

RYSZARD BUDZIŃSKI

TOMASZ ORDYSIŃSKI

Uniwersytet Szczeciński

INFORMACJA JAKO KLUCZOWY SKŁADNIK PRZEDSIĘBIORSTWA „NOWEJ GOSPODARKI”

Wstęp

Internet i jego rozwój coraz mocniej wkracza w sfery życia gospodarczego. Z całym swoim otoczeniem technologicznym stał się on fundamentem tak zwanej nowej gospodarki (*new economy* lub *e-economy*), która oznacza poszukiwanie źródeł szybszego wzrostu oraz zmian strukturalnych w rozwoju technik informatycznych, a w szczególności w szybkiej komercjalizacji Internetu, które to determinanty w swoisty sposób modyfikują działalność gospodarczą [25]. Według L.I. Nakamury, „nowa gospodarka” oznacza pogląd, w myśl którego innowacje w zaawansowanej technice i globalizacja rynków światowych zmieniły gospodarke na tyle, że trzeba myśleć o niej i działać w niej w inny sposób. Dość odmienny pogląd reprezentuje Peter Drucker, twierdząc, że nadciąga nowa rewolucja informacyjna, ale nie tam, gdzie oczekują jej uczeni, zarządzający informacją oraz cały przemysł informatyczny. To nie jest rewolucja związana z technologiami, maszynami, technikami, oprogramowaniem czy szybkością. Ta rewolucja dotyczy koncepcji. W jego poglądach siła ciężkości przyczyn zachodzących przemian została przesunięta z aspektów technologicznych na wzrost znaczenia informacji. W dobie e-ekonomii informacja jest podstawowym, niezużywalnym kapitałem zapewniającym przewagę. Jeśli przedsiębiorstwo nie ma do niej natychmiastowego, wieloaspektowego dostępu, to jej gromadzenie staje się tylko zbędnym kosztem [8]. Należy jednak pamiętać, że informacja nie może być utożsamiana z wiedzą. O ile informacja dociera dziś za pośrednictwem telewizji czy Internetu do większości ludzi naszej planety, o tyle wiedza nadal jest przywilejem mniejszości. O konieczności bardzo wyraźnego rozróżniania informacji i wie-

dzy, a także następnej kategorii – mądrości, znakomicie przypominają słowa angielskiego poety i myśliciela, laureata Nagrody Nobla, Thomasa S. Eliota:

„Gdzie się podziła nasza mądrość, którą zastąpiła wiedza,
gdzie się podziła nasza wiedza, którą zastąpiła informacja” [15].

Celem artykułu jest usystematyzowanie pojęć związanych z informacją, określenie jej miejsca w strukturze przedsiębiorstwa oraz przedstawienie nowoczesnych metod wydobywania wiedzy z danych gromadzonych przez podmioty gospodarcze. Teżą zaś, że informacja odgrywa kluczową rolę w elektronicznej fazie gospodarki, umożliwiając dogłębną analizę rynku i zachowań klientów dzięki wykorzystaniu narzędzi informatycznych. Artykuł ma charakter poznawczy. Dane i informacje w nim zawarte pochodzą z fachowej literatury i zasobów Internetu.

1. Systematyzacja pojęć związanych z informacją

Pojęcie informacja jest trudne do precyzyjnego zdefiniowania, czego dowodem jest mnogość propozycji literaturowych. Należy sądzić, że wynika to z wielości funkcji, które przypisuje się informacjom. Od prostych określeń, u podstaw których leżą treści komunikacyjne [24], przez tworzenie wiedzy [17], aż do konkluzji, że informacja jest czynnikiem sprawczym zmian struktur zdarzeń i rzeczy [3]. Trudności z przyjęciem jednoznacznej definicji informacji wynikają z przypisywania mnogości funkcji temu właśnie pojęciu przez człowieka. W prostym ujęciu informatycznym często mamy do czynienia z pomijaniem interpretacyjnego czynnika ludzkiego. Wtedy informację definiuje się jako efekt przetwarzania danych, czyli ich zbierania, analizy [19]. P. Beynon-Davies rozróżnia trzy podstawowe kategorie pojęciowe, mianowicie: dane, informacja i wiedza [4]. Za Tsitchizris i Lachovsky autor definiuje informację jako „przyrost wiedzy, który może być uzyskany na podstawie danych”. Dalsza interpretacja prowadzi autora do wielu interesujących uogólnień:

- a) dane to fakty; jednostką danych jest jeden lub kilka symboli użytych do reprezentowania czegoś;
- b) informacje to zinterpretowane dane; informacje to dane umieszczone w znaczącym kontekście;
- c) wiedza jest otrzymywana z informacji przez jej zintegrowanie z istniejącą wiedzą.

Informacja ma specyficzny status wynikający przede wszystkim z tego, kto jest jej interpretatorem. Informacja może mieć inną treść dla bezpośredniego odbiorcy, a inną dla osób niezwiązanych z nią. Inną sprawą to kompetencje odbierających informacje. Osoby o szerszej wiedzy mogą pełniej zinterpretować otrzymywane informacji niż osoby o małym zasobie wiedzy ogólnej czy wyspecjalizowanej. W tym kontekście rangi nabiera rola wykształcenia i możliwości komunikowania za pomocą współczesnych narzędzi obsługi informacyjnej przedsiębiorstwa (informatyki).

Należy wyjaśnić ważne pojęcie, jakim jest wiadomość. D.T. Dziuba [2000, s. 24] definiuje to pojęcie jako „ciąg sygnałów, mających dla danego odbiorcy (nadawcy) określone znaczenie” [6]. W hierarchii wartościowania informacji najwyższą znajduje się wiedza. Potocznie ujmując, jest to określony (zakumulowany i wzbogacony doświadczeniem) zasób informacji. Dalej autor (za B.C. Brooke-sem) przedstawia interesującą relację między wiedzą i informacją:

$$[\Delta I] + [W] \Rightarrow [W + \Delta W],$$

gdzie:

ΔI – nowa informacja,

W – dotychczasowa wiedza,

ΔW – przyrost wiedzy (na skutek informacji).

Wzbogacanie wiedzy nie jest tu prostą transformacją ilości w jakość. Duża ilość danych wcale nie świadczy, „że więcej wiemy” lub precyzyjniej – „jesteśmy mądrzejsi”. Świadczy tylko, że mamy możliwość posługiwania się szerszym zasobem danych źródłowych. W umyśle ludzkim następuje nowa interpretacja dostarczonych (i zgromadzonych) danych, w wyniku czego powstają nowe kategorie informacji. Proces ten – zwany potocznie myśleniem – składa się faz, które najczęściej polegają na identyfikacji, analizie i interpretacji. Pewne znaczenie przypisuje się tu systematyzacji informacji, głównie w kontekście szybszego dostępu do zgromadzonych zasobów informacji (w informatyce – danych) i precyzyjnej ich interpretacji. Łatwo zauważyć, że w „przekazywaniu” informacji istotną rolę odgrywa jej treść. Możemy bowiem mniejszą ilością danych lepiej opisać złożone zjawisko, jeżeli użyjemy odpowiednio przygotowanego (bardziej nośnego) systemu komunikowania się (systemu znaków). Początkiem procesu powstawania wiedzy (informacje + kontekst + doświadczenie) są celowo generowane sygnały, które możemy sformalizować za pomocą znaków. Zrozumiały dla odbiorcy ciąg

tych znaków przyjęliśmy nazywać wiadomością. Formalizacja polega tu na dokładnym określeniu alfabetu języka, znaków, których można używać, na przykład w tekstach pisanych w tym języku (programach), określeniu dopuszczanego łączenia tych znaków – gramatyki języka. Pojęcie znaczenie znaku nie może być rozpatrywane w oderwaniu interpretatora (człowieka). Dla różnych ośrodków interpretacyjnych te same znaki mogą mieć różne znaczenie.

Są cztery główne kategorie wiedzy:

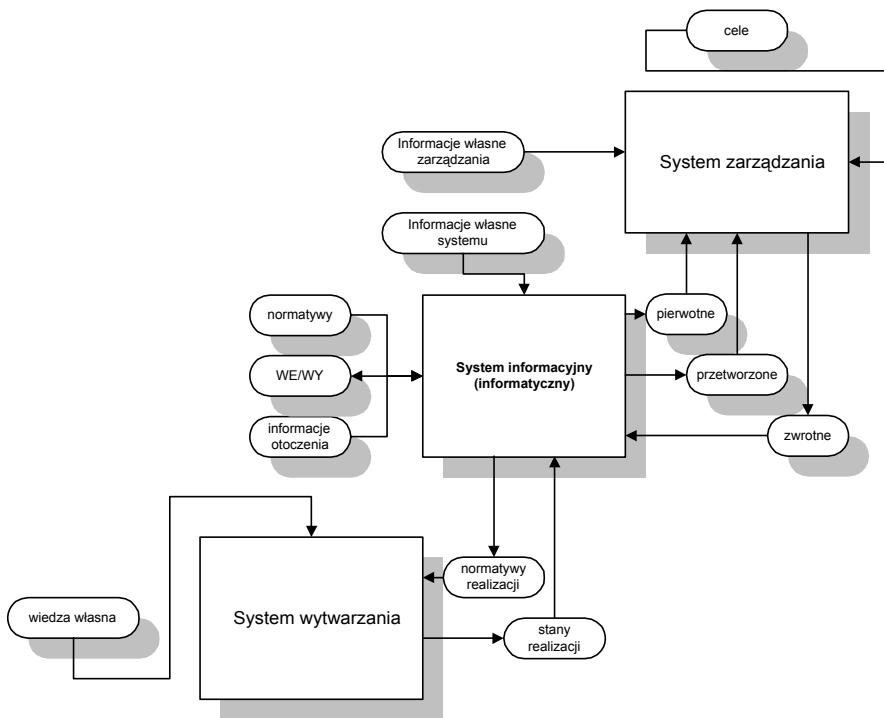
1. Wiedza typu *know-what*. Przykładem są fakty – ilu ludzi mieszka w Nowym Jorku. Wiedza ma tutaj znaczenie bliskoznaczne informacji – może być przesyłana za pomocą danych i bitów.
2. Wiedza typu *know-why*, która wyjaśnia rzeczywistość. Odnosi się ona do zasad i praw ruchu. Ten rodzaj wiedzy jest niezwykle istotny w pewnych obszarach nauki, na przykład w przemyśle chemicznym i elektronicznym. Dostęp do niej przyspiesza postęp i pozwala unikać błędów.
3. Wiedza typu *know-how* – odnosi się do umiejętności ludzi i zespołów, to znaczy do zdolności robienia czegoś. Określenie *know-how* („wiedzieć jak”) wywodzi się z sektora przemysłowego, gdzie było określeniem pewnych umiejętności i zdolności nieopisanych za pomocą patentów i licencji, lecz niezbędnych w momencie transferu technologii.
4. Wiedza typu *know-who*. Nowe produkty coraz częściej są oparte na wiedzy z wielu różnych obszarów, dziedzin i dyscyplin naukowych, co podnosi znaczenie dostępu do różnych źródeł wiedzy. *Know-who* określa posiadaczy wiedzy i opisuje ich wiedzę, a także dotyczy społecznych zdolności współpracy i komunikacji z ekspertami zewnętrznymi [17].

Podsumowując, należy powtórzyć, że wiedza różni się od informacji: „wiedza to zdolność do efektywnego działania”, podczas gdy informacja to zaledwie „znajomość rzeczy”. Informacją jest każdy czynnik, dzięki któremu ludzie lub urządzenia automatyczne mogą sprawniej działać celowo [13]. Z punktu widzenia zarządzania informacje to dane poddane obróbce w celu wykorzystania ich do podejmowania decyzji.

2. Miejsce systemu informacyjnego w przedsiębiorstwie

Przedsiębiorstwo jest celowo funkcjonującą całością (organizacją powołaną przez człowieka), mającą w swym składzie trzy zasadnicze podsystemy: zarządzający (decydenci i decyzje), wytwórczy (produkcja i usługi) i sterujący (komu-

nikacyjno-informacyjny). Właściwa współpraca tych trzech elementów decyduje o pozycji przedsiębiorstwa na rynku. Wielu autorów podkreśla, że warunkiem silnej pozycji przedsiębiorstwa na rynku jest dostęp do informacji. Można również dodać, że głównym punktem sprawnie działającej firmy jest właściciel firmy, a trafne decyzje wpływają na powodzenie wszystkich elementów przedsiębiorstwa.



Rys. 1. Miejsce systemu informacyjnego (informatycznego) w procesie zarządzania i wytwarzania

Źródło: opracowanie własne.

Podobnie można interpretować dobór profilu produkcji (lub usług) do potrzeb rynku. Coraz częściej przyjmujemy, że sukces organizacji zależy od postępów przedsiębiorstwa (system zarządzający) i usprawnianiu systemu informacyjnego. A. Nowicki wyraźnie wiąże systemem informacyjnym wszystkie elementy (podsystemy) przedsiębiorstwa (rysunek 1). Można dodać, że warunkiem sprawnego działania przedsiębiorstwa (silnej jego pozycji) jest dostęp do informacji

o realizowanej produkcji (i usługach) oraz szybka reakcja na zmiany otoczenia rynkowego. Wymagania takie współcześnie spełniają zintegrowane systemy informatyczne zarządzania klasy CIM (*Computer Integrated Manufacturing*). Przedstawiony układ organizacji procesów zarządzania nie wyczerpuje, rzecz jasna, poglądów na temat składowych elementów organizacji. Beynon-Davies wyróżnia następujące modele organizacji (i systemów informacyjnych):

- a) struktury (podsystemy) – organizacją jest rozpatrywana jako zestaw systemów działań ludzi, co oznacza, że składa się ona z kilku obszarów umożliwiających podział pracy między podsystemami pełniącymi określone funkcje (w interesie organizacji jako całości);
- b) sieci – organizacja jest rozpatrywana jako pewien nieformalny układ powstawania systemów informacyjnych, czyli równoległego projektowania, wdrażania i eksploatacji systemów społecznych i technicznych, z udziałem tak zwanych systemów miękkich;
- c) środowisko – organizacja jest rozpatrywana z punktu widzenia tak zwanego strategicznego systemu informacyjnego, umożliwiającego sprawne funkcjonowanie w powiązania z zewnętrznymi przepływami informacyjnymi, na przykład środowisko informacyjne funkcjonowania giełdy towarowo-pieniężnej;
- d) procesory informacji – organizacja jest traktowana z punktu widzenia zasadniczego mechanizmu (motoru) zarządzania, czyli podejmowania decyzji (podejście systemowe); decyzje gospodarcze tworzą najistotniejsze informacje dla trwania organizacji; elementami składowymi procesu informacyjno-decyzyjnego (konstrukcja procesora informacyjnego) jest planowanie, organizowanie, kontrolowanie, komunikacja i obsada stanowiska, wszystkie elementy tych działań opierają się na informacji i tworzą tak zwany krwiobieg organizacji;
- e) konstrukcje i producenci – organizacje, w których powstawanie systemów informacyjnych (informatycznych) wiąże się z określonym wymuszeniem komunikowania się ludzi, nie zastępując ich jednocześnie (zob. systemy klasy MRP) [4].

Na większość organizacji zasadniczy wpływ mają struktury (systemy), sieci i środowisko oraz procesory informacji. Przykładowo, strukturę funkcjonalną organizacji (na poziomie mikro) tworzą następujące podsystemy:

- a) marketing – działalność zewnętrzna przedsiębiorstwa, koncentrująca się na sprzedaży, ustalaniu cen i dystrybucji wytworzonych towarów i usług;

- b) księgowość i finanse – działalność zajmująca się obrotem kapitałowym (wraz z dokumentowaniem sposobu pozyskania i inwestowania środków pieniężnych);
- c) produkcja – obszar zajmujący się zakupami środków do produkcji, technologiami wytwarzania produktów gotowych i prowadzenia gospodarki materiałowej;
- d) badania i rozwój – poszukiwania usprawnień w zakresie technologii wytwarzania (i usług) oraz systemów informacyjnych opisujących rzeczywistość gospodarczą.

Model ten (lub podobne klasyfikacje w ujęciu systemowym) w istotny sposób wpływał (i nadal wpływa) na konstrukcję systemów informacyjnych (informatycznych) zarządzania, szczególnie w przedsiębiorstwach o niepotokowej organizacji produkcji. Pojawiają się jednak rozwiązania, w których coraz mniej zwraca się uwagę na zadania działów (gałęzi). Przykładowo, w podejściu procesowym BPR (*Business Process Reengineering*) większą wagę przywiązuje się do formalizacji międzygałęziowych przepływów informacyjnych i dopasowania procesów organizacyjnych do charakteru wykonywanej pracy. Szczególne cechy BPR to radykalizm w celu zwiększania efektywności ekonomicznej (wyniku finansowego – zyskowności). W praktyce jest to bardzo trudne do osiągnięcia ze względów personalnych i nieprzygotowania przedsiębiorstw do tego rodzaju gwałtownych zmian (w przyszłości powtarzalnych, o charakterze ciągłym).

Postępująca informatyzacja działalności gospodarczej każe traktować informację (systemy informacyjne) jako ważny element infrastruktury przedsiębiorstwa. Przykładem są tu systemy klasy MRP, które od prostych rozwiązań dla gospodarki materiałowej ewoluują w stronę silnie zintegrowanych systemów wytwarzania, dystrybucji i zarządzania kapitałami. Elementem integrującym te systemy jest informacja uprzedmiotowiona w systemie informatycznym o dużym stopniu złożoności. Rozwiązania MRP są raczej sposobem organizowania działalności przedsiębiorstwa, a nie jednorodnym narzędziem – systemem informatycznym. Jeszcze większą rolę odgrywa informacja w systemach klasy CRM (*Customer Relationship Management*). Dzięki zbieraniu danych o klientach – ich preferencjach i zachowaniach – możliwe jest lepsze dostosowywanie oferty przedsiębiorstwa. Stało się to możliwe dopiero po opracowaniu specjalistycznego oprogramowania, które po analizie zebranych informacji dostarcza praktyczną wiedzę.

3. Zarys metod ekstrakcji wiedzy z baz danych

W większości firm występuje poważny problem informacyjny związany z nadmiarem bezużytecznych danych zbieranych w systemach informatycznych i jednoczesnym brakiem możliwości dotarcia do ważnych informacji. Systemy księgowo, transakcyjne to systemy gromadzące dane dzień po dniu, wykonujące te same rutynowe działania i nastawione na masowe przetwarzanie. To właśnie one zdominowały dzisiejsze środowisko biznesowe. Nie są one jednak w stanie dostarczyć wiedzy niezbędnej do zarządzania. Powszechna potrzeba informacji zwiększa zapotrzebowanie na systemy dostarczające odpowiedzi na podstawowe pytania biznesu, nastawione na potrzeby użytkownika, zdolne wesprzeć długoterminową strategię i zapewnić konkurencyjną przewagę [5]. Obecna sytuacja na rynku, zdominowanym przez silną konkurencję, wymusza na organizacjach gospodarczych skupienie się na czterech zasadniczych obszarach:

- zwiększeniu przychodów dzięki lepszej wiedzy o wymaganiach klientów,
- lepszej obsłudze klienta,
- obniżce kosztów własnych działalności,
- zarządzaniu ryzykiem i innych ważnych aspektach zależnych od profilu przedsiębiorstwa.

Wszystkie te aspekty działalności przedsiębiorstwa mogą być wspomagane odpowiednimi systemami informatycznymi, które na podstawie danych gromadzonych przez systemy transakcyjne pozwoliłyby na efektywniejsze wykorzystanie istniejącego potencjału firmy i zdecydowane wsparcie procesu zarządzania. Zwłaszcza to ostatnie zagadnienie stało się domeną DSS (*Decision Support Systems* – systemy wspomagania decyzji), których działanie zaczyna opierać się na nowych rozwiązaniach, jakimi są hurtownie danych i metody ekstrakcji wiedzy.

Cechy danych, które były niemożliwe do osiągnięcia w tradycyjnych systemach transakcyjnych, spowodowały powstanie nowych mechanizmów i technik pozyskiwania informacji i wiedzy z danych gromadzonych w hurtowni. Do najważniejszych zaliczyć można:

- OLAP (*On-Line Analytical Processing*),
- przeszukiwanie w głąb (*Drill Down*),
- odkrywanie wiedzy (*Knowledge Data Discovery*),
- drążenie danych (*Data Mining*).

Dane zgromadzone w hurtowni danych są zoptymalizowane pod kątem ich wyszukiwania przez analityków wykorzystujących na bieżąco przetwarzanie analityczne (OLAP). W związku z tym dane są zorganizowane albo na podstawie wielowymiarowej bazy danych (MOLAP – *Multidimensional On-Line Analytical Processing*) lub relacyjnej bazy danych (ROLAP – *Relational On-Line Analytical Processing*) [20]. Systemy ROLAP charakteryzują się zdolnością do przechowywania danych o wielkiej objętości, względnie łatwą modyfikacją danych (wynikającą z zastosowanego oprogramowania i struktur danych) oraz negatywnymi cechami: złożonością struktur danych (wynikającą z konieczności relacyjnego odwzorowania zależności wielowymiarowych) i problemami z wydajnością, wynikającą z niedostosowania struktur relacyjnych do analizy wielowymiarowej. Systemy MOLAP mają, co prawda, mniejsze możliwości przechowywania danych i trudno w nich modyfikować dane (często modyfikacja danych prowadzi do przebudowy struktury wielowymiarowej), ale za to mają dużą wydajność analizy wielowymiarowej, a także naturalną reprezentację struktur wielowymiarowych. W obu przypadkach istotne jest przyśpieszenie wyszukiwania danych, tak aby analityk, który formułuje zapytanie, mógł szybko uzyskać odpowiedź. Podsumowując, rozwiązaniem to ma na celu zagwarantowanie możliwości szybkiego wykonywania aplikacji analitycznych, umożliwić łatwy dostęp do danych różnym użytkownikom przez zastosowanie intuicyjnego interfejsu użytkownika oraz umożliwić szybki czas reakcji systemu na założone pytania.

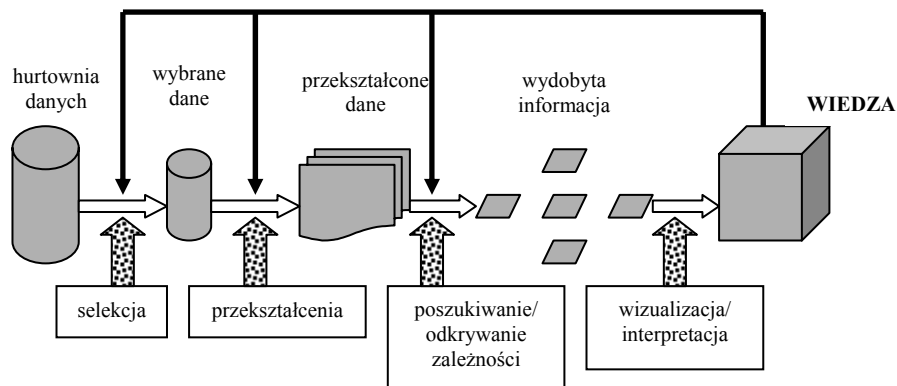
Podobnie jak w OLAP, koncepcja mechanizmu przeszukiwania w głąb opiera się na wielowymiarowej strukturze hurtowni danych. Obiekty poszczególnych wymiarów mogą być uporządkowane w pewne hierarchie, na przykład w wymiarze czasowym: dzień – tydzień – miesiąc – rok. Mechanizm przeszukiwania w głąb umożliwia dotarcie do szczegółów dotyczących faktów przez przechodzenie na coraz niższe poziomy agregacji, aż do danych o najmniejszej ziarnistości w ramach przyjętego modelu [10].

Ostatnie dwie techniki ekstrakcji wiedzy są ze sobą ściśle związane – drążenie danych (DM – *Data Mining*) jest składową odkrywania wiedzy (KDD – *Knowledge Data Discovery*). Pozyskiwanie wiedzy z baz danych jest stosunkowo młodą interdyscyplinarną dziedziną badań, łączącą ze sobą doświadczenia z dziedziny statystyki, baz danych oraz systemów uczących się i systemów odkryć. Przedmiotem badań dziedziny pozyskiwania wiedzy w bazach danych są nietrywialne procesy identyfikacji poprawnych, nowych, potencjalnie użytecznych i zrozumiałych regularności w danych bez potrzeby podawania z góry

explicite listy hipotez regularności [12]. W *Data Mining* stosuje się technologie sieci neuronowych, drzew decyzyjnych oraz standardowe techniki statystyczne do przeszukiwania dużych ilości danych. W procesie tym tworzone są modele, które przykładowo mogą służyć do przewidywania zachowań klientów [22]. Najprostszą definicją *Data Mining* jest automatyczne wykrywanie zależności w bazie danych. Przykładowo, może to być stwierdzenie, że prawdopodobieństwo prowadzenia określonego sportowego samochodu przez zamężne kobiety z dziećmi jest dwa razy większe niż przez bezdzietne mężatki. Oczywiście, przykład ten jest istotny dla producenta lub sprzedawcy samochodów, który może te informacje wykorzystać, odpowiednio kierując swoją ofertę. Samo poszukiwanie wiedzy nie jest, oczywiście, nowością – od lat statystycy przeszukiwali ręcznie zasoby baz danych w celu odnalezienia istotnych zależności. W *Data Mining* dodatkowo stosuje się techniki uczenia maszynowego i proces ten wykonuje automatycznie, wykorzystując ogromne skarbnice wiedzy, czyli hurtownie danych. Wciąż jednak uczestnictwo człowieka jest konieczne – odpowiednio wyszkolony analityk może podjąć decyzję o poprawności i użyteczności uzyskanego modelu oraz stopniu wykorzystania jego rezultatów [21]. Większość organizacji gospodarczych można śmiało określić jako „bogate w dane” z powodu ogromnych ilości danych o działalności i zasobach gromadzonych przez systemy operacyjne. Po przetworzeniu nadają się one do przedstawiania typowych faktów i wykresów, na przykład firma ma 200 klientów lub dostawca X zapewnia 60% surowca Y. Niestety, takie fakty nie dostarczają istotnej wiedzy i mogą prowadzić do przeładowania informacjami. Pomimo bogactwa danych, większość przedsiębiorstw jest więc „uboga w wiedzę”. Procesy KDD i DM mają wypełnić luki w „wiedzy” o działalności przedsiębiorstwa przez odpowiednie przetworzenie „bogactwa danych”. Odkrycie istotnych zależności w danych dotyczących przeszłych stanów organizacji gospodarczej pomoże polepszyć przyszłość dzięki wykorzystaniu pozytywnych powiązań oraz uniknąć tych niekorzystnych dla firmy. Przykładami zastosowań DM w praktycznej działalności przedsiębiorstwa są przewidywanie skali reakcji klientów na określoną formę marketingu, popytu na polisy ubezpieczeniowe w zależności od wielu czynników czy konsumpcji określonych produktów. Większość ludzi jest lepsza w wykrywaniu anomalii niż znajdowaniu związków i relacji w dużych zbiorach danych, dlatego odkrywanie wiedzy może być bardzo przydatne w działalności przedsiębiorstwa. Zamiast polegać na ludzkiej intuicji, można za pomocą odpowiedniego narzędzia wykryć, sprawdzić i wykorzystać różne powiązanie między badanymi zjawiskami (rysu-

nek 2) [11]. Popularność tego nowatorskiego rozwiązanie stale wzrasta, głównie z powodu trafności i przydatności uzyskiwanych rezultatów i coraz większej liczby dostępnych narzędzi. Przedsiębiorstwo decydując się na wprowadzenie metod KDD wraz z *Data Mining*, musi jednak rozważyć trzy podstawowe zagadnienia: metodologię, łatwość stosowania oraz reprezentację danych i skalowalność. Pierwsza pojęcie, czyli metodologia, dotyczy kroków realizacji projektu DM. Ich przestrzeganie ma na celu osiągnięcie podobnych korzyści przez przedsiębiorstwo, jakie udało się osiągnąć innym po wdrożeniu DM. Kolejność kroków ustalono następująco:

1. Analiza problemu – realizacja tego etapu ma odpowiedzieć na pytanie: czy dany problem może być badany za pomocą DM? Jeśli tak, to czy dostępne są odpowiednie dane i technologia DM oraz w jaki sposób rezultaty poszukiwań zostaną wykorzystane, biorąc pod uwagę całość rozwiązania?
2. Przygotowanie danych – etap polega na ekstrakcji odpowiednich danych i ich transformacji na wymagany format (agregacja, łączenie tabel, dodawanie pól, czyszczenie danych itd.).
3. Eksploracja danych – etap ten poprzedza moment poszukiwania powiązań i relacji między danymi. Przeprowadzana jest wizualizacja danych (tak aby użytkownik miał ich jasny obraz) oraz sprawdzanie, czy poprzednie etapy nie zawierały błędów.
4. Generowanie hipotez – za pomocą reguł wywoływania (automatycznych lub interaktywnych) oraz algorytmów odkrywania powiązań generowane są hipotezy, których poprawność jest rozważana, a następnie są one interpretowane.
5. Rozmieszczanie hipotez – etap ten polega na umieszczeniu uzyskanych hipotez w odpowiednich etapach analizy. Są one głównie stosowane w systemach SWD do generowania raportów lub filtrowania danych do dalszego przetwarzania.
6. Monitorowanie hipotez – główną przesłanką rozmieszczania hipotez jest założenie, że przyszłość przypomina przeszłość, więc hipotezy „historyczne” mogą mieć zastosowanie w przyszłych sytuacjach. Strategia ta jest jednak bezpieczna tylko w momencie stałego monitorowania hipotez „historycznych” na podstawie nowych danych i odpowiednio szybkiego wykrywania wahań. Zbyt duże odchylenia prowadzą do konieczności porzucenia dotychczasowych hipotez i poszukania nowych [1].



Rys. 2. Proces eksploracji baz danych

Źródło: C. Głowiński, *Sztuka wysokiego składowania*, „PC Kurier” 2000, nr 12.

Drugim pojęciem związanym z wprowadzaniem DM w przedsiębiorstwie jest łatwość stosowania. Wynika to z faktu, że użytkownicy tego typu programów nie muszą być ekspertami, dlatego trudności ich używania powinna być na poziomie arkusza kalkulacyjnego. Ponadto powinny one obsługiwać wszystkie etapy metodologii DM, szczegółowe dane oraz raportowanie hipotez i ich wizualizację [11]. Zwłaszcza ta ostatnia opcja jest istotna dla użytkownika. Proces DM wykrywa „ukryte” informacje w bazie danych, więc samo zrozumienie procesu może się wydać skomplikowane. Dopiero odpowiednie przedstawienie wyszukanych hipotez za pomocą wykresów, grafów i trafnych podsumowań pomoże użytkownikowi zrozumieć i przenieść odnalezioną zależność na grunt praktyczny [9]. Trzecim aspektem, istotnym dla wprowadzania rozwiązań w organizacjach gospodarczych, jest reprezentacja danych i skalowalność. Rosnący koszt przechowywania i przetwarzania danych dotyczących działalności jest już problemem firm o dowolnej wielkości – od małych do dużych. Gigabajty danych operacyjnych są obecnie normalnym zjawiskiem. Istotne jest więc, aby narzędzie służące do DM mogło obejmować tak duże zbiory danych niezależnie od platformy czy architektury systemu informatycznego w przedsiębiorstwie [16].

Podsumowując, KDD (*Knowledge Discovery in Databases*) jest nietrywialnym procesem identyfikowania ważnych, nowych, potencjalnie użytecznych i zrozumiałych wzorców w danych [16]. Proces odkrywania wiedzy polega na bardzo intensywnym współdziałaniu człowieka i maszyny. To człowiek (analityk) określa, które regularności są interesujące, nowe i zrozumiałe. Samo pozy-

skiwanie wiedzy z danych (*Data Mining*) jest jednym z kroków tego procesu, produkującym hipotezy regularności w bazie danych. Obecnie na tym kroku koncentrują się badania w dziedzinie KDD. Należy pamiętać, że odkrywanie nie jest ich prostą sekwencją, lecz raczej procesem iteracyjnym, z możliwością nawracania do poprzednich etapów praktycznie po każdym kroku [12].

4. Znaczenie informacji w nowej gospodarce

Wraz z ekspansją Internetu w sferę gospodarki wzrasta liczba przedsiębiorstw decydujących się na rozwiązania e-biznesowe w ramach systemów informacyjnych. Od jakiegoś czasu panuje moda na zarządzanie informacjami. Jednak zebrane informacje muszą być przetworzone w skuteczne działania. Portal Yahoo! zapisuje każde kliknięcie myszką przez użytkownika portalu. Dzięki temu powstaje 400 mld bajtów danych dziennie – odpowiednik 800 tys. książek. Czy znaczy to, że Yahoo! ma dużą wiedzę o swoich klientach? Firmy często dużo wiedzą o klientach, lecz ich nie znają. Muszą one zwrócić uwagę nie tylko na dane pochodzące z transakcji, lecz na ich kontekst i właściwe wykorzystanie. Wiedza jest podstawowym zasobem przedsiębiorstwa, ponieważ integruje wszystkie pozostałe zasoby [7]. W pewnym momencie jednak dostrzeżono brak narzędzi, które pozwoliłyby na pełne wykorzystanie uzyskiwanych informacji w ramach systemów wspomaganie decyzji (DSS). Lukę tę szybko zauważyli dostawcy oprogramowania, tworząc coraz nowsze i bardziej zaawansowane „inteligentne” narzędzia integrujące rozwiązania elektronicznego biznesu z istniejącymi aplikacjami wspomagającymi zarządzanie. Umożliwiają one użytkownikom wewnętrznym, partnerom handlowym oraz klientom przedsiębiorstwa natychmiastowy dostęp do pożądaných informacji, aplikacji i usług. Systemy *e-intelligence* umożliwiają między innymi:

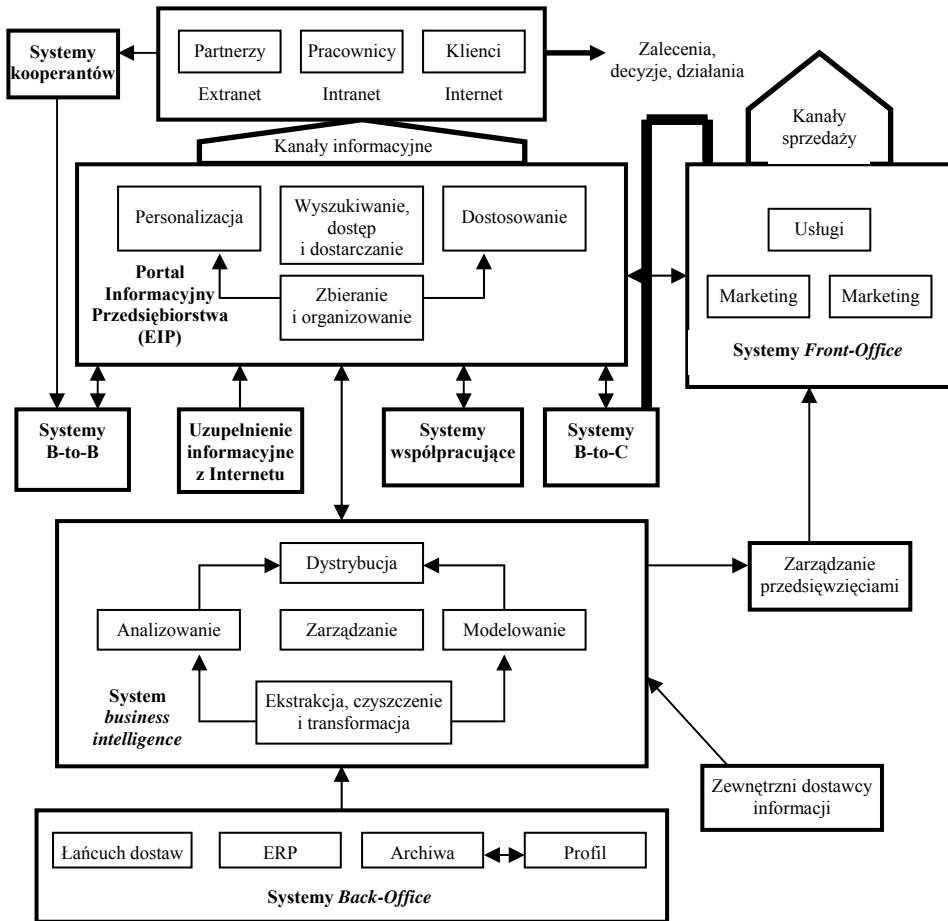
- a) integrację operacji e-biznesowych z tradycyjnym środowiskiem biznesowym, dającą użytkownikowi kompletny obraz działania przedsiębiorstwa;
- b) podejmowanie racjonalnych decyzji opartych na kompletnych informacjach uzyskiwanych z aplikacji e-biznesowych, co umożliwia użytkownikom optymalizację oferty opartej na sieci WWW (sprzedaży produktów, promocji, usług i wsparcie techniczne itp.) w taki sposób, by zaspokoić wymagania rynku oraz być konkurencyjnym na tle pozostałych przedsiębiorstw;

- c) segmentację i utworzenie profili e-klientów na podstawie analizy ich zachowań opartej na informacjach zebranych z aplikacji e-biznesowych; w ten sposób przedsiębiorstwo może „personalizować” strony WWW dla poszczególnych typów klientów i lepiej dostosować swoją ofertę do wymagań odbiorcy;
- d) rozszerzenie środowiska „inteligentnego” systemu poza przedsiębiorstwo i elektroniczną współpracę (opartą na dzieleniu informacji) z partnerami handlowymi, co powoduje optymalizację łańcucha dostaw i obniża koszty współpracy;
- e) rozszerzenie środowiska „inteligentnego” systemu poza przedsiębiorstwo i elektroniczną współpracę z klientami; udostępnienie klientowi informacji o jego zakupach oraz wiążących się z nimi korzyściami itp. z pewnością podniesie poziom jego satysfakcji, a tym samym przyczyni się do zacieśnienia więzi z przedsiębiorstwem.

Na rysunku 3 przedstawiono strukturę biznesową i technologiczną zintegrowanego e-inteligentnego środowiska. Składa się ona z dwóch głównych elementów: systemu *Business Intelligence* oraz portalu informacyjnego przedsiębiorstwa (EIP – *Enterprise Information Portal*). Przetwarzanie w systemach *Business Intelligence* wymaga zastosowania narzędzi „ekstrakcji–transformacji–załadowania” (*extract–transform–load tools*) lub własnych aplikacji przedsiębiorstwa w celu ekstrakcji danych z systemów *Back-Office*, ich transformacji i integracji z pozostałymi danymi przedsiębiorstwa. Proces ten ma zapewnić kompleksowe dane dla narzędzi wspomagania decyzji. Zwykle tak przetworzone dane umieszczane są w hurtowni danych [2]. Dalsze ich wykorzystanie odbywa się za pomocą metod ekstrakcji wiedzy opisanych w poprzednim punkcie.

Systemy *business intelligence* zaczynają cieszyć się bardzo dużą popularnością. Są bowiem ostatecznym etapem pełnego wykorzystania informacji uzyskiwanych dzięki rozwiązaniom e-biznesowym. Jednak samo narzędzie czasami nie wystarcza – jest jedynie wstępem do zarządzania wiedzą. Zaawansowane technologie „wydobycia” wiedzy oraz dzielenie się informacjami z kooperantami i klientami budzi pewien niepokój użytkowników, chodzi głównie o wiarygodność tych informacji. Konieczna jest zmiana mentalności użytkowników w dwóch aspektach:

- a) udostępniając własny dorobek, mam prawo korzystać z czyjegoś;
- b) uznanie, że cudza wiedza jest tak samo wartościowa jak własna aktywność intelektualna użytkownika.



Rys. 3. Struktura biznesowa i technologiczna zintegrowanego e-inteligentnego środowiska

Źródło: opracowania własne na podstawie [23].

Szybkość zmian warunków rynku i wiążące się z tym dostosowywanie przedsiębiorstwa do jego wymagań wręcz wymusza korzystanie z tego, co zrobili już inni. Czasami drobna modyfikacja rozwiązania, które powiodło się w przedsiębiorstwie współpracującym, umożliwi znaczne oszczędności czasu i kosztów przeprowadzanych operacji. Jednocześnie minimalizowane jest ryzyko rozwiązania i zyskiwany czas na planowanie własnych innowacji. Takie podejście jest obecnie stosowane na szczeblach zarządczych firmy Hewlett-Packard [2].

Rosnące znaczenie informacji i wiedzy w elektronicznej ekonomii zrodziło powstanie nowego pojęcia – „kapitał wiedzy”. Jego autorem jest Paul Strassman, który zauważając, że tradycyjne, księgowe metody wyceny przedsiębiorstw nie są w stanie uchwycić jego realnej wartości, określił „kapitał wiedzy” jako szacunkową wartość wszystkich środków generujących zysk, przekraczających koszty finansowania kapitałowego. Całkowita wartość firmy jest zatem równa sumie kapitału finansowego oraz „kapitału wiedzy”. Niedoskonałość tradycyjnych metod wyceny wynika z tego, że w rozwiniętych gospodarkach środki finansowe stanowią mniej niż dziesięć procent wartości największych korporacji. Stosunek zysków do tradycyjnej wartości środków finansowych (finansisci określają ją również jako „wartość księgową”) daje współczynnik (jak właśnie zwrot z inwestycji czy zwrot z kapitału) o zbyt małym nominale. W ten sposób wartość powszechnie stosowanych współczynników jest zaniżana. Według Strassmana, właściwą metodą mierzenia wydajności firmy w dobie gospodarki cyfrowej jest porównanie stosunku wzrostu trwałej wartości dodanej do kosztów zarządzania informacją (np. wydajności informatycznej) z wynikami konkurencji [18]. Według firmy konsultingowej Arthur Andersen, udział czynników pozamaterialnych w wartości giełdowej przedsiębiorstw amerykańskich wzrósł od 1978 roku do połowy lat 90. XX wieku z 5% do 80%. Ocenę tę potwierdza wyraźnie fakt, że kursy akcji nowych przedsiębiorstw na przykład z branży oprogramowania rosły nieporównywalnie szybciej niż tradycyjnych koncernów, aczkolwiek te mają znacznie większy majątek fizyczny [26].

5. Uwagi końcowe

Podsumowując, należy zauważyć, że informacja niezaprzeczalnie stała się czynnikiem produkcji. Nowe technologie ułatwiają firmom ekspansję i sprawiają, że rynki stają się bardziej konkurencyjne i efektywne. Jednocześnie w mikrodziałalności gospodarczej rośnie presja na podnoszenie efektywności, a szybkość przemian pociąga nieustannie i wywołuje ryzyko. W tych warunkach szczególnie znaczenie mają aktywa niematerialne (*intangibles*), jak na przykład siła marki, która jest wyznacznikiem renomy firmy i decyduje o zaufaniu klientów. Przed wszystkim jednak o wartości przedsiębiorstwa decyduje teraz kapitał intelektualny – jego pracownicy. Zakres i jakość wiedzy, jaką dysponuje przedsiębiorstwo, przesądza o jego przyszłości. Stale aktualizowany i powiększany zasób wiedzy ma tę zaletę, że nie traci, ale wręcz przeciwnie – zyskuje na wartości. Stosując

strategię ucieczki do przodu, firma jest w stanie utrzymać stałą przewagę konkurencyjną [14]. Autorytet z zakresu zarządzania, Peter Drucker, twierdzi, że wiek XX był czasem maksymalizowania wydajności pracowników fizycznych. Osiągnięto pięćdziesięciokrotny wzrost ich wydajności. Jego zdaniem, w zwiększaniu wydajności pracowników umysłowych jesteśmy dzisiaj w tym punkcie, w którym świat był w 1900 roku, jeśli chodzi o wydajność pracowników fizycznych. Jeśli ma rację, to trudno obecnie sobie wyobrazić, jak będzie wyglądała aktywność pracownika umysłowego za 100 lat [2]. Na razie zrozumiano jedynie to, że o wydajności pracownika, a tym samym o rosnącej wartości przedsiębiorstwa nie przesądzają szybkość pracy i pojemność jego umysłu, ale dwie inne rzeczy: organizacja pracy i dostępne narzędzia wspomagające. Obie sprowadzają się do bezproblemowego używania w swojej pracy wiedzy wytworzonej przez aplikacje komputerowe (lub czasami innych użytkowników) i zgromadzonej w różnych skarbnicach oraz do użyczenia tej wiedzy zainteresowanym podmiotom (klientom, kooperantom).

Literatura

1. Al-Attar A., *Data Mining – Beyond Algorithms*, <http://www.attar.com/tutor/mining.htm>.
2. Bartczak I., *Wiedza jak zatopiony skarb*, <http://www.cxo.pl/artykuly/20764.html>.
3. Belkin N.J., Robertson S.E., *Information Science and the Phenomenon of Information*, „Journal of the American Society for Information Science” 1976.
4. Beynon-Davies P., *Inżynieria systemów informacyjnych*, Warszawa 1999.
5. Czarnota J., Wiecka A., *Hurtownie danych w Polsce – jakie problemy rozwiązują i komu służą?* http://www.software.com.pl/pisma/software/artykuly/Software2000/Sf1_2000/hurtownie_danych.html.
6. Dziuba D.T., *Gospodarki nasycone informacją i wiedzą*, WNE UW, Warszawa 2000.
7. Fazłagić A., *Marketing a zarządzanie wiedzą*, Akademia Ekonomiczna w Poznaniu, Poznań 2000.
8. Fuglewicz P., *Zarządzanie dokumentami a strategie biznesowe XXI wieku*, Warszawa 2001.
9. Gilman M., *D.Nuggets® and Data Mining*, <http://www.data-mine.com/NugWhitePaper.htm>.
10. Gorawski M., Koziątek A., *Data Warehouse – ekstrakcja danych*, Software 1999, nr 9.

11. Karwański M., *Wiedza brana hurtem*, <http://www.sas.com/offices/europe/poland/prasa/p5.html>.
12. Kłopotek M.A., *Kierunki rozwoju systemów komputerowego wspomagania odkryć*, <http://alpha.mini.pw.edu.pl/~klopotek/suiwon/DataMining.html>.
13. Lucas R., *Econometric Policy Evaluation: A Critique*, CROCH 1976.
14. Marczuk M., *Krótką charakterystyka Nowej Gospodarki*, Wydawnictwo św. Pawła, Częstochowa 2001.
15. Morbitzer J., *Od motyki do komputera, czyli droga do społeczeństwa informacyjnego*, <http://www.wsp.krakow.pl/Plasc/konspekt/konspekt8/morbitzer8.html>.
16. Nadolny M., *Data mining and knowledge discovery*, <http://rainbow.mimuw.edu.pl/~mnadolny/datamining.html>.
17. Naisbitt J., *Megatrends: Ten New Directions Transforming Our Lives*, Warner Books, New York 1982.
18. *Nowa gospodarka to nie IT*, <http://www.dcom.pl/artykul.asp?idart=344>.
19. Stamper R.K., *Information: Mystical Fluid or a Subject for Scientific Enquiry*, „The Computer Journal” 1985.
20. Stanek S., *Hurtownia danych*, <http://figaro.ae.katowice.pl/~stanek/wareh97/warehouse.html>.
21. Thearling K., *Increasing customer value by integrating Data Mining and Campaign Management software*, <http://www3.shore.net/~kht/text/integration/integration.htm>.
22. Thearling K., *Data Mining and Customer Relationships*, <http://www3.shore.net/~kht/text/whexcerpt/whexcerpt.htm>.
23. White C., *Instant Intelligence*, <http://www.intelligententerprise.com/000410/feat1.shtml?ecommerce>.
24. Wiener N., *Cybernetics or Control and Communication in the Animal and the Machine*, MIT Press 1961.
25. Zalega T., *Kierunki rozwoju rynku pracy w epoce cywilizacji informacyjnej*, WUW, Warszawa 2002.
26. *Zarządzanie na świecie*, http://www.oditk.com.pl/zarządzanienas5_8.php3.

**INFORMATION AS KEY COMPONENT
OF ENTERPRISE „NEW ECONOMY”**

Summary

In the article information is presented as a crucial production factor of modern enterprises. There are explained basic term and definitions connected with information. Additionally there is identified a place of IT system in organization. Final part presents a role of information in „New Economy” and its share in enterprise value.

Translated by Ryszard Budziński, Tomasz Ordasiński

