoje rzeczywiste potrzeby w tej materii. Następnie, na podstawie takich porównań parami, obliczane są wagi istotności każdego z kryteriów cząstkowych.

Zwróćmy uwagę, że w liniowym modelu użyteczności każdemu z kryteriów przypisuje się stałą wagę, określającą jego istotność. Równoważne jest to założeniu, że wszystkie kryteria są niezależne, tzn. uznaje się, iż istotność danego kryterium jest stała i nie zmienia się w zależności od wartości innych kryteriów. Nie musi być to jednak zawsze zgodne z potrzebami decydenta. Łatwo wyobrazić sobie sytuację, że przy samochodach lepszych (o większej mocy silnika) cena może być dla niego mniej istotna niż w przypadku samochodów gorszych (o znacznie niższej mocy silnika).

Z tego powodu czasami stosuje się bardziej złożone modele agregacji kryteriów cząstkowych. Przykładem mogą być tu systemy kart punktowych, gdzie każdej określonej wartości kryterium (lub przedziału wartości) przydziela się (niekoniecznie w sposób liniowy) pewną liczbę punktów. Użyteczność otrzymywana jest przez zsumowanie ocen poszczególnych kryteriów.



W bardziej złożonych sytuacjach decyzyjnych agregacji kryteriów mogą dokonywać systemy ekspertowe, w których na podstawie wiedzy eksperta odpowiednie reguły oceniają finalną użyteczność dla skomplikowanych wzorców zależności między kryteriami cząstkowymi.

Jeszcze inne podejście polega na wykorzystaniu do agregacji kryteriów metod eksploracji danych. Dla określonych wzorców wartości kryteriów cząstkowych stworzony może zostać (na podstawie analizy zbioru historycznych przykładów właściwej oceny użyteczności) model użyteczności – liniowy (np. metodą regresji liniowej) lub nieliniowy (np. z wykorzystaniem sieci neuronowej).

### **3.3.3.4. Decyzje klasyfikacyjne**

Omawiane w poprzednim punkcie modele optymalizacyjne wymagają jawnego wyspecyfikowania określonej funkcji celu (bądź wielu funkcji celów), pod kątem których menedżer ocenia różne, prowadzące do rozwiązania problemu opcje decyzyjne. Jednak już w przypadku modeli wielokryterialnych napotkaliśmy na poważne problemy z tego rodzaju założeniem. Nietrudno więc wyobrazić sobie sytuację, kiedy funkcja celu okazuje się trudnym do jawnego sprecyzowania konglomeratem różnych kryteriów.

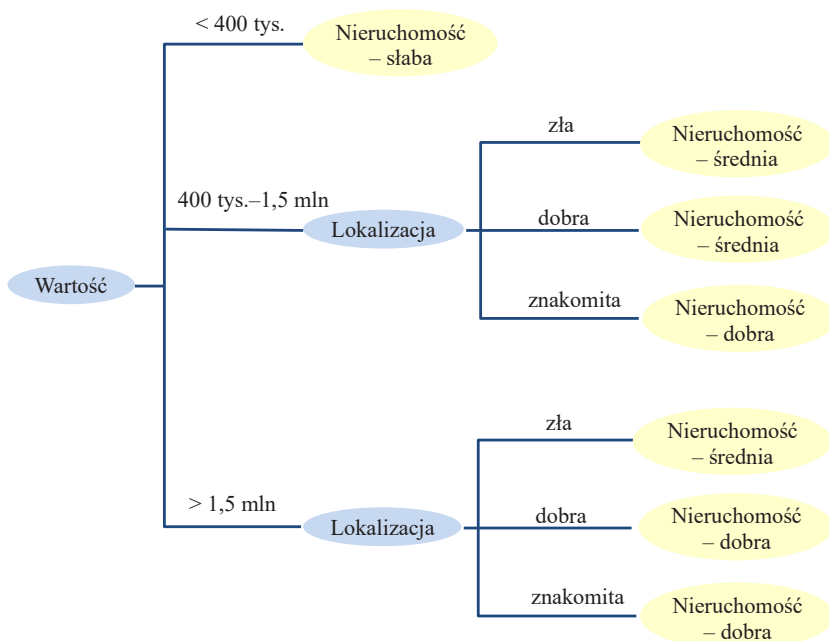
Zastanówmy się np. nad problemem decyzji kredytowej w banku. Oczywiście mamy w jej przypadku do wyboru dwie podstawowe alternatywy: przyjąć wniosek o kredyt lub go odrzucić. Czasami mogą w grę wchodzić również dodatkowe możliwości, takie jak konsultacja z przełożonym, dokładniejsze przebadanie sytuacji wnioskodawcy itp., ale dla uproszczenia pominiemy je w dalszych rozważaniach. Celem decyzji o przyznaniu kredytu jest osiągnięcie określonych zysków przy jak najmniejszym ryzyku. Analizując wniosek kredytowy, nie korzysta się jednak w sposób jawny z funkcji celu, wyznaczając jej wartość dla każdej z alternatyw i w ten sposób określając najlepszą z nich. Tutaj, na podstawie parametrów wniosku, takich jak kwota kredytu, dochody wnioskodawcy, zabezpieczenia itp., wybiera się jedną z wspomnianych wyżej alternatyw.

W tego rodzaju decyzjach funkcja celu jest więc wbudowana w sposób niejawny w logikę klasyfikacji pewnej określonej sytuacji decyzyjnej – przyporządkowuje się ją do jednej z kilku kategorii reprezentujących poszczególne alternatywy decyzyjne. W przypadku naszej decyzji kredytowej wniosek kredytowy klasyfikowany jest – na podstawie wartości

pewnych istotnych zmiennych (atrybutów) – do decyzji: „przyznać kredyt” lub „odrzuć wniosek”.

Podobnych sytuacji decyzyjnych można wymienić bardzo wiele: w grze giełdowej na podstawie cen, indeksów i wskaźników giełdowych podejmowane są decyzje transakcyjne typu „sprzedaj”, „kup”, „wstrzymaj się”; opierając się na informacjach o kliencie i jego historii zakupów decyduje się z kolei o podjęciu w stosunku do niego pewnych działań marketingowych lub też nie; rozpoznając stan przedsiębiorstwa na bazie jego wskaźników ekonomicznych otrzymuje się informację o zagrożeniu bankructwem i ewentualnej stagnacji/sprzedaży itp.

Do najczęściej wykorzystywanych modeli wspierających tego rodzaju decyzje należą drzewa decyzyjno-klasyfikacyjne. Zbudowane są one z węzłów reprezentujących poszczególne zmienne decyzyjne. Wychodzące z każdego węzła gałęzie reprezentują wartości (bądź przedziały wartości) prowadzące do odmiennych skutków decyzyjnych. Każda ścieżka kieruje ostatecznie do liścia, czyli węzła reprezentującego decyzję odpowiadającą zestawowi warunków na ścieżce (rysunek 3.3).



**Rysunek 3.3.** Drzewo decyzyjno-klasyfikacyjne dla decyzji o ocenie nieruchomości

**Źródło:** opracowanie własne.

Na rysunku 3.3 przedstawione zostało przykładowe drzewo klasyfikacyjno-decyzyjne dla decyzji odnośnie oceny nieruchomości (np. na potrzeby zabezpieczenia kredytowego). Możliwe są trzy decyzje odnośnie oceny nieruchomości: słaba, średnia i dobra. Rozważane są dwie zmienne decyzyjne: wartość nieruchomości oraz jej lokalizacja. Ścieżki w drzewie dla poszczególnych zestawów warunków nakładanych na wartości zmiennych decyzyjnych prowadzą do ostatecznej oceny nieruchomości.

Stworzenie drzewa decyzyjno-klasyfikacyjnego wymaga uporządkowania problemu decyzyjnego, dogłębnego zastanowienia się nad poszczególnymi zmiennymi decyzyjnymi i ich wartościami, na podstawie których można podjąć decyzję. Zadanie to ułatwiane jest przez fakt, że w danej chwili można skoncentrować się na określonym węźle, analizując tylko niewielki wycinek problemu. Natomiast całościowo drzewa decyzyjno-klasyfikacyjne opisują dosyć złożone wzorce zależności, o logice uwzględniającej nawet wiele różnych kryteriów decyzyjnych.

Drzewa decyzyjno-klasyfikacyjne służą nie tylko do uporządkowania problemu – umożliwiają także jego reprezentację graficzną i wizualizację, dzięki czemu decydent wie, dlaczego System Wspomagania Decyzji wskazuje taką, a nie inną decyzję. Może on przeanalizować zestaw warunków i ścieżkę w drzewie, która do niej prowadzi.

Korzystanie z drzew decyzyjno-klasyfikacyjnych okazuje się jednak trudne, gdy mamy wiele zmiennych decyzyjnych o wielu gałęziach wychodzących z kolejnych węzłów. Drzewo może bowiem stać się wtedy zawikłane, a analiza ścieżek prowadzących w nim do odpowiednich decyzji skomplikowana. W takiej sytuacji zasadne będzie wykorzystanie reguł klasyfikacyjnych, w sposób jawny definiujących wartości (lub przedziały wartości) atrybutów, dla których należy wybrać określoną alternatywę decyzyjną. W przypadku decyzji kredytowej mogą one mieć postać:

JEŻELI: *Dochody* > 2000 oraz *Wartość zabezpieczenia* > 5000

TO: Decyzja = „Przyznać kredyt”

Klasyfikatory mogą być również budowane na podstawie danych. Do najważniejszych metod wykorzystywanych w tego typu zadaniach należą: korzystająca z klasycznych modeli statystycznych regresji analiza dyskryminacyjna, klasyfikacja Bayesowska, metoda *k*-najbliższych sąsiadów,

sieci neuronowe czy też maszyny wektorów wspierających (SVM). Do zagadnień tworzenia klasyfikatorów z wykorzystaniem danych wrócimy jeszcze w rozdziale 3.4 – w części poświęconej eksploracji danych.

### **3.3.3.5. Badania symulacyjne i decyzje w warunkach niepewności**

Charakteryzując rodzaje decyzji w punkcie 3.2.2, mówiliśmy o czynniku ryzyka jako jednym z istotnych elementów wpływających na procesy decyzyjne w organizacji. Przypomnijmy w skrócie, że wyróżniliśmy pod tym względem trzy podstawowe sytuacje:

1. Warunki pewności, w których zakłada się pełną wiedzę o sytuacji decyzyjnej, zarówno pod względem informacji o wartościach zmiennych decyzyjnych i parametrów, jak i umiejętności jej przełożenia na określenie skutków decyzji. W związku z tym wszystkie potencjalne skutki decyzji są znane i można znaleźć wariant decyzji na pewno oferujący najlepsze efekty.
2. Warunki niepewności. Możliwych jest przynajmniej kilka skutków (często nazywanych stanami rzeczywistości) każdej decyzji i nie wiadomo, która z tych możliwości się zrealizuje. Nie wiemy nawet, który ze stanów rzeczywistości jest bardziej możliwy (czy prawdopodobny), a który mniej. Warunki niepewności wiążą się więc z całkowitą niewiedzą na temat możliwych skutków decyzji.
3. Warunki ryzyka, w których nadal możliwych jest kilka skutków każdej decyzji i nadal nie wiadomo, który z nich się zrealizuje. Posiadana wiedza pozwala jednak przynajmniej określić prawdopodobieństwo realizacji każdego z nich. Można więc ocenić skalę ryzyka i włączyć je do procesu podejmowania decyzji. Nie potrafimy może wskazać decyzji, która na pewno będzie najlepsza, wiadomo jednak, która najprawdopodobniej będzie najlepsza.

Podstawową metodą wspomagania decyzji w warunkach niepewności, czyli w stanie całkowitej niewiedzy na temat możliwości realizacji różnych skutków rozważanych alternatyw decyzyjnych, są po prostu badania tychże efektów, czyli analizy symulacyjne.

Symulacja to technika polegająca na przeprowadzaniu eksperymentów (zazwyczaj wspomaganych komputerowo) na modelu pewnego rzeczywistego systemu, w celu oceny dynamiki jego zachowania.

Pozwala ona więc imitować rzeczywistość, oddając ją zarówno co do stanu, jak i zachowania. Symulacja jest jednak narzędziem opisowym, a nie normatywnym, tzn. może wskazać decydentowi możliwe skutki analizowanych alternatyw, nie jest jednak w stanie zagwarantować znalezienia najlepszego wariantu decyzji. Stosuje się ją więc wtedy, gdy badany system okazuje się zbyt skomplikowany dla innych technik wspomagania decyzji, czyli kiedy stan naszej wiedzy o systemie jest niski. Symulacji powinno się jednak używać tylko, jeśli rozwiązanie analityczne staje się niemożliwe. W związku z tym rzadziej wykorzystuje się ją w przypadku, gdy wiedza o sytuacji decyzyjnej jest obszerna – dysponujemy w takim przypadku na ogół całą paletą innych metod. Natomiast symulacja często okazuje się jedynym narzędziem wspomagającym dla decyzji w warunkach niepewności, zwłaszcza o charakterze niestrukturalnym.

Symulacja to atrakcyjne narzędzie dla użytkowników; posiada wiele zalet: teoria jest dość prosta, metoda pozwala na dużą kompresję czasu, możliwe są bardzo różnorodne eksperymenty (z różnymi wariantami decyzji, w odmiennych sytuacjach) – potrafi poradzić sobie również z wieloma rodzajami problemów, może obejmować rzeczywistą ich złożoność i dostarczać ważnych informacji.

Jednakże symulacja nie gwarantuje znalezienia optymalnego rozwiązania – uzyskane wyniki mogą nie być łatwe do interpretacji. Proces tworzenia eksperymentów symulacyjnych wymaga zbudowania modelu, jest więc zwykle dosyć powolny i kosztowny. Otrzymane wnioski są na ogół specyficzne dla badanego problemu, trudno je zaadaptować do innych sytuacji. Symulacja wydaje się ponadto efektowna i łatwa do wytłumaczenia, przez co może prowadzić do przeoczenia rozwiązań analitycznych [Sharda, Delen i Turban, 2017].

W rzeczywistości proces badania symulacyjnego nie jest prosty. Wymaga on stworzenia i przetestowania modelu analizowanego systemu, zaprojektowania i przeprowadzenia eksperymentów ze stworzonym modelem, a następnie ich oceny i wyciągnięcia wniosków. Wyróżnia się trzy podstawowe rodzaje badań symulacyjnych:

- analiza wrażliwości – to proces oceny wpływu zmian (zwykle niewielkich) zmiennej wejściowej (lub parametru) na otrzymywane wyniki; zwykle stosowana jest w celu określenia wpływu niewielkich błędów w wartościach wejściowych na wyjścia systemu;

- analiza typu „co – jeżeli” (*what-if*) – polega na ocenie kształtowania się wyjść systemu w efekcie poważniejszych zmian wartości jego zmiennych lub parametrów, wynikających z różnych założeń lub warunków działania systemu;
- analiza sterowana celami – podejście „wstecz”: rozpoczyna się od zdefiniowania celów, czyli warunków nakładanych na wyjścia systemu; określa, jakie wartości musiałyby przyjmować zmienne wejściowe, aby cele te zostały spełnione.

Badania symulacyjne wymagają na ogół powtórzenia eksperymentu dla wielu przypadków symulacyjnych, definiujących różne warunki, dla których chcemy przebadać działanie systemu. Przykładowo: analizy typu „co – jeżeli” bardzo często wymuszają konieczność obliczenia wyjść systemu dla różnych zestawów wartości zmiennych wejściowych (np. skutków alternatywy decyzyjnej przy wszystkich możliwych zestawach wartości pewnych zmiennych decyzyjnych).

W sytuacjach decyzyjnych w warunkach niepewności często musimy rozważyć wiele alternatyw decyzyjnych – każda z nich może skutkować wieloma możliwymi wynikami. Liczba zmiennych decyzyjnych i parametrów modyfikowanych w eksperymentach także okazuje się niekiedy bardzo duża. Tak więc liczba przypadków i możliwych stanów systemu, które trzeba przebadać, może okazać się wręcz ogromna.

Narzędziem pozwalającym na porządkowanie przypadków symulacyjnych w logiczne wzorce są scenariusze, czyli zestawy założeń dotyczących środowiska działania określonego systemu; to opis warunków dla jednej, konkretnej sytuacji symulacyjnej. Scenariusze pozwalają więc uporządkować możliwe przypadki symulacyjne w spójne schematy i zmniejszyć ich liczbę, odrzucając przypadki nielogiczne, nieistotne – lub po prostu nieciekawe z punktu widzenia celów symulacji. Podczas procesu symulacji powtarza się eksperymenty tylko dla zdefiniowanych scenariuszy, a nie wszystkich możliwych przypadków symulacyjnych (np. wzorców wartości zmiennych wejściowych).

W praktyce zazwyczaj wykonujemy symulacje dla scenariuszy określonego typu. Należą do nich:

- scenariusze dla najgorszego możliwego przypadku,
- scenariusze dla najlepszego możliwego przypadku,
- scenariusze dla najbardziej prawdopodobnego przypadku,
- scenariusze dla przeciętnego przypadku.

Aby zilustrować wykorzystanie symulacji w decyzjach w warunkach niepewności, rozważmy następujący, prosty przykład podjęcia decyzji o wolumenie produkcji.

Ograniczenia technologiczne procesu produkcyjnego i umowy z dostawcami określają, że można produkować 200, 400 lub 600 sztuk danego wyrobu miesięcznie. Zysk z produkcji zależy od popytu. Nie posiadamy jednak jego prognozy – potrafimy jedynie określić pewne ogólne ramy, w których powinien się on mieścić: od 200 do 350 jednostek. Dysponujemy natomiast modelem dziedzinowym, pozwalającym obliczyć zysk dla danego popytu i wielkości produkcji.

Mamy więc do rozważenia trzy alternatywy decyzyjne: 200, 400 i 600. Kryterium wyboru jest zysk z produkcji: chcemy znaleźć alternatywę oferującą najwyższą jego wartość. Nie znamy jednak wielkości popytu, który jest istotnym parametrem naszej decyzji – od niego bowiem zależy wysokość przyjętego przez nas kryterium zysku. Mamy wobec tego do czynienia z decyzją w warunkach niepewności.

Aby wybrać odpowiedni wolumen produkcji, przeprowadzono badania symulacyjne zysku dla poszczególnych alternatyw i różnych możliwych wartości popytu. Ponieważ wiemy jedynie, że popyt powinien znajdować się w przedziale od 200 do 350, każda jego wielkość w tym obszarze będzie jednakowo możliwa. Postanowiono więc przyjrzeć się popytowi w całym tym przedziale, tablicując wartości zmiennej ze stałym krokiem. Przyjęto krok 30, uznając, że taka ziarnistość wystarczy do oceny zysków.

Wyniki symulacji dla utworzonych w ten sposób 18 przypadków symulacyjnych (sześć rozważanych wartości popytu i trzy wolumeny produkcji) przedstawione zostały w tabeli 3.7.

**Tabela 3.7.** Wyniki symulacji zysków z trzech wariantów produkcji dla różnych popytów

Popyt	Produkcja		
	200	400	600
200	7000	15 000	9000
230	8050	16 350	12 450
260	9100	17 700	15 900
290	10 150	19 050	19 350
320	11 200	20 400	22 800
350	12 250	21 750	26 250

**Źródło:** opracowanie własne.

Bezpośrednia analiza wyników symulacji może w niektórych przypadkach dostarczyć pewnej istotnej wiedzy dla badanego problemu decyzyjnego. Przykładowo, analizując wyniki symulacji w tabeli 3.7, widzimy, że pierwsza alternatywa decyzyjna, ustalająca wolumen produkcji na 200, przy żadnym stanie popytu nie daje najwyższego zysku – zawsze alternatywa druga lub trzecia będą od niej lepsze. Oznacza to, że przy żadnym stanie popytu ten wariant decyzji nie zostałby wybrany. Można go więc wyeliminować z dalszych rozważań.

Dalej, analizując wyniki symulacji dla drugiego i trzeciego wariantu decyzji, widać, że produkcja 400 jest lepszą decyzją przy niskich popytach, zaś przy wyższych – bardziej optymalnym rozwiązaniem okazuje się produkcja 600. Zauważamy więc jak istotnym elementem w podejmowaniu decyzji może być właściwe środowisko analityczne. Dzięki uzyskanej na podstawie badań symulacyjnych wiedzy nawet ogólna prognoza zapotrzebowania pozwoliłaby dokonać właściwego wyboru. Bardzo pomocne byłoby nawet oszacowanie, czy poziom popytu będzie niski, czy wysoki. Powyższy przykład ilustruje więc również to, jak istotne okazuje się dla procesów podejmowania decyzji całe środowisko analityczne organizacji, o którym będzie mowa w podrozdziale 3.4.

**Tabela 3.8.** Subiektywne strategie podjęcia decyzji o wolumenie produkcji

Popyt	Produkcja		
	200	400	600
200	7 000	15 000	9 000
230	8 050	16 350	12 450
260	9 100	17 700	15 900
290	10 150	19 050	19 350
320	11 200	20 400	22 800
350	12 250	21 750	26 250
$\alpha$	7 000	<b>15 000</b>	9 000
	12 250	21 750	<b>26 250</b>
	9 625	<b>18 375</b>	17 625

**Źródło:** opracowanie własne.

Obecnie przyjmujemy, że podejmujemy decyzję w warunkach całkowitej niepewności i nie ma żadnej informacji na temat możliwego poziomu zapotrzebowania. W takiej sytuacji, niestety, nie jesteśmy w stanie



w obiektywny sposób, wynikający z samego problemu, podjąć decyzji, czy wybrać alternatywę drugą, czy trzecią – i produkować 400 czy 600 jednostek rozważanego wyrobu.

Skoro brakuje obiektywnych kryteriów pozwalających na rozstrzygnięcie problemu decyzyjnego, niezbędne jest wykorzystanie kryteriów subiektywnych, opartych na czynnikach psychologicznych i intuicji decydenta. Podstawowymi czynnikami z grupy są skłonność do ponoszenia ryzyka przez decydenta (lub równoważnie: jego optymizm w ocenie sytuacji). Jeszcze raz zwróćmy uwagę, że mówimy o skłonności do ryzyka (lub optymizmie) nieopierającej się na żadnych obiektywnych przesłankach wynikających z wiedzy o problemie – nie posiadamy jej. Chodzi o czysto subiektywne odczucie, bazujące wyłącznie na psychicznym nastawieniu decydenta.

W przypadku decydenta niechętnego do ponoszenia ryzyka, pesymistycznie oceniającego rozwój sytuacji (wielkość popytu), weźmiemy pod uwagę najgorsze możliwe skutki decyzji i podejmiemy decyzję, wybierając najlepszą w takiej sytuacji alternatywę (reguła *max-min* albo Walda). Jak widać w tabeli 3.8, w naszym przykładzie byłaby to alternatywa druga, czyli 400.

Decydent skłonny do ryzyka, albo optymistycznie oceniający sytuację, rozważyłby najlepsze możliwe skutki każdej z alternatyw i wybrałby najlepszy z nich (reguła *max-max*). W naszym przykładzie będzie to alternatywa trzecia, czyli 600 (tabela 3.8).

Można wreszcie mówić o strategiach pośrednich. Wprowadźmy współczynnik skłonności do ryzyka albo optymizmu  $\alpha$ , z przedziału od 0 do 1, gdzie  $\alpha = 1$  oznacza pełną skłonność do ryzyka (optymizm), zaś  $\alpha = 0$  – zupełny jej brak, czyli pesymizm. Wówczas jako miernik wyboru alternatywy można przyjąć średnią ważoną z jej najlepszego i najgorszego skutku (reguła Hurwicza):

$$\alpha \cdot \max + (1 - \alpha) \cdot \min$$

Jak widać w tabeli 3.8, przy skłonności do ryzyka  $\alpha = 0,5$  najlepszą alternatywą jest wariant drugi, czyli produkcja 400 jednostek.

### 3.3.3.6. Modele analizy decyzyjnej i decyzje w warunkach ryzyka

Ryzyko decyzji polega na tym, że wybrana alternatywa przyniesie nieoczekiwane i gorsze od spodziewanych wyniki. Wielkość ryzyka określana jest prawdopodobieństwem takiej sytuacji. Ryzyko spowodowane

może być niepewnością co do wartości niektórych ważnych informacji lub niedoskonałością naszej wiedzy odnośnie zależności między różnymi zjawiskami w rozważanym problemie decyzyjnym.

Typową sytuacją powodującą pojawienie się elementu ryzyka będzie korzystanie w procesie podejmowania decyzji z wiedzy uzyskanej przy pomocy metod analitycznych. Zarówno ilościowe (np. modele eksploatacji danych), jak i jakościowe (analizy ekspertów) procesy analityczne oparte są nie na teorii, lecz na dociekaniach heurystycznych – z zasady stanowią więc wiedzę niedoskonałą i obciążone zostają pewnym błędem, niepewnością. Otrzymane przy ich pomocy informacje mają charakter przybliżeń, oszacowań, estymacji. W niektórych sytuacjach niepewność ta może być pomijana, jednak w większości przypadków generuje ona powstanie elementu ryzyka. Problem polega na tym, że bardzo często stanowią one jedyne źródło wiedzy, jakie decydent może wykorzystać w trakcie podejmowania decyzji.

W sytuacji podejmowania decyzji w warunkach ryzyka decydent ma również do czynienia z wieloma możliwymi skutkami wyboru. Nie wiadomo z pewnością, która alternatywa zrealizuje się w rzeczywistości, ale można określić ich prawdopodobieństwa.

W związku z tym nie sposób definitywnie odpowiedzieć na pytanie o najlepszy (pewny) wybór. Decydent jest jednak w stanie zdecydować się na tę opcję, która najprawdopodobniej będzie najlepsza. W tym celu przy wyborze alternatywy decyzyjnej w warunkach ryzyka należy kierować się kryterium najlepszej oczekiwanej wartości skutków decyzji.

W takim przypadku stosuje się przede wszystkim modele analizy decyzyjnej. Obejmują one szereg rozwiązań łączących metody analiz probabilistycznych oraz technik porządkowania i wizualizacji struktury badanego problemu decyzyjnego. Do najważniejszych podejść w tym zakresie należą diagramy wpływu (*influence diagrams*) i drzewa decyzyjne. W tej części zaprezentujemy wykorzystanie w analizie decyzyjnej drugiej z tych metod, tj. drzewa decyzyjnego. Posłużymy się przy tym następującym prostym przykładem.

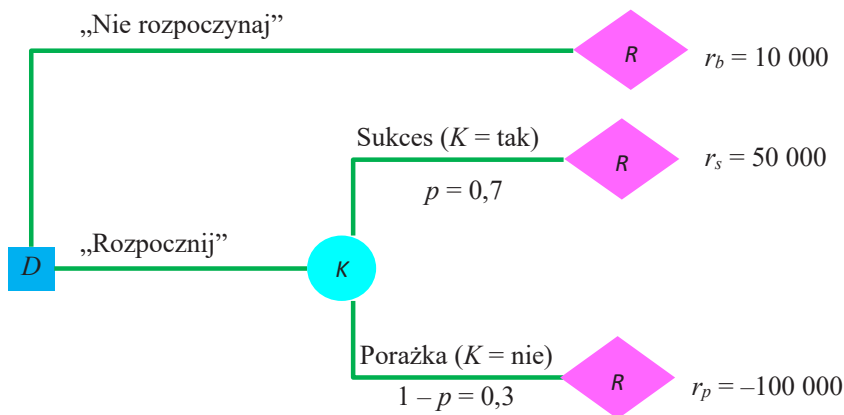
Mając nadzieję na zawarcie kontraktu na obsługę dystrybucji, rozważamy budowę centrum dystrybucyjnego. Jeżeli kontraktu nie otrzymamy, będziemy oczywiście w stanie odzyskać część zainwestowanych środków, ale szacujemy, że poniesiemy stratę w wysokości  $r_p = -100\,000$  złotych. W sytuacji, gdy podpiszemy umowę, kalkulujemy, że nasz zysk,

po odliczeniu kosztów, wyniesie  $r_s = 50\,000$  złotych. Możemy wycofać się z rozmów i wtedy środki, jakie musielibyśmy przeznaczyć na rozpoczęcie prac, inwestujemy w bezpiecznej formie (np. lokata, obligacje), dającej zysk  $r_b = 10\,000$  złotych. Prawdopodobieństwo otrzymania kontraktu oceniamy na  $p = 0,7$ .

Stosowane w analizie decyzyjnej drzewa zawierają trzy rodzaje węzłów:

- węzły decyzji – definiujące decyzje, które mają zostać podjęte w danym problemie; każda wychodząca z nich gałąź odpowiada jednej określonej możliwej alternatywie decyzyjnej i opisana jest jej nazwą;
- węzły zmiennych losowych – definiujące różnego rodzaju zjawiska losowe mające wpływ na skutki poszczególnych alternatyw decyzyjnych; wychodzące z nich gałęzie odpowiadają możliwym wartościom zmiennej i opisane są nazwą tej wartości oraz jej prawdopodobieństwem;
- węzły skutków decyzji – służące jako liście drzewa i definiujące wartość każdego możliwego skutku.

Drzewo decyzyjne zawiera liście dla każdego możliwego skutku wszystkich alternatyw decyzyjnych. Poszczególne gałęzie w drzewie łączą się w ścieżki określające sytuację, w której dany skutek jest realizowany. Drzewo decyzyjne dla sytuacji opisanej w naszym przykładzie przedstawione zostało na rysunku 3.4.



**Rysunek 3.4.** Drzewo decyzyjne dla przykładowej decyzji o rozpoczęciu budowy

**Źródło:** opracowanie własne.

Aby dokonać wyboru alternatywy decyzyjnej, obliczamy wartość oczekiwaną dla każdej gałęzi węzła  $D$ .

$$E(\text{„Nie rozpoczynaj”}) = r_b = 10\,000$$
$$E(\text{„Rozpocznij”}) = pr_s + (1 - p)r_p = 0,7 \cdot 50\,000 + 0,3(-100\,000) = 5000$$

Jak widać, wartość oczekiwana dla decyzji o rozpoczęciu budowy jest niższa niż w przypadku decyzji o rezygnacji z niej (i rezygnacji z dalszych negocjacji kontraktowych). W tym przypadku nie opłaca się zatem zaryzykować dalszych starań – powinniśmy podjąć decyzję o rezygnacji z kontraktu.

Drzewo decyzyjne dla naszego przykładu jest bardzo proste. Może ono być znacznie bardziej złożone, obejmować większą liczbę węzłów zjawisk losowych, a także węzłów decyzji (tzw. decyzje sekwencyjne). Liczba wychodzących z poszczególnych węzłów gałęzi również jest niekiedy większa. W trudniejszych przypadkach możemy być jednak zmuszeni do sięgnięcia po bardziej zaawansowane metody wspomagania decyzji w warunkach ryzyka, takie jak skomplikowane modele probabilistyczne w formie sieci Bayesowskich, symulacje Monte Carlo czy też modele teorii gier.

## 3.4. Środowisko analityczne organizacji

### 3.4.1. Struktura środowiska analitycznego organizacji

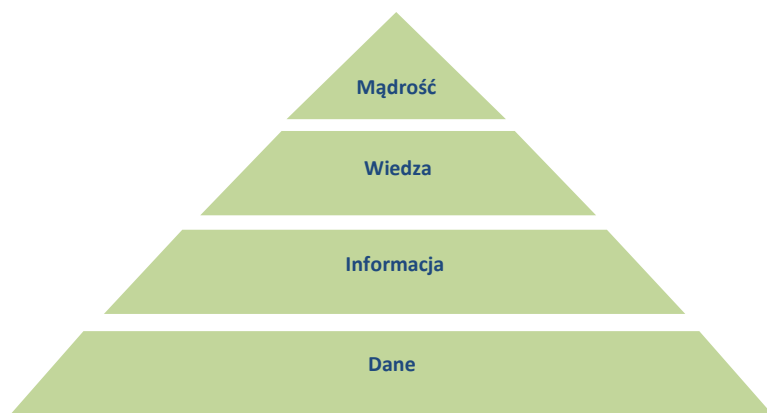
Skuteczne wsparcie informatyczne procesów decyzyjnych w organizacji w dużej mierze opiera się na wykorzystaniu posiadanych przez nią zasobów informacyjnych. Dostarczają one danych na temat istotnych faktów i zjawisk definiujących konkretne sytuacje decyzyjne, zasilając tym samym zarówno decydentów na wszystkich poziomach zarządzania, jak i używane przez nich systemy informatyczne. Stanowią one również istotny zasób analityczny, pozwalający na rozszerzanie wiedzy wykorzystywanej w podejmowaniu decyzji częściowo strukturalnych i niestrukuralnych oraz w warunkach niepewności i ryzyka.

O pierwszej z tych funkcji mówiliśmy w poprzednim rozdziale książki, charakteryzując środowisko informatyczne organizacji. Obecnie więc skupimy się przede wszystkim na drugiej – zasobach informacyjnych jako podstawie procesów analitycznych. Tradycyjnie już przyjmuje się, że różne rodzaje zasobów informacyjnych układają się w pewną piramidę, określaną często akronimem DIKW (*data-information-knowledge-wisdom*) – rysunek 3.5.

Podstawą piramidy DIKW są dane, które reprezentują surowe, nieprzetworzone rzeczy, zjawiska, fakty, procesy itp. Mogą być one rejestrowane (gromadzone), przetwarzane i transmitowane. Obsługują większość rodzajów decyzji, ale tylko najprostsze decyzje opierają się wyłącznie na danych.

Informacja to dane uporządkowane i zinterpretowane, nadające danym (surowym, nieprzetworzonym liczbom, tekstom, obrazom, dźwiękom itp.) znaczenia, co wzbogaca naszą wiedzę na określony temat. Dane zyskują wartość informacyjną dzięki odpowiedniej interpretacji. Znaczna część decyzji menedżerskich w organizacji opiera się na informacji.

Wiedza powstaje z istotnych dla odbiorcy danych i informacji, które zostały zweryfikowane w praktyce. Informacja ma charakter specyficzny, wiedza – ogólny. Innymi słowy, wiedza to informacje i dane ustrukturyzowane, przetworzone i uogólnione, odpowiednie dla danej sytuacji i możliwe do zweryfikowania empirycznie. Tworzona jest ona w wyniku procesów analitycznych, umożliwiających wykrywanie pewnych ogólnych wzorców w danych i informacjach na temat specyficznych przypadków związanych z analizowaną sytuacją.



**Rysunek 3.5.** Piramida DIKW

**Źródło:** opracowanie własne.

Dostarczanie decydentom wiedzy jest podstawowym zadaniem środowiska analitycznego organizacji. Jak wspomnieliśmy, brak wiedzy w procesie decyzyjnym powoduje rozszerzanie się obszaru niepewności. Powiększanie dostępnego decydentowi zakresu wiedzy pozwala więc tę niepewność ograniczać, a w niektórych sytuacjach nawet redukować do zera, co z kolei zmniejsza czynnik ryzyka. Okazuje się to szczególnie istotne w przypadku decyzji niestrukturalnych, których cechą, przypomnijmy, jest właśnie brak wiedzy na temat właściwych procedur ich podejmowania.

Mądrość to sfera określająca, kiedy i jak korzystać z wiedzy oraz jaka wiedza będzie odpowiednia dla danej sytuacji. Ten poziom ciągle pozostaje domeną ludzką. Systemy Wspomagania Decyzji raczej asystują w procesie dokonywania wyboru, współpracując z decydentem i pomagając mu wypracować najlepsze rozwiązanie. To decydent-człowiek jest stroną aktywną, określającą strategię tej współpracy. Tym niemniej, jak to kilkakrotnie podkreśliliśmy w poprzednich punktach bieżącego rozdziału, w pewnych wąskich dziedzinach problemowych tworzy się czasami systemy informatyczne w pełni automatyzujące złożone (częściowo strukturalne i niestrukturalne) procesy decyzyjne – dzieje się tak w przypadku wykorzystywania najbardziej zaawansowanych metod eksploracji danych i sztucznej inteligencji.

**Tabela 3.9.** Podstawowe obszary środowiska analitycznego organizacji

	<b>Opisowa</b>	<b>Predykcyjna</b>	<b>Preskrypcyjna</b>
<b>Pytania</b>	<i>Co się stało? Co się dzieje?</i>	<i>Co się stanie? Dlaczego coś się stanie?</i>	<i>Co należy zrobić? Dlaczego należy to zrobić?</i>
<b>Metody</b>	raporty biznesowe metody wizualizacji hurtownie danych	predykcyjne metody eksploracji danych analiza asocjacji prognozowanie	optymalizacja symulacja modelowanie decyzji
<b>Wyniki</b>	dobrze zdefiniowane problemy i możliwości biznesowe	przewidywania przyszłych stanów i warunków	najlepsze możliwe decyzje biznesowe

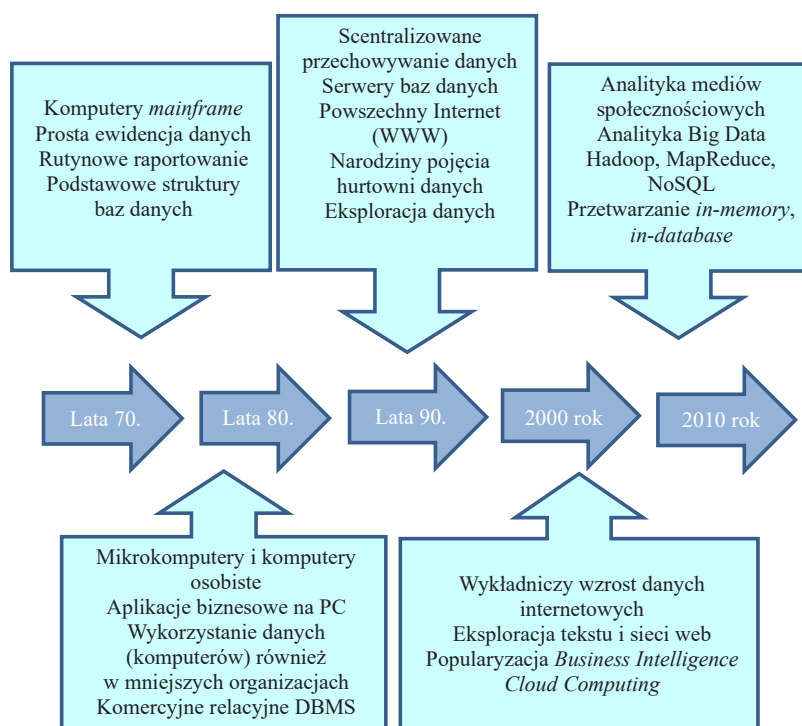
**Źródło:** opracowanie własne na podstawie [Sharda, Delen i Turban, 2017].

Służące do wspomagania decyzji w organizacji środowisko analityczne zwykle dzieli się na trzy podstawowe obszary: analitykę opisową, predykcyjną i preskrypcyjną. W tabeli 3.9 została przedstawiona ich krótka charakterystyka w formie typowych pytań, jakie są w nich stawiane,

stosowanych w każdym obszarze przykładowych metod oraz wyników, jakie chcemy otrzymać w danej sferze.

Na rysunku 3.6 zilustrowany został rozwój środowiska analitycznego organizacji na tle rozwoju informatyki. Pojawienie się komercyjnych komputerów w latach sześćdziesiątych i siedemdziesiątych XX wieku zaowocowało pierwszymi krokami w tworzeniu wsparcia informatycznego w organizacji. Związane one były przede wszystkim z możliwością informatyzacji podstawowych operacyjnych funkcji biznesowych.

Systemy transakcyjne, bo o nich tu mówimy, codziennie pomagają lub nawet samodzielnie podejmują w przedsiębiorstwie tysiące rutynowych decyzji: przede wszystkim strukturalnych, na poziomie operacyjnym. Oprócz wsparcia operacyjnego umożliwiają także ewidencję procesów biznesowych i tworzenie dokumentów pomocniczych. Z punktu widzenia środowiska analitycznego najistotniejsze jest chyba to, że systemy tego rodzaju rozpoczęły gromadzenie w organizacjach danych w formie maszynowej.



**Rysunek 3.6.** Rozwój środowiska analitycznego organizacji na tle rozwoju informatyki

**Źródło:** opracowanie własne.

Prawdziwym przełomem w tej dziedzinie było wprowadzenie scentralizowanych baz danych. Dzięki temu rozwiązaniu dane stały się strategicznym zasobem organizacji i podstawą jej środowiska analitycznego. Pierwsze koncepcje i struktury baz danych zostały opracowane już na początku lat siedemdziesiątych XX wieku, jednak ich rozwój początkowo ograniczony był dostępnością technologii komputerowej. W praktyce dopiero rewolucja mikrokomputerowa lat osiemdziesiątych spowodowała ich rozpowszechnienie w organizacjach, a w latach dziewięćdziesiątych scentralizowane serwery baz danych stały się rozwiązaniem obowiązującym niemalże wszędzie.

Zacząto więc gromadzić coraz więcej danych, dostęp do nich w organizacji stał się znacznie łatwiejszy – co, wraz z rosnącymi możliwościami sprzętu komputerowego, nakręciło spiralę potrzeb w zakresie analityki. W połowie lat dziewięćdziesiątych pojawia się taki element środowiska analitycznego, jak hurtownia danych ułatwiająca analityczne wykorzystanie zasobów informacyjnych, zgromadzonych w bazie danych.

W tym samym okresie rozpowszechniają się komercyjne aplikacje bazujące na metodach eksploracji danych. I znów – bardzo wiele metod z tej dziedziny stworzonych zostało znacznie wcześniej. Przedtem jednak były one przede wszystkim domeną badań naukowych i aplikacji prototypowych.

Termin „eksploracja danych” odnosi się do danych strukturalnych. Lata dziewięćdziesiąte XX wieku są jednak okresem intensywnego rozwoju Internetu, w którym informacja dostępna jest głównie w formie dokumentów. Jej wykorzystanie wymaga odmiennych metod analitycznych. Po roku 2000 zaczyna więc rosnąć zainteresowanie eksploracją tekstów i sieci web.

W tym samym okresie zaczyna również wzrastać zainteresowanie inną dziedziną, do tej pory również pozostającą głównie obszarem zainteresowań naukowych oraz aplikacji prototypowych – sztuczną inteligencją. Sieci neuronowe, systemy ekspertowe, logika rozmyta czy algorytmy genetyczne stają się powszechnie stosowanymi metodami komercyjnymi.

W ostatniej dekadzie, po roku 2010, środowisko analityczne organizacji uzupełniają takie elementy, jak analityka mediów społecznościowych czy zagadnienia związane z Big Data.

W dalszych punktach podrozdziału przyjrzymy się bliżej najważniejszym z wymienionych wyżej elementów środowiska analitycznego.



Przypomnijmy, że przeglądu metod analityki preskrypcyjnej dokonaliśmy w poprzednim podrozdziale 3.3; natomiast zagadnieniu sztucznej inteligencji przyjrzymy się w rozdziale 3.5.

### **3.4.2. Dane jako zasób organizacji – centralne zarządzanie danymi**

Zasadniczo można mówić o dwóch ogólnych podejściach do przechowywania danych w organizacji: w plikach lub w centralnej bazie danych (rysunek 3.7). Podejście pierwsze ma raczej charakter historyczny i wiąże się z początkowym okresem rozwoju systemów informatycznych zarządzania. Niemniej jednak, nawet w dzisiejszych czasach, pewna część zasobów informacyjnych przechowywana jest w tej formie, np. dane z arkuszy kalkulacyjnych lub innych aplikacji desktopowych, wykorzystywanych przez indywidualnych użytkowników.

Obecnie zdecydowanie dominuje wykorzystywanie do przechowywania danych scentralizowanych serwerów baz danych. Wprowadzenie takiego rozwiązania stanowiło znaczny przełom umożliwiający przekształcenie danych do zasobu całej organizacji, również z punktu widzenia celów analitycznych. Dlatego w bieżącym punkcie przyjrzymy się krótko powodowi wprowadzenia centralnego zarządzania danymi, omówimy korzyści wynikające z tego podejścia, a następnie scharakteryzujemy najczęściej wykorzystywany, relacyjny model bazy danych.

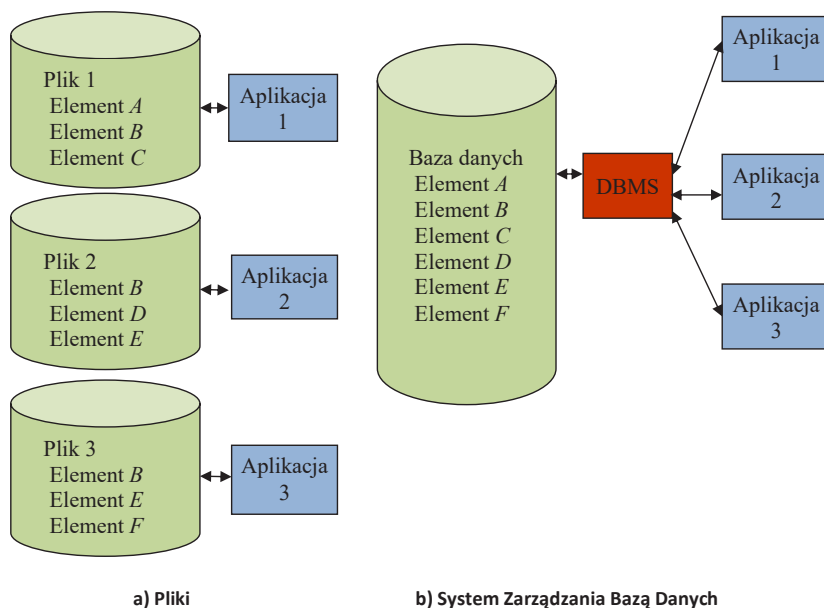
Jak wspomnieliśmy, wykorzystanie plików do przechowywania danych dominowało w początkowym okresie rozwoju systemów informatycznych zarządzania. W tradycyjnym podejściu systemy informatyczne miały bowiem charakter izolowany, tzn. każdy system złożony był ze zbioru odrębnych programów, operujących na własnym zestawie plików, przechowujących dane potrzebne do realizacji jego funkcji. Stąd np. system przetwarzania zamówień sprzedaży oraz system fakturowania stanowić mogły zupełnie oddzielne, izolowane fragmenty systemu informacyjnego organizacji.

Takie rozwiązanie powoduje powstanie całego szeregu problemów związanych przede wszystkim z brakiem integracji danych organizacji oraz przetwarzających je programów aplikacyjnych.

Dane nie stanowiły jednego wspólnego zasobu wykorzystywanego przez systemy informatyczne, ale były niejako „posiadane” odrębnie przez

każdy program (punkt *a* na rysunku 3.7). Ich współdzielenie okazywało się bardzo utrudnione – inny użytkownik (zarówno program, jak i człowiek) nie mógł często uzyskać dostępu do danych odmiennego systemu, a przynajmniej było to skomplikowane. Dlatego programy odwołujące się do tych samych obiektów musiały przechowywać swoje własne dane na ich temat – co ilustruje przypadek wykorzystywanego przez wszystkie aplikacje elementu *B* (rysunek 3.7).

Niedogodności te legły u podstaw koncepcji wyodrębnienia danych jako oddzielnego zasobu, który wykorzystywany jest i zarządzany przez całą organizację, a nie konkretną grupę użytkowników lub aplikacji (punkt *b* na rysunku 3.7). W tym celu wykorzystuje się tzw. system zarządzania bazą danych (*Data Base Management System* – DBMS) [Date, 2000; Ullman i Widom, 2000; Garcia-Molina, Ullman i Widom, 2006], stanowiący zorganizowany zbiór narzędzi umożliwiających dostęp i zarządzanie danymi.



**Rysunek 3.7.** Podejścia do przechowywania danych w organizacji

**Źródło:** opracowanie własne.

System Zarządzania Bazą Danych jest więc powłoką, która otacza dane i za pomocą której dokonują się wszystkie na nich operacje. Steruje on czymś, co nazwiemy bazą danych, jaką użytkownicy (zarówno ludzie,

jak i aplikacje) mogą przeszukiwać lub aktualizować bez odwołań czy nawet wiedzy o fizycznym sposobie przechowywania danych. Do wybranych, najistotniejszych z naszego punktu widzenia, funkcji DBMS należą:

1) Zarządzanie danymi:

- dodawanie nowych danych do bazy danych,
- usuwanie danych z bazy danych,
- modyfikowanie struktury istniejących danych,
- wstawianie nowych danych do istniejących struktur,
- aktualizowanie danych w istniejących strukturach,
- usuwanie danych z istniejących struktur,
- wyszukiwanie informacji:
  - wydobywanie danych z istniejących plików do stosowania przez użytkowników,
  - wydobywanie danych do stosowania przez programy użytkowe.

2) Zarządzanie bazą danych:

- tworzenie i monitorowanie użytkowników,
- ograniczanie dostępu do plików w bazie danych,
- monitorowanie działania bazy danych.

Zastosowanie scentralizowanej bazy danych daje organizacji szereg korzyści, z których większość wiąże się z dwoma podstawowymi elementami:

- umożliwienie dostępu do danych wielu aplikacjom i użytkownikom oraz implementacji centralnego zarządzania danymi w organizacji;
- rozdzielenie przechowywania i wykorzystania danych – niezależność danych i korzystających z nich aplikacji lub użytkowników.

Centralne zarządzanie danymi pozwala przede wszystkim na korzystanie z nich jako pewnej całości zasobów informacyjnych, opisujących stan, transakcje i zasoby organizacji.

Najważniejszą chyba korzyścią jest możliwość integracji danych. Centralna baza danych umożliwia połączenie wszystkich informacji z odrębnych plików w jeden wspólny zbiór. Pliki mają charakter izolowany: nie pozwalają na bezpośrednie powiązanie danych z dziedzin różnych systemów. Przekazanie danych między różnymi systemami informatycznymi odbywa się na ogół poprzez operacje wejścia/wyjścia systemu, np. skonfrontowanie danych o zamówieniach i stanach magazynowych wykonywane jest zwykle poprzez porównanie danych wyjściowych pochodzących z odpowiednich systemów. W przypadku centralnej bazy danych

możemy pobierać dane przekrojowo, bezpośrednio zestawiając powiązane ze sobą dane z różnych obszarów działalności organizacji.

Dane w bazie danych mogą być użytkowane przez wiele osób i aplikacji, niekiedy nawet jednocześnie. Dzięki temu łatwiejsze i bardziej efektywne staje się koordynowanie, kontrolowanie i zarządzanie danymi. Można tworzyć nowe aplikacje, korzystające z danych już przechowywanych, często bez konieczności tworzenia nowych zbiorów danych. Każdy użytkownik uzyskuje dostęp do dowolnych, przechowywanych przez organizację danych szczegółowych, które są potrzebne do realizacji jego zadań (i w związku z tym, do których posiada uprawnienia)

Centralna baza danych nie musi (i nie powinna) zawierać danych zbędnych lub zbędnie powtarzających się. W podejściu plikowym, jak wspomnieliśmy, każdy system informatyczny zarządza własnymi zbiorami danych – stąd te same dane mogą pojawiać się w wielu plikach wykorzystywanych przez różne aplikacje. Przykładowo: systemy przetwarzania zamówień i fakturowania będą przechowywać oddzielne zestawy danych z informacjami o klientach, powielając przy tym wiele wspólnych informacji.

Celem bazy danych staje się przechowywanie jednego logicznego elementu danych tylko raz, w jednym miejscu. Jeden zbiór danych klienta dostępny jest wówczas dla wszystkich systemów informacyjnych czy użytkowników. Dzięki temu np. zmiana adresu klienta wykonywana jest tylko raz, podczas gdy w przypadku funkcjonowania wielu aplikacji plikowych należy pamiętać o odpowiedniej modyfikacji adresu w każdym przechowującym dane o tym kliencie pliku.

Integralność danych polega na zapewnieniu zgodności między danymi a opisywaną przez nie rzeczywistością. Zmiany dokonane po jednej stronie tego związku powinny być dokładnie odzwierciedlone po drugiej. Dane przechowywane w różnych fragmentach bazy danych muszą być ponadto spójne i wewnętrznie niesprzeczne. W przypadku technologii plikowej za integralność danych i zachowanie ich spójności odpowiedzialne są poszczególne aplikacje – każda z nich może weryfikować dane tylko pod kątem własnej logiki działania. W przypadku centralnej bazy danych wiele reguł definiujących zasady integralności może zostać określonych na poziomie DBMS, tak aby obowiązywały one wszystkie aplikacje modyfikujące bazę danych. Dzięki temu dane mogą być weryfikowane pod względem wielu aspektów związanych z różnymi obszarami organizacji.

Centralne zarządzanie danymi pozwala na narzucenie standardów reprezentacji danych, co jest szczególnie istotne dla wymiany danych lub ich przenoszenia między systemami.

Z kolei scentralizowany charakter bazy danych umożliwia (ale również wymaga) ustanowienia dobrego systemu zabezpieczeń, zintegrowanego z ogólną polityką bezpieczeństwa organizacji. Podstawową metodą zabezpieczania danych jest określenie zbioru uprawnionych użytkowników w odniesieniu do całości lub pewnej części danych. DBMS mają wbudowane elementy, takie jak systemy uprawnień, rejestracja i uwierzytelnianie użytkowników, pozwalające na realizację tych funkcji.

Drugą istotną korzyścią wynikającą z przechowywania danych w formie scentralizowanej bazy jest rozdzielenie przechowywania i wykorzystania danych, co skutkuje większą niezależnością danych i korzystających z nich użytkowników (zarówno aplikacji, jak i ludzi). W tradycyjnych plikowych systemach informatycznych mamy często do czynienia z sytuacją, gdy „wiedza o organizacji danych i technice dostępu są wbudowane w logikę i kod aplikacji” [Date, 2000]. Aplikacja (programista) czy też korzystający z danych użytkownik musi znać fizyczne umiejscowienie i strukturę danych użytych w pliku.

W przypadku systemów informatycznych powoduje to konieczność ponoszenia często znacznych nakładów związanych z utrzymaniem programów oraz dostosowywaniem ich do ewoluującej struktury danych. Zmiana struktury danych jest zjawiskiem naturalnym – wraz z upływem czasu modyfikacjom ulegają zarówno zasoby informacyjne, jak i cele organizacji eksploatującej system. Gdy mamy do czynienia z aplikacją silnie zależną od danych, niezbędne staje się dokonanie poprawek dostosowawczych, nawet jeśli modyfikacja struktury danych nie dotyczy jej bezpośrednio.

W przypadku ludzi problem ten związany jest raczej z kwestią trudności dostępu do danych. Użytkownik, chcąc uzyskać do nich dostęp, albo musi znać fizyczne szczegóły dotyczące sposobu ich przechowywania w plikach, albo potrafić je uzyskać przy pomocy obsługujących je aplikacji. W obydwu sytuacjach wymaga to wiedzy specjalistycznej, wykraczającej zwykle poza zakres wiadomości menedżera.

Odnośnie danych przechowywanych w bazie danych możemy natomiast mówić o dwóch oddzielnych modelach:

- model logiczny – związany jest z koncepcyjną, semantyczną organizacją danych oraz związków między nimi;

- model fizyczny – odnosi się do fizycznego sposobu przechowywania danych na dyskach, taśmach lub innych urządzeniach.

Łącznikiem między logicznym i fizycznym modelem danych jest System Zarządzania Bazą Danych. Aplikacje i użytkownicy korzystać muszą wyłącznie z modelu logicznego – wiedza o sposobie przetransformowania go na widok fizyczny zawarta jest bowiem w DBMS.

Oddzielenie danych od programów, które je wykorzystują, powoduje, że zmiana pewnej części bazy danych nie zawsze musi powodować konieczność modyfikacji programów z niej korzystających i na odwrót – modyfikacja systemu aplikacyjnego nie zawsze ma wpływ na strukturę danych w bazie. Cecha ta ułatwia zarówno tworzenie aplikacji, jak i zarządzanie danymi.

Co istotniejsze, użytkownicy nie muszą znać fizycznego modelu danych – uzyskują do nich dostęp przez uniwersalny interfejs schematu logicznego. Ułatwia to nie tylko korzystanie z nich, ale i upowszechnianie wśród różnorodnych użytkowników.

Widzimy więc, że oparcie zasobów danych na scentralizowanej bazie danych, czyli centralne zarządzanie danymi, oraz uniezależnienie danych od aplikacji stanowią pewne warunki konieczne dla możliwości efektywnego wykorzystania tych zasobów w procesach analityczno-decyzyjnych w organizacji. Podsumowując nasze poprzednie rozważania, oprócz szeregu bardzo istotnych korzyści technicznych:

- daje to możliwość dostępu do dowolnych danych szczegółowych zgromadzonych przez organizację;
- umożliwia zintegrowane zaspokajanie potrzeb informacyjnych poprzez tworzenie przekrojowych raportów z różnych podsystemów organizacji;
- ułatwia dostęp do danych poprzez możliwość pominięcia ich schematu fizycznego.

Właściwości te pozwalają przekształcić dane w zasób strategiczny organizacji. Centralna baza danych stanowi niezbędny element jej środowiska analitycznego.

Obecnie wykorzystuje się kilka różnych modeli baz danych. Bez wątpienia jednak zdecydowanie najważniejszym z nich jest model relacyjny.

Podstawową strukturę danych w modelu relacyjnym stanowi tabela (w formalnym podejściu nazywana relacją) złożona z kolumn i wierszy (rekordów, krotek). Tabele relacyjne opisują zjawiska świata

rzeczywistego, takie jak rzeczy (materialne lub niematerialne) czy też wydarzenia – przy czym jedna tabela przechowuje dane na temat obiektów tego samego rodzaju. Przykładowo: tabela 3.10 zawiera dane na temat studentów. Dane dotyczące sal wykładowych muszą być gromadzone w innej tabeli.

**Tabela 3.10.** Przykładowa tabela relacyjna (przechowująca dane o studentach)

Numer albumu	Nazwisko	Imię	Rok	Grupa	ID typu studiów	...
1248390/S	Kowalski	Adam	2	3	1	...
4382350/S	Nowak	Juliusz	1	1	2	...
2114345/S	Malinowski	Jan	2	3	2	...
1823839/S	Jankowski	Leopold	3	1	3	...

**Źródło:** opracowanie własne.

Każdy wiersz tabeli zawiera dane dotyczące jednego konkretnego obiektu, a każda kolumna – jednej określonej cechy wszystkich przechowywanych obiektów. W przypadku większości tabel jedną z cech (kolumn) musi być jednoznaczny identyfikator obiektu w każdym wierszu. W terminologii baz danych nazywany jest on kluczem głównym.

Podsumowując, jeśli chcemy uzyskać dostęp do przechowywanych w jakiejś tabeli obiektów, musimy: znać nazwę właściwej tabeli, umieć wskazać, które wiersze z niej są nam potrzebne oraz znać nazwy kolumn zawierających poszukiwane dane dla wybranych wierszy.

W praktyce zazwyczaj obiekty świata rzeczywistego łączą się między sobą w różnorodne powiązania. Przykładowo: studenci uczęszczają na określone zajęcia; u danego dostawcy przedsiębiorstwo dokonuje szeregu zakupów; na fakturze wyspecyfikowane są określone towary; w każdym budynku znajduje się zazwyczaj szereg pomieszczeń itp. Często więc potrzebujemy informacji nie z jednej, ale z dwóch lub większej liczby tabel przechowujących dane o powiązanych ze sobą obiektach.

Wyróżnić możemy trzy podstawowe rodzaje powiązań między tabelami. Mówimy, że między dwiema tabelami istnieje powiązanie jeden-do-jednego, jeśli pojedynczemu rekordowi (krotce, wierszowi) z tabeli pierwszej przyporządkowany jest co najwyżej jeden rekord z drugiej, i na odwrót – pojedynczemu rekordowi z drugiej tabeli może odpowiadać co najwyżej jeden rekord z tabeli pierwszej.

**Tabela 3.11.** Przykład powiązania jeden-do-jednego

Kod wykładowcy	Nazwisko	Imię	Jednostka	...
KI38	Banach	Stefan	Katedra Informatyki	...
KI45	Kopernik	Mikołaj	Katedra Informatyki	...
KI34	Steinhaus	Hugon	Katedra Informatyki	...

Kod wykładowcy	Płaca zasadnicza	Dodatek stażowy	...
KI45	1934	10%	...
KI38	1890	15%	...
KI34	2433	15%	...

**Źródło:** opracowanie własne.

Przykład powiązania jeden-do-jednego przedstawiony został w tabeli 3.11. Dla każdego wiersza danych wykładowcy w tabeli pierwszej istnieje dokładnie jeden wiersz określający składniki jego wynagrodzenia w tabeli drugiej. Zwróćmy uwagę, że powiązanie między wierszami tabel następuje poprzez jednakowe wartości danych występujące w jednej z kolumn (zazwyczaj identyfikatorów, kluczy głównych) w obu tabelach, w tym przypadku w kolumnie „kod wykładowcy”.

Mówimy, że między dwoma tabelami istnieje powiązanie jeden-do-wielu, jeśli pojedynczemu rekordowi (wierszowi) z pierwszej tabeli może odpowiadać jeden rekord lub więcej z tabeli drugiej, podczas gdy pojedynczemu rekordowi z tabeli drugiej odpowiada najwyżej jeden rekord tabeli pierwszej. Jest to najczęściej występujący przypadek powiązania między tabelami.

Przykład takiego powiązania przedstawiony został w tabeli 3.12. Każdy z wykładowców może mieć wiele zajęć, ale konkretne zajęcia prowadzone są przez jednego wykładowcę. Również w tym przypadku powiązanie następuje poprzez jednakowe wartości danych we wspólnej kolumnie: tabela druga musi zawierać kolumnę odpowiadającą kluczowi głównemu tabeli pierwszej – w tym przypadku jest to „kod wykładowcy” (nazwiemy go kluczem obcym). Jak widzimy w tabeli 3.12, wykładowca o identyfikatorze KI38 prowadzi dwa zajęcia, dla których wartość klucza obcego jest także równa KI38.



**Tabela 3.12.** Przykład powiązania jeden-do-wielu

Kod wykładowcy	Nazwisko	Imię	Jednostka	...
KI38	Banach	Stefan	Katedra Informatyki	...
KI45	Kopernik	Mikołaj	Katedra Informatyki	...
KI34	Steinhaus	Hugon	Katedra Informatyki	...

Kod przedmiotu	Nazwa przedmiotu	Kod wykładowcy	Dzień tygodnia	Godzina	Rok	Grupa	...
238	Wstęp do informatyki	→ KI38	wtorek	11.30	2	3	...
239	Projektowanie i programowanie	→ KI38	wtorek	13.15	1	1	...
345	Systemy informatyczne zarządzania	KI45	środa	11.30	2	3	...
334	Sieci komputerowe	KI34	piątek	16.30	3	1	...

**Źródło:** opracowanie własne.

Mówimy, że między dwoma tabelami występuje powiązanie wiele-do-wielu, jeśli pojedynczemu rekordowi jednej z nich może odpowiadać jeden rekord lub więcej z drugiej i na odwrót – pojedynczy rekord z drugiej powiązany jest z jednym rekordem (lub więcej) z tabeli pierwszej.

Przykład tego rodzaju powiązania przedstawiony został w tabeli 3.13. Każdy student uczestniczy w wielu zajęciach, z kolei na każde zajęcia zapisuje się wielu studentów. Zwróćmy jednak uwagę, że tabele te nie mają wspólnej kolumny, która mogłaby je ze sobą powiązać. Ponieważ w zajęciach uczestniczy wielu studentów, trudno by nam było dodać identyfikatory ich wszystkich do wiersza zajęć. I na odwrót – ponieważ student bierze udział w różnych zajęciach, pracochłonne będzie dodawanie ich identyfikatorów w wierszu studenta.

Dlatego powiązanie wiele-do-wielu definiuje się na ogół z wykorzystaniem dodatkowej tabeli łączącej, której każdy wiersz zawiera parę powiązanych identyfikatorów. Jeśli chcemy np. znaleźć wszystkich studentów uczęszczających na dane zajęcia, powiedzmy „Wstęp do informatyki” (kod 238), znajdujemy wszystkie wiersze zawierające ten kod w tabeli „Rozkład zajęć”. Odpowiadające im numery indeksów pozwolą zlokalizować dane studentów.

**Tabela 3.13.** Przykład powiązania wiele-do-wielu

Numer albumu	Nazwisko	Imię	Wiek	Rok	Grupa	ID typu studiów	...
1248390/S	Mickiewicz	Adam	20	2	3	1	...
4382350/S	Słowacki	Juliusz	19	1	1	2	...
2114345/S	Kochanowski	Jan	20	2	3	2	...
1823839/S	Staff	Leopold	22	3	1	3	...
5682441/S	Rej	Mikołaj	22	1	2	1	...

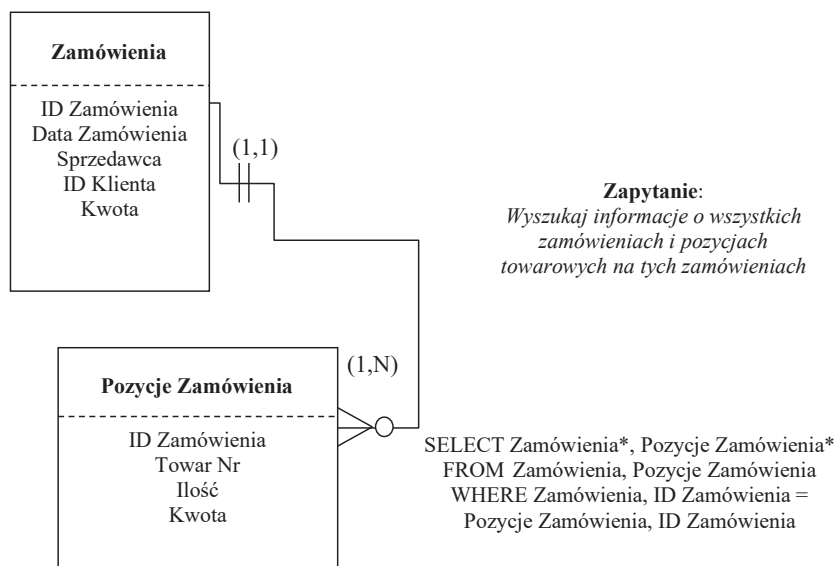
Kod przedmiotu	Nazwa przedmiotu	Kod wykładowcy	Dzień tygodnia	Godzina	Rok	Grupa	...
238	Wstęp do informatyki	KI38	wtorek	11.30	2	3	...
239	Projektowanie i programowanie	KI38	wtorek	13.15	1	1	...
345	Systemy informatyczne zarządzania	KI45	środa	11.30	2	3	...
334	Sieci komputerowe	KI34	piątek	16.30	3	1	...

Rozkład zajęć (tabela łącząca)	Numer albumu	Kod przedmiotu
	1248390/S	238
	1248390/S	345
	1248390/S	334
	4382350/S	238
	2114345/S	345
	2114345/S	334
	1823839/S	239
	5682441/S	334

**Źródło:** opracowanie własne.

Podsumowując, jeśli chcemy uzyskać dostęp do danych obiektów przechowywanych w kilku tabelach, musimy: znać nazwy tych tabel, umieć zdefiniować powiązania między nimi (w niektórych przypadkach ten element może być określony w bazie danych), umieć wskazać, które wiersze z tych tabel są nam potrzebne, a także znać nazwy kolumn zawierających poszukiwane dane dla wybranych wierszy.

Operacje wyszukiwania w bazie danych często określa się jako „zapytania”. Wykonuje się je przy pomocy specjalnego języka, nazywanego SQL [Bartkiewicz i Jabłoński, 2005]. Przykładowe zapytanie przedstawione zostało na rysunku 3.8. Znajomość języka SQL, chociaż przydatna, nie jest jednak niezbędna. Istnieje wiele narzędzi pozwalających tworzyć zapytania w formie graficznej, przy wskazaniu jedynie potrzebnych w konkretnym przypadku elementów.



**Rysunek 3.8.** Przykładowe zapytanie do bazy danych

**Źródło:** opracowanie własne na podstawie [Bartkiewicz i Jabłoński, 2005].

### 3.4.3. Hurtownie danych

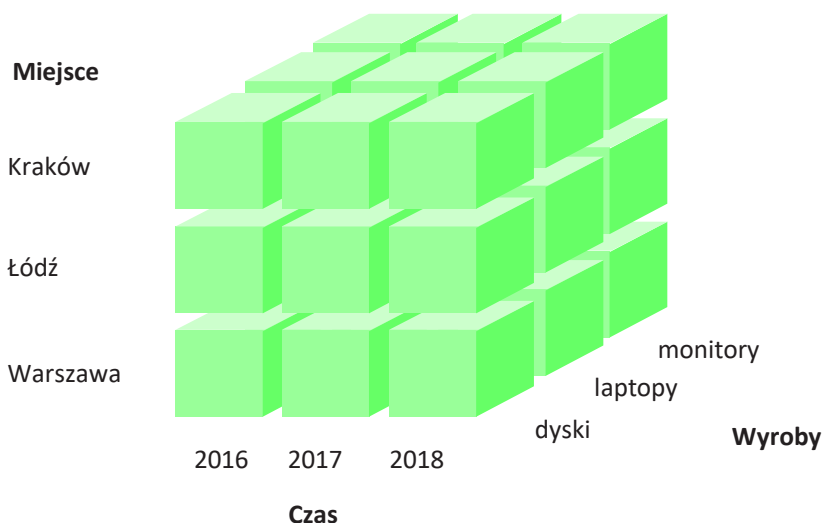
Hurtownia danych stanowi rozwinięcie idei bazy danych, dołączając do niej narzędzia ułatwiające rozwiniętą analizę danych.

Koncepcja hurtowni danych opiera się na spostrzeżeniu, że jednym z wąskich gardeł możliwości wykorzystania danych w procesach analitycznych jest konieczność ich wcześniejszej agregacji. Jeśli np. potrzebne okazuje się sporządzenie prognozy miesięcznej sprzedaży, to do jej stworzenia niezbędne staje się zagregowanie danych historycznych z poszcze-

gólnych transakcji w układzie miesięcznym. A teraz powiedzmy, że celem jest analiza danych nie tylko w rozbiciu na miesiące, ale jednocześnie z uwzględnieniem odbiorców, sprzedawanych towarów, miejsc sprzedaży itp. Przygotowanie niezbędnych agregacji danych może być bardzo kosztowne – czasowo i obliczeniowo.

Hurtownia danych składa się więc z:

- warstwy danych szczegółowych – normalna baza danych przechowująca dane organizacji, nazywana często od sposobu ich przetwarzania jako warstwa OLTP (*On-Line Transaction Processing*);
- warstwy danych zagregowanych wspomagających przetwarzanie analityczne, nazywana OLAP (*On-Line Analytical Processing*).



**Rysunek 3.9.** Trójwymiarowa kostka OLAP

**Źródło:** opracowanie własne.

Dane zagregowane w warstwie analitycznej organizowane są przy użyciu specjalnych metadanych, określanych jako kostki (mogą być wielowymiarowe) OLAP, zawierające dane szczegółowe zagregowane zgodnie z zadanymi kryteriami, nazywanymi „wymiarami”. Przez analogię można je rozumieć jako pewne rozwinięcie koncepcji tabeli przestawnej w arkuszu kalkulacyjnym dla wielu wymiarów i znacznie większej liczby danych.

Na rysunku 3.9 przedstawiony został przykład trójwymiarowej kostki OLAP zbudowanej dla danych dotyczących sprzedaży, zagregowanych pod względem czasu, miejsca oraz wyrobu. Każdy sześciąt na rysunku reprezentuje szczegółowe dane transakcyjne zsumowane według podanych kryteriów. Oczywiście w zależności od potrzeb kostki OLAP mogą posiadać znacznie większą liczbę wymiarów. Przykładowo, oprócz trzech przedstawionych wyżej, dane dotyczące sprzedaży można zorganizować, zaprezentować i analizować według klienta, sprzedawcy, wielkości transakcji itp.

Ogólnie rzecz ujmując, hurtownie danych umożliwiają wielowymiarową prezentację danych według kryteriów, takich jak:

- obiekty – produkty, sprzedawcy, segmenty rynku, jednostki biznesowe, lokalizacje geograficzne, kanały dystrybucji, kraje, branże;
- miary (rzeczywiste) – pieniądze, wielkość sprzedaży, liczba osób, zysk z zapasów;
- okresy, czas – dobowy, tygodniowy, miesięczny, kwartalny lub roczny.

Ponadto w hurtowni danych może znajdować się wiele kostek dla różnych zmiennych, zestawów wymiarów itp.

System zarządzania hurtownią danych zapewnia utrzymywanie kostek OLAP w stanie aktualnym. Zazwyczaj definiuje się również pewne podstawowe operacje na nich działające, takie jak:

- selekcja – wybór grup danych interesujących analityka; np. dla kostki na rysunku 3.9 dane o sprzedaży dysków z 2018 roku z Warszawy i Krakowa;
- projekcja – zmniejszenie liczby wymiarów (agregacja danych w likwidowanych wymiarach); np. uzyskanie danych w dwóch wymiarach, w rozbiciu na miejsce i czas sprzedaży (likwidujemy więc wymiar wyrobów, czyli dane dla wszystkich towarów muszą zostać zagregowane);
- wycinanie – połączenie selekcji z projekcją;
- zwijanie (agregacja) – czyli uogólnienie, zwiększanie poziomu agregacji danych; np. dla wymiaru miejsca sprzedaży z poziomu miast do poziomu województw;
- rozwijanie – uszczegóławianie, zmniejszenie poziomu agregacji danych; np. dla wymiaru czasu z poziomu lat na poziom miesięcy w każdym roku;
- sortowanie – tworzenie uporządkowanych rankingów.

### 3.4.4. Eksploracja danych

#### 3.4.4.1. Charakterystyka procesu eksploracji danych

Problematyka analizy danych i pozyskiwania z nich wiedzy nie jest nowym zagadnieniem. W przeszłości była ona jednak raczej przedmiotem zainteresowania środowisk naukowych niż praktyków biznesu. Choć dawniej powstawały – i to dosyć licznie – różne aplikacje, to ostatnie kilka, kilkanaście lat stanowi pod tym względem zupełnie nową epokę. Niemal wszystkie organizacje sięgają po metody zaawansowanej analityki danych, zaś wykorzystanie narzędzi opartych na zastosowaniu czasami bardzo zaawansowanych metod stało się w zasadzie czymś powszechnym.

Eksploracja danych (*data mining*) to proces polegający na wykorzystaniu metod analizy statystycznej oraz uczenia maszynowego do znajdowania niewidocznych przy bezpośredniej obserwacji, nowych, poprawnych, potencjalnie użytecznych zależności i prawidłowości w strukturalnych zbiorach danych. Inne nazwy to: drążenie danych, pozyskiwanie wiedzy, wydobywanie danych, ekstrakcja danych.

Zwróćmy uwagę na kilka istotnych cech eksploracji danych, wynikających wprost lub pośrednio z powyższej definicji [Sharda, Delen i Turban, 2017]. Określenie „proces” implikuje, że eksploracja danych obejmuje wiele iteracyjnych kroków. Stosuje ona nietrywialne metody, co oznacza, że w grę wchodzi poszukiwanie typu eksperymentalnego lub wnioskowanie – nie jest to zatem tak proste, jak obliczenie predefiniowanych wielkości. Ponadto celem eksploracji danych jest znalezienie nowych wzorców – takich, które nie były wcześniej znane użytkownikowi w kontekście analizowanego systemu. Muszą być one poprawne, z wystarczającym stopniem pewności prawdziwe (również dla nowych, niewykorzystanych w procesie analizy danych) oraz potencjalnie użyteczne – czyli przynosić pewne korzyści użytkownikowi lub zadaniu.

Można by zadać sobie pytanie, co stoi za obecnie obserwowanym sukcesem eksploracji danych. Dlaczego organizacje tak chętnie sięgają po te, wcale przecież niełatwe w tworzeniu i eksploatacji, narzędzia analityczne? Odpowiedź na to pytanie nie będzie specjalnie zaskakująca: po prostu mamy obecnie do czynienia z kombinacją potrzeb użytkowników oraz poziomu rozwoju samej technologii – wspólnie tworzą one potężny bodziec wpływający na falę zastosowań systemów eksploracji danych.

Podstawowym czynnikiem stojącym za szybkim rozwojem zastosowań narzędzi eksploracji danych są przede wszystkim potrzeby. Organizacje dosyć rozpaczliwie poszukują metod, które pomogłyby im w opamięnieniu złożoności współczesnego świata, charakteryzującego się coraz intensywniejszą konkurencją w skali globalnej, napędzaną stale zmieniającymi się preferencjami i żądaniami klientów, na coraz bardziej nasyconym rynku. Ponadto nabyte z kilkudziesięciu już lat rozwoju informatyki doświadczenia sprawiają, że współcześni menedżerowie dobrze rozumieją wartość kryjącą się w zasobach informacyjnych i zdają sobie sprawę z możliwości, jakie się za nimi kryją.

Jednakże współczesne infrastruktury informatyczne w organizacjach oferują znaczną dostępność wysokiej jakości danych o klientach, sprzedawcach, transakcjach, Internecie itp. Zasoby te są coraz większe, skonsolidowane i zintegrowane w repozytoriach baz danych oraz często hurtowniach danych. Mamy także do czynienia z gwałtownym wzrostem możliwości przetwarzania, przechowywania i dostępu do danych, związanym z rozwojem sprzętu komputerowego, oprogramowania oraz infrastruktury sieciowej i telekomunikacyjnej. Nie bez znaczenia jest przy tym znacząca redukcja kosztów sprzętu oraz oprogramowania do przechowywania i przetwarzania danych.

Oczywiście przedstawioną wyżej listę należy uzupełnić o kwestię opracowania nowych metod analitycznych. Szczególnie owocna pod tym względem była końcówka XX wieku, kiedy stworzone zostały główne metody stosowane obecnie w eksploracji danych. Okres po roku 2000 cechowały natomiast intensywne prace nad badaniem właściwości tych metod, ich rozwojem oraz zagadnieniami implementacji w praktycznych warunkach [Bengio, Courville i Goodfellow, 2018].

Do podstawowych ogólnych zadań eksploracji danych należą [Cichosz, 2000; Larose 2006 i 2008; Morzy, 2019]:

- 1) predykcja – określanie przyszłych wartości niektórych zmiennych lub wartości bieżących, których nie można bezpośrednio zaobserwować ani pomierzyć; mówimy tu o dwóch podstawowych rodzajach zadań:
  - analiza regresji – wykorzystanie procedur statystycznych do szacowania zależności między pewnymi opisywanymi przez dane zmiennymi; próbuje ona znaleźć funkcję modelującą tego rodzaju zależność z jak najmniejszym błędem;

- klasyfikacja – problem określania dla zjawisk opisanych wielowymiarowymi wzorcami cech (zmiennych), jednej (bądź kilku) kategorii, do której (których) należy je zakwalifikować (ze skończonego, predefiniowanego z góry zbioru kategorii);
- 2) grupowanie danych (analiza skupień) – zadanie polegające na wykrywaniu w zbiorach wielowymiarowych danych, grup danych podobnych do siebie; wykrywanie i nadawanie zbiorom danych struktury – rozbijanie ich na homogeniczne segmenty;
- 3) analiza asocjacji – wykrywanie związków między zmiennymi reprezentowanymi przez dane; opiera się na badaniu podobieństwa i częstości współwystępowania wartości zmiennych:
  - najważniejsze zastosowanie i źródło pochodzenia – analiza koszyka rynkowego;
  - analiza połączeń – automatycznie wykrywanie powiązań między wieloma interesującymi obiektami, takich jak łącza między stronami internetowymi i związki między grupami autorów publikacji akademickich;
  - eksploracja sekwencji – badanie związków między zdarzeniami pod względem kolejności ich występowania, w celu identyfikacji powiązań w czasie;
- 4) wizualizacja – występuje często w powiązaniu z innymi zadaniami eksploracji danych jako narzędzie prezentacji wizualnej wykrytych zależności i prawidłowości; może jednak być stosowana także jako samodzielne zadanie – prezentacji wielowymiarowych danych (opisanych wieloma cechami-zmiennymi) w formie jedno-, dwu- lub trójwymiarowych wizualizacji, ułatwiających człowiekowi ich analizę i wykorzystanie.

Utworzenie modelu eksploracji danych zwykle wymaga wykonania kilku standardowych kroków [Sharda, Delen i Turban, 2017]:

- 1) pierwszym krokiem jest zgromadzenie danych – wykorzystanie infrastruktury informatycznej i telekomunikacyjnej do zebrania jak największej liczby obserwacji danych; interesują nas przy tym tylko dane istotne, spełniające określone warunki z punktu widzenia wykonywanej analizy; na tym etapie trzeba więc zazwyczaj dokonać także wyboru właściwych danych spośród zasobów informacyjnych organizacji (np. z bazy danych), następnie dane są integrowane w jeden scalony zbiór – co wymaga ujednolicenia ich formatu i postaci,



przeliczeń doprowadzających do porównywalności (np. co do jednostek, czasów gromadzenia, mas badania itp.);

- 2) zebrane dane muszą zostać odpowiednio przygotowane do dalszego przetwarzania przy pomocy określonej metody eksploracji – kluczowe okazuje się tu na ogół ich oczyszczenie, np. poprzez uzupełnienie brakujących wartości, redukcję poziomu szumów, eliminację niezgodności w danych itp.; w wielu przypadkach dane muszą zostać poddane różnym transformacjom, takim jak: normalizacja do określonego przedziału wartości, dyskretyzacja danych lub ich agregacja, przeliczenie w celu stworzenia nowych syntetycznych atrybutów itp.; dane mogą być również poddawane redukcji co do liczby zmiennych lub przypadków (obserwacji);
- 3) kolejnym krokiem jest podział danych – w większości sytuacji dane muszą zostać podzielone na kilka podzbiorów; niemal zawsze ze zbioru danych wybieramy tzw. zbiór treningowy (zbiór danych wykorzystywanych do budowy modelu eksploracji) oraz zbiór testowy (do oceny działania stworzonego modelu); powinny to być niezależne zbiory danych, ponieważ model eksploracji jest na ogół dopasowywany do danych ze zbioru treningowego i może dla nich dawać zbyt optymistyczne wyniki działania; w przypadku wielu metod eksploracji niezbędne staje się również przygotowanie trzeciego zbioru danych, tzw. zbioru walidacyjnego, wykorzystywanego do roboczej niezależnej oceny działania modelu w trakcie jego tworzenia i na potrzeby tego procesu;
- 4) po przygotowaniu danych można przystąpić do budowy modelu – czyli wykorzystania podzbioru zebranych danych do analizy przy pomocy określonego algorytmu eksploracji, pozwalającego na identyfikację występujących w nich wzorców;
- 5) walidacja i testowanie modelu – ocena działania modelu przy pomocy walidacyjnego i testowego podzbioru danych; przez walidację rozumiemy bieżącą ocenę modelu w trakcie jego tworzenia; testowanie natomiast polega na ostatecznej ocenie modelu po zakończeniu jego budowy;
- 6) wykorzystanie modelu – zastosowanie modelu w kontekście rozmaitych sytuacji decyzyjnych jako elementu całości rozwiązania analitycznego.

Na koniec bieżącego punktu przyjrzymy się wybranym (spośród wymienionych) metodom realizacji danych zadań.

### 3.4.4.2. Analiza regresji

Metody regresyjne stosujemy w sytuacji, gdy nie jesteśmy w stanie bezpośrednio zaobserwować, pomierzyć wartości pewnej zmiennej liczbowej  $y$ , ale wiemy, że na tę nieznaną zmienną wpływ mają inne czynniki, określone przez zmienne  $x_1, \dots, x_n$ , których wartości potrafimy określić. Zadanie regresji polega na identyfikacji zależności między tymi zmiennymi  $y = f(x_1, \dots, x_n)$ . Zmienną  $y$  nazywamy zwykle zmienną zależną albo wyjściową, zaś  $x_1, \dots, x_n$  – zmiennymi niezależnymi albo wejściowymi tworzonego modelu.

Pamiętajmy przy tym, że odzwierciedlane przez dane zależności niemal zawsze mają charakter przybliżony. Nawet jeśli obserwujemy dwa dokładnie takie same wzorce wartości zmiennych niezależnych, może towarzyszyć im nieco inna wartość zmiennej zależnej. Wynika to z faktu, że zmienne  $x_1, \dots, x_n$  reprezentują tylko najważniejsze czynniki wpływające na zmienną  $y$ . Wszystkich niemal nigdy nie jesteśmy w stanie uwzględnić.

Zadanie regresji polega więc na predykcji średniej, przybliżonej wartości zmiennej  $y$  (wartości oczekiwanej tej zmiennej) dla danego zestawu wartości zmiennych wejściowych  $x_1, \dots, x_n$ . Z założenia więc musimy się liczyć z tym, że będziemy potrafili przewidzieć wartość zmiennej  $y$  z pewnym błędem. Błąd i będące jego skutkiem niepewność oraz ryzyko są elementami, które należy niemal zawsze brać pod uwagę przy wykorzystaniu metod analitycznych opartych na obserwacji.

Typowy sposób postępowania w przypadku modeli regresji polega na tym, że ogólny kształt funkcji modelu  $f$  (np. funkcja liniowa, logistyczna itp.) dobierany jest zazwyczaj z góry, na bazie pewnych przesłanek teoretycznych, tak aby odpowiadał on znanemu typowi zależności między zmiennymi. Następnie dopasowujemy dokładniej model do danych treningowych, szacując na ich podstawie parametry funkcji danego typu. Mówimy wówczas o modelach dedukcyjnych, regresji liniowej lub określonego typu, krzywoliniowej.

W sytuacji, gdy nie potrafimy z góry określić kształtu funkcji  $f$ , możemy zastosować modele, których postać funkcjonalna okazuje się na tyle bogata, że po oszacowaniu parametrów są one w stanie przybliżać dowolny kształt zależności między zmiennymi (tzw. właściwość uniwersalnej aproksymacji). Mówimy wówczas o modelach indukcyjnych, regresji nieliniowej. Modele są jednak znacznie trudniejsze w zastosowaniu,

wymagają także większych ilości danych oraz większych nakładów obliczeniowych na ich budowę.

Najczęściej wykorzystywane są liniowe modele regresyjne, w których przyjmujemy, że funkcja modelu  $f$  ma charakter funkcji liniowej, czyli:

$$y = b_0 + b_1x_1 + \dots + b_nx_n$$

Zależności liniowe (albo zbliżone do liniowych) bardzo często występują w świecie rzeczywistym.

Parametry funkcji  $f$ ,  $b_0, b_1, \dots, b_n$ , określane są na ogół przy użyciu procedury uczenia statystycznego opartej na metodzie najmniejszych kwadratów. Wykorzystywany jest w niej (to ogólna zasada dla modeli regresji) zbiór treningowy danych, składający się z zestawu wzorców wartości zmiennych wejściowych i odpowiadającej im wartości zmiennej wyjściowej:

$$D = \{(x_{1k}, \dots, x_{nk}), y_k\}, k = 1, \dots, N$$

Przykładowy zbiór danych treningowych do budowy modelu regresji przedstawiony został po lewej stronie tabeli 3.14. Naszym celem jest w tym przypadku utworzenie modelu predykcji o równaniu  $y = b_0 + b_1x_1 + b_2x_2$ , gdzie zmienną wyjściową (zależną, objaśnianą)  $y$  jest zużycie paliwa, a zmiennymi wejściowymi (niezależnymi, objaśniającymi) będą z kolei:

- $x_1$  – liczba kilometrów do przejechania w mieście,
- $x_2$  – liczba kilometrów do przejechania poza miastem.

Na zużycie paliwa wpływa z pewnością znacznie więcej zjawisk, np. pogoda, sytuacja na drodze, ewentualne zdarzenia nadzwyczajne itp. Niektórych czynników nawet nie jesteśmy przecież w stanie przewidzieć z góry – trudno powiedzieć, czy podczas przewozu przytrafi się (lub nie) wypadek, czy utkniemy w korku. Chcemy więc prognozować przeciętne zużycie paliwa, jakiego możemy się spodziewać przy przejechaniu określonej liczby kilometrów w mieście i poza miastem.

Do znalezienia parametrów  $b_0, b_1, b_2$  wykorzystać chcemy dane historyczne o zużyciu paliwa w trakcie np. realizacji dostaw do klientów w przeszłości. Jeśli spojrzymy na nasz zbiór treningowy w tabeli 3.14, to w każdym wierszu widzimy zarejestrowaną dla pewnej dostawy liczbę przejechanych kilometrów w mieście i poza miastem oraz odpowiadające im, rzeczywiście zarejestrowane kiedyś dla tej dostawy, zużycie paliwa.

Przykładowo: informacje w pierwszym wierszu mówią nam, że kiedyś wykonana została dostawa, w czasie której przejechano 200 kilometrów w mieście i 500 poza miastem. Zużyto wówczas 44 litry paliwa.

**Tabela 3.14.** Przykładowy zbiór treningowy dla modelu regresji i odpowiadający mu układ równań – służący do wyznaczenia parametrów modelu liniowego

Zużycie paliwa	Miasto (km)	Teren (km)	
44	200	500	$b_0 + 200b_1 + 500b_2 = 44$
50	250	450	$b_0 + 250b_1 + 450b_2 = 50$
53	620	100	$b_0 + 620b_1 + 100b_2 = 53$
54	400	280	$b_0 + 400b_1 + 280b_2 = 54$
52	350	300	$b_0 + 350b_1 + 300b_2 = 52$
...	...	...	...

**Źródło:** opracowanie własne.

Dla każdej obserwacji (wiersza) danych ze zbioru treningowego można utworzyć równanie liniowe ze względu na niewiadome. Równania odpowiadające poszczególnym wzorcom danych z tabeli 3.14 zostały przedstawione po jej prawej stronie. Aby znaleźć parametry funkcji liniowej modelu, musimy więc rozwiązać taki właśnie układ równań liniowych.

Niestety układ ten nie ma rozwiązań. Weźmy jednak pod uwagę, że niczego innego się nie spodziewaliśmy. Gdyby układ miał rozwiązanie, oznaczałoby to, że udało nam się znaleźć dokładną, funkcyjną zależność między zmiennymi wejściowymi i zmienną wyjściową modelu. A przecież wiemy, że tak nie jest. Nie uwzględniamy wszystkich czynników wpływających za zużycie paliwa, tylko te najważniejsze.

Aby określić przybliżoną zależność między zmiennymi, znajdujemy wartości niewiadomych  $b_0$ ,  $b_1$ ,  $b_2$  stanowiące najlepsze przybliżone rozwiązanie układu równań. W tym celu zazwyczaj stosuje się metodę najmniejszych kwadratów.

Dla zbioru treningowego  $D = \{(x_{1k}, x_{2k}), y_k\}$ ,  $k = 1, \dots, N$ , błąd  $k$ -tego równania dla określonych wartości niewiadomych  $b_0$ ,  $b_1$ ,  $b_2$  jest równy różnicy między prawą i lewą stroną tego równania  $y_k - (b_0 + b_1x_{1k} + b_2x_{2k})$ . Błąd kwadratowy całego układu jest równy sumie kwadratów błędów poszczególnych równań:

$$E(b_0, b_1, b_2) = \sum_{k=1}^N (y_k - (b_0 + b_1 x_{1k} + b_2 x_{2k}))^2$$

Naszym celem jest więc znalezienie takich wartości współczynników  $b_0, b_1, b_2$ , dla których błąd całego układu równań  $E(b_0, b_1, b_2)$  będzie jak najmniejszy.

Istnieje kilka algorytmów szczegółowych pozwalających na znalezienie minimum błędu. Najczęściej stosowane podejście polega na rozwiązaniu tzw. układu równań normalnych. Błąd układu  $E(b_0, b_1, b_2)$  jest funkcją kwadratową – ma więc jednoznaczne minimum, które można wyznaczyć z warunku zerowania się pochodnej. Warunek ten pozwala nam na utworzenie (w tym przypadku) trzech równań, tzw. równań normalnych:

$$\frac{\partial E(b_0, b_1, b_2)}{\partial b_i} = 0, \quad i = 1, 2, 3$$

Ponieważ pochodną funkcji kwadratowej jest funkcja liniowa, więc równania normalne mają charakter liniowy. Układ ten stosunkowo łatwo zatem rozwiązać, co pozwala na wyznaczenie optymalnych wartości parametrów modelu liniowego  $b_0, b_1, b_2$ .

### 3.4.4.3. Klasyfikatory

Zagadnienie klasyfikacji, z punktu widzenia jego ogólnej koncepcji, jest w zasadzie podobne do zadania regresji. Również tu staramy się wyjaśnić zachowanie niektórych zjawisk poprzez wpływ innych. Jednak w przeciwieństwie do zagadnień regresji, w zadaniu klasyfikacji zmienna wyjściowa jest dyskretna. Zbiór jej wartości okazuje się skończony, a wartości te określone są w formie pewnych kategorii (nominalnych lub porządkowych). Ta, wydawałoby się niewielka, różnica w sformułowaniu problemu powoduje, że metody stosowane do tworzenia modeli klasyfikacyjnych są nieco odmienne od metod regresji.

W zadaniu klasyfikacji mamy do czynienia z obiektami (rzeczami, zjawiskami lub sytuacjami) opisanymi przy pomocy wielu zmiennych wejściowych  $x_1, \dots, x_n$ , z których każda reprezentuje pewną cechę lub

właściwość tych obiektów. Dla uproszczenia zapiszmy je w formie wielowymiarowej zmiennej wejściowej  $\mathbf{x} = (x_1, \dots, x_n)$ . Naszym celem będzie znalezienie modelu (odwzorowania) klasyfikatora  $f$ , przypisującego obiektowi (opisanemu konkretnym zestawem wartości zmiennych wejściowych  $\mathbf{x}$ ) jedną z kategorii (klas) zmiennej wyjściowej  $y$  z pewnego zbioru  $C = \{c_1, \dots, c_m\}$ .

O klasyfikacji wspominaliśmy już w punkcie 3.3.3.4, omawiając decyzje klasyfikacyjne. Wymieniano tam szereg przykładów zastosowania klasyfikacji we wspomaganiu decyzji, m.in. zadanie kwalifikacji wniosku kredytowego – określonego przy pomocy szeregu zmiennych (cech wniosku) opisujących sytuację finansową wnioskodawcy, jego historię kredytową, charakterystyki zabezpieczenia i inne ewentualne informacje – do jednej z dwóch kategorii:  $C = \{\text{„przysnać kredyt”, „odrzuć wniosek”}\}$ . Innym, już wspomnianym, przykładem klasyfikacji jest decyzja giełdowa: na podstawie cen, indeksów i wskaźników giełdowych należy przyporządkować sytuację do jednej z decyzji typu: „sprzedaj”, „kup”, „wstrzymaj się”.

Klasyfikatory budowane mogą być przy pomocy różnych metod, również o charakterze symboliczno-logicznym. W przypadku klasyfikacji jako zadania eksploracji danych mówi się o tworzeniu klasyfikatorów na podstawie danych. Aby zbudować tego rodzaju klasyfikator, potrzebny jest zestaw przykładów treningowych klasyfikowanych obiektów, dla których znamy poprawną kategorię wyjściową, jaka każdemu z tych obiektów powinna zostać przypisana. Naszym celem będzie stworzenie modelu jak najtrafniej przyporządkowującego do właściwej kategorii nowe, niewystępujące w zbiorze treningowym obiekty.

Dla potrzeb eksploracji danych stworzonych zostało wiele metod, umożliwiających tworzenie klasyfikatorów. Oparte są one na bardzo różnorodnych podejściach i podstawach matematycznych. W książce o charakterze ogólnym, takiej jak ta, niemożliwe (a nawet niecelowe) jest przedstawienie nawet w miarę pełnego katalogu tych metod. Czytelników odsyłamy do pozycji poświęconych zagadnieniom eksploracji danych [Larose, 2006 i 2008; Morzy, 2019], także tych bardziej zaawansowanych [Cichosz, 2000; Bengio, Courville i Goodfellow, 2018]. Obecnie przyjrzymy się tylko kilku wybranym metodom, które mają przedstawić różnorodność istniejących podejść do

zagadnienia klasyfikacji oraz podkreślić pewne istotne elementy wiążące się z tym zadaniem.

Modele klasyfikacyjne oparte są na pewnej funkcji dyskryminacyjnej, rozróżniającej obiekty należące do odmiennych klas. Szereg metod klasyfikacji stara się modelować tę funkcję bezpośrednio, na podstawie przykładów treningowych. Jedno z najważniejszych podejść polega na zastosowaniu liniowej funkcji dyskryminacyjnej.

Aby pokazać pewne ważne właściwości modeli klasyfikacyjnych, rozważmy zadanie klasyfikacji binarnej, tzn. o dwóch klasach wyjściowych (oznaczymy je przez  $A$  oraz  $B$ ). Zbiór danych treningowych dla tego rodzaju zagadnienia zazwyczaj ma postać:

$$D = \{(x_{1k}, \dots, x_{nk}), y_k\}, y_k \in \{-1, 1\}, k = 1, \dots, N$$

gdzie:

- $\mathbf{x}_k = (x_{1k}, \dots, x_{nk})$  – zestaw cech wejściowych pewnego  $k$ -tego przykładowego obiektu,
- $y_k = 1$  – oznacza, że obiekt ten należy do klasy  $A$ ,
- $y_k = -1$  – oznacza, że obiekt należy do klasy  $B$ .

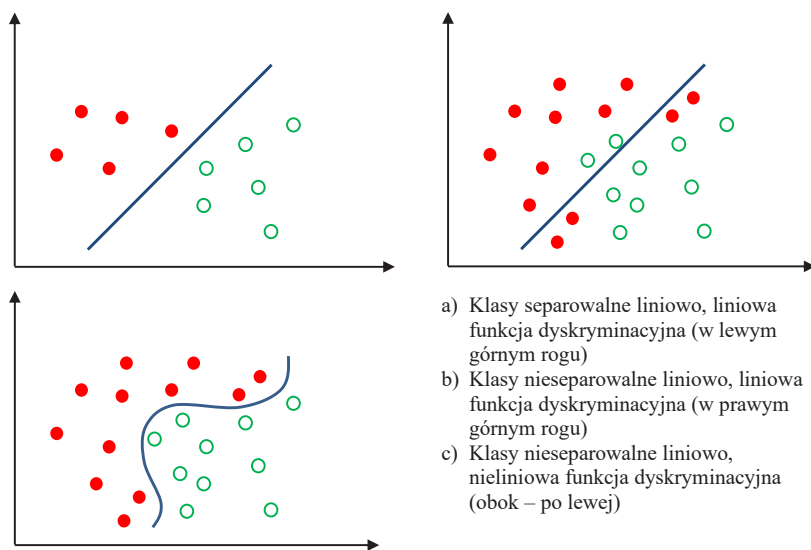
Budujemy model liniowy (korzystając np. z metody najmniejszych kwadratów, opisaną w poprzednim punkcie):

$$y = b_0 + b_1 x_1 + \dots + b_n x_n$$

Reguła klasyfikacyjna dla nowego obiektu, opisanego zestawem cech wejściowych  $\mathbf{x} = (x_1, \dots, x_n)$ , polega na obliczeniu wartości wyjścia modelu  $y = b_0 + b_1 x_1 + \dots + b_n x_n$  dla tego nowego obiektu, a następnie:

- jeżeli  $y > 0$  – przypisujemy nowy obiekt do klasy  $A$ .
- jeżeli  $y < 0$  – przypisujemy nowy obiekt do klasy  $B$ .

Zwróćmy uwagę, że funkcja dyskryminacyjna modelu rozdziela (separuje) obiekty należące do klasy  $A$  oraz do klasy  $B$  w przestrzeni wejściowej. Wszystkie obiekty, dla których wektor cech wejściowych  $\mathbf{x}$  daje dodatnie wartości funkcji dyskryminacyjnej modelu, zaliczane będą do jednej klasy, ujemne – do drugiej. Innymi słowy, funkcja dyskryminacyjna wyznacza pewne obszary w przestrzeni wejść, związane z poszczególnymi klasami.



- a) Klasy separowalne liniowo, liniowa funkcja dyskryminacyjna (w lewym górnym rogu)
- b) Klasy nieseparowalne liniowo, liniowa funkcja dyskryminacyjna (w prawym górnym rogu)
- c) Klasy nieseparowalne liniowo, nieliniowa funkcja dyskryminacyjna (obok – po lewej)

**Rysunek 3.10.** Separowalność klas i funkcja dyskryminacyjna modelu

**Źródło:** opracowanie własne.

Model liniowy funkcji dyskryminacyjnej będzie działał poprawnie tylko wtedy, gdy wzorce danych reprezentujących obiekty różnych klas w przestrzeni wejść dadzą się rozdzielić przy pomocy hiperpłaszczyzny wyznaczonej przez tę funkcję liniową. Mówimy wtedy o zbiorach danych separowalnych liniowo. Przykład takiej sytuacji dla dwuwymiarowej przestrzeni wejściowej (dwóch cech wejściowych) przedstawiony został na rysunku 3.10 (podpunkt *a*).

Oczywiście, jeżeli zbiory danych dla poszczególnych klas nie są separowalne liniowo, zastosowanie liniowej funkcji dyskryminacyjnej skutkuje błędami klasyfikacji (rysunek 3.10 podpunkt *b*), wynikającymi z niewłaściwych założeń co do kształtu funkcji separującej. Będą to błędy dodatkowe, niewynikające z komponentu losowego związanego z modelowanymi zjawiskami. W sytuacji, gdy dodatkowe pogorszenie jakości modelu jest dla nas akceptowalne, można zastosować tego rodzaju przybliżenie w analizie. Jeśli jednak nie godzimy się na zwiększony błąd modelu, niezbędne staje się zastosowanie klasyfikatorów o nieliniowej funkcji dyskryminacyjnej (rysunek 3.10 podpunkt *c*), które potrafią separować dowolne obszary dla poszczególnych klas. W tym celu

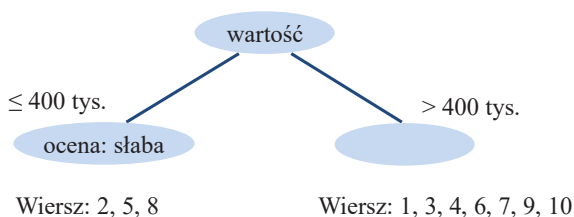


wykorzystane mogą zostać np. oparte na nieliniowej regresji modele neuronowe lub inne nieliniowe techniki modelowania.

Pamiętajmy, że wszystkie klasyfikatory definiują funkcję dyskryminacyjną, separującą poszczególne klasy. Natomiast nie w każdej metodzie modeluje się ją bezpośrednio w sposób jawny. Do popularnych modeli klasyfikacyjnych należą drzewa decyzyjne (klasyfikacyjno-decyzyjne), o których mówiliśmy już w punkcie 3.3.3.4.

Klasyfikatory w formie drzewa decyzyjnego mogą być również budowane na podstawie danych. Istnieje cały szereg wykorzystywanych w tym celu metod eksploracji danych. W zasadzie wszystkie one opierają się na pewnym wspólnym schemacie działania, różniąc się między sobą przede wszystkim sposobem jego implementacji. Przyjrzyjmy się ogólnej idei tego schematu na wybranym przykładzie.

wniosek	lokalizacja	wartość (tys.)	ocena
1	zła	850	średnia
2	średnia	400	słaba
3	średnia	1600	dobra
4	dobra	800	dobra
5	zła	350	słaba
6	dobra	2150	dobra
7	dobra	900	dobra
8	średnia	300	słaba
9	średnia	1100	średnia
10	zła	1800	średnia



**Rysunek 3.11.** Generowanie drzewa decyzyjnego na podstawie danych

**Źródło:** opracowanie własne.

W punkcie 3.3.3.4 na rysunku 3.3 przedstawione zostało przykładowe drzewo decyzyjne, służące do oceny nieruchomości (z użyciem kategorii: słaba, średnia, dobra) na potrzeby decyzji kredytowej. Klasyfikator ten wykorzystuje dwie cechy wejściowe: lokalizacja nieruchomości (ocenia-

na w formie kategorii: zła, średnia, dobra) oraz wartość nieruchomości (w tysiącach złotych). Zbiór danych treningowych do budowy tego rodzaju drzewa miałby więc postać:

$$D = \{(lokalizacja_k, wartosc_k), ocena_k\}, k = 1, \dots, N$$

gdzie:  $ocena_k \in \{\text{słaba, średnia, dobra}\}$ .

Przykładowy zbiór treningowy przedstawiony został na rysunku 3.11.

Algorytmy budowy drzew decyzyjnych wykonują szereg kroków związanych z analizą kolejnych węzłów drzewa. Obejmują one kilka standardowych operacji. W pierwszym kroku określana jest zmienna, którą będzie reprezentował dany węzeł, oraz podział zbioru jej wartości na wychodzące z niego gałęzie. Następnie zbiór przykładów treningowych rozdzielony zostaje na podzbiory odpowiadające poszczególnym wartościom tej zmiennej i gałęziom wychodzącym z węzła. Dalej algorytm przystępuje do analizy podwęzłów drzewa w poszczególnych gałęziach, wykorzystując podzbiory danych treningowych z nimi związane.

Podział następuje przy tym tylko w sytuacji, gdy nie jest spełnione kryterium zatrzymania algorytmu w danej gałęzi drzewa.

W przykładzie na rysunku 3.11 przyjmijmy, że w pierwszym węźle, korzeniu tworzonego drzewa, algorytm określił, iż będzie on reprezentował zmienną *wartość*, zaś jej wartości zostaną podzielone na dwa przedziały: poniżej i powyżej (lub równej) 400 tysięcy złotych. W pierwszej gałęzi, odpowiadającej warunkowi  $wartość \leq 400$ , znalazły się wiersze danych: 2, 5, 8, a w drugiej – pozostałe. Algorytm przejdzie do analizy węzła dla pierwszej gałęzi.

Ponieważ związane z tą gałęzią wzorce danych: {2, 5, 8} dają taką samą wartość oceny nieruchomości: „słaba”, najprawdopodobniej zadziała tu kryterium zatrzymania, które zakończy tworzenie drzewa w tej gałęzi. Algorytm wróci o poziom wyżej i przejdzie do analizy drugiej gałęzi.

Ponieważ przykłady dotyczące drugiej gałęzi: {1, 3, 4, 6, 7, 9, 10} prowadzą do różnych ocen nieruchomości, algorytm na podstawie analizy tego podzbioru wybierze zmienną, podzieli jej wartości, rozdzieli dane na podwęzły i przejdzie do ich analizy itp.

Algorytmy generowania drzew decyzyjnych, po stworzeniu wstępnego drzewa odpowiadającego zbiorowi przykładów treningowych, zwykle wykonują jeszcze operację tzw. przycinania drzewa (*pruning*). W jej

wyniku z drzewa usuwane są gałęzie nieistotne, tworzone przez dane zaszumione i odstające. Gałęzie te są charakterystyczne wyłącznie dla konkretnego zestawu danych i nie odzwierciedlają zależności ogólnych z modelowanej dziedziny. Pozostawienie ich upośledzałoby zdolności generalizacji (klasyfikacji nowych obiektów, niewystępujących w danych treningowych) klasyfikatora.

Szczegółowe algorytmy generowania drzew decyzyjnych na podstawie danych różnią się między sobą przede wszystkim implementacją wyżej wymienionych operacji: kryteriami i sposobem wykonywania podziału, warunkami zatrzymania algorytmu oraz metodami przycinania drzewa. Wśród najważniejszych metod z tego zakresu wymienić można: algorytmy CART, ID3, C4.5, C5.

Klasyfikatory w postaci drzew decyzyjnych modelują funkcję dyskryminacyjną separującą klasy wyjściowe w formie zestawu logicznych warunków, które muszą być spełnione, aby można było zakwalifikować obiekt do danej kategorii. Jeszcze inne podejście wykorzystuje metoda  $k$ -najbliższych sąsiadów ( $k$ -Nearest Neighbors –  $k$ -NN).

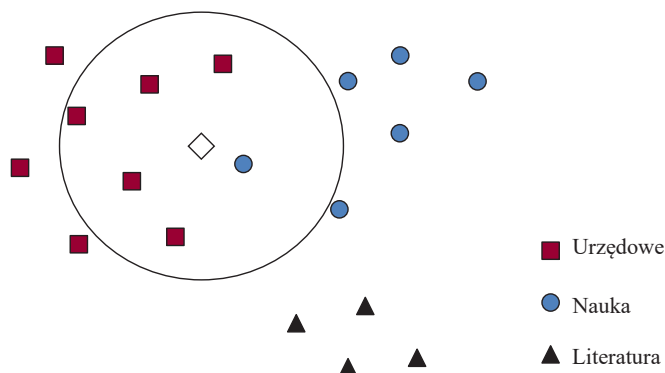
Metoda  $k$ -najbliższych sąsiadów jest nieskomplikowaną, ale bardzo skuteczną metodą klasyfikacji, opartą na historii (przykładach). Wyniki otrzymywane przy pomocy klasyfikatorów tego rodzaju są zwykle bardzo dobre; wszystkie obliczenia wykonuje się w zasadzie w trakcie klasyfikacji nowego obiektu, a nie uczenia – czasami takie podejście określa się jako tzw. leniwe uczenie (*lazy learning*).

Samą procedurę klasyfikacji zreasumować można następująco: dla nowego obiektu, opisanego zestawem wartości cech wejściowych  $\mathbf{x} = (x_{k1}, \dots, x_{kn})$ :

- 1) obliczamy podobieństwo wzorca  $\mathbf{x}$  do każdego wektora cech przykładów ze zbioru treningowego  $D$ ;
- 2) znajdujemy  $k$  najbardziej podobnych przykładów treningowych i analizujemy kategorie, do których one należą;
- 3) przypisujemy nowy obiekt  $\mathbf{x}$  do kategorii reprezentowanej przez większość z najbardziej podobnych przykładów;
- 4) alternatywnie możemy obliczyć rozkład prawdopodobieństwa  $P(c_i|\mathbf{x}) = n_i/k$ , gdzie  $n_i$  jest liczbą przykładów treningowych (spośród  $k$  najbliższych) należących do klasy  $c_i$ .

Prosty przykład klasyfikacji metodą  $k$ -najbliższych sąsiadów przedstawiony został na rysunku 3.12. Naszym celem jest w tym przypadku klasyfikacja dokumentów (opisanych pewnymi cechami) do jednej

z trzech kategorii: urzędowe, naukowe oraz literackie. Przyjmijmy, że przykłady treningowe dla poszczególnych klas rozkładają się tak, jak przedstawiliśmy to na rysunku. Klasyfikujemy nowy dokument, oznaczony niewypełnionym rombem. Przyjmujemy liczbę najbliższych sąsiadów  $k = 6$ .



**Rysunek 3.12.** Przykład klasyfikacji metodą  $k$ -najbliższych sąsiadów

**Źródło:** opracowanie własne.

Jak widać, spośród sześciu dokumentów – najbardziej podobnych pod względem cech do dokumentu nowego (znajdują się one w zaznaczonym kole – jego otoczeniu) – pięć należy do kategorii *Urzędowe*, zaś jeden do kategorii *Nauka*. Można więc zakwalifikować nowy obiekt do kategorii dokumentów urzędowych lub wyznaczyć prawdopodobieństwa jego przynależności do poszczególnych klas:

$$P(\text{Urzędowe}) = 5/6, P(\text{Nauka}) = 1/6, P(\text{Literatura}) = 0$$

Podstawowym parametrem metody jest liczba sąsiadów  $k$ . Należy go dobrać dla konkretnego przypadku. Przyjęcie zbyt małej liczby najbliższych sąsiadów spowodowałoby, że o wyniku klasyfikacji mogłyby decydować specyficzne, obciążone szumem pojedyncze przykłady. Przykładowo: dla  $k = 1$  (tzw. metoda najbliższego sąsiada) na naszym rysunku najbliższym sąsiadem będzie dokument o charakterze naukowym. Zakwalifikowanie nowego obiektu do tej kategorii wydaje się jednak błędne – położony jest on raczej w obszarze elementów urzędowych.

Z kolei przyjęcie zbyt dużej wartości  $k$  spowodowałoby, że promień zaznaczonego koła – otoczenia najbardziej podobnych dokumentów – mógłby zbyt mocno wzrosnąć. O wyniku klasyfikacji decydowałyby obiekty położone daleko od nowo klasyfikowanego – czyli mało do niego podobne.

Jak widać, metoda  $k$ -najbliższych sąsiadów modeluje funkcję dyskryminacyjną lokalnie, w najbliższym otoczeniu nowo klasyfikowanego obiektu, korzystając z najbardziej podobnych przykładów treningowych. Klasyfikator tego rodzaju potrafi rozdzielać zbiory danych separowane nawet bardzo złożonymi, nieliniowymi zależnościami. Wyniki otrzymywane przy jego pomocy są często bardzo dobre. Aby poprawnie działać, wymaga on jednakże stosunkowo dużych zbiorów danych treningowych. Ponadto proces klasyfikacji nowego obiektu jest dosyć długi – konieczne staje się tu przeanalizowanie całego zbioru treningowego, co okazuje się często nieakceptowalne, zwłaszcza w przypadku, kiedy musimy szybko poklasyfikować wiele nowych obiektów.

#### **3.4.4.4. Analiza skupień (grupowanie danych)**

Grupowanie (analiza skupień, segmentacja) – to proces rozdziału zbioru obiektów (opisanego wielowymiarowymi wektorami wartości poszczególnych atrybutów – cech) na klasy (grupy, klastry, segmenty) obiektów, przy czym:

- obiekty w tej samej grupie powinny być jak najbardziej do siebie podobne,
- obiekty należące do różnych grup powinny jak najbardziej się między sobą różnić.

W przeciwieństwie do klasyfikacji grupowanie jest procesem nienadzorowanym (wykorzystuje uczenie się bez nadzoru). Zauważmy, że zarówno w przypadku regresji, jak i klasyfikacji zbiory treningowe składały się z par danych obejmujących pewne obserwowane w przeszłości wejścia modelu oraz odpowiadającą im wartość wyjściową. Algorytmy uczące działają tak, by dla wejść treningowych model dawał wyniki jak najbliższe treningowym wyjściom. Dane treningowe dla zadania grupowania nie obejmują predefiniowanych wartości wyjściowych i składają się tylko z wzorców opisujących grupowane obiekty, tj.  $D = \{(x_{k1}, \dots, x_{kn})\}, k = 1, \dots, N$ . Algorytm uczący ma przeanalizować powiązania między obiektami w surowym zbiorze danych.

Grupowanie danych służy automatycznej identyfikacji naturalnych skupień (segmentacji) rozmaitych obiektów (rzeczy zjawisk, interakcji). Jako typowe przykłady zastosowań grupowania danych wymienić można [Sharda, Delen i Turban, 2017]:

- identyfikację naturalnych segmentów klientów;
- identyfikację reguł przypisywania nowych przypadków do klas w celach targetowania/diagnostyki;
- dostarczanie charakterystyki, definicji, oznakowania dla pewnych zbiorowości;
- zmniejszanie rozmiarów i złożoności problemów w przypadku innych metod eksploracji danych;
- identyfikację wartości odstających w określonej dziedzinie (np. wykrywanie rzadko występujących zdarzeń).

Większość metod grupowania wymaga użycia pewnej miary odległości do obliczenia bliskości (podobieństwa) między parami obiektów.

Z punktu widzenia zależności między właściwościami a skupieniami metody grupowania możemy podzielić na:

- grupowanie monotetyczne – wszystkie elementy grupy współdzielą te same właściwości; np. przy grupowaniu tematycznym dokumentów, wszystkim elementom należącym do danej grupy tematycznej można przypisać jedno słowo kluczowe (frazę), określające ich temat;
- grupowanie politetyczne – wszystkie elementy grupy współdzielą większość właściwości; np. dokumenty należące do danej grupy tematycznej współdzielą większość słów kluczowych, ale istnieją pewne różnice indywidualne; temat grupy musi być zazwyczaj opisany kilkoma słowami kluczowymi.

Z kolei z punktu widzenia zależności między obiektami i grupami wyróżniamy:

- grupowanie rozłączne (ostre, twarde) – każdy z obiektów przydzielany jest jednoznacznie do jednej tylko grupy; nie dopuszcza się tu niepewności przy klasyfikacji elementu do grupy oraz przynależności dokumentu do kilku grup;
- grupowanie miękkie (*soft*) – pozwala na przynależność obiektu do kilku klas; zazwyczaj metody tej grupy bazują na rozkładzie prawdopodobieństwa (lub w przypadku rozmytym – możliwościach), pozwalającym na klasyfikację dokumentu do poszczególnych klas.

Wreszcie, z punktu widzenia zależności między grupami, można mówić o grupowaniu:

- podziałowym – wszystkie skupienia są niezależne od siebie, zazwyczaj rozłączne (jak w grupowaniu twardym);
- hierarchicznym – skupienia tworzą hierarchię, w której skupienia na niższym poziomie zawierają się w skupieniach na poziomie wyższym.

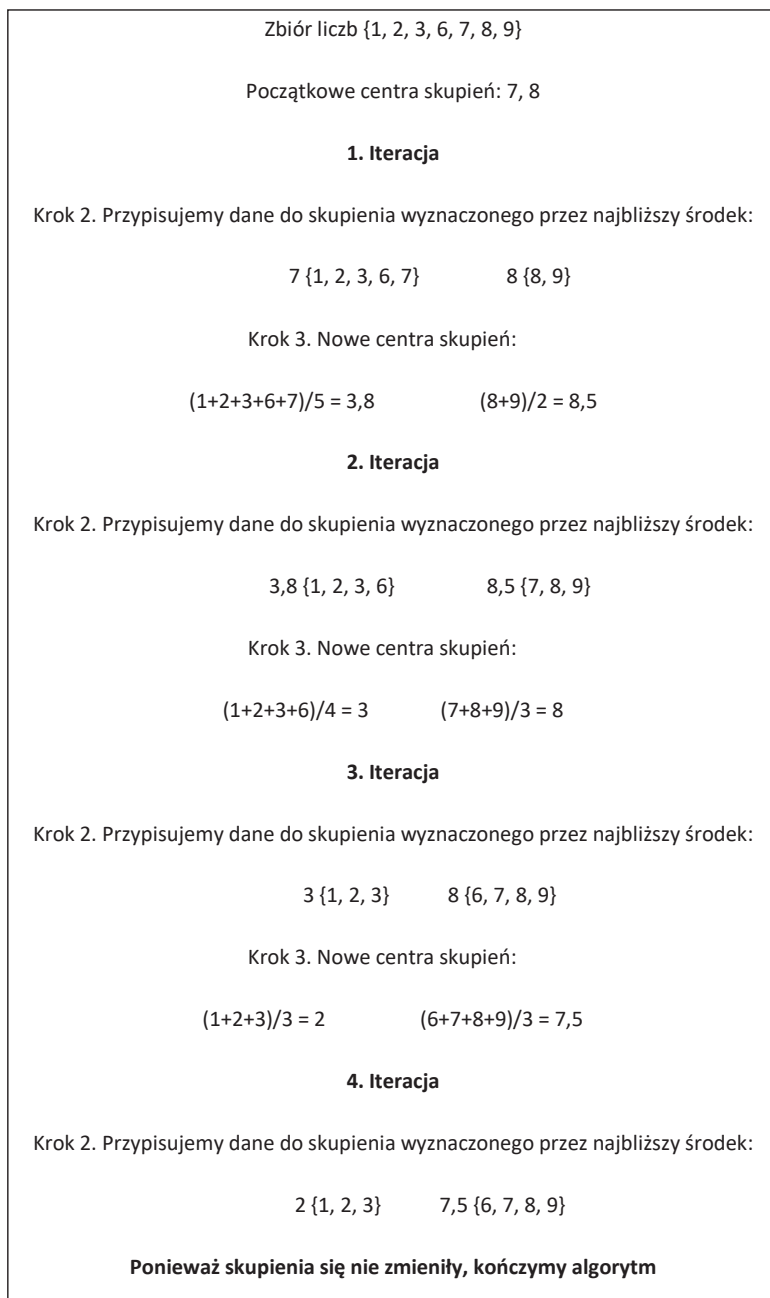
Te wydawałoby się dosyć złożone wymagania, jakie stawiamy metodom grupowania, często można osiągnąć zaskakująco prostymi środkami. Przykładowo: najpopularniejszym algorytmem grupowania podziałowego jest algorytm *c*-środków. Jego ogólny schemat można zreasumować w następujący sposób:

- krok 0: ustalamy z góry *c* – liczbę skupień, które chcemy otrzymać;
- krok 1: losowo generujemy *c* losowych wzorców jako początkowe centra skupień;
- krok 2: przypisujemy każdy wzorzec ze zbioru danych treningowych do najbliższego centrum skupienia;
- krok 3: ponownie obliczamy nowe centra skupień.

Ostatnie dwa kroki są powtarzane dopóki nie zostanie spełnione przyjęte kryterium zbieżności (zwykle przyjmuje się kryterium, aby przypisanie danych do skupień stało się stabilne, czyli nie zmieniało się w kolejnych iteracjach).

Jak widzimy, algorytm *c*-środków jest wyjątkowo prosty. Aby pokazać, że będzie on również skuteczny w działaniu, zastanówmy się nad prostym przykładem liczbowym, przedstawionym na rysunku 3.13. Dany jest zbiór liczb: {1, 2, 3, 6, 7, 8, 9}, który w sposób naturalny rozkłada się na dwa skupienia: {1, 2, 3} oraz {6, 7, 8, 9}. Analizując przykład na rysunku, można zauważyć, że algorytm *c*-środków pozwoli je uformować, nawet jeśli początkowe centra skupień będą dobrane zupełnie nietrafnie, np. 7, 8.

Metody grupowania podziałowego tworzą pewien konkretny podział zbioru danych na skupienia. Zazwyczaj musimy z góry założyć albo ich liczbę (tak jak w algorytmie *c*-środków), albo maksymalnie dopuszczalne rozmiary skupienia. Algorytmy grupowania hierarchicznego pozwalają uzyskać strukturę skupień, poczynawszy od dużej liczby drobnych (na najniższym poziomie oddzielne skupienia dla każdego obiektu) aż do niewielkiej liczby najbardziej ogólnych (na poziomie najwyższym jedno ogromne skupienie obejmujące wszystkie dane).



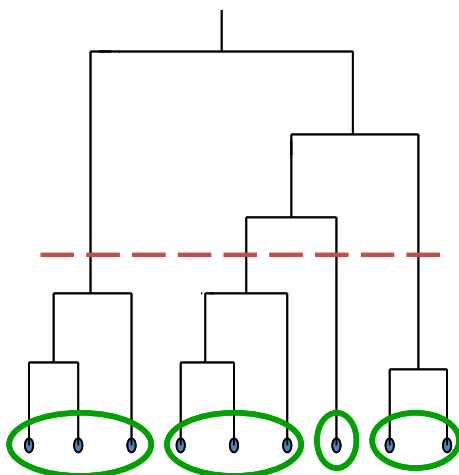
**Rysunek 3.13.** Przykład działania algorytmu c-środków

**Źródło:** opracowanie własne.



Wyniki działania algorytmu grupowania hierarchicznego prezentowane są zazwyczaj przy pomocy dendrogramu, czyli drzewa definiującego hierarchię skupień. Obejmuje ono informacje o zawieraniu się grup na każdym z poziomów. Przykład dendrogramu przedstawiony został na rysunku 3.14.

Jak widać, grupowanie hierarchiczne nie pokazuje nam podziału zbioru danych na jednym poziomie, ale zależności hierarchiczne występujące w całej strukturze skupień. Podział na skupienia na wybranym określonym poziomie otrzymać można poprzez obcięcie dendrogramu na tym poziomie. Przykładowo: na rysunku 3.14 przeprowadzony został (za pomocą przerywanej linii) podział na skupienia zaznaczone okręgami.



**Rysunek 3.14.** Przykład dendrogramu skupień (na poziomie obcięcia, wyznaczonego przez przerywaną poziomą linię, otrzymujemy podział na skupienia zaznaczone okręgami)

**Źródło:** opracowanie własne.

Można mówić o dwóch podejściach do grupowania hierarchicznego:

- metody rozdzielające (*top-down*) – rozpoczynają od jednego zawierającego cały zbiór danych skupienia, następnie iteracyjnie rozdzielają najmniej podobne obserwacje (obiekty) na osobne grupy, dopóki każdy obiekt nie znajdzie się w osobnym skupieniu;
- metody aglomeracyjne (*bottom-up*) – rozpoczynają od tylu grup, ile jest przykładów (obiektów) w zbiorze danych, a następnie iteracyjnie łączą je, tworząc coraz większe skupienia, aż do uformowania jednego wielkiego skupienia obejmującego wszystkie dane.

Częściej wykorzystywane są metody grupowania aglomeracyjnego. Podstawowym algorytmem grupowania hierarchicznego jest Hierarchiczne Grupowanie Aglomeracyjne (*Hierarchical Agglomerative Clustering* – HAC). Można go zreasumować następująco:

- rozpoczynamy od umieszczenia każdego wzorca danych w odrębnej grupie;
- powtarzamy kolejne operacje, dopóki nie pozostanie wyłącznie jedno skupienie:
  - wśród skupień, jakie uformowane zostały w bieżącym kroku, znajdujemy dwa skupienia najbardziej podobne do siebie (najbliższe sobie)  $c_i, c_j$ ;
  - zastępujemy  $c_i$  i  $c_j$  jednym nowym skupieniem  $c_i \cup c_j$ .

Można zastanawiać się, jak zdefiniować bliskość (podobieństwo) skupień wektorów danych. W przypadku pojedynczych obiektów, reprezentowanych przez pojedyncze wektory cech, stosujemy zazwyczaj oferowane przez matematykę miary odległości. Dwa obiekty są podobne, jeśli odległość ich wektorów cech jest niewielka. W przypadku skupień nie istnieje tak jednoznaczna metoda określania ich podobieństwa (odległości). Stosuje się kilka różnych podejść w tym zakresie:

- metoda pojedynczego połączenia (*single-link*) – podobieństwo dwóch skupień równe jest podobieństwu pary najbardziej podobnych (najbliższych) obiektów z obu grup; o podobieństwie skupień decyduje więc jedna wybrana para obiektów z obu z nich (pojedyncze połączenie);
- metoda całkowitego połączenia (*complete-link*) – podobieństwo dwóch skupień równe jest podobieństwu pary najmniej podobnych (najdalszych) obiektów z obu grup; o podobieństwie grup decydują wszystkie obiekty z obu skupień (całkowite połączenie);
- metoda centroidu – podobieństwo dwóch skupień równe jest podobieństwu ich centrów (prototypów);
- metoda średniego połączenia (*average-link*) – podobieństwo dwóch skupień równe jest średniemu podobieństwu wszystkich możliwych par obserwacji z obu grup.

Użycie różnych metod może przynosić nieco odmienne wyniki grupowania. Przykładowo: metoda pojedynczego połączenia daje zazwyczaj w efekcie długie i wąskie skupienia – z powodu efektu łańcucha. Metoda całkowitego połączenia tworzy z kolei bardziej skupione, sferyczne grupy (skupienia).

### 3.4.4.5. Analiza reguł asocjacyjnych

Analiza reguł asocjacyjnych jest dosyć popularną dziedziną eksploracji danych w biznesie. Znajduje ona interesujące powiązania (powinowactwa) między elementami (przedmiotami lub zdarzeniami). Czasami określa się ją również, od pierwszego z zastosowań, które legło u źródła rozwoju tej dziedziny, jako analizę koszyka produktów rynkowych.

Zadaniem analizy asocjacji jest przebadanie zbioru transakcji, z których każda składa się z zestawu (sekwencji) pewnej liczby elementów (*items*) współwystępujących w tej transakcji. Transakcjami tymi mogą być np. transakcje handlowe produktów zakupionych przez klienta w sklepie, ale niekoniecznie. Przez transakcje rozumieć będziemy ogólnie zbiory pewnych elementów (obiektów lub zdarzeń), np. wymieniane przez klientów zestawy cech najlepszego (dla nich) telewizora czy też sekwencje (ścieżki) kliknięć internautów na łącza w obrębie danej witryny itp. Celem analizy jest wykrycie najczęściej występujących powiązań między poszczególnymi elementami.

Jako przykłady częstych zastosowań eksploracji reguł asocjacyjnych w przedsiębiorstwie wymienić można [Sharda, Delen i Turban, 2017]:

- analizę transakcji sprzedaży – badając zestawy produktów zakupionych przez poszczególnych klientów, chcemy wykryć, które z elementów kupowane są najczęściej razem; uzyskane w ten sposób powiązania mogą zostać wykorzystane w celu poprawy rozmieszczenia produktów w sklepie lub magazynie czy też do zarządzania promocjami (np. unikanie jednoczesnych promocji na dwa najczęściej kupowane razem produkty) itp.;
- analizę transakcji kartami kredytowymi – analiza zestawów produktów zwykle kupowanych za pomocą określonej karty kredytowej może być istotnym orężem w walce z wykrywaniem nieuczciwych (lub co najmniej podejrzanych) transakcji dokonanych na numer tej karty;
- analizę usług bankowych – zestawy usług często wykorzystywanych razem przez klientów mogą być użyte do rekomendacji ich nowym klientom;
- analizę produktów ubezpieczeniowych – często wykorzystywane razem zestawy produktów ubezpieczeniowych mogą stanowić

podstawę do tworzenia standardowych planów ubezpieczeniowych w ofercie firmy, rekomendacji klientom dodatkowych produktów ubezpieczeniowych, a także do wykrywania nietypowych kombinacji roszczeń ubezpieczeniowych, będących oznaką oszustwa.

Oczywiście nie interesują nas wszystkie powiązania między elementami badanych transakcji. Wiele z takich związków ma charakter przypadkowy i incydentalny. Wiedza o takich relacjach nie byłaby specjalnie przydatna do podejmowania decyzji związanych w jakiś sposób z tymi transakcjami oraz ich elementami. Przez reguły asocjacyjne rozumiemy powiązania istotne, które mogą w miarę często występować również w przyszłości czy w innym, interesującym nas aspekcie.

Reguła asocjacyjna ma postać:  $X \Rightarrow Y$ , gdzie poprzednik  $X$  oraz następnik  $Y$  są rozłącznymi podzbiorami badanego zbioru elementów. Do oceny istotności danej asocjacji zazwyczaj wykorzystuje się dwie charakterystyki:

1) wsparcie (*support*) – jak często  $X$  i  $Y$  występują razem:

$$\text{wsparcie} = P(X \cap Y) = \frac{\text{liczba transakcji zawierających } X \text{ i } Y}{\text{liczba wszystkich transakcji}}$$

2) ufność (*confidence*) – jak często  $Y$  występuje po  $X$ :

$$\text{ufność} = P(Y|X) = \frac{P(X \cap Y)}{P(X)} = \frac{\text{liczba transakcji zawierających } X \text{ i } Y}{\text{liczba transakcji zawierających } X}$$

Najlepsze reguły asocjacyjne to takie, które mają zarówno wysokie wsparcie, jak i poziom ufności. Jakie w praktyce przyjąć wartości progowe dla tych mierników? To zależy od analizowanego zagadnienia. Przykładowo: jeśli chcemy ustalić, które produkty z supermarketu są kupowane razem, minimalny poziom wsparcia mógłby zostać przyjęty jako 20%, a minimalny poziom zaufania w wysokości 70%. Jednocześnie, przy wykrywaniu oszustw lub działalności terrorystycznej, należałoby obniżyć minimalny poziom wsparcia do wartości 1% lub mniej, ponieważ stosunkowo niewiele transakcji jest nieuczciwych lub związanych z terroryzmem [Larose, 2006].

Ogólny schemat wykrywania reguł asocjacyjnych obejmuje zazwyczaj dwa podstawowe kroki:

- 1) w pierwszym z nich należy znaleźć wszystkie tzw. częste zbiory elementów (*frequent itemsets*); przez częstość zbioru elementów rozumiemy liczbę transakcji zawierających ten zbiór, a częsty zbiór zdarzeń to taki, którego częstość przekracza pewien założony poziom  $\Phi$ ;
- 2) w drugim kroku z częstych zbiorów elementów generujemy reguły asocjacyjne spełniające minimalne poziomy wsparcia i ufności.

Podstawową metodą generowania reguł asocjacyjnych jest algorytm *A priori*. Najtrudniejszym problemem jest wygenerowanie wszystkich częstych zbiorów elementów – algorytm *A priori* stosuje w tej fazie podejście oddolne: częste zbiory rozszerzane są o jeden element naraz. Rozpoczyna się od wygenerowania wszystkich możliwych częstych zbiorów jednoelementowych, następnie rozmiar częstych zbiorów zwiększa się ze zbiorów jednoelementowych do zbiorów dwuelementowych, trzelementowych itp.

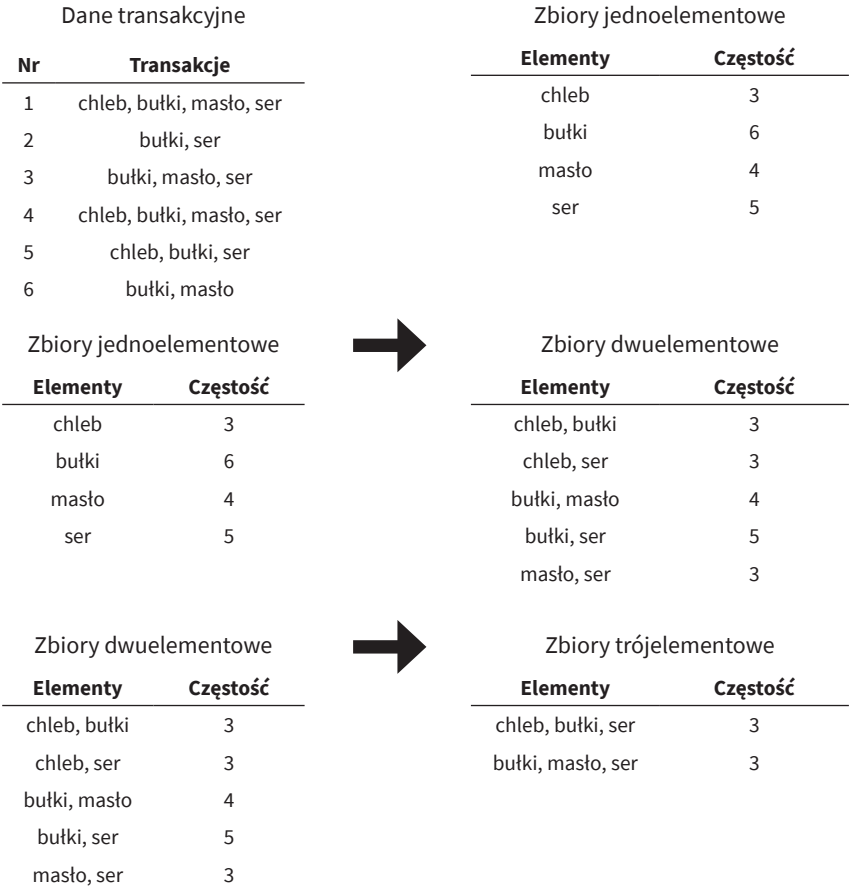
Do redukcji liczby analizowanych zbiorów elementów wykorzystuje się przy tym tzw. właściwość *A priori*: jeśli zbiór elementów  $Z$  nie jest częsty, wówczas dla dowolnego elementu  $A$ ,  $Z \cup A$  także nie będzie częsty.

Przykład sposobu tworzenia częstych zbiorów elementów przy pomocy algorytmu *A priori* przedstawiony został na rysunku 3.15. Zakładamy przy tym graniczny poziom częstości zbioru elementów  $\Phi = 3$ .

W pierwszym kroku ze zbioru danych transakcyjnych generujemy wszystkie częste zbiory jednoelementowe. W naszym przykładzie w jego skład wchodzi po prostu wszystkie elementy. Każdy z nich spełnia założony próg częstości  $\Phi = 3$ .

Drugi krok polega na utworzeniu wszystkich częstych zbiorów dwuelementowych. Tworzymy je na podstawie zbiorów jednoelementowych otrzymanych w poprzednim kroku. Bierzymy pierwszy zbiór jednoelementowy {chleb} i próbujemy rozszerzyć go o każdy możliwy dodatkowy element: {chleb, bułki}, {chleb, masło} i {chleb, ser}. Zauważmy, że próba rozszerzenia {chleb, masło} zakończyła się niepowodzeniem. Łatwo sprawdzić, że częstość takiego zbioru elementów wynosi 2 – jest więc mniejsza od założonego progu. Następnie próbujemy rozszerzyć zbiór jednoelementowy {bułki}. W tym przypadku mamy tylko dwa możliwe rozszerzenia: {bułki, masło} i {bułki, ser} – zbiór elementów {bułki,

chleb} był już rozważany wcześniej (kolejność elementów w zbiorze jest nieistotna). Oba te zbiory elementów są częste. Pozostaje ostatni możliwy zbiór dwuelementowy {masło, ser}. On także ma charakter częsty.



**Rysunek 3.15.** Przykład tworzenia częstych zbiorów elementów przy pomocy algorytmu *A priori* ( $\Phi = 3$ )

**Źródło:** opracowanie własne.

W trzecim kroku próbujemy rozszerzyć zbiory dwuelementowe do trójelementowych. Pierwszy zbiór dwuelementowy {chleb, butki} można rozszerzyć do {chleb, butki, ser}. Drugie możliwe rozszerzenie {chleb, masło, ser} ma zbyt niską częstość – 2. Zbioru {chleb, ser} nie można już rozszerzyć. Nie ma elementów położonych dalej niż {ser}, które moglibyśmy dołączyć – wszystkie poprzedzające były już w jakiś sposób rozważone

wcześniej. Przykładowo: zbiór {chleb, ser, bułki} przed chwilą analizowaliśmy. Jedyne rozszerzenie, jakie możliwe jest do uzyskania, to dodanie do zbioru {bułki, masło} elementu {ser}. I taki zbiór trójelementowy jest częsty. Zauważmy, że próbujemy rozszerzać tylko zbiory dwuelementowe powstałe w poprzednim kroku. Nie ma sensu analizować możliwości rozszerzenia odrzuconego wcześniej zbioru {chleb, masło}. Ponieważ nie był on częsty, jego rozszerzenia tym bardziej nie mogą być częste. Nie musimy ich w ogóle sprawdzać.

W tym miejscu algorytm *A priori* kończy tworzenie częstych zbiorów elementów, ponieważ nie da się utworzyć żadnego częstego zbioru czte-roelementowego.

Następny etap polega na przeanalizowaniu wszystkich znalezionych częstych zbiorów elementów i próbie utworzenia na ich podstawie reguł asocjacyjnych. Zwróćmy uwagę, że dany zbiór elementów może dać w wyniku kilka reguł asocjacyjnych. Weźmy pierwszy z trójelementowych zbiorów na rysunku 3.15, czyli zbiór {chleb, bułki, ser}. Przyjmijmy, że interesują nas reguły o wsparciu co najmniej 50% i ufności co najmniej 70%. Założmy również, że chcemy uzyskać reguły, które w następniku mają tylko jeden element (często przyjmowane założenie).

Zbiór elementów {chleb, bułki, ser} może potencjalnie dać trzy tego rodzaju reguły:

chleb, bułki $\Rightarrow$ ser	wsparcie = $3/6 = 50\%$	ufność = $3/3 = 100\%$
chleb, ser $\Rightarrow$ bułki	wsparcie = $3/6 = 50\%$	ufność = $3/3 = 100\%$
bułki, ser $\Rightarrow$ chleb	wsparcie = $3/6 = 50\%$	ufność = $3/5 = 60\%$

Jak widać, reguła trzecia ma zbyt niską ufność, więc zostanie odrzucona.

### 3.4.5. Eksploracja tekstu i sieci Internet

#### 3.4.5.1. Eksploracja tekstu (informacji)

Podobnie jak w przypadku eksploracji danych, modele eksploracji tekstu mają za zadanie poszukiwanie nowych i użytecznych wzorców. Różnica polega na naturze poddawanych procesowi analizy zasobów informacyjnych.

Źródła danych mają ściśle wyznaczoną strukturę, np. relacyjne bazy danych, pliki rekordów itp. W tego rodzaju źródłach zasoby informacyjne podzielone są na określone elementy o znanym, jednorodnym znaczeniu. Semantyka każdego elementu danych jest ściśle określona, czyli znany jest schemat tej struktury. Wiadomo, czego każdy z elementów danych dotyczy – przykładowo: w relacyjnej bazie w tabeli o nazwie *Studenti* w kolumnie *Imię* muszą znajdować się imiona studentów, a nie, powiedzmy, nazwy tradycyjnych potraw regionalnych.

Strukturalny charakter danych jest bardzo istotny dla sprawnego ich przetwarzania przez programy komputerowe. Teksty to informacje przeznaczone dla człowieka. Ludzie potrafią przetwarzać informację niestukturalną, mającą postać ogólnie nazywaną „dokumentem”. Dokumenty to przede wszystkim teksty, strony webowe i multimedia (grafika, dźwięk, film). Nie mają one ścisłej struktury semantycznej ani określonego schematu przechowywanej informacji. Dokument to zwykle spory zasób różnorodnych informacji, pozbawiony uporządkowanego podziału na jednorodne semantycznie, stosunkowo proste elementy danych.

Niemniej jednak szacuje się, że nawet do 85–90% wszystkich danych w organizacjach posiada nieuporządkowaną formę dokumentów. Zasoby te podwajają się przy tym co osiemnaście miesięcy. Wykorzystanie źródeł informacji niestukturalnej nie jest więc opcją, ale niezbędnym elementem utrzymania przewagi konkurencyjnej [Sharda, Delen i Turban, 2017].

Przez eksplorację tekstu rozumie się więc półautomatyczny proces pozyskiwania wiedzy z nieustrukturyzowanych źródeł danych. Do najważniejszych zadań eksploracji tekstów należą m.in.:

- wyszukiwanie informacji – wyszukiwanie dokumentów tekstowych związanych z daną potrzebą informacyjną użytkownika, określaną zazwyczaj w postaci zestawu słów kluczowych;
- klasyfikacja tematyczna dokumentów tekstowych (katalogowanie) – przypisywanie dokumentu do jednej (bądź kilku) grup tematycznych, np. w celu łatwiejszego przeglądania kolekcji dokumentów;
- filtrowanie dokumentów – specjalny rodzaj klasyfikacji dokumentów tekstowych zazwyczaj do jednej z dwóch kategorii, np. spam/nie spam, dokument związany z określonym tematem/niezwiązany, dokument, którym może być zainteresowany dany użytkownik/niezainteresowany itp.;



- wspomaganie przeglądania w obszernych, złożonych przestrzeniach informacyjnych – organizowanie skojarzeniowe zbiorów informacji (np. z wykorzystaniem algorytmów grupowania dokumentów), tworzenie rozwiązań ułatwiających nawigację w tego rodzaju kolekcjach (generowanie podpowiedzi skrótów, połączeń skojarzeniowych między częściami dokumentów, wizualizacja kolekcji dokumentów itp.);
- ekstrakcja wiedzy i informacji – rozpoznawanie w tekście określonych nazwanych elementów (*named entity recognition*), wyszukiwanie ich i wyodrębnianie;
- wnioskowanie na podstawie bazy informacji – systemy odpowiedzi na zapytania, rozumowania opartego na przypadkach; znajdowanie dokumentów tekstowych zawierających informacje związane z pytaniem użytkownika, analiza ich i ekstrakcja niezbędnych informacji oraz tworzenie na ich podstawie konkretnej odpowiedzi;
- sumaryzacja dokumentów i zbiorów dokumentów – tworzenie streszczeń (tekstowych lub graficznych) zawartości pojedynczych dokumentów tekstowych, grup dokumentów oraz całych kolekcji;
- generowanie powiązań między dokumentami – znajdowanie powiązanych ze sobą fragmentów informacji w różnych dokumentach i tworzenie między nimi łączy hipertekstowych;
- automatyczne tłumaczenie – automatyczne generowanie tekstów odpowiadających danemu dokumentowi w innych, wybranych językach.

Właściwie wszystkie te zadania dla ich jak najlepszej realizacji wymagają czegoś w rodzaju „zrozumienia” tekstu. Dokumenty tekstowe mają charakter niestrukuralny, dlatego ludzie bez problemu są w stanie zrozumieć ich treść – jest to natomiast ogromne wyzwanie dla systemów informatycznych. Ponadto język naturalny, w którym zapisana jest treść dokumentu, ma charakter nieprecyzyjny, słowa okazują się wieloznaczne, ich sens wynika często z kontekstu. Prawdziwe zrozumienie dokumentu wymaga również w wielu przypadkach znacznej wiedzy z danego obszaru.

Wszystkie te elementy powodują, że stworzenie programów komputerowych rozumiejących zawartość dokumentów z informacją niestrukuralną stanowi wciąż niezrealizowane zadanie, wymagające wielu dalszych prac łączących elementy sztucznej inteligencji, analizy języka naturalnego, lingwistyki obliczeniowej i szeregu pokrewnych dziedzin.

Obecnie stosowane metody należy uznać za stosunkowo początkowe stadium rozwoju eksploracji tekstów, a ich wyniki są nadal dalekie od satysfakcjonujących.

Pomimo tego, że okazują się niedoskonałe, metody eksploracji tekstów dostarczają istotnych dla organizacji informacji. W większości stosowanych obecnie algorytmów treść analizowanych dokumentów reprezentowana jest przez pewne przydzielone im zestawy odpowiednich słów kluczowych (tzw. termów). Proces przypisywania słów kluczowych dokumentowi nazywamy indeksowaniem – może mieć on charakter ręczny (indeksowanie przez człowieka-redaktora) lub automatyczny (na podstawie słów występujących w tekście).

W większości metod eksploracji tekstów podstawową strukturą służącą do reprezentacji zbioru dokumentów tekstowych (określanego często jako „kolekcja dokumentów” albo „korpus”) jest tzw. macierz termów-dokumentów [Markov i Larose, 2009; Liu, 2011; Jo, 2019]. Zawiera ona tyle wierszy, ile znajduje się wszystkich słów kluczowych w słowniku oraz tyle kolumn, ile mamy dokumentów w kolekcji. Element macierzy przyjmuje wartość 1, jeśli dany term (słowo kluczowe) występuje w opisie danego dokumentu, lub 0, gdy nie występuje. Macierz termów-dokumentów może mieć więc olbrzymie rozmiary i zazwyczaj przechowywana jest w postaci specjalnej struktury danych, nazywanej indeksem odwrotnym.

W tabeli 3.15 przedstawiony został przykład macierzy termów-dokumentów dla kolekcji dwunastu elementów, które zostały zaindeksowane w sumie trzynastoma słowami kluczowymi. Pierwszy dokument zaindeksowano czterema słowami kluczowymi: {*Blaster*, *Gwiazda*, *Nadprzestrzeń*, *Podróż*}. Reprezentująca go pierwsza kolumna ma więc postać: (1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0). Jedyne występują w wierszu pierwszym (ponieważ słowo kluczowe *Blaster* jest pierwsze w naszym słowniku), czwartym (słowo kluczowe *Gwiazda* znajduje się na czwartej pozycji), ósmym (*Nadprzestrzeń* znajduje się na ósmym miejscu) oraz jedenastym (to z kolei numer w słowniku terminu *Podróż*). W pozostałych wierszach pierwszej kolumny występują wartości zerowe. Podobnie drugi dokument zaindeksowany został termami: {*Gwiazda*, *Kowboj*, *Podróż*, *Rewolwerowiec*}. Reprezentująca go druga kolumna przyjmuje postać: (0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0). Analogiczna zasada obowiązuje w przypadku pozostałych dokumentów.

**Tabela 3.15.** Przykład obliczania podobieństwa nowego wzorca słów kluczowych do istniejących dokumentów

Term\Dokument	1	2	3	4	5	6	7	8	9	10	11	12	
<i>Blaster</i>	1	0	0	0	1	0	0	0	0	0	0	1	1
<i>Czarna dziura</i>	0	0	0	1	1	0	1	0	0	1	0	0	0
<i>Grawitacja</i>	0	0	1	1	0	1	1	0	0	1	0	0	0
<i>Gwiazda</i>	1	1	0	0	1	1	0	1	0	1	1	1	1
<i>Indianie</i>	0	0	1	0	0	0	0	0	1	0	1	0	0
<i>Kosmos</i>	0	0	0	1	1	1	1	1	0	1	1	1	1
<i>Kowboj</i>	0	1	1	0	0	0	0	0	1	0	0	1	1
<i>Nadprzestrzeń</i>	1	0	0	0	1	0	1	0	0	0	0	1	1
<i>Obserwacja</i>	0	0	1	1	0	1	0	1	0	0	0	0	0
<i>Planeta</i>	0	0	0	1	0	0	0	1	1	1	0	0	0
<i>Podróż</i>	1	1	0	0	0	1	0	1	1	0	1	0	0
<i>Rewolwerowiec</i>	0	1	1	0	0	0	0	0	1	0	1	0	0
<i>Teleskop</i>	0	0	0	1	0	1	1	0	0	1	1	1	1
	3	2	1	3	4	4	4	3	1	4	3	3	
	0,61	0,41	0,18	0,50	0,73	0,67	0,73	0,55	0,18	0,67	0,50	0,50	

**Źródło:** opracowanie własne.

Większość metod eksploracji tekstu wymaga konieczności obliczenia podobieństwa pewnego nowego wzorca słów kluczowych (reprezentującego np. zapytanie użytkownika lub inny dokument) do dokumentów przechowywanych w macierzy termów-dokumentów. Podobieństwo to określane jest na podstawie liczby słów kluczowych występujących jednocześnie (współwystępujących) w opisie (kolumnie) dokumentu i w nowym wzorcu. Najczęściej wykorzystuje się tzw. współczynnik podobieństwa kosinusoidalnego [Jo, 2019]. Jeżeli oznaczmy przez  $Q$  zbiór termów reprezentujących nowy wzorec, zaś przez  $D$  zbiór termów wykorzystanych w opisie dokumentu, to współczynnik kosinus jest równy liczbie słów kluczowych współwystępujących w dokumencie i w nowym wzorcu, znormalizowanej przez długość (liczbę słów kluczowych) opisu dokumentu i nowego wzorca:

$$\frac{|Q \cap D|}{\sqrt{|D|} \sqrt{|Q|}}$$

Przyjmijmy, że nowy wzorec, powiedzmy zapytanie użytkownika  $\mathbf{q}$ , złożony jest z następujących słów kluczowych: {*Blaster, Gwiazda, Kosmos, Kowboj, Nadprzestrzeń, Teleskop*}, co odpowiada wektorowi: (1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1). W tabeli 3.15 wektor ten przedstawiony został w osobnej kolumnie po prawej stronie. W osobnej dolnej kolumnie wyszczególniono liczby współwystępujących termów oraz podobieństwo każdego z dokumentów kolekcji  $\mathbf{d}_j$ ,  $j = 1, \dots, 12$ , do zadanego przez użytkownika zapytania.

W przypadku pierwszego dokumentu  $\mathbf{d}_1$  łatwo zauważyć, że występują w nim trzy termy obecne również w zapytaniu  $\mathbf{q}$  – {*Blaster, Gwiazda, Nadprzestrzeń*}. W zapytaniu  $\mathbf{q}$  znajduje się łącznie sześć słów kluczowych, w opisie dokumentu  $\mathbf{d}_1$  – cztery; tak więc ich podobieństwo jest równe:

$$\text{sim}(\mathbf{q}, \mathbf{d}_1) = \frac{3}{\sqrt{4}\sqrt{6}} \approx 0,61$$

Jak widać w tabeli 3.15, okazuje się to zgodne z podobieństwem pierwszego dokumentu (pierwsza komórka osobnej dolnej części tabeli).

Podobnie wygląda kwestia drugiego dokumentu  $\mathbf{d}_2$  – współwystępują w nim i w zapytaniu  $\mathbf{q}$  tylko dwa słowa kluczowe: {*Gwiazda, Kowboj*}. Wobec tego podobieństwo tego dokumentu do zapytania  $\mathbf{q}$  jest znacznie niższe (druga komórka osobnej dolnej części tabeli):

$$\text{sim}(\mathbf{q}, \mathbf{d}_2) = \frac{2}{\sqrt{4}\sqrt{6}} \approx 0,41$$

W analogiczny sposób wyznaczylibyśmy podobieństwa do zapytania  $\mathbf{q}$  wszystkich dokumentów  $\mathbf{d}_j$ ,  $j = 1, \dots, 12$ . Następnie możemy uporządkować dokumenty według obliczonych podobieństw i zwrócić taki ranking dokumentów jako wynik zapytania  $\mathbf{q}$ . W tabeli 3.15 widać, że najbardziej podobne są dokumenty 5 i 7, następnie dokumenty 6 i 10, potem 1 itp. W tej więc kolejności powinny one zostać przedstawione użytkownikowi, który wpisał zapytanie  $\mathbf{q}$  w wyszukiwarce.

Przedstawiona w tabeli 3.15 macierz termów/dokumentów ma charakter binarny (zero-jedynkowy). W praktyce zazwyczaj przyjmuje się,

że słowa kluczowe (termy) opisują zawartość dokumentu z pewną wagą, określającą ich znaczenie dla możliwości dyskryminacji różnych dokumentów. W takim przypadku elementy macierzy termów/dokumentów mają wartości rzeczywiste. Procedury obliczania podobieństwa między wzorcami słów kluczowych są jednak bardzo zbliżone do przedstawionych w przykładzie.

### **3.4.5.2. Eksploracja sieci Internet**

Internet to obecnie największe repozytorium danych na świecie. Jest on przy tym istotnym źródłem danych, dającym olbrzymie możliwości zbudowania przewagi konkurencyjnej na rynku. Należy jednak zwrócić uwagę, że Internet nie jest źródłem danych łatwym do wykorzystania.

Dane w sieci są rozproszone na miliardach serwerów, co już samo w sobie powoduje ogromne problemy. Brak jest przy tym koordynacji i uporządkowania zarówno struktury sieci, jak i jej zawartości. Internet podlega nieustannemu wzrostowi i ma charakter bardzo dynamiczny; tempo przyrostu ilości informacji jest naprawdę wysokie. Jednocześnie zasoby informacyjne w sieci mają charakter ulotny – wiele dokumentów zmienia się nagle lub znika (np. martwe linki).

Osobnym problemem są rozmiary Internetu oraz dostępnych w nim zasobów informacyjnych. Znajdują się tam miliardy dokumentów o rozproszonej lokalizacji. W dodatku sieć jest zasobem redundantnym. Wiele dokumentów powtarza się albo, co gorsza, prawie powtarza – w różnych wersjach. Brak przy tym zupełnie kontroli wersyjności – czasami nawet trudno ustalić, która wersja dokumentu jest bardziej aktualna.

Zasoby informacyjne w Internecie mają bardzo różnorodny charakter. W większości są to niestrukuralne dokumenty (tekst, strony HTML, dokumenty multimedialne), ale znajdziemy tu również źródła strukturalne (bazy danych). W wielu jednak przypadkach zawarte w nich dane strukturalne dostępne są za pośrednictwem niestrukuralnych stron HTML. Dosyć często spotykanym w Internecie rodzajem zasobów informacyjnych okazują się źródła pośrednie, częściowo strukturalne, takie jak: dokumenty XML albo opisane przy pomocy uporządkowanych metadanych dokumenty tekstowe czy multimedialne.

Problemem staje się także niejednorodność zasobów informacyjnych w sieci. Mogą mieć one różne formaty (czysty tekst, dokumenty PDF,

DOC itp.), być zapisane w wielu językach, z wykorzystaniem rozmaitych zbiorów znaków itp.

Odrębnym zagadnieniem jest jakość informacji w Internecie. Sieć to środowisko samopublikacyjne, panuje tu znaczna demokratyzacja publikacji. Rozproszony charakter generowania zawartości, jej łączenia między dokumentami, powoduje brak kontroli edytorskiej (stąd słaba jakość języka, wyglądu, rysunków) oraz rzetelności informacji – może być ona prawdziwa, fałszywa, czasem sprzeczna.

Wyżej wymienione cechy Internetu, a także wiele innych, związanych nie tylko z charakterem publikowanej w nim informacji, ale również ze sposobem działania samej sieci jako takiej i komunikowania się w niej komputerów, powodują, że jest on zasobem informacyjnym dającym ogromne możliwości, ale również stwarzającym problemy, zwłaszcza jeśli chodzi o jego wykorzystanie.

Zagadnienia eksploracji sieci Internet (*web mining*) dzielone są zwykle na trzy podstawowe grupy [Markov i Larose, 2009; Liu, 2011]:

- eksploracja zawartości sieci (*web content mining*),
- eksploracja struktury sieci (*web structure mining*),
- eksploracja wzorców wykorzystania sieci (*web usage mining*).

Eksploracja zawartości sieci wykorzystuje zazwyczaj opisane w poprzednim punkcie metody eksploracji tekstu do analizy zawartości dostępnych w Internecie niestrukturalnych dokumentów tekstowych, stron HTML itp. Przykładowo: typowe wyszukiwarki internetowe indeksują strony webowe przy pomocy zestawów słów kluczowych (termów), a następnie tworzą rankingi najbardziej podobnych dokumentów dla zapytań zadawanych przez użytkowników również w formie grup słów kluczowych.

Eksploracja struktury sieci ma dwa podstawowe zadania:

- wykrywanie wzorców połączeń między stronami w sieci, łąkami hipertekstowymi;
- eksploracja struktury samych dokumentów internetowych, wykrywanie tego, który fragment dokumentu, jaką informację zawiera, czego ona dotyczy itp.

Najważniejszym algorytmem z tej dziedziny jest PageRank, wykorzystywany w wyszukiwarce Google do tworzenia rankingu znalezionych dokumentów (obok oceny tematycznej dokumentu – podobieństwa do zapytania). Analizuje on strukturę łączy między stronami i wykorzystuje uzyskane informacje do oceny wiarygodności (tzw. autorytetu) danego dokumentu.

Z kolei eksploracja wzorców wykorzystania sieci zajmuje się przede wszystkim pozyskiwaniem informacji z danych generowanych przez użytkowników odwiedzających witryny oraz z wykonywanych przez nich operacji. Metody tej grupy analizują głównie:

- dzienniki (logi) serwerów webowych lub funkcjonujących na tych serwerach aplikacji obsługujących użytkowników; *cookies* przechowujące stan aplikacji po stronie klienta HTTP;
- dane na poziomie aplikacji, takie jak: profile użytkowników, charakterystyki ich zachowania i historie wykonywanych operacji;
- dane i metadane dotyczące samej witryny, takie jak: atrybuty stron, zawartości witryny.

Celem eksploracji jest zdobycie wiedzy na temat m.in. wzorców wykorzystania witryny, profili użytkowników, profili stron, profili wizyt, danych biznesowych na temat klientów itp. Do najważniejszych zadań eksploracji wzorców wykorzystania sieci należą [Sharda, Delen i Turban, 2017]:

- analiza użyteczności witryny (czyli w jaki sposób użytkownicy ją wykorzystują) – wyświetlanych stron (np. które wyświetlane są częściej, które rzadziej), czasu spędzanego w witrynie, ściąganych z niej treści, map kliknięć (rozkładu częstości kliknięć na poszczególne elementy witryny), ścieżek kliknięć itp.;
- analiza źródeł ruchu sieciowego (czyli skąd pochodzą użytkownicy odwiedzający witrynę) – z innych witryn linkujących do analizowanej lub z wyszukiwarek, użytkownicy bezpośredni (np. wchodzący za pomocą zakładek lub bezpośrednio wpisujący adres w przeglądarce) czy ruch powstały w wyniku kampanii promocyjnych (z bannerów, łączy promocyjnych, e-maili itp.);
- badanie profili użytkownika (czyli pewnych cech specyficznych, charakteryzujących danego użytkownika) – analiza słów kluczowych wykorzystywanych do przeszukiwania witryny, treści, z których on korzysta, danych geo-demograficznych, czasu dnia odwiedzin, profili stron docelowych;
- charakterystyka wizyt użytkowników – wykrywanie nowych użytkowników, użytkowników powracających, stopnia kompletności wypełniania formularzy, statystyki różnego rodzaju działań (np. zakupów, rejestracji, zgłoszeń, zapisów itp.), statystyki opuszczenia witryny (liczba odwiedzonych stron, z której strony nastąpiło wyjście itp.).

### 3.4.6. Analityka Big Data

Big Data to najnowsza grupa metod analizy danych wykorzystywanych we współczesnej infrastrukturze informatycznej. Ma ona zastosowanie wszędzie tam, gdzie dużej ilości danych cyfrowych towarzyszy potrzeba zdobywania nowych informacji lub wiedzy. Szczególnie ważne jest także zwiększenie się stopnia dostępności Internetu i usług świadczonych drogą elektroniczną, które są naturalnie dostosowane do korzystania z danych.

Tak więc o potrzebie zastosowania środowisk Big Data można mówić w sytuacji, gdy analiza przeprowadzana jest:

- dla dużej ilości danych – co uniemożliwia ich przetwarzanie przy użyciu tradycyjnych, szeroko dostępnych metod;
- dla źródeł danych o rozproszonym charakterze, często znajdujących się w sieci rozległej, a więc wysoce heterogenicznych, nie tylko o charakterze SQL, ale także nieustrukturyzowanych, o dużej zmienności itp.

Główne elementy systemu Big Data to znane nam już z wcześniej omówionych metod:

- techniki analizy danych, takie jak: uczenie maszynowe i niestrukturalne przetwarzanie danych,
- *Business Intelligence*, przetwarzanie danych w chmurze obliczeniowej i bazach danych,
- wizualizacje, takie jak: wykresy i inne formy prezentacji danych.

Natomiast wzmiankowane wyżej elementy środowiska danych powodują, że rozwiązania Big Data muszą odpowiadać również na szereg wyzwań z nimi związanych. Wśród nich, jako najważniejsze, można wymienić:

- kwestie rozmiaru danych – od środowiska Big Data oczekuje się możliwości przechwytywania, przechowywania i przetwarzania ogromnej ilości danych na czas; wymaga to specjalnego doboru i implementacji wykorzystywanych metod analitycznych;
- kwestie integracji danych – jak wspomniano, źródła danych do analiz Big Data leżą w sieci Internet, a nie w bazie danych organizacji; środowisko Big Data musi więc umożliwiać szybkie (i ekonomiczne) łączenie danych; dane muszą być często przetwarzane w trakcie ich przechwytywania (tzw. analiza strumieniowa);



- kwestie zarządzania danymi – ponieważ źródła danych do analiz Big Data leżą poza bazą danych organizacji, nie są więc przez nią kontrolowane; środowisko Big Data musi więc rozwiązywać takie kwestie, jak bezpieczeństwo, prywatność dostępu itp.

Środowiska Big Data muszą więc wykorzystywać metody przetwarzania i analizy danych, przy uwzględnieniu ogromnych rozmiarów danych i potężnych nakładów obliczeń niezbędnych do ich wykorzystania. Wymaga to użycia specjalnych struktur danych, takich jak np. macierze rzadkie, wykonania na nich operacji i przede wszystkim szerokiego zastosowania równoległości wykonywanych działań. Operacje na danych muszą być rozdzielane na współbieżne komponenty i rozpraszane na wielu komputerach. Przykładem tego typu narzędzia jest opracowane i spopularyzowane przez firmę Google środowisko MapReduce.

Drugim niezbędnym elementem środowiska Big Data jest platforma do przechowywania, integracji i zarządzania ogromnymi ilościami rozproszonych, nieuporządkowanych danych o różnych formatach i strukturach, również tych niestrukturalnych. Sztandarowym przykładem tego rodzaju środowiska jest Hadoop, będący częścią Apache Software Foundation.

## 3.5. Narzędzia sztucznej inteligencji

### 3.5.1. Pojęcie sztucznej inteligencji

Zdefiniowanie pojęcia „inteligencja” wydaje się bardzo proste i intuicyjne, niemniej jednak zawsze sprawiało ono problem specjalistom zajmującym się problematyką kognitywistyczną. Sami również nie jesteśmy w stanie zrozumieć, na czym polegają procesy przebiegające w naszej psychice, tym bardziej należy spodziewać się trudności ze zdefiniowaniem tych zjawisk u maszyn.

Jedynymi znanymi nam istotami inteligentnymi są ludzie, więc pojęcie inteligencji można omawiać tylko w odniesieniu do człowieka. Istnieją dwa podstawowe podejścia związane z tą koncepcją:

- definiowanie inteligencji w powiązaniu z procesami myślowymi i rozumowaniem człowieka, innymi słowy, „aby komputery myślały jak ludzie”;
- definiowanie inteligencji pod kątem zachowania, „aby komputery zachowywały się jak ludzie”.

W związku z pierwszym podejściem określa się zazwyczaj szereg cech, oznak myślenia inteligentnego. Wśród nich wymieniamy [Sharda, Delen i Turban 2017]:

- uczenie się lub rozumienie na podstawie doświadczenia;
- możliwość wykrycia niejednoznacznych lub sprzecznych komunikatów;
- szybkie i elastyczne dostosowywanie się do nowych sytuacji;
- rozwiązywanie problemów z wykorzystaniem podejścia opartego na rozumowaniu;
- radzenie sobie w skomplikowanych sytuacjach;
- rozumowanie i wnioskowanie w sposób racjonalny;
- stosowanie i pozyskiwanie wiedzy do manipulacji otoczeniem;
- rozpoznawanie względnego znaczenia różnych elementów danej sytuacji.

W drugim przypadku tworzone są pewne testy oceniające zachowanie maszyn (systemów informatycznych). Najbardziej chyba znanym jest test inteligencji sformułowany przez Alana Turinga – według matematyka komputer można uznać za inteligentny tylko wtedy, gdy osoba przeprowadzająca wywiad, „rozmawiając” zarówno z niewidocznym dla niej człowiekiem, jak i niewidocznym komputerem, nie jest w stanie ustalić, który z nich jest kim.

Sztuczną inteligencję moglibyśmy więc zdefiniować jako dziedzinę nauki, której celem jest stworzenie myślącej maszyny. W praktyce jednak obecne badania w tej dziedzinie, a już z pewnością w sferze zarządzania, stawiają sobie znacznie mniej ambitne zadania. Systemy tej kategorii mają więc nie tyle wykazywać się myśleniem, co po prostu rozwiązywać bądź wspomagać człowieka w rozwiązywaniu skomplikowanych problemów, z którymi nie można poradzić sobie przy pomocy klasycznych, algorytmicznych środków.

Z tego punktu widzenia i na potrzeby niniejszej książki zdefiniujemy więc sztuczną inteligencję jako kategorię systemów informatycznych starających się naśladować (albo przynajmniej czerpać z nich inspirację) sposoby rozwiązywania problemów stosowane przez człowieka.

Motywacją zdecydowanej większości współczesnych badań nad sztuczną inteligencją jest zatem nie tyle jakieś abstrakcyjne dążenie do inteligentnego działania systemu informatycznego, co po prostu konieczność skutecznego rozwiązywania coraz trudniejszych problemów pojawiających się w sferze zarządzania organizacją. Od systemów informatycznych,

asystujących początkowo przy stosunkowo prostych, rutynowych decyzjach, obecnie wymaga się wspomagania złożonych decyzji pozostających wcześniej wyłącznie w sferze ludzkiej. W tym celu sięga się zatem po nowe metody oparte na naśladownictwie tego typu działań.

Systemy sztucznej inteligencji imitują więc aktywność człowieka. Należy jednak wyraźnie stwierdzić, że w praktyce są to jedynie dosyć odległe inspiracje. Systemy inteligentne wykorzystują pragmatyczne techniki modelowania oparte m. in. na logice, statystyce i innych dziedzinach nauki. Pojęcie sztucznej inteligencji obejmuje cały szereg węższych dziedzin, wykorzystujących różne aspekty ludzkiej działalności. W praktyce najbardziej rozwiniętymi gałęziami są:

- systemy ekspertowe (systemy z bazą wiedzy),
- sztuczne sieci neuronowe (SSN),
- systemy z logiką rozmytą,
- algorytmy genetyczne (GA).

Pokrótkie je omówmy. Systemy ekspertowe (systemy z bazą wiedzy) starają się modelować wiedzę człowieka-eksperta i sposób jego wnioskowania na wysokim poziomie, przy wykorzystaniu podejść symbolicznych, opartych na logice. Używają one wiedzy zgromadzonej w bazie wiedzy, symulując proces rozumowania człowieka za pomocą podsystemu wnioskującego. Systemy te rozwiązują złożone problemy na podstawie analizy baz wiedzy, a nie realizacji określonego algorytmu, jak dzieje się to w przypadku programów tradycyjnych. Wiedza przechowywana w systemie ekspertowym może mieć zarówno charakter zależności wynikających z teorii danej dziedziny (np. w formie reguł) między poszczególnymi elementami doświadczenia eksperta lub uzyskanych poprzez analizę danych, jak również pewnych ogólnych relacji semantycznych, takich jak: generalizacja–specjalizacja (samochód jest maszyną, człowiek jest ssakiem itp.), część–całość (silnik to część samochodu, biurko to element wyposażenia pokoju itp.).

Sztuczne sieci neuronowe wzorują się na niskopoziomowym, biologicznym działaniu układu nerwowego. Starają się modelować sposób działania poszczególnych neuronów, uczenia i przechowywania wiedzy w mózgu człowieka. Generalnie mają one charakter raczej ilościowy niż symboliczny. Sztuczna sieć neuronowa składa się z wzajemnie połączonych prostych elementów przetwarzających informacje, zwanych neuronami, jednostkami lub węzłami. Połączeniom między elementami

przyporządkowane są współczynniki wagowe wyznaczające siłę powiązań i tworzące zbiór parametrów modelu. Cała wiedza sieci o sposobie rozwiązania danego problemu przechowywana jest w jej wewnętrznych odwzorowanych, definiowanych przez wartości wag – i może być przywołana w procesie reakcji na określony sygnał. Współczynniki wagowe są przydzielone albo wyznaczone w procesie treningu (uczenia sieci), zmierzającym do nauczenia SSN identyfikowania wzorców albo odwzorowania przekształceń.

Systemy z logiką rozmytą modelują nieprecyzyjne pojęcia, jakimi posługują się w swoim myśleniu ludzie. Można je uznać za pewien rodzaj pomostu między dwoma poprzednimi podejściami. Z jednej bowiem strony mogą być traktowane jako narzędzie do modelowania nieprecyzji w systemach ekspertowych, z drugiej zaś – interpretowane jako struktury zbliżone do sieci neuronowych. Pojęcie „zbioru rozmytego” jest uogólnieniem pojęcia zwykłego zbioru, polegającym na dopuszczeniu, aby jego funkcja przynależności przyjmowała obok stanów krańcowych (0, 1) również wartości pośrednie.

Algorytmy genetyczne (GA) to algorytmy do znajdowania optymalnych rozwiązań z wykorzystaniem symulacji mechanizmów ewolucji. Zwróćmy uwagę, że w sensie dosłownym nie do końca spełniają one przyjętą definicję sztucznej inteligencji. Próbuje bowiem naśladować nie inteligentne zachowanie człowieka, ale mechanizmy ewolucyjne. Dlatego wydziela je się czasami (obok innych podobnych metod, takich jak algorytmy mrówkowe, czy też rojowe) do równoległej grupy obliczeń naturalnych. Ponieważ jednak zadanie wyszukiwania rozwiązania jest istotnym aspektem inteligencji, a w zastosowaniach algorytmy genetyczne bardzo często łączy się z sieciami neuronowymi czy logiką rozmytą, dlatego ze względów praktycznych wielu badaczy traktuje „obliczenia naturalne” jako część sztucznej inteligencji.

W praktyce zastosowań metody sztucznej inteligencji, jak zresztą jasno wynika z powyższych rozważań, nie stanowią zupełnie wydzielonego i wyizolowanego tworu. Są one naturalnym rozwinięciem elementów środowiska analitycznego organizacji. Systemy ekspertowe stanowią zaawansowaną formę modeli decyzyjnych, dyskutowanych w punkcie 3.3 – podobnie algorytmy genetyczne. Sztuczne sieci neuronowe i systemy z logiką rozmytą na ogół stosowane są jako zaawansowane metody eksploracji danych, które zostały przedstawione w punkcie 3.4.

## 3.5.2. Systemy ekspertowe

### 3.5.2.1. Ogólna charakterystyka systemów ekspertowych

Systemy ekspertowe (*Expert Systems* – ES) to komputerowe systemy informacyjne, które próbują naśladować procesy rozumowania eksperta i zdobywać wiedzę w zakresie rozwiązywania określonych problemów. ES wykorzystują ją, aby osiągnąć wysoki poziom decyzji w wąsko zdefiniowanej dziedzinie problemowej.

Ekspert to osoba posiadająca specjalną wiedzę, zdolność osądu, doświadczenie i umiejętności, aby zastosować je w działaniu, w celu udzielania rzetelnych porad i rozwiązywania złożonych problemów w wąsko określonym obszarze [Sharda, Delen i Turban, 2017]. Ekspert wie, które fakty są ważne, a także rozumie i wyjaśnia zachodzące pomiędzy nimi relacje. Jest on w stanie rozpoznać i sformułować problem, szybko i poprawnie go rozwiązać, a także wyjaśnić rozwiązanie.

W związku z tym jako podstawowe cechy systemów ekspertowych można wymienić [Zieliński, 2000; Jabłoński i Bartkiewicz, 2006]:

- gromadzenie w systemie kompletnej wiedzy z danej dziedziny oraz możliwość jej ciągłej aktualizacji,
- umiejętność naśladowania sposobu rozumowania człowieka-eksperta, a co za tym idzie – oferowanie rad i wariantowanie decyzji,
- zdolność wyjaśniania przeprowadzonego toku rozumowania dla przyjętych rozwiązań.

Wiedza eksperta, a więc i systemu ekspertowego, powinna obejmować zazwyczaj szereg elementów dotyczących rozwiązywanego problemu. Wśród najważniejszych znajdują się [Sharda, Delen i Turban, 2017]:

- teorie dotyczące dziedziny problemu,
- zasady i procedury dotyczące ogólnej dziedziny problemu,
- heurystyki dotyczące postępowania w danej sytuacji problemowej,
- globalne strategie rozwiązywania tego rodzaju problemów,
- fakty dotyczące obszaru problemowego,
- metawiedza (tj. wiedza o wiedzy).

W skład typowego ES wchodzi: baza wiedzy, podsystem wnioskujący, mechanizmy objaśniające oraz podsystem gromadzenia wiedzy. Omówimy je w kolejnych punktach bieżącego podrozdziału.

### 3.5.2.2. Baza wiedzy systemu ekspertowego

Baza wiedzy przechowuje wiedzę pozyskaną na potrzeby systemu od eksperta, tj. podstawowe fakty charakterystyczne dla dziedziny, w której działa system, oraz opis związków i zależności między nimi, odzwierciedlający prawidła proceduralne i doświadczenie potrzebne w działaniu ES. Wiedza gromadzona w bazie wiedzy decyduje w dużym stopniu o możliwościach systemu ekspertowego [Sroka, 1994].

Jak zostało wspomniane, wiedza przechowywana w systemie ekspertowym może mieć charakter zależności przyczynowo-skutkowych między poszczególnymi elementami doświadczenia eksperta lub elementami uzyskanymi w wyniku analizy danych powiązań (asocjacji), a także pewnych ogólnych zależności semantycznych, takich jak: generalizacja–specjalizacja (samochód jest maszyną, człowiek jest ssakiem itp.), część–całość (silnik to część samochodu, biurko to element wyposażenia pokoju itp.).

Zazwyczaj wymienia się następujące kategorie wiedzy [Sharda, Delen i Turban, 2017]:

- fakty opisujące charakterystykę konkretnej sytuacji problemowej (lub podstawy faktów) i teorię obszaru problemowego (wiedza deklaratywna);
- heurystyki lub reguły (lub asocjacje) reprezentujące wiedzę ekspercką w celu rozwiązania określonych problemów w danej dziedzinie (wiedza proceduralna);
- relacje między pojęciami – uwzględniające słowa i inne symbole reprezentujące pojęcia, ich znaczenie i zasady użycia, relacje między nimi oraz metody manipulacji (wiedza semantyczna);
- informacje autobiograficzne lub eksperymentalne zorganizowane w formie przypadków (wiedza epizodyczna);
- ogólna wiedza na temat rozwiązywania problemów i podejmowania decyzji (lub metawiedza – o metodach wnioskowania w systemie).

Istnieje cały szereg metod reprezentacji wiedzy stosowanych w systemach ekspertowych. Do najważniejszych z nich zaliczyć można:

- metody logiki formalnej (rachunek zdań, rachunek predykatów – form zdaniowych);
- reguły i fakty;
- ramy;
- sieci semantyczne;
- scenariusze.

### 3.5.2.3. Podsystem wnioskujący systemu ekspertowego

Podsystem wnioskujący to „mózg” systemu ekspertowego. Zapewnia prowadzenie procesu rozumowania na podstawie wiedzy znajdującej się w bazie wiedzy i informacji wprowadzonych przez użytkownika podczas konsultacji – w celu sformułowania odpowiednich wniosków.

Proces poszukiwania stosownej wiedzy, niezbędnej dla rozwiązania stawianego problemu oraz dopasowanej do sytuacji określonej przez użytkownika, ma więc kluczowe znaczenie dla pracy systemu ekspertowego. Strategie stosowane w tym zakresie mogą wykorzystywać wiele wzorców, ale większość z nich zasadza się na jednej z dwóch podstawowych koncepcji [Zieliński, 2000]:

- wnioskowaniu progresywnym (wnioskowanie wprzód – *forward chaining*) startującym od znanych warunków i zmierzającym do określenia celu (tzw. sterowanie danymi);
- wnioskowaniu regresywnym (wnioskowanie wstecz – *backward chaining*) startującym z wymaganych celów i działającym wstecz do koniecznych warunków (tzw. sterowanie celami).

Jeśli po sformułowaniu wymagań dotyczących poziomu nakładów bezpieczeństwa i innych parametrów możliwych inwestycji analizujemy je i wybieramy te inwestycje, które naszym warunkom odpowiadają, to mamy do czynienia z rozumowaniem wprzód. Gdy odwrotnie – najpierw konstruujemy listę interesujących nas inwestycji, potem badamy, jakie każda z nich powoduje wymagania odnośnie poszczególnych parametrów, a następnie akceptujemy je lub nie – to rozumujemy wstecz.

Pamiętać również należy, że wnioskowanie w systemie ekspertowym jest zazwyczaj procesem wielokrokowym. Znajdująca się w bazie ES wiedza układa się zazwyczaj w pewne hierarchiczne struktury o charakterze drzewiastym. Poszczególne kroki wnioskowania rozszerzają wiedzę systemu o stany pewnych hipotez pośrednich. Dopiero po pewnej liczbie kroków sporządzona zostaje finalna ekspertyza. Również z tego punktu widzenia można mówić o dwóch podstawowych strategiach przeszukiwania bazy wiedzy.

Przeszukiwanie wszerz rozwija stan wyjściowy określony punktem początkowym, tj. generuje wszystkie możliwe hipotezy pośrednie w stosunku do stanu wyjściowego wiedzy, wyszukując wszystkie możliwe, znajdujące się w bazie asocjacje. Jeśli nie osiągniemy hipotez docelowych, to



generujemy następny poziom, kontynuując proces do osiągnięcia finalnej ekspertyzy. Strategia ta może być teoretycznie zastosowana do rozwiązania szeregu problemów. W praktyce jednak należy zauważyć, że liczba możliwości wygenerowanych na każdym poziomie rośnie wykładniczo w miarę kontynuacji poszukiwań, co powoduje wykładniczy wzrost czasu działania komputera, a czasem również wzrost obszaru pamięci.

Przeszukiwanie w głąb wybiera drogę, którą podąża się poprzez wszystkie poziomy do chwili osiągnięcia punktu docelowego. W danym kroku wnioskujemy więc na podstawie pojedynczej asocjacji w bazie wiedzy, określając stan pojedynczej nowej hipotezy. Jeśli nie jest to hipoteza docelowa, to w następnym kroku, również korzystając z nowej wiedzy, wyszukujemy w bazie kolejną asocjację. Proces ten powtarzamy aż do osiągnięcia finalnej ekspertyzy. W stosunku do poprzedniej strategii poszukiwanie w głąb wymaga mniejszych zasobów pamięci, ale wadą jej jest możliwość niezalezienia rozwiązania, nawet gdy ono istnieje.

Przeszukiwanie wszczepia pełną możliwość implementacji zarówno wnioskowania wprzód, jak i wstecz. Jak już wspomniano, jego wdrożenie przy dużych i skomplikowanych bazach wiedzy może sprawiać problemy na skutek efektu tzw. eksplozji kombinatorycznej, czyli lawinowego narastania liczby zależności, jakie musi sprawdzić system w każdym kolejnym kroku wnioskowania. Przeszukiwanie w głąb – przy konieczności osiągnięcia (podczas wnioskowania wprzód) lub potwierdzenia (podczas wnioskowania wstecz) większej liczby celów lub w przypadku ugrzęźnięcia w ślepej ścieżce wnioskowania, nieprowadzącej do żadnego z celów – może wymuszać konieczność kilkukrotnego powtórzenia procesu, z wykorzystaniem mechanizmów wycofywania się z poprzednio rozważanej ścieżki. Wymaga to zastosowania mieszanej strategii wnioskowania, wykonującej naprzemiennie fazy rozumowania wstecz i wprzód. Takie postępowanie nazywamy wnioskowaniem z nawrotami.

W skomplikowanych sytuacjach decyzyjnych, wspomaganych przez ES, wiedza eksperta może mieć charakter niepewny lub niekompletny. Wśród powodów tego stanu rzeczy wymieniamy:

- niepewność wiedzy wynikającą z faktu, że ekspert rozważa jedynie pewne aspekty problemu,
- niepewność danych opisujących rozważane zjawiska,
- niepełną informację będącą punktem wyjścia do podjęcia decyzji,
- stochastyczny charakter szeregu zagadnień,



- nieprecyzję językową,
- trudność w łączeniu opinii różnych ekspertów.

Reprezentacja wiedzy oraz mechanizmy wnioskowania oparte na klasycznej dwuwartościowej logice arystotelesowskiej nie dają narzędzi dla wyrażenia niepewności wiedzy wejściowej dostarczonej przez użytkownika oraz jej propagacji w kolejnych krokach wnioskowania za pośrednictwem niepewnych związków w bazie wiedzy. Bardziej złożone systemy ekspertowe wykorzystują więc mechanizmy wnioskowania oparte na logikach wielowartościowych. Do najbardziej znanych metod wyrażania niepewności w ES można zaliczyć:

- metody probabilistyczne oparte na prawdopodobieństwie warunkowym oraz twierdzeniu Bayesa,
- miary ufności oparte na teorii Dempstera-Shafera,
- czynniki pewności,
- logikę rozmytą.

### 3.5.2.4. Mechanizmy objaśniające

Ekspert-człowiek potrafi objaśniać swoje działania – podobne wymagania stawiać więc będziemy systemom ekspertowym. Procedura objaśniająca dotyczyć powinna nie tylko rekomendacji uzyskiwanych od systemu ekspertowego, ale również innych akcji przez niego podejmowanych. Objaśnienia czynią działanie systemu bardziej zrozumiałym, co ma istotne znaczenie dla zbudowania zaufania użytkownika do proponowanej decyzji: uważa się, że system, który nie potrafi wytłumaczyć użytkownikowi działania, jest niewiarygodny. Podsystem objaśniający pozwala ponadto na wyjaśnienie użytkownikowi nieoczekiwanych dla niego sytuacji, do których doprowadzić może konsultacja, jak również zapoznawać go ze strukturą przechowywanej w bazie wiedzy.

Zazwyczaj w systemach ekspertowych mamy do czynienia z dwoma typowymi kategoriami mechanizmów objaśniających:

- objaśnienia typu „jak” – pozwalają odpowiedzieć na pytanie o sposób osiągnięcia konkretnej konkluzji będącej wynikiem działania systemu ekspertowego; zazwyczaj ES w odpowiedzi podaje ścieżkę wnioskowania, czyli łańcuch asocjacji w bazie wiedzy wykorzystany do wywnioskowania danej konkluzji;
- objaśnienia typu „dlaczego” – pozwalają odpowiedzieć na pytanie o cel informacji, których podania system żąda od użytkownika.

### 3.5.2.5. Podsystem gromadzenia wiedzy

Podsystem gromadzenia wiedzy umożliwia aktualizację i rozszerzanie bazy wiedzy ES. Jego zadaniem jest pozyskiwanie (akwizycja) wiedzy z różnych źródeł, takich jak: eksperci, książki, filmy, komputerowe bazy danych, obrazki, mapy, obserwacje, działania itp. Osobę gromadzącą wiedzę od ekspertów (i z innych dostępnych źródeł) i przekształcającą tę wiedzę do postaci zgodnej ze sposobem jej przedstawienia (reprezentacji) nazywamy w bazie wiedzy inżynierem wiedzy.

Proces pozyskiwania wiedzy okazuje się niejednokrotnie najtrudniejszym etapem w tworzeniu ES i stanowi wąskie gardło w budowie dużych systemów ekspertowych, w których liczba stosowanych reguł przekracza kilkaset. W procesie tym podstawowym elementem jest tzw. artykulacja wiedzy, czyli przekazywanie jej przez eksperta inżynierowi wiedzy – to jego umiejętności współpracy z ekspertami decydują o jakości tworzonej bazy wiedzy.

W procesie pozyskiwania wiedzy na ogół wyróżnia się pięć etapów [Zieliński, 2000; Sharda, Delen i Turban, 2017; Flasiński, 2011]:

- identyfikacja – w tym etapie inżynier wiedzy i ekspert (eksperci) identyfikują problem i jego główne charakterystyki, tj. określają zakres projektowanego ES oraz celowość i możliwość jego opracowania;
- konceptualizacja – w tej fazie inżynier wiedzy i ekspert ustalają dostępność danych, koncepcję ES, wybierają sposób reprezentacji wiedzy i mechanizm propagacji jej niepewności;
- formalizacja – to etap, w którym inżynier wiedzy pozyskuje wiedzę od eksperta i organizuje ją zgodnie z przyjętym mechanizmem reprezentacji;
- implementacja – w tej fazie następuje zakodowanie wiedzy w bazie oraz stworzenie prototypowego systemu, weryfikuje się tu narzędzia wspomagające i w razie potrzeby dokonuje ich wymiany;
- testowanie na różnych przykładach – to etap oceny prototypowego systemu celem wykrycia błędów i słabych miejsc w bazie wiedzy.

Jak wspomniano, wiedza na potrzeby systemu ekspertowego pozyskiwana może być nie tylko od eksperta, ale również z innych źródeł. Rozwój informatyki spowodował obfite nagromadzenie w organizacjach różnorodnych danych, stanowiących cenne źródło dla menedżerów. Powoduje to, że w chwili obecnej coraz większe znaczenie mają metody

automatycznego pozyskiwania wiedzy, wykorzystujące różnorodne indukcyjne techniki uczenia maszynowego do eksploracji danych i ekstrakcji wiedzy. Obszerny przegląd tego typu metod czytelnik znajdzie w literaturze przedmiotu [Cichosz, 2000; Bengio, Courville i Goodfellow, 2018].

### **3.5.3. Sztuczne sieci neuronowe**

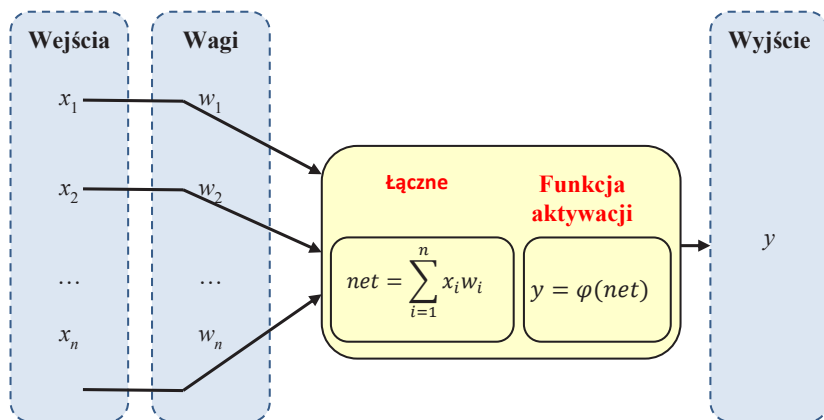
#### **3.5.3.1. Pojęcie i rodzaje sztucznych sieci neuronowych**

Sztuczna Sieć Neuronowa (SSN) złożona jest z wzajemnie połączonych prostych elementów przetwarzających informacje, zwanych neuronami, jednostkami lub węzłami. Połączeniom między elementami przyporządkowane są współczynniki wagowe wyznaczające siłę powiązań i tworzące zbiór parametrów modelu. Cała wiedza sieci o sposobie rozwiązania danego problemu przechowywana jest w jej wewnętrznych odwzorowaniach, definiowanych przez wartości wag – może być ona przywołana w procesie reakcji na określony sygnał (wzorzec wartości wejściowych). Współczynniki wagowe przydziela się lub wyznacza w procesie treningowym (uczenia sieci), zmierzającym do nauczenia SSN identyfikowania wzorców albo odwzorowania przekształceń [Zieliński, 2000; Jabłoński i Bartkiewicz, 2006].

Inspiracją do stworzenia tej klasy systemów była budowa mózgu ludzkiego. Ten skomplikowany, gromadzący i przetwarzający informacje układ w wielu dziedzinach działa lepiej i sprawniej od najlepszych nawet komputerów. Struktura sieci neuronowej oraz sposób rozwiązywania przez nią zadań przypominają zasadę działania systemu nerwowego. Zasada działania elementów przetwarzających sieci wzorowana jest na działaniu rzeczywistych neuronów. Sygnały wejściowe do komórki nerwowej wprowadzane są poprzez sieć wypustek, tzw. dendrytów. Wyjście stanowi jedno długie włókno, zwane aksonem, rozgałęziające się na końcu na szereg mniejszych włókienek. Łączą się one poprzez synapsy z dendrytami innych komórek nerwowych. Neuron przesyła więc sygnał do wielu komórek, ale jego wartość jest dla wszystkich taka sama.

Należy jednak zauważyć, że inspiracje biologiczne, aczkolwiek istotne, dotyczą jedynie ogólnych zasad funkcjonowania SSN. Zgodnie z tym, co powiedziano na temat naturalnych komórek nerwowych, przyjmuje się, że sztuczny neuron jest układem przetwarzającym o wielu wejściach

i jednym wyjściu. Schemat jego działania opiera się na modelu zaproponowanym w 1943 roku przez Warrena S. McCullocha i Waltera H. Pittsa [McCulloch i Pitts, 1943] – rysunek 3.16.



**Rysunek 3.16.** Model sztucznego neuronu

**Źródło:** opracowanie własne.

Neurony sztucznej sieci neuronowej połączone są między sobą. Podobnie jak w przypadku sieci naturalnej, sygnał wyjściowy sztucznego neuronu może być przekazywany na wejściach wielu innych neuronów, przy czym zazwyczaj połączeniom tym nadaje się pewną strukturę. Jednostki sieci neuronowej grupuje się w większe zespoły zwane warstwami. Struktura wewnętrzna, wraz z określeniem sposobu propagacji sygnału między neuronami, tworzy tzw. architekturę sieci neuronowej. Wyróżnić możemy jej trzy podstawowe rodzaje:

- sieci jednokierunkowe,
- sieci rekurencyjne,
- sieci komórkowe.

W przypadku sieci jednokierunkowych można powiedzieć, że struktura ich połączeń stanowi acykliczny graf skierowany. Sieci jednokierunkowe mają wyraźnie wyróżnione neurony wejściowe (przyjmujące informacje z zewnątrz) i wyjściowe (przesyłające przetworzoną informację na zewnątrz). Sygnał przekazywany jest zawsze do przodu: z warstwy wejściowej, poprzez jednostki ukryte, aż do warstwy wyjściowej – bez rekurencyjnych połączeń wstecznych. Dla dowolnego neuronu wartości wejść

nie zależą w żaden sposób (bezpośredni czy też pośredni) od jego stanu, czyli wartości wyjściowej. Typowym przykładem takiej sieci są wielowarstwowe sieci perceptronowe i sieci z funkcjami o bazie radialnej.

W przeciwieństwie do sieci jednokierunkowych w sieciach rekurencyjnych dopuszczamy występowanie cykli. Sygnał wyjściowy neuronu może więc bezpośrednio lub za pośrednictwem innych węzłów zostać przekazany na jego wejście. Dynamika działania tego typu sieci jest znacznie bardziej skomplikowana niż w przypadku sieci jednokierunkowych. W sieci rekurencyjnej jednokrotne pobudzenie sieci przez sygnał wejściowy powoduje wielokrotną aktywację wszystkich lub tylko części neuronów, w procesie tzw. relaksacji sieci. Dla jej poprawnego działania należy więc zapewnić dodatkowy warunek stabilności. Pobudzona sieć musi w skończonym czasie osiągać stan stabilny, w którym wartości neuronów dla danego wejścia pozostaną stałe. Dopiero wówczas możliwe staje się określenie wartości wyjścia. Przykładem sieci rekurencyjnej mogą być sieci Hopfielda.

W sieciach komórkowych wprowadza się dodatkowo pojęcie „sąsiedztwa węzłów”. Połączone między sobą są tu tylko jednostki znajdujące się w jego obrębie. Charakter tych powiązań może być różny, zależny od konkretnego przypadku. Przykładem tego typu sieci są neuronowe sieci komórkowe (*cellular neural networks*). Do tej kategorii zaliczyć można również sieci SOM Kohonena.

### 3.5.3.2. Charakterystyka zastosowań sieci neuronowych

Sieci neuronowe stanowią bezpośrednie rozwinięcie środowiska analitycznego w organizacji – dostarczają bowiem zaawansowanych narzędzi eksploracji danych, tekstów sieci. Obszary zastosowań SSN pokrywają się więc z zastosowaniami tych dziedzin [Zieliński, 2000]:

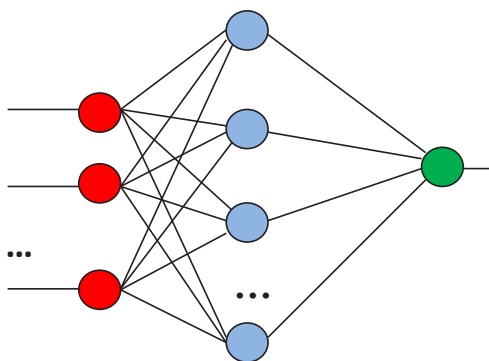
- przypominanie polegające na odzyskiwaniu (albo interpretowaniu) zmagazynowanych w SSN informacji, obliczaniu wyjścia dla danego wejścia;
- skojarzenie, które może być realizowane w następujących wariantach: skojarzenie uszkodzonego (zdeformowanego) wejścia (albo wywołania) z najbliższym przechowywanym wzorcem, skojarzenie między parą wzorców, diagnostyka, analiza;
- klasyfikacja, która realizowana jest poprzez podział zbioru wejściowego na klasy lub kategorie i skojarzenie każdego wejścia z kategorią;

- rozpoznawanie rozumiane jako klasyfikowanie wejścia pomimo tego, że nie odpowiada ono żadnemu z przechowywanych wzorców;
- estymacja, czyli realizacja następujących zadań: aproksymacja, interpolacja, filtrowanie, predykcja, prognozowanie;
- optymalizacja, w tym rozwiązywanie liniowych i nieliniowych równań;
- sterowanie realizowane inteligentnie bez konieczności opracowania modelu, oparte wyłącznie na doświadczeniu.

Sieci neuronowe wykorzystywane są więc w takich zadaniach eksploracji danych, jak prognozowanie (regresja), klasyfikacja czy grupowanie danych.

W zadaniach regresji (zob. punkt 3.4.4.2) wykorzystywane są najczęściej sieci jednokierunkowe (rzadziej rekurencyjne), przede wszystkim wielowarstwowe sieci perceptronowe i sieci z funkcjami o bazie radialnej. Przykład typowej sieci perceptronowej dla tego rodzaju zastosowań przedstawiony został na rysunku 3.17. Zazwyczaj składa się ona z trzech kolejno połączonych ze sobą warstw neuronów:

- warstwy wejściowej (elementy po lewej stronie rysunku 3.17) – której neurony reprezentują poszczególne zmienne wejściowe  $x_1, \dots, x_n$ ;
- warstwy ukrytej (elementy środkowe rysunku 3.17);
- warstwy wyjściowej (elementy po prawej stronie rysunku 3.17) – w większości przypadków złożonej z jednego neuronu reprezentującego zmienną wyjściową  $y$ .



**Rysunek 3.17.** Przykład typowej wielowarstwowej sieci perceptronowej dla zadania regresji

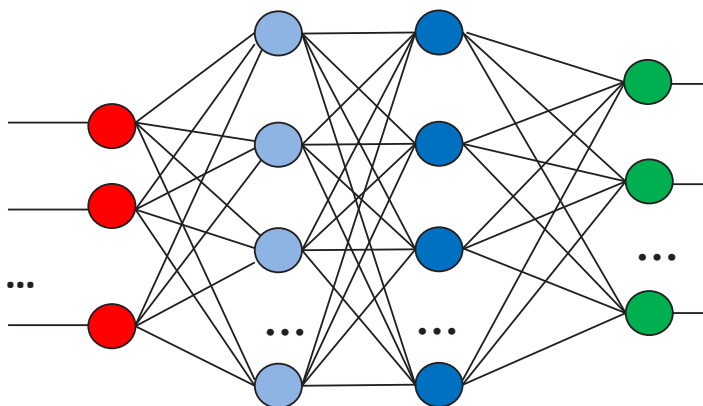
**Źródło:** opracowanie własne.

Sieci neuronowe w rodzaju przedstawionych na rysunku 3.17 są w stanie modelować dowolne zależności ciągłe (posiadają tzw. właściwość

uniwersalnej aproksymacji). Nie musimy więc z góry wybierać jednego modelu do odwzorowania np. zależności logistycznej, a innego dla zależności np. wielomianowej, co w wielu zadaniach jest bardzo trudne. Do modelowania obu tych funkcji może zostać wykorzystany ten sam model sieci neuronowej – pod warunkiem, że posiada on dostatecznie dużą liczbę neuronów w warstwie ukrytej.

W zadaniach klasyfikacji (zob. punkt 3.4.4.3) również typowymi stosowanymi modelami sieci neuronowych są wielowarstwowe perceptrony. Do tego zadania jednak stosuje się na ogół sieci z dwiema warstwami ukrytymi (rysunek 3.18). Warstwa wyjściowa sieci dla klasyfikacji z wieloma kategoriami wyjściowymi zawiera zazwyczaj kilka neuronów, z których każdy odpowiada jednej klasie wyjściowej.

Jak pokazywaliśmy na rysunku 3.10, klasyfikatory separują dane należące do różnych klas. Sieci neuronowe z dwiema (co najmniej) warstwami ukrytymi (rysunek 3.18) są w stanie separować dowolne zbiory danych. Podobnie jak w przypadku zadania regresji, nie musimy więc z góry decydować, jaki model sieci neuronowej powinien zostać zastosowany do modelowania określonej funkcji dyskryminującej.



**Rysunek 3.18.** Przykład typowej wielowarstwowej sieci perceptronowej dla zadania klasyfikacji

**Źródło:** opracowanie własne.

Innymi często wykorzystywanymi klasyfikatorami, zaliczanymi zazwyczaj do sieci neuronowych, są modele SVM (*Support Vector Machines*) – maszyny wektorów wspierających.

Sieci neuronowe stosowane są również w zadaniach analizy skupień (grupowania danych) – zob. punkt 3.4.4.4. Typowymi, stosowanymi w tej dziedzinie modelami neuronowymi są sieci Kohonena.

### 3.5.3.3. Tworzenie modeli sieci neuronowych

Sieci neuronowe należą do kategorii systemów uczących się. W ich działaniu wyodrębnić można dwie fazy: treningową oraz reakcji na określony bodziec zewnętrzny. Model rozwiązywania nie musi być znany *a priori* – budowany jest przez sieć w procesie uczenia, na podstawie dostarczonych danych treningowych. Rozwiązania analityczne oparte na sieciach neuronowych tworzone są więc zgodnie z ogólnym schematem budowy modeli eksploracji danych [Sharda, Delen i Turban, 2017]:

- należy zebrać, zorganizować i przygotować dane według zasad określonych w punkcie 3.4.4.1;
- podzielić dane na zbiór treningowy, walidacyjny i testowy;
- określić strukturę i architekturę sieci;
- wybrać algorytm uczenia sieci;
- przeprowadzić proces uczenia sieci;
- przetestować nauczoną sieć;
- wykorzystać sieć dla nowych, nieznanych wcześniej przypadków.

Jak widać, w systemach neuronowych istotną funkcję pełni proces uczenia sieci. Polega on na modyfikacji (najczęściej w procesie iteracyjnym) współczynników wagowych połączeń jej elementów. Ze względu na sposób prowadzenia treningu zazwyczaj wyróżnia się dwie grupy algorytmów uczących:

- uczenie nadzorowane (z nauczycielem) – dane treningowe zawierają zestaw sygnałów wejściowych sieci oraz poprawnych na nie reakcji; uczenie polega na takiej modyfikacji wag, aby rzeczywiste wyjścia były jak najbliższe wartościom pożądanym; jeżeli w czasie treningu nie prezentujemy sieci dokładnej wartości pożądanego wyjścia, a jedynie informację, czy reaguje ona prawidłowo, to mamy do czynienia ze specjalnym przypadkiem uczenia nadzorowanego – tzw. uczeniem ze wzmocnieniem;
- uczenie bez nadzoru – w procesie uczenia sieć neuronowa nie otrzymuje żadnej informacji na temat pożądaných reakcji; dane treningowe obejmują jedynie zbiór sygnałów wejściowych, a sieć ma za



zadanie samodzielnie zanalizować zależności i korelacje w zbiorze treningowym; tego typu sieci nazywamy samoorganizującymi (*self-organising networks*) lub autoasocjacyjnymi.

### 3.5.4. Systemy z logiką rozmytą

#### 3.5.4.1. Nieprecyzja lingwistyczna i zbiory rozmyte

Ponieważ metody sztucznej inteligencji próbują opisywać rzeczywistość, naśladując sposób rozumowania człowieka, nie mogą więc przejść do porządku dziennego wobec naturalnej nieprecyzji zjawisk rzeczywistego świata. Nieprecyzja ta może być związana z ich kształtem, położeniem, kolorem, powierzchnią lub nawet z semantyką opisującą, czym one są. Rozważmy następujące stwierdzenia [Zieliński, 2000]:

- Udział przedsiębiorstwa w rynku jest duży.
- Większość ekspertów stwierdziło, że transakcja jest bardzo ryzykowna.
- Cena towaru znacznie przekracza tysiąc złotych.
- Stan zapasów magazynowych jest prawie zerowy.
- Obroty na dzisiejszej sesji były znacznie wyższe niż wczoraj.
- W przyszłym roku poziom sprzedaży powinien wzrosnąć około 5%.
- W ciągu kilku następnych miesięcy inflacja powinna wyraźnie zmaleć.
- Współczynnik strat jest niewysoki.

Wszystkie te zdania opisują fakty i zjawiska zawierające poważny ładunek nieprecyzji. Jak bowiem zdefiniować takie określenia, jak: „duży”, „znacznie wyższe” czy też „bardzo ryzykowna”? Co to znaczy, że współczynnik strat jest niewysoki? Czy wynosi on 10%, czy może 15%? Ludzie są jednak w stanie interpretować powyższe stwierdzenia i wykorzystywać tak sformułowaną wiedzę do rozwiązywania stawianych przed nimi problemów. Zauważmy, że nieprecyzja ta nie ma nic wspólnego z niepewnością stwierdzeń – niepewność zdania wiąże się z faktem, że nie można definitywnie określić jego prawdziwości (lub fałszywości). Nieprecyzja dotyczy natomiast niemożności dostatecznie dokładnego określenia wartości wszystkich występujących w nim zmiennych. Zdanie precyzyjne może być więc niepewne, a kompletnie pewne sformułowanie bywa nieprecyzyjne.

ne. Stwierdzenie, że współczynnik strat jest niewysoki może być przecież całkowicie pewne. Problem polega na ustaleniu, coż to właściwie w tym przypadku znaczy „niewysoki” i jaka jest wartość współczynnika strat. Ten rodzaj nieprecyzji nazywany jest zwykle „rozmyciem” [tamże].

W przeciwieństwie do zagadnienia modelowania niepewności, w którym wykorzystywane są zwykle metody oparte na rachunku prawdopodobieństwa, konwencjonalne podejścia do reprezentacji wiedzy nie dostarczają odpowiednich środków dla przedstawienia pojęć rozmytych. W związku z tym Lotfi A. Zadeh wprowadził w 1965 roku pojęcie „zbioru rozmytego” (*fuzzy set*). W zbiorze rozmytym przynależność elementów określona jest przez pewną funkcję  $\mu_A$ , stwierdzającą, w jakim stopniu – z przedziału  $[0, 1]$  element przynależy do zbioru  $A$ . Funkcję  $\mu_A$  nazywamy funkcją przynależności zbioru  $A$ .

Pojęcie zbioru rozmytego jest uogólnieniem pojęcia zwykłego zbioru, polegającym na dopuszczeniu stopniowego przejścia między całkowitą przynależnością ( $\mu_A(x) = 1$ ) i nieprzynależnością ( $(x) = 0$ ) danego elementu  $x$ . Elementy mogą należeć do zbioru również w pewnym określonym stopniu, np. 0,5 czy 0,7.

Zwróćmy uwagę, że funkcja przynależności zbioru rozmytego nie ma nic wspólnego z funkcją rozkładu prawdopodobieństwa przynależności. Całkowita masa prawdopodobieństwa w rozkładzie musi być równa 1. Przykładowo: dwa różne elementy nie mogłyby należeć do zbioru z prawdopodobieństwem 1. Natomiast w przypadku zbioru rozmytego wiele elementów może się w nim znaleźć ze stopniem przynależności równym 1.

Zbiory rozmyte w wielu przypadkach mogą znacznie lepiej modelować pojęcia świata rzeczywistego niż tradycyjne zbiory, mocno związane z klasyczną logiką arystotelesowską, dopuszczającą tylko dwa stany rzeczywistości: „prawda” lub „fałsz”. Logika klasyczna już od dawna atakowana była jako zbyt uproszczona w stosunku do pełnej złożoności świata rzeczywistego. Wskazywał na to choćby już współczesny Arystotelesowi (IV wiek p.n.e.) grecki filozof Eubulides z Miletu. Stworzył on zestaw paradoksów wskazujących na niespójności w binarnym (arystotelesowskim) opisie świata. Przykładem może być tu choćby tzw. paradoks łysego człowieka:

- człowiek z głową pełną włosów w oczywisty sposób nie jest łysy;
- usunięcie pojedynczego włosa nie zmienia człowieka bez łysiny w łysiego.

Te dwa zdania opisujące pewien element rzeczywistości należy uznać za prawdziwe. A jednak oczywiste jest, że kontynuacja procesu usuwania pojedynczych włosów musi ostatecznie doprowadzić do powstania łysiny. Paradoks ten powstaje w wyniku próby użycia klasycznych narzędzi logicznych do opisanego pojęcia rozmytego. Nie jest bowiem możliwe określenie dokładnej ilości włosów, które są konieczne, aby określić człowieka jako nie-łysego.

Natomiast opis z wykorzystaniem logiki rozmytej nie zawiera już tego rodzaju sprzeczności logicznej:

- dana liczba włosów  $x$  należy do pojęcia „nie-łyse” w stopniu  $\mu_{nie-ly-sy}(x)$  – być może nawet równym 1, tj. pełna przynależność;
- usunięcie jednego włosa oznacza, że pozostała liczba należy do tej koncepcji tylko w stopniach  $\mu_{nie-ly-sy}(x - 1)$  i:  $\mu_{nie-ly-sy}(x - 1) \leq \mu_{nie-ly-sy}(x)$ ;
- kontynuacja tego procesu jeszcze bardziej obniży stopień członkostwa i ostatecznie musi doprowadzić do łysiny, tj. liczby włosów  $x$ , dla których  $\mu_{nie-ly-sy}(x) = 0$ .

Innym przykładem jest tzw. paradoks kłamcy. Pewien człowiek twierdzi: „To, co teraz mówię, jest kłamstwem”. Jeśli stwierdzenie jest prawdziwe, to kłamie, nawet jeśli jest ono prawdą. Gdy natomiast zdanie jest kłamstwem, to tak naprawdę nie kłamie, nawet jeśli wypowiedź jest fałszywa. Zatem jeżeli mówca kłamie, mówi prawdę – i na odwrót.

Opisując tę samą sytuację z punktu widzenia logiki rozmytej, można powiedzieć, że mówca do pewnego stopnia mówi prawdę i do pewnego stopnia kłamie.

### 3.5.4.2. Systemy z logiką rozmytą

Systemy z logiką rozmytą zaliczyć można do systemów ekspertowych. Ich baza wiedzy składa się z grup reguł opisujących zależności między zmiennymi [tamże]:

$$\begin{array}{ll} \text{JEŻELI: } V_1 \text{ jest } A_1^1 \dots V_n \text{ jest } A_n^1 & \text{TO: } U \text{ jest } B^1 \\ \text{JEŻELI: } V_1 \text{ jest } A_1^K \dots V_n \text{ jest } A_n^K & \text{TO: } U \text{ jest } B^K \end{array}$$

gdzie: wartości zmiennych wejściowych  $V_1, \dots, V_n$  oraz zmiennej wyjściowej  $U$  określone są w sposób nieprecyzyjny, tzn.  $A_1^1, \dots, A_n^1, \dots, A_1^K, \dots, A_n^K$ ;  $B^1, \dots, B^K$  są zbiorami rozmytymi. Zmienne przyjmujące wartości rozmyte nazywane są zmiennymi lingwistycznymi.

Rozważmy przykładowy, bardzo prosty system rozmyty, modelujący zależność między kosztami, sprzedażą i zyskiem. Baza wiedzy systemu zawierać może dwie następujące reguły rozmyte:

JEŻELI koszty są *średnie* i sprzedaż jest *duża*, TO zysk jest *wysoki*  
JEŻELI koszty są *średnie* i sprzedaż jest *średnia*, TO zysk jest *średni*

Powyższe reguły wydają się zdroworozsądkowe i trywialne, a przede wszystkim zbyt nieprecyzyjne, by przy ich pomocy uzyskać jakiegokolwiek istotne wnioski. W przypadku systemów rozmytych to jednak sytuacja normalna. Najważniejszą kwestią związaną z wykorzystaniem logiki rozmytej okazuje się to, że umożliwia ona budowę i działanie modelu, nawet jeśli leżąca u jego podstaw wiedza jest zbyt mało precyzyjna, aby można ją było sformalizować w inny sposób [tamże]. Nie bez znaczenia staje się przy tym fakt, że ludzie w swoim rozumowaniu posługują się kategoriami i związkami rozmytymi. Gdy np. podchodzimy do drzwi, nie oceniamy naszej odległości od nich w sposób precyzyjny: półtora metra, jeden metr, 80 centymetrów, 40 centymetrów itp. Nie mierzymy przecież odległości żadną miarką – po prostu precyzyjnych jej oszacowań nie mamy. Rozumujemy w kategoriach: jestem zbyt daleko, więc podchodzę bliżej; znalazłem się dostatecznie blisko, zatrzymuję się i otwieram drzwi.

Próba ujęcia tego typu wiedzy w postaci ścisłego formalizmu matematycznego prowadzić może do poważnej komplikacji modelu. Jak widać z naszego przykładu, wyrażane w sposób rozmyty asocjacje myślowe eksperta są często zaskakująco proste i oczywiste. Nawet w przypadku, gdy dysponujemy dostateczną wiedzą na temat rozwiązywanego problemu, by zastosować precyzyjne metody jego modelowania system rozmyty często oferuje rozwiązanie znacznie prostsze, a przy tym działające zaskakująco dobrze.

Pomimo że system rozmyty ma strukturę podobną do systemu ekspertowego, wnioskowanie prowadzone jest tu na zupełnie odmiennych zasadach, według tzw. reguł wnioskowania rozmytego. W przeciwieństwie do klasycznego systemu ekspertowego w modelu rozmytym uaktywniane są wszystkie reguły. Dla każdej z nich:

- na podstawie stopnia dopasowania poszczególnych wejść do warunków w regule obliczany jest stopień prawdziwości poprzednika;

- w efekcie korelacji poprzednika z następnikiem znajdujący jest zbiór rozmyty będący wynikiem działania reguły; przy czym zbiór ten tym bardziej powinien odpowiadać konkluzji reguły, im bardziej dane wejściowe dopasowane są do jej warunków.

Wyniki działania pojedynczych reguł scalane są w jeden rozmyty zbiór wyjściowy.

### 3.5.5. Algorytmy genetyczne

Jeśli liczba alternatyw decyzyjnych zidentyfikowanych podczas fazy projektowania jest nieskończona (lub bardzo duża), faza wyboru w procesie podejmowania decyzji (rozwiązywania problemu) wymaga przeszukania wszystkich możliwych decyzji, które mogą rozwiązać problem. Jak wspomnieliśmy w punkcie 3.3.3.2 poświęconym modelom optymalizacyjnym, niezbędne staje się wówczas zastosowanie metod optymalizacyjnych realizujących proces przeszukiwania. Niemożliwe jest bowiem enumeratywne sprawdzenie wszystkich możliwych wariantów decyzji.

Ogólny schemat przeszukiwania wykonywanego przez modele optymalizacyjne można scharakteryzować następująco: rozpoczyna się od pewnego przybliżenia początkowego rozwiązania problemu (wariantu decyzji), a następnie, dopóki możliwe okazuje się polepszenie rozwiązania (w sensie wartości funkcji celu – kryterium – decyzji), powtarzane są następujące operacje:

- generowany jest nowy kandydat na rozwiązanie;
- obliczana jest wartość funkcji celu i jeśli nowe rozwiązanie okazuje się pod tym względem lepsze od poprzedniego, zastępujemy stare rozwiązanie nowym.

Poszczególne metody optymalizacji różnią się między sobą przede wszystkim pierwszym krokiem. Jak wspomnieliśmy w punkcie 3.3.3.2, do najważniejszych metod optymalizacyjnych należą metody heurystyczne. Heurystyki wykorzystywane do wyboru nowego kandydata na rozwiązanie mogą być różne. Do najczęściej stosowanych należą te oparte na obserwacji różnych mechanizmów natury. Używające ich metody przeszukiwania zazwyczaj zaliczane są do dziedziny sztucznej inteligencji, chociaż nie starają się naśladować inteligentnego działania eksperta-człowieka, lecz ślepe siły natury. Z tego powodu czasami wydzielane są

one do odrębnej grupy, określanej jako metody obliczeń naturalnych. Wśród najważniejszych w tej grupie znajdziemy:

- algorytmy genetyczne – wykorzystujące heurystyki oparte na działaniu mechanizmów ewolucji;
- algorytmy mrówkowe (*ant colony optimization*) – wykorzystujące heurystyki oparte na obserwacji zachowania mrówek;
- algorytmy rojowe (*particle swarm optimization*) – wykorzystujące heurystyki oparte na obserwacji rojów owadów.

Wśród powyższych metod bez wątpienia najszerzej stosowanym rozwiązaniem są algorytmy genetyczne.

Algorytm genetyczny jest efektywną, niezależną od dziedziny zastosowań heurystyczną metodą przeszukiwania, możliwą do wykorzystania w szerokim spektrum dziedzin problemowych. Do jego działania niezbędna staje się tylko sposobność wyznaczania funkcji celu dla dowolnego badanego rozwiązania. Nie ma potrzeby obliczania innych informacji, np. dotyczących pochodnych. W przeciwieństwie do najbardziej efektywnych metod gradientowych algorytm genetyczny może być stosowany również w sytuacjach, gdy funkcja celu jest nieróżniczkowalna, a nawet nieciągła. Kosztem jest zazwyczaj znacznie wolniejsze tempo działania.

Algorytm genetyczny korzysta z heurystyki (a w zasadzie meta-heurystyki) opartej na naśladownictwie biologicznych procesów ewolucji. Główna idea to: przetrwanie najsilniejszych, posuwanie się w kierunku coraz lepszych rozwiązań, pozwalając, aby tylko najzdolniejsi rodzice tworzyli przyszłe pokolenia. Zasadniczo algorytm ten wpisuje się w podany wyżej ogólny schemat metod przeszukiwania, z tym, że wykorzystuje nie pojedyncze rozwiązania, lecz ich grupy, nazywane populacjami. Kolejne kroki przybierają formę dwóch operacji:

- 1) nowi kandydaci na rozwiązania tworzeni są przy pomocy operatorów genetycznych:
  - krzyżowanie – polega na mieszaniu fragmentów rozwiązań z poprzedniej populacji;
  - mutacja – polega na losowej zmianie wybranego fragmentu rozwiązania;
- 2) proces oceny nowych kandydatów na rozwiązania i wyboru ich do nowej populacji odbywa się na zasadzie selekcji ewolucyjnej: do nowej populacji kandydaci na rozwiązania przechodzą z prawdopodobieństwem proporcjonalnym do wartości funkcji celu.

Zwróćmy uwagę, że w drugim kroku selekcja oparta jest na mechanizmach probabilistycznych. Oznacza to, że z niewielkim prawdopodobieństwem (czyli dosyć rzadko) do następnej populacji może zostać wybrane rozwiązanie nieco gorsze, co pozwala na uniknięcie problemu utknięcia algorytmu optymalizacji w minimum lokalnym.

O przydatności algorytmów genetycznych może świadczyć fakt, że możliwość użycia mechanizmu ewolucyjnego optymalizacji została włączona do popularnego dodatku Solver w arkuszu kalkulacyjnym Excel. Pamiętajmy jednak, że podejście to powinno być stosowane w przypadku „trudnych” zadań optymalizacyjnych o nieróżniczkowalnych, a zazwyczaj nieciągłych funkcjach celu. Przy różniczkowalnych funkcjach celu raczej lepszym wyjściem będzie zastosowanie szybszych metod gradientowych.

### 3.6. Podsumowanie

Współczesne środowisko organizacji w znaczny sposób utrudnia proces podejmowania decyzji [Sharda, Delen i Turban, 2017]. Decydenci dysponują jednak coraz doskonalszymi środkami technicznymi, korzystają z coraz bogatszych zasobów informacyjnych, stosując przy tym coraz lepsze środki ich przeszukiwania oraz analizy. Skutkuje to powstawaniem coraz większej liczby alternatyw decyzyjnych, pozostających do wyboru nie tylko przez menedżera wyższego szczebla. Wspomaganie tych procesów narzędziami sztucznej inteligencji w praktyce, zwłaszcza w sferze zarządzania, jest bardzo ambitnym zadaniem. Systemy tej kategorii, które zdefiniowane zostały w niniejszej monografii jako kategoria systemów informatycznych starających się naśladować sposoby rozwiązywania problemów stosowane przez ludzi, muszą zatem nie tyle wykazywać się myśleniem, co po prostu rozwiązywać bądź wspierać człowieka w rozwiązywaniu skomplikowanych problemów, z którymi nie można poradzić sobie przy pomocy klasycznych, algorytmicznych środków.

## Literatura

Bartkiewicz W. (2013), *Modelowanie niepewności krótkoterminowego popytu na energię elektryczną z wykorzystaniem sieci neuronowych i neuronowo-rozmytych*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.



- Bartkiewicz W., Jabłoński W.J. (2005), *Systemy informatyczne zarządzania. Wprowadzenie do technologii przetwarzania danych*, Wydawnictwo Wyższej Szkoły Humanistyczno-Ekonomicznej we Włocławku, Włocławek.
- Bengio Y., Courville A., Goodfellow I. (2018), *Deep Learning. Systemy uczące się*, Wydawnictwo Naukowe PWN, Warszawa.
- Cichosz P. (2000), *Systemy uczące się*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Date C.J. (2000), *Wprowadzenie do systemów baz danych*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Flasiński M. (2011), *Wstęp do sztucznej inteligencji*, Wydawnictwo Naukowe PWN, Warszawa.
- Garcia-Molina H., Ullman J.D., Widom J. (2006), *Systemy baz danych – Pełny wykład*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Jabłoński W.J., Bartkiewicz W. (2006), *Systemy informatyczne zarządzania. Klasyfikacja i charakterystyka systemów*, Wydawnictwo Kujawsko-Pomorskiej Szkoły Wyższej w Bydgoszczy, Bydgoszcz.
- Jo H. (2019), *Text Mining. Concepts, Implementation, and Big Data Challenge*, Springer, New York.
- Krawczyk S. (2001), *Metody ilościowe w planowaniu działalności przedsiębiorstwa*, Wydawnictwo C.H. Beck, Warszawa.
- Larose D.T. (2006), *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa.
- Larose D.T. (2008), *Metody i modele eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa.
- Laudon K.C., Laudon J.P. (1993), *Business Information Systems – a problem solving approach*, The Dryden Press, New York.
- Liu B. (2011), *Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data*, wyd. 2, Springer, New York.
- Markov Z., Larose D.T. (2009), *Eksploracja zasobów internetowych*, Wydawnictwo Naukowe PWN, Warszawa.
- Marshall K.T., Oliver R.M. (1995), *Decision making and forecasting*, McGraw-Hill, New York.
- McCulloch W.S., Pitts W.H. (1943), *A logical calculus of the ideas immanent in nervous activity*, „The Bulletin of Mathematical Biophysics”, nr 5.
- Morzy T. (2019), *Eksploracja danych*, Wydawnictwo Naukowe PWN, Warszawa.
- Papińska-Kacperek J. (red.), (2008), *Społeczeństwo informacyjne*, Wydawnictwo Naukowe PWN, Warszawa.
- Radzikowski W. (1990), *Komputerowe systemy wspomagania decyzji*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Sharda R.E., Delen D., Turban E. (2017), *Analytics, Data Science, and Artificial Intelligence: Systems for Decision Support*, wyd. 11, Pearson, London.
- Sroka H. (1994), *Systemy ekspertowe. Komputerowe wspomaganie decyzji w zarządzaniu i finansach*, Akademia Ekonomiczna w Katowicach, Katowice.
- Stair R.M. (1992), *Principles of information systems – A managerial approach*, Thomson Publishing, Stamford.
- Ullman J.D., Widom J. (2000), *Podstawowy wykład z systemów baz danych*, Wydawnictwo Naukowo-Techniczne, Warszawa.
- Zieliński J.S. (red.), (2000), *Inteligentne systemy w zarządzaniu – Teoria i praktyka*, Wydawnictwo Naukowe PWN, Warszawa.





# Analiza przypadku 1

Firma zatrudniająca 300 pracowników produkuje grille i brykiety, ma swoją siedzibę w jednym z amerykańskich miast. Jej produkty są zarówno sprzedawane na rynku krajowym, jaki i eksportowane. Do wspomagania procesów biznesowych firma przez wiele lat wykorzystywała system Sage ERP w modelu *on-premise*. Kilka lat temu zdecydowała o zmianie systemu na Epicor i implementacji w modelu chmury prywatnej. Inne aplikacje, takie jak: oprogramowanie biurowe, poczta elektroniczna, system składania zamówień i sprzedaży elektronicznej, zostały wdrożone w modelu chmury publicznej. Usługi tych aplikacji oferowali różni dostawcy.

Firma wyszła z założenia, że wykorzystanie aplikacji w środowisku chmury obliczeniowej pozwoli na większą elastyczność całego wsparcia IT, przy jednoczesnej redukcji kosztów zatrudnienia i utrzymywania systemów. Firma zatrudniała 10 osób w dziale IT. Obecnie około 75% czasu spędzają oni na rozwiązywaniu problemów integracji systemów i platform, które wykorzystują pracownicy organizacji. Rozproszenie aplikacji i dostawców sprawia, że w przyjętym rozwiązaniu występuje około 30 punktów integracji, a awarie wpływają na frustrację użytkowników. Kierownik działu IT jest zdania, że źródłem kłopotów był brak przygotowania odpowiedniej strategii integracji.

**Źródło:** opracowanie na podstawie [searcherp.techtarget.com/feature/ERP-hybrid-cloud-architecture-causes-integration-headaches](https://searcherp.techtarget.com/feature/ERP-hybrid-cloud-architecture-causes-integration-headaches) [dostęp: 25.02.2020].

## Pytania

1. Jaki poziom wsparcia biznesu przez IT ma firma?
2. Co firma powinna wziąć pod uwagę przy wyborze systemu ERP?
3. Co stanowi główny problem w przyjętym rozwiązaniu?
4. Jakie bariery integracji aplikacji napotkała i może napotkać firma?
5. Jak rozwiązać istniejący problem?



## Analiza przypadku 2

Szef firmy New York's Veritas Studio Wines, położonej przy przeprawie promowej, mawia, że im więcej konsumuje się wina, tym więcej prawdy wypływa w rozmowach. Wiele przedsiębiorstw dowiaduje się, że „prawdę”, a raczej wiedzę, można znaleźć za pomocą analityki dużych zbiorów danych oraz odpowiedniego nadzoru nad IT, co pozwala im uzyskać prawdziwie istotne informacje w nowy i innowacyjny sposób.

Analizy Big Data są postrzegane jako domena dużych przedsiębiorstw, ponieważ zwykle tylko one posiadały zasoby IT do zarządzania i analizy tak dużej ilości danych. Obecna oferta rynkowa aplikacji biznesowych i modele jej świadczenia zmieniły ten stan rzeczy, czego przykładem jest firma Veritas Studio Wines. Mały, niezależny sklep z winami w Nowym Jorku był w stanie wykorzystać narzędzia do analiz Big Data, by odkryć pewne „prawdy” na temat swoich klientów, prowadzące do lepszej obsługi i poprawy efektywności prowadzonego biznesu.

Firma Veritas Studio Wines wykorzystwała rozwiązanie SAP Digital Consumer Insight (DCI) pozwalające na przechowywanie, analizę i zarządzanie danymi w modelu „dane jako usługa” (*Data as a Service* – DaaS).

Koncepcja biznesu firmy była dość prosta – miał to być sklep z dobrej jakości winami, prowadzony w przyjaznej atmosferze. Konkurując w rejonie z kilkoma innymi sklepami z winami i sklepami monopolowymi, firma Veritas Studio Wines musiała zaoferować coś, co zachęciłoby klientów i pozwoliło uzyskać im pozytywne doświadczenie. Konkurenci oferowali zbliżony asortyment, a ich sklepy i wyposażenie były podobnie – proponowały duże przestrzenie z fluorescencyjnym oświetleniem. Przedsiębiorstwo zdecydowało się zatem stworzyć ciepłe i przyjazne otoczenie dla osób, które chcą odkrywać świat wina i mieć pewność, że kiedy przyjdą do sklepu, to wyjdą z dobrym, odpowiadającym ich gustom trunkiem. Firma postanowiła bazować na zdobywaniu zaufania klientów poprzez dopasowane do ich preferencji rekomendacji odpowiedniego wina – właściciel uważa, że najgorsze, co może się zdarzyć, kiedy kupuje się butelkę wina, to sytuacja, gdy po powrocie do domu, otworzeniu i degustacji okazuje się, że jest nieodpowiadanie lub wywołuje poczucie przepełnienia.

Firma rozpoczęła poszukiwania odpowiednich rozwiązań informatycznych. Jedna z ofert napłynęła od przedsiębiorstwa SAP. Jak wiele mniejszych firm, Veritas Studio Wines postrzegała SAP jako dostawcę, który wdraża złożone systemy biznesowe dla dużych przedsiębiorstw i była sceptyczna, co do możliwości korzystania z zaproponowanego rozwiązania SAP DCI.

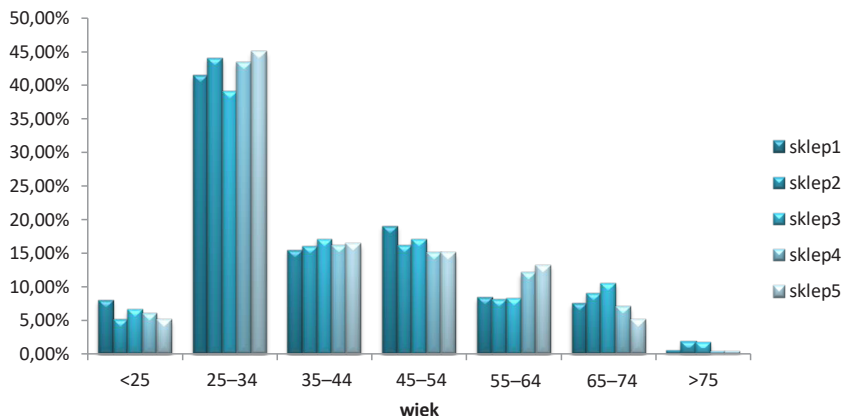
SAP Digital Consumer Insight (SAP DCI) oferuje aplikacje mobilne do analiz danych w modelu usługi pozwalającej także małym firmom analizować dane osiągalne zwykle dla znacznie większych przedsiębiorstw. Wykorzystuje pozyskane z telefonów komórkowych informacje na temat lokalizacji użytkownika oraz dane demograficzne – i umieszcza je w raportach, które firmy mogą przeglądać, aby dowiedzieć się więcej o swoich klientach lub konkurentach.

Tego rodzaju analizy można zakupić na stronie w sklepie internetowym SAP w postaci pojedynczego zestawienia lub pakietu pięciu raportów. Użytkownik wyznacza lokalizację (która może dotyczyć całego miasta lub ściśle określonej części, np. sklepu) oraz trzydniowy zasięg czasu przeszłego, np. ostatni kwartał. Raporty są dostępne w różnych formatach graficznych. Informacje w nich zawarte pozbawione zostały jakichkolwiek danych identyfikacyjnych klientów i obejmują takie informacje, jak: wiek użytkowników telefonów komórkowych (z wyjątkiem osób w wieku poniżej osiemnastu lat), płeć, rodzaj używanego urządzenia i system operacyjny (np. iOS9 lub Android). Dane te dają użytkownikowi bardzo dobrą podstawę do analiz potencjalnych klientów w określonym czasie i miejscu oraz pomagają w działaniach marketingowych i sprzedażowych.

Do tej pory firma zbierała informacje o nastawieniu klientów z recenzji na Facebooku czy w Google, a dane o sprzedaży z systemu wspierającego procesy sprzedaży o nazwie Bindo POS. Podczas gdy dane w mediach społecznościowych były ograniczone, te z systemu Bindo POS okazywały się bardzo istotne, gdyż zawierały informacje o codziennej sprzedaży (niemniej jednak dane demograficzne zawarte w raportach SAP DCI mogły dostarczyć informacji o lokalnej społeczności, których firma nie uzyskiwałaby nigdzie indziej).

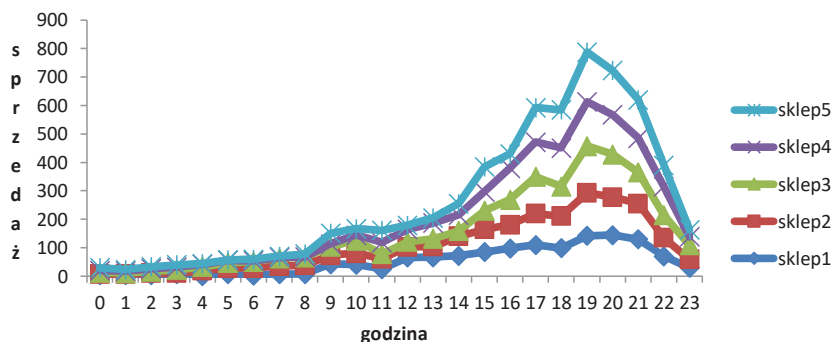
Veritas Studio Wines miała już informacje na temat tego, że duży odsetek jej klientów pochodzi z najbliższego otoczenia – dużego apartamentowca po drugiej stronie ulicy – więc postanowiła dowiedzieć się więcej na temat lokalizacji swoich konkurentów. Zamiast mierzyć dane ze swojej lokalizacji, która znajduje się na bocznej ulicy, zwróciła większą uwagę na

konkurentów usytuowanych przy głównych ulicach. SAP DCI pobiera dane z telefonów komórkowych w określonej lokalizacji, w ciągu określonego trzydniowego okresu. Do zakupionych analiz wybrała okres weekendu.



**Rysunek A2.1.** Graficzna analiza danych otrzymanego raportu

**Źródło:** opracowanie własne na podstawie [searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-see-truth-about-customers](https://searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-see-truth-about-customers) [dostęp: 15.02.2020].



**Rys A2.2.** Graficzna analiza danych otrzymanego raportu

**Źródło:** opracowanie własne na podstawie [searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-see-truth-about-customers](https://searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-see-truth-about-customers) [dostęp: 15.02.2020].

Zamówione raporty dostarczyły trzech istotnych spostrzeżeń. Po pierwsze firma zauważyła, że w okolicy jest dużo starszych osób, mieszkającym tam od pokoleń – jednak większość ruchu pieszego w okolicy

generują osoby w wieku od 25 do 34 lat (rysunek A2.1). Po drugie większość potencjalnych klientów jest aktywna późnym popołudniem lub wczesnym wieczorem (rysunek A2.2) – później niż sądził na początku menedżer firmy. Po trzecie większość kupujących pochodzi z sąsiedztwa, podczas gdy menedżer zakładał, że wielu z nich mieszka w innych częściach miasta – raport wykazał również, że znaczna liczba potencjalnych klientów pochodzi z dzielnicy położonej po przeciwnej stronie rzeki.

W rezultacie, po analizie otrzymanych danych, menedżer zmienił godziny otwarcia sklepu, co poprawiło obsługę klienta, ponieważ w tym samym czasie obsługa składała się z dwóch osób. Przed analizą firma miała zamiar uruchomić trzecią zmianę, aby przedłużyć godziny pracy, ale nie musiała tego robić, znając godziny natężenia ruchu, co dało oszczędności około 28 000 USD – zrezygnowano bowiem z zatrudnienia nowego pracownika.

Kolejna decyzja, jaką podjęto, to ograniczenie działań marketingowych do najbliższego sąsiedztwa, z wyjątkiem dzielnicy położonej po przeciwnej stronie rzeki – niektórzy klienci wchodzili do sklepu przed udaniem się do terminalu promowego. Obecnie firma posiada pracownika udającego się do terminalu promowego i rozdającego ulotki, co stanowi nową metodę promocyjną.

Ponieważ raport SAP pokazał, że większość klientów to ludzie młodzi, wprowadzono do sklepu prostą i domową atmosferę, co pomogło Veritas Studio Wines odróżnić się od konkurentów.

Menedżer firmy jest przekonany, że analizy Big Data mogą przynieść korzyści także małej firmie, ale ostrzegł, że może być to trudniejsze niż się wydaje. Wiele firm, w tym SAP, ma przed sobą wiele wyzwań związanych z pozyskiwaniem małych przedsiębiorstw dla swoich rozwiązań. Jednym z nich jest to, że pracownicy z małych firm są zwykle bardzo zajęci i mają wiele obowiązków. Istotną sprawą okazuje się również to, co firma chce analizować i co zrobić po uzyskaniu raportów, a także czego może się nauczyć z otrzymanych graficznych analiz danych.

**Źródło:** opracowanie na podstawie [searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-look-for-truth-about-customers](https://searchsap.techtarget.com/feature/SAP-data-as-a-service-helps-wine-shop-look-for-truth-about-customers) [dostęp: 15.02.2020].

## Pytania

1. Jakie systemy wykorzystuje Veritas Studio Wines? Jaki ma do nich dostęp?
2. Jakie technologie wykorzystują te systemy?
3. Jakie dane (i z jakich źródeł) systemy te analizują? Jakie korzyści przyniosła firmie analiza danych?
4. Jakie inne analizy (i na podstawie jakich danych) może prowadzić Veritas Studio Wines?
5. Jaki wpływ na biznes miało wykorzystanie technologii IT?





# Analiza przypadku 3

## Ryzyko wykorzystania niepewnej informacji

W tej analizie przypadku uwaga czytelnika zostanie zwrócona na pewną istotną kwestię, często występującą w praktyce, a nie zawsze dostatecznie docenianą przez decydentów. Dyskutowane w rozdziale trzecim metody analityczne w większości sytuacji mają charakter dedukcyjny lub indukcyjny (więcej o modelach algorytmicznych, dedukcyjnych i indukcyjnych zob. punkt 3.3.2). Oznacza to, że modelują one badane zagadnienia w sposób przybliżony na podstawie obserwacji, a nie dokładnego ich zrozumienia. Dotyczy to zarówno metod ilościowych (wykorzystujących dane), jak i symboliczno-logicznych (takich jak systemy ekspertowe).

W związku z powyższym, informacje otrzymywane przy użyciu tych metod w większości przypadków również mają charakter przybliżony i obciążone są niepewnością, co przekłada się na niepewność decyzji podejmowanych na ich podstawie. W wielu sytuacjach stosowane metody analityczne umożliwiają ocenę tej niepewności, pozwalając na skalkulowanie ryzyka decyzji.

Aspekt ten często pomijany jest przez decydentów, mających tendencję do przyjmowania wyników analiz jako pewnych i dokładnych, co może prowadzić do błędnych decyzji. Rozważmy następujący przykład.

Firma „Promox” zamierza przeprowadzić w kilkunastu kolejnych miesiącach inwestycję związaną z kontaktami z odbiorcami. Wymagać ona będzie ponoszenia dodatkowych kosztów w wysokości pięciu tysięcy złotych w każdym miesiącu. Oceniono jednak, że inwestycja ta zwiększy miesięczne zyski ze sprzedaży o kwotę dziesięciu tysięcy złotych – przy uwzględnieniu również kosztów samej akcji promocyjnej.

Przedstawiona sytuacja decyzyjna jest bardzo prosta. Należy rozważyć dwa możliwe warianty decyzji:

- $d_1$ : przeprowadzić rozważaną inwestycję,
- $d_2$ : zrezygnować z niej.

Jak wskazano w punkcie 3.2.3, aby właściwie wybrać wariant decyzji, jaki zostanie później wprowadzony w życie, należy określić pewne

kryterium (funkcję celu), pod kątem którego oceniane są brane pod uwagę możliwości, a następnie wyznaczane wartości tego kryterium dla wariantu  $d_1$  oraz  $d_2$ . W naszym przypadku naturalnym kryterium jest skutek finansowy każdego z rozważanych wariantów decyzji.

Zysk przy wyborze wariantu  $d_1$  wynosi dziesięć tysięcy złotych, podczas gdy w przypadku wariantu  $d_2$  – zero (ponieważ sytuacja się nie zmienia, nic nie wydajemy ani nie zarabiamy). Dla oceny wariantu drugiego można ewentualnie wziąć pod uwagę zysk z comiesięcznej inwestycji kwoty pięciu tysięcy złotych w bezpieczne formy inwestowania, np. oprocentowanie lokaty bankowej lub bezpiecznych obligacji. Przyjmijmy, dla uproszczenia, że jest on na tyle niewielki, że możemy go pominąć.

Wybór wariantu decyzji jest w tym przypadku prosty. Zrealizujemy wariant dający wyższe wartości zysku. Ponieważ  $\text{zysk}(d_1) = 10\ 000$ , zaś  $\text{zysk}(d_2) = 0$ , wybierzemy oczywiście wariant pierwszy, czyli zdecydujemy się na przeprowadzenie rozważanej inwestycji.

Funkcją systemów wspomagania decyzji przy tego rodzaju zagadnieniach decyzyjnych jest przede wszystkim dostarczanie informacji oraz wiedzy dziedzinowej – w przypadku firmy „Promox” informacji o istotnych zmiennych decyzyjnych niezbędnych do obliczenia przychodów i kosztów związanych z inwestycją, wiedzy dziedzinowej w postaci modeli finansowych wyznaczających wartości kryterium (zysku) czy też innych niezbędnych wielkości pośrednich.

Spróbujmy teraz skomplikować nieco sytuację. Otóż firma „Promox”, analizując skutki realizacji rozważanej inwestycji, stwierdziła, że w danym miesiącu założone zyski na poziomie dziesięciu tysięcy złotych zostaną osiągnięte, ale tylko w przypadku, gdy sprzedaż w tym miesiącu będzie dostatecznie duża. Jeśli okaże się zbyt mała, to w danym miesiącu poniesiemy koszty inwestycji – pięć tysięcy złotych – nie osiągając żadnych zysków. Próg opłacalności inwestycji wyznaczony został na poziomie siedemdziesięciu tysięcy złotych.

Pojawiła się więc nowa istotna zmienna decyzyjna, niezbędna do określenia skutków poszczególnych wariantów decyzji, a zwłaszcza wariantu  $d_1$ : poziom miesięcznej sprzedaży. Przyjmijmy jednak, iż zastosowane zostały odpowiednie metody analityczne (na przykład oparte na metodach eksploracji danych – zob. punkt 3.4.4, lub sieciach neuronowych – zob. punkt 3.5.3) do oszacowania prognozy miesięcznej sprzedaży. Zgodnie z prognozą w okresie inwestycji prognozowany poziom mie-

sięicznej sprzedaży jest mniej więcej stały i wynosi sześćdziesiąt sześć tysięcy złotych.

Jak widać, prognoza wydaje się być niekorzystna. W obecnej sytuacji  $\text{zysk}(d_1) = -5\,000$ , zaś  $\text{zysk}(d_2) = 0$ , a więc raczej skłonni jesteśmy wybrać wariant drugi decyzji, czyli z rozważanej inwestycji zrezygnować.

Jednak taki sposób postępowania, choć dosyć powszechny, nie zawsze jest niewłaściwy. Pomijamy przez to pewne istotne czynniki wpływające na wybór najlepszej decyzji. Problem polega na tym, że podejmując decyzję w ten sposób, traktujemy prognozę jako fakt niewątpliwy i pewny. Oczywiście to nieprawda. Każda prognoza obarczona jest błędem. Nasza prognoza określa jedynie wartość oczekiwaną miesięcznej sprzedaży firmy „Promox”. Faktyczna realizacja tego procesu kształtować powinna się w pobliżu prognozy – w jednym miesiącu będzie nieco wyższa, w innym nieco niższa, ale niemal nigdy nie będzie jej dokładnie równa.

Skoro rzeczywista miesięczna sprzedaż firmy „Promox” w kolejnych miesiącach będzie jedynie oscylować wokół wartości sześćdziesięciu sześciu tysięcy złotych, to niewykluczone, że czasami przekroczy ona próg opłacalności siedemdziesięciu tysięcy złotych. Być może nawet będzie się to zdarzało na tyle często, że jednak inwestycja okaże się dla firmy „Promox” opłacalna.

Aby właściwie określić, który wariant decyzji wybrać, należy oszacować, jak często (czyli z jakim prawdopodobieństwem) przekroczymy – bądź nie – próg opłacalności planowanej inwestycji. Innymi słowy, w sytuacji, gdy nie potrafimy z pewnością ocenić skutków naszych decyzji, musimy wziąć pod uwagę nie tylko możliwe efekty, ale także prawdopodobieństwo ich wystąpienia, czyli ryzyko decyzji.

W naszym przypadku ryzyko bierze się, jak powiedzieliśmy, z niepewnego charakteru informacji, jaką wykorzystujemy w procesie wyboru sposobu postępowania – opieramy się tu na prognozie. Aby je ocenić, musimy oszacować niepewność prognozy. W tym celu niezbędne jest wzięcie pod uwagę możliwego błędu prognozy.

Firma „Promox”, tworząc system prognozy sprzedaży, zadbała o to, by oprócz samej prognozowanej wartości podawał on istotne informacje dodatkowe: o błędzie. Powiedzmy, że w interesującym nas przypadku średni błąd kwadratowy prognozy wynosi dwanaście tysięcy złotych. Średni błąd kwadratowy stanowi oszacowanie średniego odchylenia między prognozą a wartością faktyczną, jakiego możemy się spodziewać w danym przypadku, czyli tzw. odchylenia standardowego prognozy.

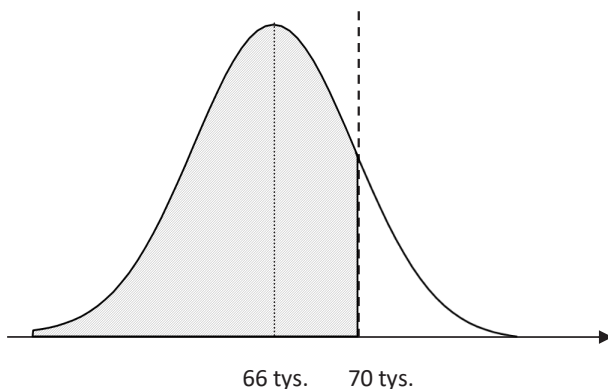
Przyjmijmy dalej, że sporządzający system prognozy sprzedaży analityk w trakcie weryfikacji modelu potwierdził, że jego błąd ma rozkład normalny. Nie jest to absolutnie założenie na wyrost. Przy tego rodzaju prognozach i generalnie przy prognozach wartości przeciętnych różnych zjawisk (zdecydowana większość prognoz, którymi zazwyczaj posługujemy w zarządzaniu) błąd dobrze zbudowanego modelu prognostycznego powinien mieć rozkład zbliżony do normalnego.

Podsumowując ostatni fragment przykładu, na podstawie posiadanych informacji można powiedzieć, że rozkład wartości prognozowanej miesięcznej sprzedaży, powstały w wyniku błędu modelu prognostycznego, będzie miał charakter rozkładu normalnego o wartości oczekiwanej równej prognozie i odchyleniu standardowym określonym przez odchylenie standardowe prognozy, czyli  $N(66\ 000, 12\ 000)$ .

Aby obliczyć prawdopodobieństwo, że sprzedaż miesięczna firmy „Promox” nie przekroczy poziomu opłacalności inwestycji 70 000 złotych (rysunek A3.1), należy wyznaczyć wartość dystrybuanty rozkładu  $N(66\ 000, 12\ 000)$ , w punkcie 70 000. Można to zrobić przy wykorzystaniu choćby funkcji arkusza kalkulacyjnego:

$$\Pr(\text{Sprzedaż} \leq 70\ 000) \approx 0,6$$

Oczywiście prawdopodobieństwo, że sprzedaż przekroczy próg opłacalności 70 000 złotych wynosi w tym przypadku około 0,4.



**Rysunek A3.1.** Prawdopodobieństwo, że przy prognozie 66 000 złotych, sprzedaż nie przekroczy poziomu 70 000 złotych.

**Źródło:** opracowanie własne.

Wybierając wariant decyzji  $d_1$  („zainwestuj”), nie potrafimy na pewno podać skutku naszej decyzji. Możemy go określić jedynie z dokładnością co do prawdopodobieństwa: zarobimy 10 000 złotych z prawdopodobieństwem 0,4; stracimy 5000 złotych z prawdopodobieństwem 0,6. Istnieje więc aż 60% ryzyko, że decydując się na ten wariant, stracimy 5000 złotych, co wydaje się potwierdzać, że prognoza jest niekorzystna i powinniśmy wybrać wariant decyzji  $d_2$ , czyli zrezygnować z inwestycji.

Obiektywną metodą wyboru wariantu decyzji w takich warunkach (tzw. warunki ryzyka) jest jednak kryterium wartości oczekiwanej skutku decyzji. Otóż nie jesteśmy w stanie z pewnością określić efektów wariantu  $d_2$ . Możemy jednak wyznaczyć wartość oczekiwaną tego skutku, ważąc możliwe wartości prawdopodobieństwami ich realizacji:

$$10\,000 \cdot 0,4 + (-5000) \cdot 0,6 = 1000$$

W ten sposób dochodzimy do dosyć zaskakującej może dla czytelnika konkluzji. Wartość oczekiwana skutku wariantu decyzji  $d_1$  wynosi tysiąc złotych, podczas gdy dla wariantu  $d_2$  – zero złotych (wartość oczekiwana ze stałej równa jest tej stałej). Powinniśmy wobec tego wybrać decyzję  $d_1$  i zainwestować! Pomimo, wydawałoby się, niekorzystnej prognozy i dużego ryzyka straty.

Na pierwszy rzut oka powyższa konkluzja wydaje się dziwna i niezbyt zgodna z intuicją. Inwestować przy 60% ryzyku straty? A jednak, to jest najprawdopodobniej najlepsze rozwiązanie – w podanych warunkach powinniśmy wybrać pierwszy wariant decyzji (obiektywnie, ważąc zyski i straty oraz prawdopodobieństwa ich uzyskania). Powody takiego wyboru szczególnie dobrze widać w naszym przykładzie, ponieważ skutki rozważanej decyzji, wynikające z niej zyski i straty, realizować się będą wielokrotnie, w kolejnych miesiącach. Prognoza jest rzeczywiście niekorzystna. W dłuższym horyzoncie czasu, powiedzmy, dziesięciomiesięcznym, aż sześć razy stracimy, a tylko cztery razy zyskamy. Ale stracimy aż sześć razy po pięć tysięcy złotych – w sumie 30 000 złotych, zaś zyskamy tylko cztery razy, ale aż po 10 000 złotych, czyli razem aż 40 000 złotych. W tym przypadku opłaca się zatem zaryzykować.



# Zakończenie

Zastosowanie nowych technologii i przetwarzanie w chmurze obliczeniowej to interdyscyplinarne pole badawcze, które stymuluje jednocześnie ewolucję technologiczną i biznesową. Implementacja rozwiązań opartych na modelu chmury obliczeniowej powoduje zmiany paradygmatu zarówno infrastruktury IT, jak i infrastruktury biznesowej, szczególnie w zakresie wydajności IT i sprawności biznesowej organizacji, m.in. poprzez usługi chmury obliczeniowej umożliwiające szybkie otrzymanie wyników analiz masowo napływających danych.

Kreowanie nowych strategii, zmiany w modelach biznesowych, poprawa efektywności pracy i relacji z partnerami wymagają odpowiednich, coraz bardziej złożonych decyzji, podejmowanych w coraz szybszym tempie współczesnego świata i wytwarzanej przez niego presji środowiskowej. Wbudowanie elementów decyzyjnych w aplikacje biznesowe jest jednym z elementów, który pozwala współczesnej organizacji sprostać wyzwaniom rynku.

Organizacje od lat wykorzystują różne aplikacje do wspomagania procesów biznesowych. Systemy ERP ciągle pełnią kluczową rolę, zwłaszcza w tradycyjnym obszarze funkcji *back-office*, związanym z zarządzaniem finansami, magazynem czy produkcją. Głównymi wyzwaniami pozostają dla nich integracja aplikacji, przepływ danych, przygotowanie interfejsów do aplikacji analitycznych. Sprawność i umiejętność korzystania z technologii informatycznych warunkuje istnienie organizacji. Organizacja, której szkielet gromadzenia danych gospodarczych i wspomagania zarządzania procesami oraz podejmowaniem decyzji stanowi system ERP – wdrożony w chmurze, z wbudowanymi systemami decyzyjnymi zapewniającymi funkcje analityczne *online* i przetwarzanie danych w pamięci (*in-memory processing*) – notuje poprawę wydajności i jakości pracy. Systemy te uczą się, ulepszają i dokonują predykcji, korzystając z uczenia maszynowego. Ich struktura umożliwia szybkie wykorzystanie tych możliwości w całej firmie, które zapewniają nie tylko utrzymanie się na rynku, ale i osiągnięcie przewagi konkurencyjnej.



Rozwój technologii multimedialnej, serwisów społecznościowych, usług urządzeń mobilnych, analityki i rozwiązań chmury obliczeniowej zmienia funkcjonalność aplikacji biznesowych, pozwalając na większą elastyczność i zwinność oraz lepszą konwergencję IT i biznesu. Dużego znaczenia dla transformacji cyfrowej, zwłaszcza w przedsiębiorstwach produkcyjnych, nabierają przedstawione w pozycjach cyklu technologie: przemysłowy Internet Rzeczy, cyfrowe odpowiedniki (*Digital Twin*), wirtualna i rozszerzona rzeczywistość, Big Data, automatyzacja i robotyka czy Blockchain. Podłączenie ogromnej liczby urządzeń inteligentnych, ich wzajemna komunikacja, możliwość analiz Big Data, szybkie podejmowanie decyzji w warunkach niepewności i ryzyka wpływają bowiem na zwiększenie produktywności, wydajności i efektywności dóbr oraz usług, a także stymulują wprowadzanie innowacyjnych propozycji i nowych modeli biznesu.

Mapa drogowa innowacji dla inteligentnych systemów biznesowych obejmuje możliwości wykorzystania wielu nowych technologii, np. technologii Blockchain do prowadzenia ksiąg rachunkowych czy Internetu Rzeczy do budowy nowych globalnych cyfrowych modeli biznesowych. Doświadczenie użytkownika w całym procesie kontaktu z organizacją pełni obecnie kluczową funkcję dla utrzymania relacji z klientem. Inteligentny system wspomagający pracę biznesu musi zatem stawiać je na pierwszym miejscu i – w połączeniu z podejściem wykorzystania aplikacji mobilnych – oferować nowe rodzaje interfejsów umożliwiających interakcję za pomocą systemów głosowych. Z każdym rokiem wzrasta stopień automatyzacji i stosowania narzędzi sztucznej inteligencji w procesach biznesowych organizacji (temu tematowi będzie poświęcona kolejna pozycja serii).

Podążając za cytowanymi w rozdziale drugim niniejszego studium autorami, trzeba zauważyć, że model platformy powoli zastępuje tradycyjne teorie ekonomiczne oparte na organizacjach, firmach i rynkach. Świat zamkniętych, scentralizowanych i hierarchicznych korporacji jest wypierany. Internet, algorytmy, oceny *online* i sztuczna inteligencja zapewniają natychmiastowy dostęp do wszystkich rodzajów informacji i dają właścicielom platform prawie nieograniczone możliwości łączenia użytkowników, co ułatwia im angażowanie się w ciągłe innowacje. Współtworzenie wartości i osiąganie zysków są obecnie rozwijane poprzez tworzenie inteligentnych platform, a nie przez statyczne, hierarchiczne zarządzanie

pracownikami i produktami, wspierane tradycyjnymi systemami informatycznymi. Wiele przedsiębiorstw o ugruntowanej pozycji również dostrzega możliwości tego nowego stylu prowadzenia działalności, ale mimo to nadal bazuje na istniejących strukturach, procesach, procedurach i systemach. Wejście w erę cyfrową jest ogromnym wyzwaniem nawet dla firm o ugruntowanej pozycji na rynku, ponieważ wymaga fundamentalnej transformacji ich organizacji i działalności oraz zmiany narzędzi je wspierających.

