

Bazy Danych

Jesień 2018

Baza danych

- Zbiór danych lub jakichkolwiek elementów i materiałów zgromadzonych według określonej systematyki, indywidualnie dostępnych w jakikolwiek sposób, w tym środkami elektronicznymi. Za: Ustawa o prawie autorskim i prawach pokrewnych z 4 II 94.
- Indywidualny dostęp (bezpośredni). Do każdej informacji zapewniony jest oddzielny dostęp, tak aby nie trzeba było odtwarzać wszystkich danych zawartych w bazie. Bazą nie są zatem utwory audiowizualne, fonogramy.

Cele wykorzystywania BD

- **Wspomaga działanie firmy.**
- **Dostarcza model funkcjonowania firmy wspomagający analizę sytuacji firmy i proces decyzyjny.**

Baza Danych (BD)

- Jest centralną przechowywalnią zasobów informacyjnych wyróżnionej jednostki organizacyjnej, tematyki, także osób prywatnych.
- Zarządza informacją i koordynuje dostęp do niej.
- Umożliwia: odszukiwanie, analizowanie, manipulowanie, wyświetlanie i drukowanie dużych ilości danych.

Podstawowe funkcje bazy danych

- **przechowywanie danych**,
- **importowanie danych** (możliwość przeniesienia zasobów informacyjnych, całości lub części z innych źródeł informacji, np. z innych baz, plików),
- **wprowadzanie danych** (zazwyczaj złożona procedura pobierania i wstępnej kontroli wprowadzanych informacji),
- **wyszukiwanie danych** (funkcja wyszukiwania informacji – podobna do stosowanych w wyszukiwarkach internetowych).

Funkcje BD cd.

- „**wybierania**” (filtry, kwerendy) danych. **Filtry** umożliwiają wybrać z całej bazy tylko te części rekordów, które spełniają określone kryteria, np. z kadrowej bazy pracowników wybór pracowników w wieku 30 – 50 lat. **Kwerendy** zawężają zasób informacyjny bazy tylko do potrzebnych informacji oraz umożliwiają dokonywania obliczeń na wybranych zasobach, które mogą być zbierane z wielu tabel jednocześnie,
- **eksportowania danych** (możliwość przeniesienia zasobów informacyjnych do zewnętrznych baz danych, plików itp. – analogicznie jak przy imporcie),
- zapewnienie **dostępu do danych** poprzez dedykowaną aplikację kliencką, interfejs sieciowy itp.

Cechy baz danych

- Trwałość – dane muszą być przechowywane przez pewien czas.
- Zgodność z rzeczywistością.

Inne cechy BD

- **Kontrolowanie replikacji** – jeden fakt odzwierciedlany w BD powinien być reprezentowany tylko raz. Należy wykluczyć by np. nazwa tego samego produktu wraz z jego opisem występowała w bazie wielokrotnie.
- **Spójny system reprezentacji danych** wykluczający m.in. wielokrotne wprowadzanie danych opisujących ten sam element bazy (np. wspomniany produkt),
- **Współbieżny dostęp do BD przez wielu użytkowników**, który zapewnia jednoczesną, bezkonfliktową pracę użytkowników odczytujących zasoby bazy,
- **Ochrona danych**, ta cecha baz dotyczy bezpieczeństwa zasobów bazy, a w szczególności wykluczenia nieautoryzowanej zmiany zasobów, nieautoryzowanego dostępu, niekontrolowanego kopiowania zasobów i fizycznego bezpieczeństwa zasobów (np. przez odpowiednio częste robienie backupów),
- **Niezależność danych** – zmiany danych w jednym rekordzie ani wykonywanie jakichkolwiek użytkowych funkcji korzystania z zawartych tam informacji, nie wpływają na treść innych rekordów. Przykładem tego jest gwarancja, iż zmiana nazwiska jednej osoby nie wpłynie na zmianę nazwiska innej osoby zarejestrowanej w bazie (w praktyce cecha ta, mimo pozorów błahości, nie jest banalna).

Principia baz danych:

- **informacja w centrum uwagi** - aplikacje powinny być budowane wokół bazy danych, nie odwrotnie. Dane są dobrem wspólnym,
- **eliminacja nadmiarowości** - informacja w jednym miejscu, łatwa do aktualizacji,
- **szybkie, najlepiej bezpośrednie gromadzenie informacji** ze źródeł danych, informacja wprowadzana do BD natychmiast po jej powstaniu,
- **baza danych powinna być obrazem rzeczywistości** tzn. im dane bardziej aktualne tym baza wierniej oddaje rzeczywistość.

Informacje w jednym miejscu

- łatwiej zmienić dane („za jednym zamachem”),
- spójność danych (na ten sam temat),
- oszczędność miejsca na dysku,
- (powiązania, zależności różnych rekordów).

Wyzwanie do usprawnienia organizacji firmy.

Zarządzanie dokumentami

- Nie przesyłanie dokumentów: nadawca – odbiorca.
- Zarządzanie dostępem do wszystkich dokumentów zgromadzonych w jednym miejscu.

Projekt BD dla firmy = informacyjny model firmy

Wyzwaniem dla większości organizacji nie jest rozwiązanie problemów z góry, ale rozpoznanie na tyle wcześnie by mogły być rozwiązane odpowiednio wcześnie. Jeśli brakuje czasu to albo zadawane były złe pytania, albo udzielane złe odpowiedzi.

Niektóre funkcje Access'a

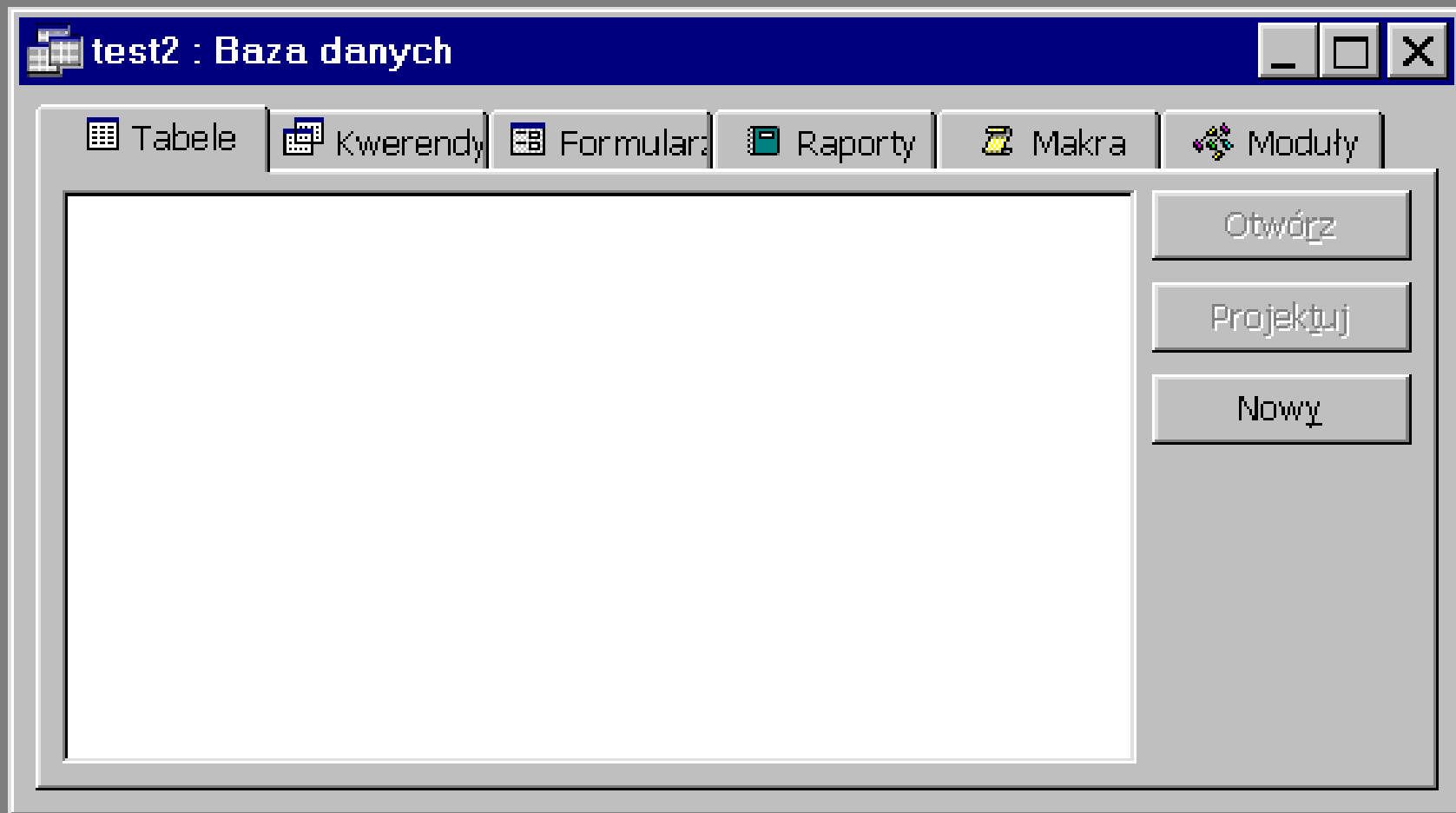
- Tworzenie tabeli w widoku Arkusz danych
- Wprowadzanie informacji do arkusza (dodawanie, usuwanie i edycja rekordów)
- Relacje pomiędzy tabelami
- Kreatory tabel
- Ustalanie formatów pól
- Sortowanie, filtry i kwerendy

Przykład fragmentu tabeli i interpretacji pojęć: rekord i pole

Nazwy pól	Nazwisko	Imię	Data urodzenia	e-mail	ltp.
rekordy	Abacki	Andrzej	01-11-1987	a.abacki@serwer	
	Babacki	Andrzej	19-06-1986	b.babacki@serwer	

Pola rekordów

Tworzenie bazy w Access'ie



Widok arkusz danych

Microsoft Access - [Tabela1 : Tabela]

Plik Edycja Widok Wstaw Format Rekordy Narzędzia Okno Pomoc

Pole1 Arial CE 10 B I U

	Pole1	Pole2	Pole3	Pole
▶				

Nazwisko Arial CE 10

	Nazwisko	Pole2
▶		

Ustalanie formatów pól

- Data/Godzina, typ danych
- Liczba i Walutowy, typy danych
- Tekst i Memo, typy danych
- Tak/Nie, typ danych

The image shows a screenshot of a software configuration dialog box. The dialog has two tabs: 'Ogólne' (General) and 'Oдноśnik' (Reference). The 'Ogólne' tab is active. The settings are as follows:

Rozmiar pola	50
Format	[Dropdown menu]
Maska wprowadzania	
Tytuł	
Wartość domyślna	
Reguła poprawności	
Komunikat o błędzie	
Wymagane	Nie
Zerowa długość dozwolona	Nie
Indeksowane	Nie

At the bottom of the dialog, there is a status bar with the text: 'Widok Projekt F6 = Przełącz okienka F1 = Pomoc'. Below the dialog is a Windows taskbar with the Start button, several application icons, and a system tray containing 'PointCast Network v...' and 'Micros'.

Formularz i raport

formPelny : Formularz

Tytuły



Tytuł | DataUmowy | Wydawnictwo | WarunkiFilm | WarunkiPrawa | OfertyDruku | D

Tytuł:

LbaStron:

ISBN:

RokWyd:

UwagiTytuł:

Rejestr_pelny

Tytuł

TerminyPlat

UwagiDruka

ISBN

UwagiTytuł

CenaEgzDruka

ZaplacDataDruka

OfertaDodruku

UwagiWydawnict

CenaK4

DataK4

Zintegrowane bazy danych

Stare założenia

**Aplikacje są najlepiej
budowane w
wyspecjalizowany sposób.
Każdy oddział najlepiej zna
swoje potrzeby**

**Otrzymanie informacji
wymaga czasu**

Nowe założenia

Aplikacje nadal
specjalistyczne ale dane
są wspólnymi zasobami

Dowolna informacja jest
dostępna przez Sieć.
Uniwersalna BD, zawsze
aktualna

Relacyjny model bazy danych

- Baza danych jest zbiorem relacji.
- Pojedyncza relacja jest dwuwymiarową tabelą złożoną z kolumn i wierszy.

ID agenta	Imię	Nazwisko	rejon
12234	Adam	Abacki	Wola
3456	Zenek	Babacki	Żerań

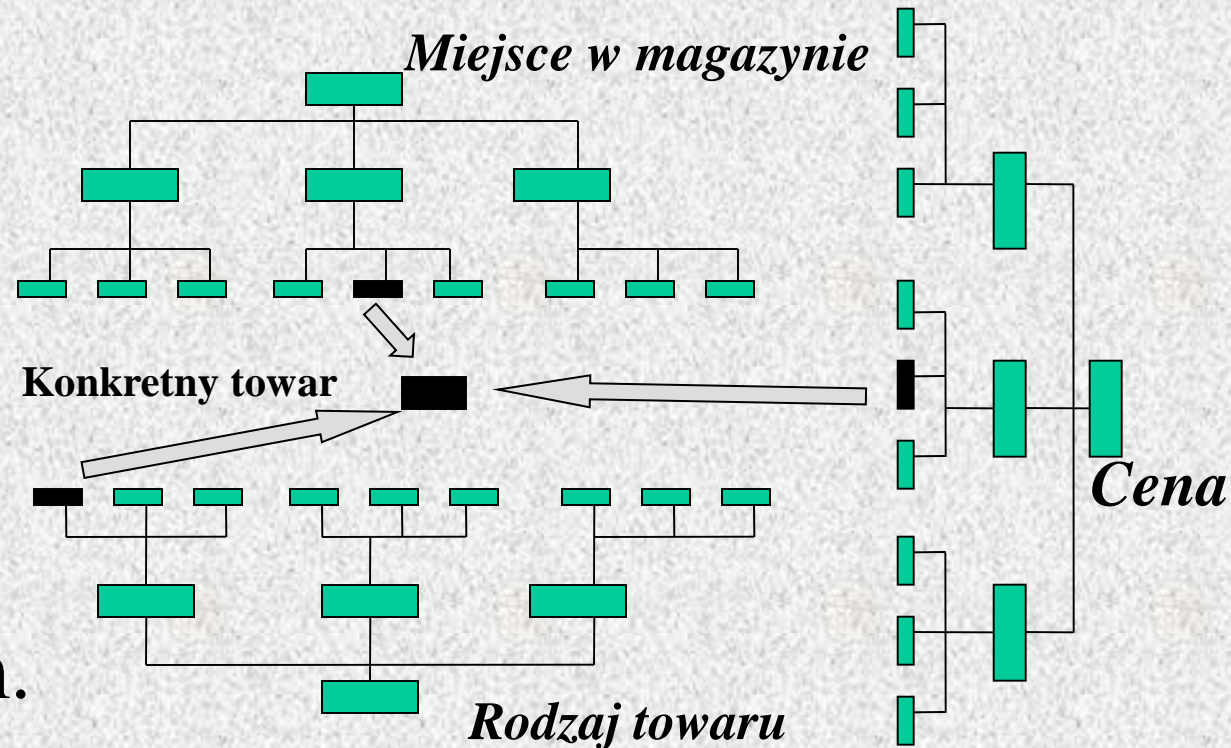
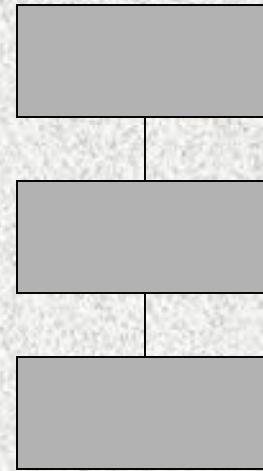
Agenci

Klienci

ID klienta	Imię	Nazwisko	ID agenta
4567	Waldemar	Cabacki	12234
3456	Wojtek	Dabacki	3456

Typy BD

- **Hierarchiczne** - każdy rekord powiązany jedynie z jednym nadrzędnym rekordem.



- **Sieciowe** - każdy rekord może być powiązany z dowolnym innym.

Relacyjne BD (RBD)

- Plik=tablica, rekord=wiersz, pole=kolumna.
- Relacyjne BD w przeciwieństwie do sieciowych i hierarchicznych pozwalają „w locie” określić zależności pomiędzy pierwotnie niezależnymi plikami.
- Użytkownik może łatwo zdefiniować zależność między dwiema tablicami (w sieciowych a priori) .
- Wcześniej każda RBD wyposażona w SQL.

MySQL

- **Relacyjny system baz danych.**
- **Obsługuje zapytania w SQL.**

SQL

Strukturalny język zapytań - SQL
(structured query language)

Relacyjne bazy danych

Oracle, DB2

- posługują się językiem zapytań SQL,
- posiadają wbudowane języki wewnętrzne tworzenia nowych procedur zarządzania bazami.

Hurtownie danych

- Są to zorientowane tematycznie, integralne i uwzględniające wymiar czasowy zbiory danych wykorzystywane w procesie podejmowania decyzji.
- Dane gromadzone w hurtowni pochodzą z wielu źródeł i pozwalają na szybką i efektywną analizę zawartości zgromadzonych informacji.

Hurtownie - zastosowania

- Stosowane są do tego celu narzędzia analityczne hurtowni danych, m.in. służące zadawaniu zapytań i generowaniu raportów, analitycznemu przetwarzaniu online.
- Z hurtownią współpracują takie systemy jak na przykład: – system sprawozdawczości wewnętrznej, dedykowany do zarządzania ryzykiem oraz oceny rentowności departamentów, klientów i produktów firmy/korporacji, – system umożliwiający analizę bilansu, rachunku wyników, analizy wskaźnikowe oraz wspomaganie procesów planowania i kontroli, – system sprawozdań branżowych – wspomaga tworzenie.

Prawo

- Bazy danych podlegają ochronie – ustawa z 27.07.2001 o ochronie baz danych.
- Ochrona prawno-autorska bazy dotyczy jej struktury, nie zawartości.

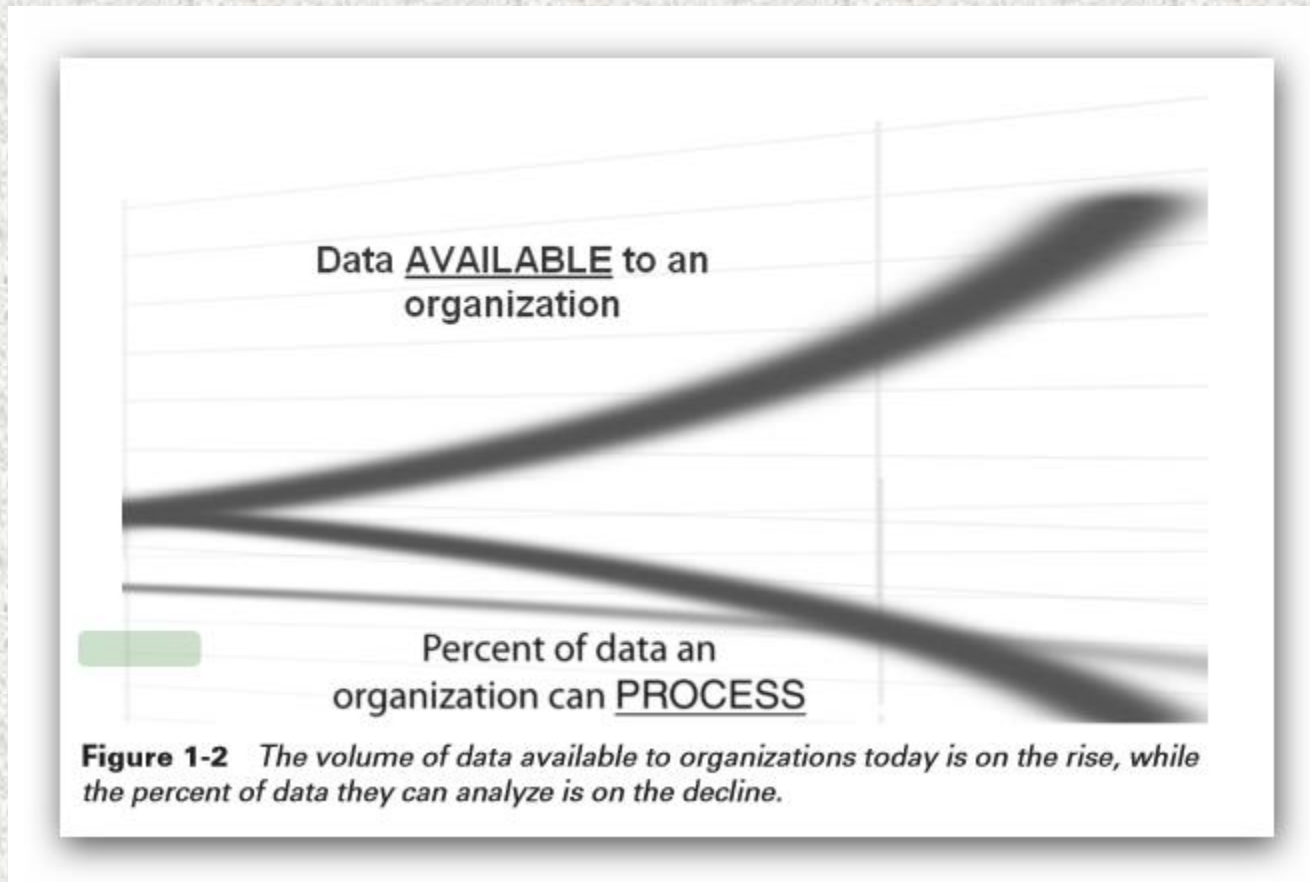
Big Data – Ekstremalne bazy

- "Big data" – bazy danych przekraczające możliwości typowych narzędzi programowych baz danych w zakresie zbierania, utrzymywania, zarządzania i analizowania.
- Dane uzyskiwane z komputerów stacjonarnych, mobilnych, komórek, sensorów (RFID) i innych urządzeń wspomagających operowanie informacjami.
- Dane obejmujące sektor prywatny i publiczny (np. opieka zdrowotna).

Big data vs hurtownie/bazy

- Gromadzenie danych w bazach kosztowne, ustrukturyzowane, czasochłonne, wymagające zabiegów kategoryzacji, czyszczenia, wyboru.
- Dane w Bazach uproszczone, mniej wartościowe od BD.
- Big Data obejmuje dane szybko zmienne, w ruchu – streams computing, które zazwyczaj nie trafiają do baz.

Wartość rekordów w bazach jest znacznie większa od elementów BD.



Streams computing – BigInsights IBMa

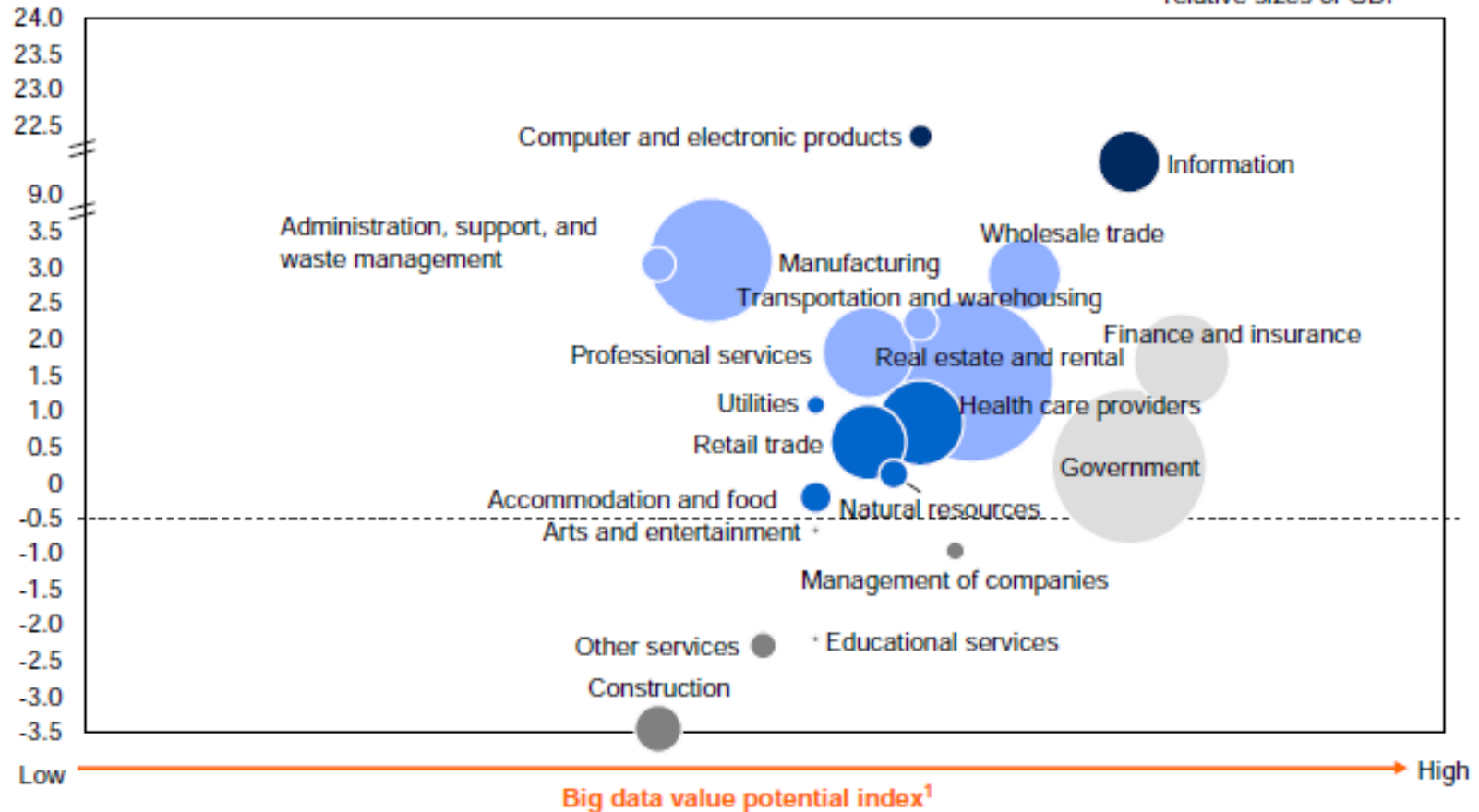
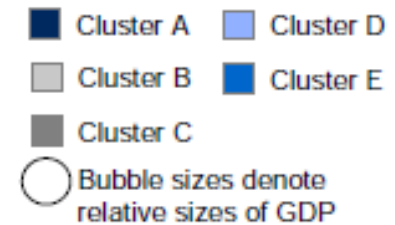
- Analizowanie danych w ruchu, z mikroopóźnieniami, nie tradycyjnie – gromadzenie w bazach i statyczna analiza.

Poprawa Bezpieczeństwa

- Egzekwowanie, obrona, nadzór i cyber obrona z użyciem streamingu zapewniają duże możliwości poprawy egzekwowania prawa i zwiększenia poziomu ochrony, jak i oferuje nieskończony potencjał, kiedy mowa o typach aplikacji, które mogą być budowane w tej przestrzeni, jak np. rozpoznawanie imienia w czasie rzeczywistym, analiza tożsamości, sytuacyjna świadomość aplikacji, wielofunkcyjny nadzór, wykrywanie cyber ochrony, wire taps, nadzór wideo i rozpoznawanie twarzy. Firmy mogą również wykorzystywać streaming analityczny do wykrywania i zapobiegania cyber atakom poprzez streaming sieci i innych danych systemowych, powstrzymując w ten sposób włamania albo wykrywać szkodliwą aktywność gdziekolwiek w jej sieci.

Some sectors are positioned for greater gains from the use of big data

Historical productivity growth in the United States, 2000–08
%



1 See appendix for detailed definitions and metrics used for value potential index.
SOURCE: US Bureau of Labor Statistics; McKinsey Global Institute analysis

Big data - korzyści

- Dla rozwoju ekonomii.
- Zwiększenie produkcji i konkurencyjności.
- Sektor publiczny bliżej konsumentów.
- Oszczędności, np. w opiece zdrowotnej 8% (USA).

Big data - wyzwania

- Tworzenie transparentności zasobów.
- Eksperymentowanie w celu odkrycia potrzeb, wskazania zmienności i poprawy wydajności.
- Segmentacja populacji w celu dostosowania działań.
- Zastępowanie/uzupełnianie ludzkich decyzji stosownymi algorytmami.
- Innowacyjność opracowywania nowych modeli biznesowych, produktów i usług.