

Współczesna Gospodarka



Contemporary Economy
Electronic Scientific Journal
www.wspolczesnagospodarka.pl

Vol. 7 Issue 2 (2016) 1-19
ISSN 2082-677X

INŻYNIERIA WIEDZY STYMULATOREM PRZEDSIĘBIORCZOŚCI W DOBIE NOWEJ GOSPODARKI

Krzysztof Ficoń

Streszczenie

W artykule przedstawiono informacyjne wyzwania przedsiębiorstw nowej gospodarki w zakresie efektywnego wykorzystania zasobów informacyjnych zgromadzonych w systemach teleinformatycznych. Współczesne przedsiębiorstwa zalewane są lawiną informacyjną w postaci elementarnych, bezkontekstowych danych źródłowych o niewielkiej przydatności w procesie podejmowania decyzji biznesowych. Zjawisku temu mają przeciwdziałać nowoczesne technologie ekstrakcji danych i kompresji informacji rozwijane m.in. w ramach inżynierii wiedzy (*Knowledge Engineering*). Szczegółowo zostały omówione trzy wybrane metody OLAP (*On-Line Analytical Processing*), DM (*Data Mining*) i BI (*Business Intelligence*) coraz częściej stosowane na wirtualnych platformach *Cloud Computing*. Inteligentna selekcja informacji źródłowej i dedykowana personalizacja wiedzy biznesowej stanowi ogromne wyzwanie pod adresem całego sektora IT i wymaga radykalnych rozwiązań koncepcyjno-technologicznych.

*Gdzie się podziła nasza mądrość, którą zastąpiła wiedza,
gdzie się podziła nasza wiedza, którą zastąpiła informacja?*
Thomas S. Eliot¹

Słowa kluczowe: dane, informacja, informatyka, inteligencja, inżynieria, wiedza.

Wstęp

Od przełomu XX i XXI wieku świat żyje w erze rozwiniętego społeczeństwa informacyjnego, którego głównym paradygmatem jest symboliczna informacja,

¹ Thomas Stearns Eliot (1888 - 1965) – amerykańsko-brytyjski poeta, dramaturg i eseista. Laureat Nagrody Nobla w dziedzinie literatury za rok 1948.

prakseologiczna wiedza i utylitarna technologia. Zgodnie z tofflerowską koncepcją trzech fal rozwoju społeczno-cywilizacyjnego² obejmującą chronologicznie erę społeczeństwa agrarnego, industrialnego i informacyjnego, aktualnie punkt ciężkości został przeniesiony z materii (ziemia), energii (kapitał) na symboliczną informację, której zasoby są dynamicznie rozwijane i praktycznie nieograniczone. Bardzo trafnie ideę społeczeństwa informacyjnego ujął D. Bell, sywierzając: „...najistotniejsze w społeczeństwie poprzemysłowym jest to, że wiedza i informacja staje się źródłem strategii i przemian społeczeństwa, czyli tym samym co praca i kapitał w społeczeństwie przemysłowym”³.

Idea społeczeństwa informacyjnego zakwitła w Japonii jako *Johoka Shakai* na początku lat 60. XX w. za sprawą Tadao Umesao⁴. Pozbawiony niemal wszelkich bogactw naturalnych kraj skierował swój wysiłek intelektualny na niematerialną informację, kreatywną wiedzę i pragmatyczną technologię, inwestując w kapitał intelektualny. Komercjalizacja idei *Johoka Shakai* nastąpiła jednak na kontynencie amerykańskim, początkowo w nurcie badań militarnych, w ramach programu budowy ogólnoswiatowej sieci łączności ARPA-NET (*Advanced Research Projects Agency Network*) dla potrzeb dowodzenia siłami strategicznymi USA na wypadek totalnej wojny światowej (atomowej). Dla zagwarantowania odpowiednio sprawnej i niezawodnej łączności, pozbawionej krytycznych węzłów centralnych została zaprojektowana w połowie lat 60. XX w. elastyczna, rozległa i rozproszona sieć komputerowa ARPA-NET. Zasadniczym walorem użytkowym tej sieci wyróżniającym ją spośród wielu innych sieci komputerowych jest genialny protokół transmisji danych TCP/IP (*Transmission Control Protocol/Internet Protocol*) gwarantujący wysoką sprawność techniczną oraz niezawodność i bezpieczeństwo transferu danych.

Stosunkowo szybko potencjalne możliwości sieci ARPA-NET na nowo odkryli amerykańscy naukowcy, nauczyciele akademicki i studenci, bo już pod koniec lat 60. XX w. przodujące uniwersytety połączyła kontynentalna sieć komputerowa INTER-NET, a koordynatorem całego projektu był słynny amerykański instytut technologiczny MIT (*Massachusetts Institute of Technology*). Po latach zastanawiające jest, że zarówno amerykański, jak też światowy biznes dość długo, bo aż do końca lat 70. XX w. z wielką rezerwą odnosił się do możliwości komercyjnego wykorzystania Internetu dla potrzeb gospodarczych. Nieufność ta wynikała głównie z otwartości sieci Internet i powszechnego dostępu różnych użytkowników do otwartych danych biznesowych, co godziło w odwieczną zasadę tajemnicy handlowej.

W początkowym okresie rozwoju Internetu także wielkie światowe firmy komputerowe takie jak, np. IBM (*International Business Machines*), NCR (*National Cash Register*), HP (*Hewlett-Packard*), *Apple Computer*, *Sun Microsystems*, *Digital Equipment Corporation*, *Compaq Computer Corporation* traktowały Internet bardziej jako chwilową modę i rozrywkę, niż komercyjną technologię dla całego sektora IT (*Information Technology*). Pogląd ten zmienił dopiero Bill Gates i jego *Microsoft Corporation*, wprowadzając na rynek kolejne wersje graficznego systemu operacyjnego WINDOWS, który zrewolucjonizował kontakt z komputerem i wprowadził komputery pod strzechy⁵. Ogromny sukces Windows'a wynikał głównie z wielkiej liczby użytkowników pracujących coraz częściej w sieci Internet. Na masowe upowszechnienie technologii komputerowej oprócz systemu operacyjnego Windows miała wpływ także sieć Internet, a przede wszystkim ogromny postęp naukowo-techniczny w obszarze mikroelektroniki (układy scalone, dedykowane chipy, mikroprocesory, pamięci masowe, komputery osobiste, urządzenia mobilne, w tym telefony komórkowe itp.).

² A. Toffler, *Trzecia fala*. Wyd. Kurpisz, Warszawa 1997. (*The Third Wave*, USA 1980).

³ D. Bell, *Technika łączności*. Przegląd Zagranicznej Literatury Progностycznej. Seria Komitet PAN Polska 2000.

⁴ T. Umesao, *Joho sangyuo ron*. "Hoso Asahi" 1963.

⁵ Przełomowym osiągnięciem był wprowadzony w roku 1995 przez firmę Microsoft system operacyjny Windows 95, który w dodatku PLUS zawierał przeglądarkę internetową Internet Explorer.

Omawiając rozwój Internetu nie sposób nie wspomnieć o genialnym programiście z CERN (*European Organization for Nuclear Research*) - *Timothy Berners-Lee*, który opracował metajęzyk HTML (*Hyper Text Markup Language*). Dał on początek rewelacyjnej przeglądarce linkowej WWW (*World Wide Web*), umożliwiającej łatwą penetrację zasobów informacyjnych sieci Internet. Niekończąca się pajęczyna WWW pozwala dziś każdemu użytkownikowi na bezkresne surfowanie niemal po wszystkich jego zasobach informacyjnych.

Zaawansowana i powszechnie dostępna zminiaturyzowana technologia komputerowa stała się platformą sprzętową Internetu, a tym samym podstawowym narzędziem społeczeństwa informacyjnego, którego "medium roboczym" jest niematerialna i nieograniczona informacja, która odpowiednio przetworzona, właśnie za pomocą komputerów, pozwala na budowę piramidy kreatywnej wiedzy i prakseologicznej technologii. Dla potrzeb efektywnego wykorzystania możliwości komercyjnych Internetu rozwinęła się m.in. inżynieria wiedzy (*Knowledge Engineering*), jako praktyczne narzędzie eksploracji lawiny informacyjnej zgromadzonej m.in. w przepastnych zasobach współczesnego Internetu.

Komercyjny *Business*, choć dość późno odkrył fenomen Internetu, skorzystał z jego dobrodziejstwa zadziwiająco skutecznie i wszechstronnie. Po pierwsze, Internet stał się podstawą rozwoju nowej formacji społeczno-gospodarczej jako globalizacji, a po drugie, w dobie obecnej każdy biznes - mały i duży, bez względu na skalę i zakres prowadzonej działalności gospodarczej musi funkcjonować w sieci Internet. Bezpośrednim efektem tych ścisłych związków biznesu z Internetem było powstanie tzw. gospodarki elektronicznej w postaci *e-Economy*, czy *e-Business*, odnoszonej także do kategorii nowej gospodarki *New-Economy*.

Kluczowym dobrem gospodarczym nowej gospodarki jest symboliczna informacja i budowana na jej podstawie specjalistyczna wiedza biznesowa gwarantująca odpowiednią konkurencyjność e-przedsiębiorstw na globalnych rynkach. W dobie nowej gospodarki prowadzenie działalności gospodarczej odbywa się w skali globalnej wioski⁶ na otwartych księgach finansowych. Bardzo dobitnie o strategicznym znaczeniu informacji w działalności biznesowej wyraził się Bill Gates, kiedy stwierdził, że "...najbardziej znaczącym sposobem odróżniania twojej firmy od konkurencji, najlepszą metodą na pozostawienie tłumu w tyle jest odpowiednie korzystanie z informacji"⁷. Informacja i budowana na jej podstawie wiedza jest dziś najważniejszym zasobem strategicznym każdego przedsiębiorstwa oddzielnie i całej gospodarki globalnej rozpatrywanej w ujęciu holistycznym. Niematerialna informacja w XXI wieku stała się substytutem wszelkich dóbr gospodarczych, a potencjał intelektualny pracowników wyznacznikiem pozycji rynkowej firmy i niezastąpionym atrybutem konkurencyjności rynkowej.

Dla potrzeb niniejszej pracy pojęcie inżynierii wiedzy zostało zawężone do jej narzędziowych aplikacji odpowiedzialnych za integralność i spójność danych (bazy i hurtownie danych), za szybkie i skuteczne wyszukiwanie żądanych informacji (metody ekstrakcji informacji) oraz za inteligentne opracowywanie syntetycznych analiz i sprawozdań (*Business Intelligence*), w szczególności za elastyczną i semantyczną personalizację obsługiwanych przez sektor IT zgłoszeń i żądań użytkowników (*Cloud Computing*), w tym przypadku dotyczących działalności biznesowej.

⁶ M. Mc Luhan, *Zrozumieć media: Przedłużenia człowieka*. WN-T Warszawa 2004, (The Extensions of Man, 1964).

⁷ B. Gates, *Biznes szybki jak myśl*. Prószyński i S-ka, Warszawa 2001, s. 25.

1. Triada - dane, informacja, wiedza

Kardynalne znaczenie dla nowej gospodarki, nazywanej też gospodarką opartą na wiedzy ma pojęcie informacji, rozpatrywane nie tylko w aspekcie komputerowym (informatycznym). Naukowe podstawy ilościowej teorii informacji dla potrzeb rodzącej się telekomunikacji zostały wypracowane przez C. Shannona w roku 1948⁸. W tym sensie elementarną jednostką informacji jest bit potrzebny do zakodowania elementarnej wiadomości, które z dwóch możliwych zdarzeń np. prawda lub fałsz aktualnie wystąpiło. Teoria informacji została także wykorzystana w pracach N. Wienera przy formułowaniu podstaw rodzącej się cybernetyki - nauki o sterowaniu w organizmach żywych i urządzeniach technicznych. Intuicyjnie w potocznym rozumieniu informacją nazywamy dane (w dowolnej postaci - numerycznej, tekstowej, graficznej, dźwiękowej, smakowej, dotykowej itp.), dzięki którym zmniejsza się stopień niewiedzy odbiorcy oraz takie które wnoszą do jego świadomości element nowości lub prakseologicznej użyteczności.

Według *Słownika informatycznego* informacja to „...wiadomość dotycząca fragmentu rzeczywistości lub go zastępująca, czynnik fizyczny powodujący taką wiadomość, możliwy do utrwalenia i podatny na przetwarzanie”⁹. Natomiast *Leksykon naukowo-techniczny* pojęcie informacji określa jako: „...każdy czynnik, dzięki któremu obiekt odbierający go (człowiek, organizm żywy, organizacja, urządzenie automatyczne) może polepszyć swoją znajomość otoczenia i bardziej sprawnie przeprowadzić celową działalność”¹⁰. Z pojęciem informacji ściśle wiąże się entropia informacyjna, która symbolizuje systemowe nieuporządkowanie, które można zmienić za pomocą dodatkowej informacji pobieranej z otoczenia. Jest ona ściśle związana z ilością informacji zawartej w odebranych komunikacie. Formalnie entropia informacyjna to „...miara nieokreśloności zdarzeń stanowiących źródła informacji przy określonym stanie niewiedzy o tych zjawiskach”¹¹. W tym kontekście można interpretować wypowiedź C. Shannon - ojca teorii informacji kiedy stwierdził, że „...informacja jest wszystkim co osłabia niepewność”¹².

Dokładne wyjaśnienie interdyscyplinarnego i bardzo popularnego pojęcia informacji wymaga zdefiniowania pewnego kontekstu związanego z pierwotnym pojęciem danych. Tradycyjnie w literaturze występuje znana powszechnie 3-poziomowa piramida informacyjna obejmująca sekwencyjnie: dane, informacje, wiedzę (rys.1). W hierarchii wartościowania informacji poziom elementarny stanowią bezkontekstowe, niemianowane dane, natomiast najwyżej znajduje się syntetyczna, kreatywna wiedza. Najczęściej piramidę tą wieńczy mądrość, jako symbol perfekcyjnej wiedzy i wysokiej indywidualnej inteligencji człowieka lub zbiorowej inteligencji, np. przedsiębiorstwa.

Uogólniając powszechnie występujące definicje danych, informacji i wiedzy można sformułować następujące konkluzje:

- dane to elementarne (suche) fakty, których jednostką jest jeden lub kilka symboli użytych do opisu czegoś lub kogoś,
- informacja to zinterpretowane w określonym kontekście dane, będące efektem celowego przetwarzania (kompilacji),
- wiedza reprezentuje pewną superpozycję uogólnionej informacji nałożonej na istniejący dotychczas stan metawiedzy.

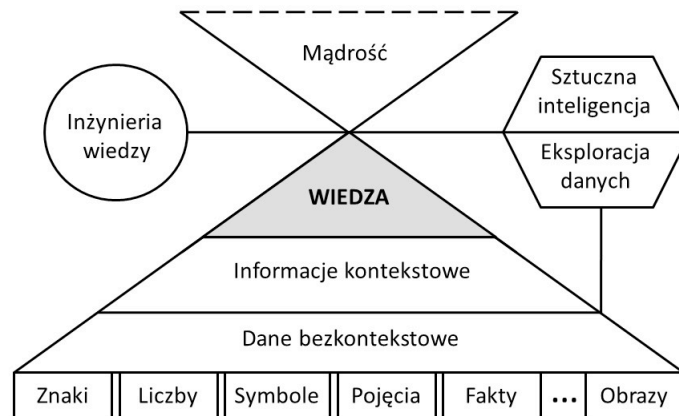
⁸ C.E. Shannon, *A Mathematical Theory of Communication*. “The Bell System Technical Journal”, Vol. 27, 1948.

⁹ Z. Płoński, *Słownik informatyczny*. EUROPA Wrocław 2003, s. 96.

¹⁰ *Leksykon naukowo-techniczny*. WNT Warszawa 1984, s. 306.

¹¹ M. Kempisty (red.), *Słownik cybernetyczny*. WP Warszawa 1973, s. 105.

¹² C.E. Shannon, *A Mathematical Theory of Communication...*, op. cit.



Rysunek 1. Miejsce inżynierii wiedzy w piramidzie informacyjnej

Źródło: Opracowanie własne.

Wiedzę można interpretować jako pewien obiekt budowlany (gmach), który zbudowany jest z elementarnych cegiełek reprezentowanych przez wszelkiego rodzaju dane. Najczęściej obiekty budowlane mają bardzo wyszukaną konstrukcję architektoniczną i składają się z różnych struktur, takich jak: fundamenty, ściany, sufity, elewacja, okna, drzwi itp. które będziemy utożsamiać z odpowiednimi strukturami (systemami) informacji. Informacja jest więc celowo preparowaną kompozycją danych przygotowaną na potrzeby gmachu wiedzy. Poszczególne struktury informacyjne tworzone są bardzo często na bazie pewnych standardów i dotychczasowego doświadczenia. Podobnie symboliczny gmach wiedzy tworzymy nie zawsze od przysłowiowego zera, a korzystamy z pewnej wiedzy historycznej, osiągnięć i doświadczeń innych architektów i budowniczych (podmiotów). Wiedza naukowa musi spełniać kategoryczny postulat ciągłości i trwałości tworzonych praw, zasad i teorii, weryfikowanych na etapie praktyki, najczęściej społecznej¹³.

2. Formalizacja piramidy informacyjnej

Uogólniając dotychczasowe rozważania pojęcie triady: dane – informacja - wiedza można zrekapitulować w sposób następujący:

- Dane jest to zbiór różnych atrybutów ilościowo-jakościowych zebranych ze względu na określone kryterium np. merytoryczne, formalne.
- Informacja jest to metasystem celowo wyodrębnionych danych ze względu na semantyczne kryterium kontekstowe.
- Wiedza jest to logicznie uporządkowany system informacji, heurystyk i różnych artefaktów reprezentujący stan rozwoju cywilizacji.

Formalizując te trzy poziomy piramidy informacyjnej możemy zaproponować następujący zapis:

$$\Pi: X \subseteq Y \subseteq Z \cup \Pi = f(Z(Y(X))) \quad (1)$$

gdzie:

Π - piramida informacyjna.

X - zbiór elementarnych danych,

¹³ Syntetycznym uogólnieniem tej tezy jest słynny napis widniejący na budynku Kongresu USA: "Knowledge is Power". Sztyld ten doskonale komponuje ze znanym stwierdzeniem Heraklita z Efezu: *Wiedza jest drugim słowem dla tych, którzy ją posiadają.*

Y - zbiór kontekstowej informacji,
 Z - subzbiór wiedzy.

Elementarne pojęcie danych X rozumiane jest jako zbiór bezkontekstowych atrybutów, zapisanych najczęściej w strukturze pewnej implementacji komputerowej, np. w postaci bazy danych (*Data Base*). Jak definiuje J. Trajer dane to tzw. „...suche fakty stanowiące bezpośrednie odwzorowanie pewnego obszaru rzeczywistości, albo jeden lub kilka symboli użytych do reprezentowania czegoś”¹⁴, co symbolicznie zapiszemy jako:

$$X = \{x_i; \quad i = \overline{1, I}\} \quad (2)$$

gdzie:

$x_i \in X$ - elementarna dana należąca do wyodrębnionego zbioru danych X .

Informacja powstaje jako rezultat porządkowania i integrowania danych, które w ten sposób nabierają znaczenia i stają się wartościowe dla potencjalnego użytkownika. W aspekcie ilościowym informację (Y) można zdefiniować jako przyrost pierwotnego potencjału informacyjnego dzięki dodatkowym danym:

$$Y(t) = |Y_0 + \Delta X_{Y_K} + E_{Y_K}| \quad (3)$$

gdzie:

$Y(t)$ - rzeczywista wielkość informacji w chwili t ,

Y_0 - pierwotny (wyjściowy) potencjał informacji,

ΔX_{Y_K} - przyrost informacji z tytułu pozyskania dodatkowych danych kontekstowych, stosownie do aktualnych potrzeb,

E_{Y_K} - czynnik zdroworoządkowej heurzy, dotychczasowego doświadczenia, w zakresie wykorzystania danych do tworzenia struktur informacyjnych.

Wiedzę (Z) w aspekcie ilościowym można zdefiniować jako przyrost pierwotnego potencjału wiedzy na skutek pozyskania dodatkowych informacji:

$$Z(t) = |Z_0 + \Delta Y_{Z_K} + E_{Z_K}| \quad (4)$$

gdzie:

$Z(t)$ - rzeczywista wielkość wiedzy w chwili t ,

Z_0 - pierwotny (wyjściowy) potencjał wiedzy,

ΔY_{Z_K} - przyrost wiedzy z tytułu pozyskania dodatkowych informacji kontekstowych, stosownie do aktualnych potrzeb,

E_{Z_K} - czynnik zdroworoządkowej heurzy, dotychczasowego doświadczenia w zakresie wykorzystania informacji w procesie generowania wiedzy.

Proces wzbogacania informacji (3) lub wiedzy (4) nie jest prostym działaniem arytmetycznym, powodującym automatyczne przechodzenie ilości w jakość, dlatego we wzorach (3) i (4) występuje symbol modulo. Proces transformacji danych w informację (3) lub informacji w wiedzę (4) jest skomplikowaną procedurą niejawną i heurystyczną, co dodatkowo obrazuje udział czynnika heurzy E_{Y_K} oraz E_{Z_K} .

Odpowiednio duża ilość danych ΔX_{Y_K} wcale nie musi oznaczać relatywnie bardziej wartościowej informacji lub jej proporcjonalnego przyrostu. Przyrost informacji zależy przede wszystkim od relacji kontekstowych między docelowym obrazem informacyjnym, a aktualnie wybranym zestawem danych. Jeszcze bardziej subtelne kryteria występują w procesie generowania wiedzy, o jakości której wcale nie musi decydować duży przyrost informacji ΔY_{Z_K} . W przypadku wiedzy bardzo ważny jest postulat synergii, czyli semantycznej interpretacji zgromadzonych informacji dającej niekiedy zupełnie odmienne efekty końcowe. Ten sam zestaw informacji w umyśle człowieka, w zależności od okoliczności może być rozmaicie

¹⁴ J. Trajer, A. Paszek, S. Iwan, *Zarządzanie wiedzą*. PWE Warszawa 2012, s. 295.

interpretowany, co prowadzi do nowych pokładów wiedzy. W budowaniu gmachu wiedzy niezmiernie istotne są subtelne kryteria i czynniki heurystyczne, jakimi operuje podmiot operujący zbiorem informacji w procesie tworzenia wiedzy.

3. Ranga i znaczenie systemów informatycznych

Informatyka do działalności biznesowej, zarówno w skali mikro, jak też makroekonomicznej została włączona w formie systemów informatycznych, obejmujących formalnie tradycyjne systemy informacyjne (decyzyjne) przedsiębiorstw, w których procesy informacyjno-decyzyjne zostały usprawnione poprzez automatyczne przetwarzanie danych (informacji) za pomocą nowoczesnych technologii komputerowych. Pierwotny system informacyjny można określić jako: „...wielopoziomą strukturę, która pozwala użytkownikowi tego systemu na transformację określonych informacji wejścia na pożądaną informację wyjścia, za pomocą odpowiednich procedur i modeli”¹⁵. Zdaniem Z. Bubnickiego wtórny informatyczny system zarządzania (*Management Information System*) można rozumieć jako: „...taki system zarządzania w którym niektóre funkcje zarządzania polegające na gromadzeniu i przetwarzaniu informacji oraz wyznaczaniu decyzji realizowane są za pomocą komputera, czyli zarządzanie jest częściowo skomputeryzowane”¹⁶. Jeszcze inna definicja system informatyczny określa jako: „... wydzielona część systemu informacyjnego, która została poddana komputeryzacji”¹⁷. Formalnie system informatyczny (*SI*) można zdefiniować za pomocą uporządkowanej piątki:

$$SI = \langle U, I, S, H, N \rangle \quad (5)$$

gdzie:

U - użytkownicy, *I* - informacje, *S* - *Software*, *H* - *Hardware*, *N* - *Netware*.

Współczesne systemy informatyczne obsługujące przedsiębiorstwo zbierają codziennie, z reguły automatycznie ogromne ilości źródłowej informacji transakcyjnej, operacyjnej, rynkowej i niemal natychmiast przystępują do jej rutynowego gromadzenia i przetwarzania. Często jest to informacja multimedialna o atrakcyjnych walorach użytkowych. Wyniki tego przetwarzania są zapisywane w odpowiednich formatach na właściwych nośnikach, które podlegają sukcesywnej aktualizacji, aż do spełnienia określonych warunków przechowywania, dotyczących np. terminu.

Standardy nowej gospodarki wymagają od przedsiębiorstw nieustannej aktualizacji baz i hurtowni danych i ciągłego monitorowania i raportowania podstawowych wskaźników biznesowych. Aktualna i wiarygodna informacja jest podstawą operatywnej sprawozdawczości, a także planowania rozwoju firmy w bliższym i dalszym horyzoncie czasowym. Rzetelna i wyczerpująca informacja jest niezbędna przy ocenie rynkowej pozycji firmy oraz szacowaniu poziomu konkurencyjności rynkowej. Dostęp do właściwej informacji jest substytutem wielu zasobów materialnych i niematerialnych, gwarantem przewagi rynkowej i bezpiecznej linii strategii rozwojowej firmy. W wysoce konkurencyjnym środowisku rynkowym natychmiastowy dostęp do strategicznej i operacyjnej informacji w różnym stopniu przetworzonej i udostępnianej w sposób syntetyczny jest absolutnie niezbędny. Największej troski "informacyjnej" wymagają przykładowo takie obszary działalności biznesowej przedsiębiorstw jak¹⁸:

– bezpośrednie kontakty z klientami,

¹⁵ J. Kisielnicki, H. Soroka, *Systemy informatyczne biznesu*. AW Placet, Warszawa 1995, s.19.

¹⁶ Z. Bubnicki, *Podstawy informatycznych systemów zarządzania*. PW Wrocław 1993, s. 22.

¹⁷ E. Kolbusz, (red.), *Inżynieria systemów informatycznych w e-gospodarce*. PWE Warszawa 2005. s. 52

¹⁸ E. Kolbusz, (red.), *Inżynieria systemów informatycznych...*, op. cit., s. 104.

- bieżąca współpraca z partnerami rynkowymi,
- monitorowanie i obsługa zamówień i zaopatrzenia,
- prognozowanie popytu rynkowego,
- obsługa kanałów dystrybucji,
- sterowanie poziomem zapasów,
- monitorowanie płynności finansowej firmy,
- zarządzanie procesami produkcyjnymi,
- zarządzanie personelem,
- rozwój i modyfikacja produktu,
- optymalizacja kosztów działalności gospodarczej,
- prowadzenie badań i analiz marketingowych,
- ocena konkurencyjności rynkowej firmy,
- symulowanie różnych sytuacji rynkowych,
- analiza ryzyka prowadzonej działalności rynkowej.

Wszystkie te obszary są dziś szeroko wspomagane przez systemy informatyczne - dedykowane, standardowe, zintegrowane funkcjonujące w strukturze teleinformatycznych aplikacji sieciowych, głównie internetowych. Systemy te wypracowują dopuszczalne w danych warunkach brzegowych warianty decyzji biznesowych, które powinny usprawnić zarządzanie menedżerskie na różnych szczeblach i poziomach kierowania firmą. Do tego celu wykorzystuje się specjalistyczne systemy informatyczne klasy DSS (*Decision Support Systems*) ewoluujące dziś w stronę systemów ekspertowych (ES - *Expert Systems*), należących do aplikacyjnych produktów sztucznej inteligencji¹⁹.

4. Zjawisko lawiny i smogu informacyjnego

W dobie społeczeństwa informacyjnego użytkowników technologii komputerowych zaskakują dwa odrębne zjawiska - lawina informacyjna i smog informacyjny. Lawina informacyjna to nieustanny strumień informacji (pożytecznej i niepożytecznej) zalewający wszystkie sensory, także zmysłowe i mentalne oraz serwery, komputery, urządzenia mobilne itp. płynący z różnych źródeł, nie zawsze oczekiwanych. W efekcie następuje niekontrolowany i wykładniczy przyrost informacji, zwłaszcza wizualnej, charakterystycznej dla społeczeństwa technokratycznego. Postępujący proces wizualizacji powoduje odchodzenie od myślenia symbolicznego i abstrakcyjnego i powrót do pierwotnego w dziejach ludzkości myślenia obrazkowego i obiektowego²⁰. Skala tego problemu w społeczeństwie informacyjnym potęguje się niemal z dnia na dzień, następuje swoiste przesylenie informacją, a możliwości jego zwalczania ciągle są niewielkie.

Smog informacyjny J. Janowski definiuje jako „...całokształt negatywnych skutków informacji nadmiernie atakujących człowieka”²¹. Jest to niemożliwa do skonsumowania ilość informacji, przy niskim poziomie jej istotności i braku zapotrzebowania na informacje dokładniejsze. Zjawisko to wynika z przewagi podaży nad popytem, czyli dynamicznie kształtującej się lawiny informacyjnej, generowanej przez niekontrolowane źródła informacji. Smog informacyjny polega na celowym lub przypadkowym zaśmiecaniu naszych komputerów różnymi informacjami, o charakterze mało merytorycznym.

Jedno i drugie zjawisko jest bardzo groźne dla współczesnego *e-Businessu*, którego permanentną cechą jest nieustanna aktywność sieciowa i związana z tym podatność na różne

¹⁹ K. Ficoń, *Sztuczna inteligencja. Nie tylko dla humanistów*. BEL Studio Warszawa 2014.

²⁰ Doskonałym przykładem tego trendu jest graficzny system operacyjny Windows, który bezmyślnym ruchem przysłowiowego keiuka otwiera zupełnie nieoczekiwane przestrzenie informacyjne.

²¹ J. Janowski, *Technologia informacyjna dla prawników i administratywistów*. Difin Warszawa 2009, s. 65.

aktywności i patologie sieciowe. Jak zauważa R. Tadeusiewicz żywiłowo konsumowany Internet jest groźnym medium pogłębiającym smog informacyjny niemal we wszystkich jego wymiarach - informacyjnym, mentalnym, społecznym, gospodarczym, a także cywilizacyjnym²². Asymetria informacyjna w sieci polega na tym, że łatwiej do niej dotrzeć niż ją zrozumieć i ogarnąć, i jest bardzo charakterystyczna dla współczesnego biznesu i wszystkich form jego działalności.

Trudno przesądzić *ad hoc* o większej szkodliwości jednego lub drugiego zjawiska, gdyż zależy ona od bardzo wielu czynników, a przede wszystkim od techniczno-organizacyjnej podatności przedsiębiorstwa na te procesy. Panuje dość powszechne przekonanie, że pomimo bogactwa danych większość przedsiębiorstw jest uboga w wiedzę, choć jest ona kapitałem strategicznym każdej organizacji. Dodatkowo jakość pozyskanej informacji biznesowej pozostawia wiele do życzenia i tylko częściowo spełnia oczekiwania decydentów i menedżerów związane z koniecznością podejmowania decyzji.

Szczególnym balastem informacyjnym cechują się klasyczne, transakcyjne systemy informatyczne, operujące na dużych z reguły relacyjnych lub obiektowych bazach niskoprzetworzonych danych, ciągle dominujące w dziedzinie informatycznych systemów zarządzania. Ograniczone możliwości zaawansowanej eksploracji danych powodują, że w grupie tych systemów tylko niewielka część danych jest rzeczywiście przeglądana i analizowana oraz wykorzystywana w praktyce dla potrzeb zarządzania menedżerskiego.

5. Pojęcie i atrybuty inżynierii wiedzy

Najogólniej inżynieria (*Engineering*) – to sztuka tworzenia rzeczy (wizji, koncepcji, projektów) nowych i jednocześnie pragmatyczna działalność intelektualna i eksperymentalna (badawcza), polegająca na modelowaniu, projektowaniu, konstruowaniu, modyfikacji i utrzymaniu użytecznych i efektywnych rozwiązań dla rozwiązania praktycznych problemów²³. Zadaniem najbardziej ogólnej inżynierii systemów w opinii W. Bojarskiego „...jest tworzenie oraz organizowanie takich obiektów systemowych i systemów nominalnych, które są zadawalająco bliskie systemom maksymalnym”²⁴.

Tytułowa inżynieria wiedzy (*Knowledge Engineering*) w opinii W. Traczyka to „...dział sztucznej inteligencji obejmujący metody i systemy służące do uzyskania użytecznych informacji na podstawie dostępnej wiedzy i konkretnych danych”²⁵. Leksykalne ujęcie inżynierii wiedzy definiuje jako: „...wykorzystanie wiedzy i doświadczenia eksperta oraz przekładanie ich na reguły i struktury danych, które mogą być wykorzystane przez komputer”²⁶. Inżynieria wiedzy zajmuje się identyfikacją i opisywaniem wiedzy i systemami wykorzystującymi wiedzę do automatycznego wnioskowania dla wspomagania diagnostyki, zarządzania, planowania itp. Głównym jej zadaniem jest stawianie hipotez i wnioskowanie oraz kreowanie nowych innowacyjnych teorii (pomysłów, aplikacji). Korzeni inżynierii wiedzy można się dopatrzeć w takich dyscyplinach filozoficznych jak: epistemologia (teoria poznania), ontologia (teoria bytu i rzeczywistości), logika (reguły poprawnego myślenia). Inżynierię wiedzy od klasycznych dziedzin inżynierskich odróżnia niematerialny charakter produktu, jakim jest wiedza. W początkowym okresie swojego rozwoju inżynieria wiedzy zajmowała się niemal wyłącznie systemami ekspertowymi, tzn. systemami, w których wiedzę ekspertów w danej dziedzinie reprezentowano zazwyczaj w regułowej bazie wiedzy, a przetwarzanie ograniczało się do sformalizowanego wnioskowania logicznego.

²² R. Tadeusiewicz, *Spółeczność Internetu*. AOW EXIT Warszawa 2002, s. 120.

²³ K. Ficoń, *Inżynieria szanse i wyzwania XXI wieku*. Kwartalnik BELLONA, nr 2(661)/2010.

²⁴ W. Bojarski, *Podstawa analizy inżynierii systemowej*. PWN Warszawa 1984, s. 47.

²⁵ W. Traczyk, *Inżynieria wiedzy*. AOW EXIT, Warszawa 2010, s. 5.

²⁶ J. Penc, *Leksykon biznesu*. AW Placet Warszawa 1997, s. 177.

Zasadniczym terminem w inżynierii wiedzy jest interdyscyplinarne pojęcie wiedzy, definiowanej na różne sposoby praktycznie dla potrzeb wszystkich dyscyplin naukowych. W tym miejscu należy wykazać należytą refleksję wobec ostrzeżeń Leszka Kołakowskiego kiedy mówi: „...z definicją wiedzy jak dotąd nikomu nie udało się uporać i zapewne nigdy to się nie uda”²⁷. Zasadnym wydaje się być stwierdzenie, że wiedza to ogół wiadomości i umiejętności w jakiejś dziedzinie, zdobytych dzięki permanentnemu uczeniu i doświadczeniu preferujący udział praktyki społecznej. Bardzo ważna jest też konieczność logicznej interpretacji wiedzy, stawiania hipotez i wyciągania odpowiednich wniosków. W tym sensie wiedza to selektywnie wybrane informacje, połączone z doświadczeniem, interpretacją i wnioskowaniem. Kolejnym obszarem odniesienia jest teoria podejmowania decyzji, gdzie pojęcie wiedzy oznacza zdolność podejmowania rozsądnych decyzji dla prowadzenia efektywnych akcji. Wiedza i wynikające z niej decyzje są podstawą skutecznego działania w praktyce, determinując jednocześnie pozycje społeczną podmiotu (np. gospodarczego) podejmującego decyzje.

Wiedza jest klasyfikowana na wiele sposobów, zależnie od potrzeb, a najbardziej podstawowa sformalizowana klasyfikacja stosowana m.in. w inżynierii i informatyce dzieli wiedzę na trzy jej rodzaje:²⁸

- wiedza proceduralna - zestaw uporządkowanych funkcji, czynności, programów opisujących sposób rozwiązania problemu,
- wiedza deklaratywna - zbiór użytecznych narzędzi potrzebnych do rozwiązania problemu (reguły logiczne: jeśli ... to...),
- wiedza strukturalna - uporządkowany zbiór pojęć obiektów lub kategorii oraz relacji między nimi.

Inne kryteria dzielą wiedzę m.in. na wiedzę ogólną i szczególną, w tym dziedzinową - według obszaru zainteresowań. Szczególnym rodzajem wiedzy jest wiedza ukryta zawarta w ludzkich umysłach, która najczęściej ewoluje w stronę wiedzy jawnej (publicznej). Dychotomicznym podziałem jest wiedza indywidualna ściśle związana z inteligencją człowieka i wiedza zbiorowa bazująca na społecznym jej znaczeniu. Bardzo ważną prakseologiczną kategorią wiedzy jest rozmyta wiedza heurystyczna, zdroworozsądkowa oparta na doświadczeniu, intuicji, której przeciwieństwem jest ścisła wiedza logiczna i faktograficzna. Niekiedy wyodrębnia się wiedzę jakościową, bazującą na kryteriach opisowych i wiedzę ilościową odnoszoną do obszaru nauk ścisłych i statystycznych.

Dla potrzeb utylitarnej inżynierii wiedzy atrakcyjny jest system klasyfikacyjny wiedzy oparty na kryteriach prakseologicznych preferujących celowość, użyteczność i skuteczność działania. Według tego podejścia wyróżniamy cztery kategorie wiedzy (rys. 2):

- wiedza typu *know-what* jest bliskoznaczna informacji niższego poziomu i różni się tylko doprecyzowaniem kontekstowym,
- wiedza typu *know-why* wyjaśnia istotę zjawiska, procesu lub determinuje przyczyny, co wpływa na bezpieczeństwo,
- wiedza typu *know-how* precyzuje standardy zachowań ludzi w zakresie umiejętności wykonywania czegoś,
- wiedza typu *know-who* określa posiadaczy wiedzy i opisuje ich wiedzę, np. w określonej dyscyplinie lub dziedzinie.

Powyższy system klasyfikacyjny doskonale obrazuje ukryty potencjał wiedzy rozumiany jako zdolność do efektywnego działania na bazie pozyskanych informacji, użytecznych heurystyk i dotychczasowego doświadczenia. Potencjał wiedzy jest niezbędny do podejmowania racjonalnych wyborów i decyzji do działania, przede wszystkim praktycznego.

²⁷ Pomimo ostrzeżeń Mistrza obszerny przegląd różnych definicji wiedzy można znaleźć w pracy: A. Szewczyk (red.), *Spółeczeństwo informacyjne. Problemy rozwoju*. Difin Warszawa s. 175-177.

²⁸ W. Traczyk, *Inżynieria wiedzy ... op. cit.*, s. 8.



Rysunek 2. Przykładowe systemy klasyfikacyjne wiedzy

Źródło: Opracowanie własne.

Pojęcie prakseologicznej inżynierii wiedzy jest odnoszone do dwóch obszarów aplikacyjnych dotyczących wspomagania sztucznej inteligencji, głównie w kontekście uczenia oraz wydobywania i przetwarzania wiedzy ukrytej w dużych zbiorach informacyjnych. W tym drugim obszarze użytecznym aparatem narzędziowym inżynierii wiedzy są nowoczesne metody eksploracji wiedzy, bazujące na zaawansowanych aplikacjach komputerowych wykorzystujących dorobek m.in. inżynierii programowania, statystyki matematycznej, teorii grafów, sieci semantycznych, lingwistyki, kognitywistyki a przede wszystkim rozlicznych metod sztucznej inteligencji, w tym: sztuczne sieci neuronowe, algorytmy ewolucyjne, teorię zbiorów rozmytych, modele wieloagentowe oraz systemy ekspertowe.

6. Komputerowe metody eksploracji danych i odkrywania wiedzy

W zarządzaniu menedżerskim współczesnym przedsiębiorstwem dostęp do właściwej informacji jest głównym wyznacznikiem trafności podejmowanych decyzji. Informacje, która są wartościowe z punktu widzenia zarządzania przedsiębiorstwem najczęściej nie są dostępne wprost z jego zasobów informacyjnych, także komputerowych. Często bowiem są wynikiem dogłębnych analiz i logicznego wnioskowania na podstawie dostępnych danych, a więc muszą być odpowiednio przygotowane. Uzyskanie w ten sposób informacji, które mogą być przydatne w realizacji procesów decyzyjnych nie zawsze jest łatwe i proste. W tym kontekście bardzo aktualna wydaje się być prorocza metafora R. Tadeusiewicza, kiedy mówi, że: "Od komputerów oczekiwaliśmy fontanny wiedzy, a dostaliśmy potop danych"²⁹. Tymczasem menedżerowie muszą uwzględniać dane o coraz szerszych rozmiarach i zakresie merytorycznym, a analiza musi być coraz bardziej wnikliwa i obejmować coraz większą liczbę kryteriów. Dlatego jednym z czynników decydujących o efektywności działań decyzyjnych jest zastosowanie nowoczesnych narzędzi wspomagających analizę danych i preparację użytecznych w danym momencie informacji, a docelowo oryginalnej wiedzy.

Podstawowym źródłem danych na potrzeby zarządzania operatywnego w przedsiębiorstwie jest bieżąca komputerowa ewidencja zdarzeń gospodarczych, (transakcji),

²⁹ R. Tadeusiewicz, <http://docplayer.pl/311649-Ryszard-tadeusiewicz-akademia-gorniczo-hutnicza-laboratorium-biocybernetyki.html> (07.12.2015r.).

które są niezbędne dla celów sprawozdawczości obligatoryjnej, a także do prawidłowego funkcjonowania przedsiębiorstwa zarówno na poziomie operacyjnym, jak też strategicznym. Aby sprostać wysokim wymaganiom rynkowym nowoczesne systemy informatyczne obsługujące sferę zarządzania menedżerskiego muszą spełniać szereg dodatkowych kryteriów organizacyjno-funkcjonalnych gwarantujących wysoką ich elastyczność, użyteczność i sprawność działania. Wobec wspomnianych zagrożeń związanych z lawiną informacyjną i smogiem informacyjnym systemy te muszą być jednocześnie zdolne do inteligentnej selekcji danych źródłowych oraz generowania adekwatnych do rzeczywistych potrzeb informacji operacyjnych, co wiąże się z problematyką semantycznego selekcjonowania i pozyskiwania informacji i wiedzy z systemowych zasobów informacyjnych, takich jak: bazy danych, hurtownie danych czy bazy wiedzy.

Terminy „eksploracja danych” i „odkrywanie wiedzy w bazach danych” są często stosowane wymiennie, choć w rzeczywistości drugi termin posiada dużo szersze znaczenie. Odkrywanie wiedzy obejmuje cały proces akwizycji wiedzy, począwszy od selekcji danych źródłowych, poprzez czyszczenie, transformację, kompresję danych, odkrywanie wzorców, a skończywszy na ocenie odkrytych wzorców. Zgodnie z tą definicją eksploracja danych oznacza zastosowanie konkretnego algorytmu odkrywania wzorców na wybranych danych źródłowych i stanowi jeden z etapów składowych całego procesu odkrywania wiedzy.

Ciągła presja użytkowników na selektywny i inteligentny dostęp do informacji źródłowej wymusił na producentach systemów informatycznych wprowadzenie wielu mechanizmów i technik pozyskiwania ściśle określonej informacji i spójnej wiedzy biznesowej. Ogromne ilości informacji gospodarczej przepływające przez serwery nowej gospodarki wcale nie ułatwiają podjęcia trafnych decyzji, a niekiedy dodatkowo komplikują ten proces, ponieważ człowiek nie jest w stanie ich wszystkich ogarnąć i wyciągnąć sensownych konkluzji. Z pomocą przychodzi nowa dziedzina informatyki - eksploracja (ekstrakcja) danych³⁰, której idea polega na wykorzystaniu szybkości komputera do efektywnego przeszukiwania ogromnych zbiorów danych i znajdowania ukrytych dla człowieka (właśnie z uwagi na ograniczone możliwości czasowe) prawidłowości w zasobach informacyjnych komputerowych baz danych, hurtowniach danych, a także w bazach wiedzy. Przykładami takich rozwiązań są:

- OLAP (*On-Line Analytical Processing*) - umożliwia budowanie analitycznych baz danych, które pozwalają na logiczne agregowanie danych i ukierunkowane ich analizowanie według pewnych kryteriów,
- DM (*Data Mining*) - oznacza drążenie danych prowadzące do odkrywania nowych treści i użytecznych informacji zawartych w danych,
- BI (*Business Intelligence*) - zwana też analityką biznesową, pozwala na wypracowanie i generowanie różnych wariantów decyzyjnych, które umożliwiają praktyczne działanie.

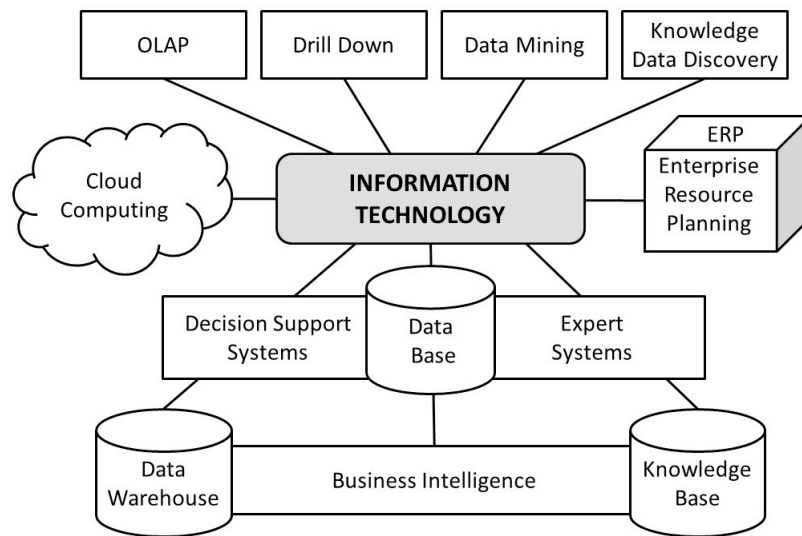
W tej dziedzinie wykorzystywane są jeszcze inne bardziej narzędziowe technologie, takie jak np.: DD (*Drill Down*) - przeszukiwanie w głąb, czy KDD (*Knowledge Data Discovery*) - odkrywanie wiedzy.

Technologia Drill Down znaczy ogólnie "wejście w bardziej szczegółowe rozwinięcie tego co widać na wierzchu" - może to być lista, tablica rozwijana, okno czy co zwykła zakładka. Drill Down to przechodzenie na coraz niższy poziom metodą wielopoziomowego, zagnieżdżonego menu obsługiwanego przez odpowiednie aplikacje. Przykładowo zakładka "Nr zamówienia" na formularzu "Rozliczenie faktury" przenosi nas na ekran "Zamówienia", z którego można wchodzić coraz głębiej w rozmaite szczegóły tej transakcji.

Odkrywanie wiedzy w bazach danych - *Knowledge Data Discovery* jest nietrywialnym procesem identyfikacji ważnych, nowych potencjalnie użytecznych i zrozumiałych wzorców

³⁰ Na określenie eksploracji danych używa się także takich pojęć jak: drążenie danych, pozyskiwanie wiedzy, wydobywanie danych, czy ekstrakcja danych.

wyszukiwanych w zbiorach danych³¹. Najczęściej proces odkrywania wiedzy polega na intensywnym współdziałaniu komputera z ekspertem za pomocą specjalistycznych interfejsów, ponieważ tylko użytkownik potrafi poprawnie ocenić wartość i realną użyteczność odkrytej wiedzy. Z reguły aplikacje *Drill Down* są wykorzystywane w systemach drążenia wiedzy KDD jako pomocnicze narzędzia programowe.



Rysunek. 3. Technologia informatyczna w strategii eksploracji danych

Źródło: Opracowanie własne.

Przedmiotem dalszej analizy będą głównie technologie OLAP, *Data Mining* i *Business Intelligence*, jako wiodące w procesach ekstrakcji danych i odkrywania kolejnych pokładów nowej wiedzy.

7. Aplikacje OLAP

Zarządzanie współczesnym przedsiębiorstwem wiąże się z koniecznością dostępu do coraz bardziej przetworzonych, a zarazem wyrafinowanych i syntetycznych informacji. Dotychczasowe narzędzia ich pozyskiwania, takie jak transakcyjne systemy finansowo-księgowe, systemy ERP, arkusze kalkulacyjne czy generatory raportów nie są w stanie w pełni obsłużyć tych potrzeb. Naprzeciw tym wyzwaniom wychodzi zaś technologia OLAP (*On-Line Analytical Processing*)³². Model OLAP to specjalistyczne oprogramowanie bazujące na zoptymalizowanych hurtowniach danych i zaawansowanej inżynierii programowania wspierające podejmowanie decyzji, np. biznesowych.

Na modelowy proces przetwarzania danych w systemie OLAP składają się następujące czynności³³: ekstrakcja danych źródłowych, oczyszczanie danych, dostosowanie danych, agregacja wielowymiarowa, składowanie w bazie OLAP, odczyt (eksploracja analityczna) i zapis (symulacje, budżetowanie i planowanie) danych przez użytkowników końcowych. Proces ekstrakcji ma na celu pozyskanie danych z zewnętrznych baz danych do centralnej bazy OLAP.

³¹ G. Piatetsky-Shapiro and W. J. Frawley. *Knowledge Discovery in Databases*. AAAI/MIT Press, 1991.

³² Historia OLAP datuje się od roku 1993 gdy E. Codd opublikował pierwszy opis modelu technologii OLAP, obejmujący 12 charakterystycznych reguł (funkcji). W roku 1995 w wyniku wielu dyskusji model ten został rozszerzony do 18 funkcji.

³³ D. Chase, G. Spofford, E. Thomsen, *Microsoft OLAP Solutions*. John Willey & Sons, New York, 1999.

Oczyszczanie polega na usunięciu błędów i niespójności powstałych na etapie ewidencji w źródłowych systemach informatycznych. Dostosowanie może obejmować takie czynności jak: selekcja danych, kategoryzacja, tłumaczenie słownikowe oraz scalanie danych z heterogenicznych źródeł. Kategoryzacja polega na wyodrębnieniu w zbiorach danych pewnych charakterystycznych cech ze względu na wybrane kryterium. Etap scalania obejmuje stosowanie różnych metod przetwarzania na potrzeby integracji danych pochodzących z różnych, często rozbieżnych systemów ewidencyjnych. Tak przygotowane dane są przesyłane do wielowymiarowych struktur OLAP, zwanych kostkami (*Cubes OLAP*), w których umieszczone są zagregowane wartości według wcześniej wskazanych kryteriów agregacji.

Technologia OLAP pozwala użytkownikowi szybko i syntetycznie analizować, redagować i edytować duże zbiory informacji zorganizowane w wielowymiarowych widokach i strukturach hierarchicznych. Najbardziej popularne technologie ROLAP (*Relational OLAP*) oparte na relacyjnych bazach danych charakteryzują się zdolnością do przechowywania danych o dużej objętości, pozwalające na szybką ich modyfikację pod kątem łatwego dostępu do wyselekcjonowanej tematycznie (statystycznie) informacji.

Tradycyjne systemy OLAP są często nazywane wielowymiarowymi MOLAP (*Multidimensional OLAP*), gdyż przekształcają masowe dane transakcje na wielowymiarowe kostki OLAP. Kostka OLAP *Cube* przechowuje dane w sposób zbliżony do struktur wielowymiarowych arkuszy kalkulacyjnych, co pozwala na pokonanie ograniczeń relacyjnych baz danych. Dzięki temu daje użytkownikowi możliwość manipulowania i analizowania danych z różnych punktów widzenia. Zastosowanie wielowymiarowych kostek OLAP wspomaga sposób prezentacji informacji i umożliwia elastyczny i intuicyjny dostęp użytkownikom do wielowymiarowego modelu danych, prezentowanego niekiedy w formie graficznej.

Oprogramowanie OLAP wykorzystujące intuicyjny interfejs użytkownika, umożliwia szybki i syntetyczny dostęp do żądanych informacji różnym użytkownikom minimalizując czas reakcji systemu na zadawane pytania. Istotną korzyścią zastosowania technologii OLAP w ujęciu problemów rachunkowości zarządczej jest wielowymiarowe podejście do reprezentacji informacji (*Multidimensional View*), dzięki czemu modele analityczne zyskują na przejrzystości i intuicyjności, co ma bezpośredni wpływ na efektywność procesu podejmowania decyzji. Przykładowo aplikacje OLAP są często używane do wykonywania skomplikowanych analiz różnych trendów np. sprzedaży, a także do wielokryterialnych analiz finansowych wykonywanych na bazie hurtowni danych w ramach inżynierii finansowej. Są bardzo przydatne do wstępnego przeglądania zbioru danych przez analityka we wstępnej fazie analiz statystycznych³⁴.

Narzędzia raportowania OLAP, pomimo swoich niewątpliwych zalet: wielowymiarowości i wielopoziomowości (umożliwiającej przechodzenie od ogółu do szczegółu i z powrotem) oraz pomimo analiz zapewniających wgląd w dane z różnych perspektyw badawczych, wynikających z potrzeb różnych użytkowników, okazały się nie w pełni wystarczające. Zaawansowane metody eksploracji danych stanowią analityczne rozszerzenie technik OLAP i prezentują podejście do analizy danych oparte raczej nie na raportowaniu, ale na zdobywaniu nowej wiedzy. Środowisko typu *OLAP* pozwala na prostą analizę baz danych, które umożliwiają wielowymiarową obserwację zagregowanych wartości wybranych atrybutów jednej lub wielu połączonych relacji. Metodologia *OLAP* zakłada, że użytkownik przygotowuje pewną hipotezę, której poprawność weryfikuje korzystając z narzędzi *OLAP* (np. *Oracle Express Server*).

³⁴ E. Kolbusz, (red.), *Inżynieria systemów informatycznych...*, op. cit., s. 232.

8. Technologia Data Mining

Technologia *Data Mining* (DM) jest jednym z etapów odkrywania ukrytej wiedzy w komputerowych bazach danych³⁵. Istota DM polega na wykorzystaniu szybkości komputerów do znajdowania ukrytych dla człowieka zależności i prawidłowości w ogromnych zbiorach (hurtowniach) danych. Dlatego najprostszą definicją DM jest określenie - automatyczne wykrywanie poszukiwanych zależności w bazach danych. Bardziej formalnie *Data Mining* jest definiowane jako „...nauka zajmująca się wydobywaniem informacji z dużych zbiorów danych lub baz danych”³⁶. Kolejna definicja mówi, że *Data Mining* to: „...nietrywialne wydobywanie ukrytej, poprzednio nieznannej i potencjalnie użytecznej informacji z danych”³⁷. Jeszcze inna definicja pod pojęciem *Data Mining* precyzuje „...proces przekształcania danych w użyteczną wiedzę, począwszy od wprowadzenia danych lub pobrania ich z zewnętrznego źródła do utworzenia wynikowego raportu”³⁸.

W ujęciu praktycznym *Data Mining* jest oryginalną aplikacją softwarową umożliwiającą uzyskanie takich zależności, jakie nie były uprzednio wprowadzone do bazy danych. Stanowi więc rodzaj języka manipulowania na danych zgromadzonych w zasobach komputerowych. Mechanizmy działania DM wywodzą się z teorii statystyki i zasad uczenia maszynowego. W dużym uproszczeniu cały proces odkrywania wiedzy DM sprowadza się do tego, że wybrane metody uczenia są używane do odnalezienia regularności w danych i zapisania ich w odpowiedniej postaci, zaś metody statystyczne służą do oceny ich jakości. Technologia DM wykorzystuje m.in. modele sztucznych sieci neuronowych, różne struktury drzew decyzyjnych oraz statystyczne techniki do efektywnego przeszukiwania dużych zbiorów danych. Główne kierunki wykorzystania technologii DM to obecnie biznes, nauka i internet. Przykładem aplikacji biznesowych DM może być zadanie polegające na wykryciu, jakie towary w danym sklepie są najczęściej kupowane w pakietach konsumenckich lub jakie cechy wyróżniają kierowców powodujących najczęściej wypadki drogowe? Zaawansowane technologie *Data Mining* są dziś szeroko stosowane w eksploracji zasobów internetowych, jako *Web Mining*. Oprócz tradycyjnych technik, takich jak pozyskiwanie informacji z danych na podstawie pewnych reguł i zależności umożliwiają one uzyskanie szybkich odpowiedzi na spersonalizowane zapytania kierowane do sieci.

9. Systemy Business Intelligence

Business Intelligence (BI) jest pojęciem o bardzo szerokim i elastycznym znaczeniu i obejmuje zaawansowane aplikacje informatyczne bazujące na hurtowniach danych i specjalizujące się w tzw. analityce biznesowej³⁹. Przykładowo termin BI może być określony jako: „...narzędzie informatyczne, którego zadaniem jest dostarczanie uporządkowanej i zrozumiałej informacji zarządczej wspomagającej podejmowanie decyzji na różnych szczeblach zarządzania firmą”⁴⁰. Zdaniem A. Januszewskiego *Business Intelligence* to: „...szerokie połączenie aplikacji i technologii umożliwiających zbieranie, scalanie, gromadzenie, selekcję,

³⁵ http://data-mining.wyklady.org/wyklad/297_eksploracja-danych-data-mining.html.

³⁶ D. Hand, H. Mannila, P. Smyth: *Principles of Data Mining*, MIT Press, Cambridge, MA, 2001.

³⁷ W. Frawley, G. Piatetsky-Shapiro, C. Matheus: *Knowledge Discovery in Databases: An Overview*. AI Magazine, 1998.

³⁸ M.J.A. Berry, G.S. Linoff: *Mastering data mining*. New York: Wiley, 2000.

³⁹ Pojęcie *Business Intelligence* zostało wprowadzone przez Gartner Group w 1989 r jako zestaw koncepcji i metodyk mających na celu usprawnienie podejmowanych decyzji biznesowych przez użycie systemów opartych na danych, informacji i wiedzy.

⁴⁰ J. Trejter, A. Paszek, S. Iwan, *Zarządzanie wiedzą...*, op. cit., s. 295.

analizę i czytelną prezentację pochodzących z różnych źródeł informacji (a także wiedzy) biznesowej na potrzeby konkretnych dziedzin działalności gospodarczej"⁴¹.

Systemy BI zajmują się przekształcaniem elementarnych danych w syntetyczne informacje, a w dalszej kolejności informacji w realną wiedzę biznesową, która może być bezpośrednio wykorzystana, np. do inteligentnego zwiększenia konkurencyjności przedsiębiorstwa. *Business Intelligence* to szeroka dziedzina informatyki ekonomicznej obejmująca aplikacje i technologie służące gromadzeniu i analizie danych w celu wspomagania procesu podejmowania decyzji biznesowych⁴². Narzędzia BI umożliwiają sprawne i szybkie tworzenie raportów, analiz oraz sprawozdań na potrzeby wewnętrzne firmy oraz dla instytucji zewnętrznych. Platforma *Business Intelligence* wspomaga menedżerów podmiotów gospodarczych w podejmowaniu kluczowych decyzji, a pozostałym pracownikom daje możliwość podejmowania optymalnych działań w oparciu o rzetelną informację i merytoryczną wiedzę. Skuteczność działania w obszarze *Business Intelligence* wymaga dogłębnej znajomości wszystkich czynników i zmiennych wpływających na biznes. Narzędzia i systemy *Business Intelligence* znacznie usprawniają pracę analityków finansowych, menedżerów biznesowych czy kadry zarządzającej firmą. Pozwalają głównie menedżerom na zajmowanie się swoimi faktycznymi zadaniami – efektywnym zarządzaniem firmą, a nie żmudnym przygotowywaniem multimedialnych raportów i pracochłonną analizą, np. trendów sprzedaży czy konkurencyjności rynkowej.

M. Flasiński wdrożenie systemów klasy *Business Intelligence* widzi na dwóch poziomach - podstawowym i zaawansowanym⁴³. Na poziomie podstawowym systemy BI przeprowadzają analizy biznesowe w oparciu o metody statystyczne, modele ekonometryczne i badania operacyjne, a ich rezultatem jest syntetyczna informacja zarządcza w postaci zestawień, raportów i sprawozdań. Na poziomie zaawansowanym celem wdrożenia BI jest wspomaganie procesu podejmowania decyzji menedżerskich poprzez zastosowanie metod sztucznej inteligencji, a ich efektem może być odkrywanie przyczyn zaistniałej sytuacji biznesowej, prognozowanie dalszego jej rozwoju, czy symulacja możliwych skutków podejmowanych decyzji lub sugerowanie optymalnych opcji decyzji zarządczych. Typowymi metodami sztucznej inteligencji stosowanymi do wyżej wymienionych celów są: analiza skupisk, rozpoznawanie wzorców, sztuczne sieci neuronowe drzewa decyzyjne, systemy regułowe i ekspertowe⁴⁴.

Efektem zastosowania narzędzi BI jest dostępność do szybkiej informacji na temat najważniejszych wskaźników firmy, takich, jak dane o klientach, konkurencji, partnerach biznesowych, sytuacji ekonomicznej i operacjach wewnętrznych. Przykładowo technologie BI pozwalają na uzyskanie jakościowych analiz i syntetycznych ocen w zakresie takich obszarów, jak np.: miesięczne trendy sprzedaży (zakupów), wielokryterialne rankingi kontrahentów rynkowych, analiza cyklu życia wyrobu, najczęściej kupowane produkty (usługi), rotacja zapasów magazynowych, analiza zatorów i płynności finansowej. Jednocześnie należy pamiętać, że przygotowanie elastycznych, inteligentnych analiz, raportów i sprawozdań wymaga szybkiego i krótkotrwałego dostępu do dużych zasobów informacyjnych, dlatego nowoczesne rozwiązania *Business Intelligence* ewoluują w stronę chmury obliczeniowej (*Cloud Computing*), gdzie duże zasoby informacyjne są współdzielone z różnymi użytkownikami i każdy może z nich korzystać w odpowiednim do jego potrzeb optymalnym wymiarze merytorycznym, czasowym, technicznym⁴⁵.

⁴¹ A. Januszewski, Funkcjonalność informatycznych systemów zarządzania. Systemy Business Intelligence T.2. WN PWN Warszawa 2013, s. 6.

⁴² J. Surma, *Business Intelligence. Systemy wspomaganie decyzji biznesowych*. WN PWN Warszawa 2013.

⁴³ M. Flasiński, *Wstęp do sztucznej inteligencji*. WN PWN SA Warszawa 2011.

⁴⁴ K. Ficoń, *Sztuczna inteligencja. Nie tylko dla humanistów*. BEL Studio Warszawa 2013.

⁴⁵ J. Rosenberg, A. Mateos, *Chmura obliczeniowa. Rozwiązania dla biznesu*. Helion S.A. Gliwice 2011.

Zakończenie

W dobie społeczeństwa informacyjnego opartego w dużej części na technologii komputerowej firmy posiadają ogromne ilości elementarnych, bezkontekstowych danych, gromadzonych mniej lub bardziej świadomie w swoich systemach informatycznych, choć nie potrafią ich efektywnie wykorzystać dla swoich biznesowych potrzeb. Paradoks ten ma niestety tendencję wzrostową. Według agencji IDC⁴⁶ w latach 1999-2003 wytworzono i zapisano więcej informacji niż kiedykolwiek wcześniej, tj. od początku ludzkości do roku 1998. Baza danych systemu Retail Link wykorzystywana przez sieć amerykańskich marketów Wal-Mart gromadzi dziennie informację o ponad 20 milionach transakcji handlowych⁴⁷. Z kolei IBM twierdzi, że firmy podejmują strategiczne decyzje biznesowe przy wykorzystaniu statystycznie zaledwie 7% zgromadzonych w swoich zasobach danych⁴⁸. Nasze zdolności analizowania i rozumienia tak dużych wolumenów danych i lawinowo rosnących strumieni informacyjnych są obecnie dużo mniejsze od realnych, głównie technicznych możliwości ich zbierania i przechowywania.

Wobec losowo narastającej lawiny informacyjnej trend ten będzie jeszcze bardziej się pogłębiał. Stąd pilna konieczność, po pierwsze, opanowania zjawiska lawiny i smogu informacyjnego, a po drugie, selektywnego i natychmiastowego dostępu do żądanej informacji, aby bieżąco kompilować ją do postaci kreatywnej wiedzy niezbędnej w procesie podejmowania racjonalnych decyzji m.in. menedżerskich. Szybkość dostępu do komputerowo skompilowanej informacji kontekstowej i jakość generowanych klastrów zadaniowo dedykowanej wiedzy stanowić będzie o konkurencyjności realizowanej strategii biznesowej. Celowi temu mają służyć inteligentne technologie komputerowe, zaczynające się na etapie drażenia elementarnych, bezkontekstowych danych do postaci kontekstowej, użytecznej informacji, pozwalającej na efektywne kompilowanie jej do poziomu selektywnej wiedzy biznesowej. Potężna moc obliczeniowa współczesnych komputerów oraz gigantyczne zasoby informacyjne współczesnych serwerów bazodanowych, sieciowych pozwalają z optymizmem patrzeć na dalszy rozwój nowej gospodarki, efektywnie wykorzystującej m.in. potencjalne walory biznesowe inżynierii wiedzy.

Potencjalnym sojusznikiem nowej gospodarki jest technologia chmury cybernetycznej, która z jednej strony udostępnia potężne zasoby informacyjne światowych serwerów bazodanowych penetrowanych za pomocą nowoczesnych technologii informatycznych, np. klasy *Business Intelligence*, z drugiej, dzięki impulsowemu współdzieleniu i współużytkowaniu tej zaawansowanej technologii pozwala na ograniczenie kosztów eksploatacji drogich aplikacji softwarowych i równie kosztownej infrastruktury hardwarowej. Wielka rzesza potencjalnych użytkowników technologii *Cloud Computing* stymuluje w sposób naturalny najlepsze rozwiązania w zakresie utylitarnej inżynierii wiedzy pozwalające na efektywne użytkowanie ogólnoswiatowych zasobów informacyjnych zgodnie z rzeczywistymi potrzebami nowej gospodarki, w której szybki dostęp do realnej, wysokiej jakości i przede wszystkim użytecznej wiedzy biznesowej urasta do rangi pierwszoplanowej.

Reasumując można stwierdzić, że postawiona na wstępie teza typująca inżynierię wiedzy, a konkretnie jedną z jej narzędziowych aplikacji - metody eksploracji danych i odkrywania wiedzy do roli stymulatora przedsiębiorczości w dobie nowej gospodarki wydaje się być potwierdzona.

⁴⁶ M.E. Burke, *The Phenomenal Power of Business Intelligence: Managerial Skills for the 21st Century*, Europa Publications Limited, 2003.

⁴⁷ Piatetsky-Shapiro G., Smyth P., Uthurusamy R., *Advances in Knowledge Discovery and Data Mining*, ed. Fayyad U.M., MIT Press, 1996.

⁴⁸ M. Biere, *Business Intelligence for the Enterprise*, IBM Press, Prentice Hall, 2003.

Literatura

1. Bell D., *Technika łączności. Przegląd Zagranicznej Literatury Progностycznej*. Seria Komitet PAN Polska 2000
2. Berry M. J. A., Linoff G.S., *Mastering data mining*. New York: Wiley, 2000
3. Biere M., *Business Intelligence for the Enterprise*, IBM Press, Prentice Hall, 2003
4. Bojarski W., *Podstawy analizy inżynierii systemowej*. PWN Warszawa 1984
5. Bubnicki Z., *Podstawy informatycznych systemów zarządzania*. PW Wrocław 1993
6. Burke M.E., *The Phenomenal Power of Business Intelligence: Managerial Skills for the 21st Century*, Europa Publications Limited, 2003
7. Ficoń K., *Funkcjonowanie przedsiębiorstw wirtualnych w cyberprzestrzeni gospodarczej*. PK Koszalin 2010
8. Ficoń K., *Inżynieria szanse i wyzwania XXI wieku*, Kwartalnik BELLONA, nr 2(661)/2010
9. Ficoń K., *Systemy informatyczne przedsiębiorstw*. BEL Studio Warszawa 2001
10. Ficoń K., *Sztuczna inteligencja. Nie tylko dla humanistów*. BEL Studio Warszawa 2013
11. Flasiński M., *Wstęp do sztucznej inteligencji*. WN PWN SA Warszawa 2011
12. Frawley W., Piatetsky-Shapiro G., Matheus C., *Knowledge Discovery in Databases: An Overview*. AI Magazine, 1998
13. Gates B., *Biznes szybki jak myśl*. Prószyński i S-ka, Warszawa 2001
14. Hand D., Mannila H. Smith P., *Eksploracja danych*, WNT, Warszawa 2005
15. Janowski J., *Technologia informacyjna dla prawników i administratywistów*. Difin Warszawa 2009
16. Januszewski A., *Funkcjonalność informatycznych systemów zarządzania. Systemy Business Intelligence*. WN PWN Warszawa 2013
17. Kisielnicki J., Soroka H., *Systemy informatyczne biznesu*. AW Placet, Warszawa 1995
18. Kolbusz E., (red.), *Inżynieria systemów informatycznych w e-gospodarce*. PWE Warszawa 2005
19. Larose D.T., *Odkrywanie wiedzy z danych*, Wyd. Nauk. PWN, Warszawa 2005
20. Mc Luhan M., *Zrozumieć media: Przedłużenia człowieka*. WNT Warszawa 2004
21. Nycz M., *Hurtownie danych i Business Intelligence w organizacji*. AE Wrocław 2012
22. Piatetsky-Shapiro G., Frawley W.J., *Knowledge Discovery in Databases*. AAAI/MIT Press, 1991
23. Piatetsky-Shapiro G., Smyth P., Uthurusamy R.; *Advances in Knowledge Discovery and Data Mining*, ed. Fayyad U.M., MIT Press, 1996
24. Rosenberg J., Mateos A., *Chmura obliczeniowa. Rozwiązania dla biznesu*. Helion S.A. Gliwice 2011
25. Shannon C.E., *A Mathematical Theory of Communication*. "The Bell System Technical Journal", Vol. 27, 1948
26. Surma J., *Business Intelligence. Systemy wspomaganie decyzji biznesowych*. WN PWN Warszawa 2013
27. Tadeusiewicz R., *Spoleczność Internetu*. AOW EXIT, Warszawa 2002
28. Toffler A., *Trzecia fala*. Wyd. Kurpisz, Warszawa 1997
29. Traczyk W., *Inżynieria wiedzy*. AOW EXIT, Warszawa 2010
30. Trajer J., Paszek A., Iwan S., *Zarządzanie wiedzą*. PWE Warszawa 2012
31. Umesao T., *Joho sangyō ron*. "Hoso Asahi" 1963

ENGINEERING KNOWLEDGE PACEMAKER ENTREPRENEURSHIP IN THE AGE OF THE NEW ECONOMY

Summary

The article presents information challenges of the new economy enterprises in the field of efficient use of resources in information stored in the information and communication systems. Modern companies are inundated with an avalanche of information in the form of elementary context-source data with little usefulness in making business decisions. This phenomenon is to counteract modern technologies, data extraction and compression of information among developed within the engineering know-between (Knowledge Engineering). Were discussed in detail three selected methods of OLAP (On-Line Analytical Processing), DM (Data Mining) and BI (Business Intelligence) is increasingly used on virtual platforms Cloud Computing. Intelligent selection of information source and dedicated personalization business knowledge is a huge challenge to the entire sector IT and requires radical solutions conceptual and technological.

Keywords: data, information, computer, intelligence, engineering, knowledge.

Prof. zw. dr hab. Krzysztof Ficoń
Akademia Marynarki Wojennej
81-103 Gdynia ul. inż. J. Śmidowicza 69
e-mail: krzysztof.ficon@wp.pl