

Zasoby informacyjne

*Dotkliwie odczuwa się
nadmiar wiedzy
przy równoczesnym
niedoborze mądrości
Antoni Kępiński*

<http://www.archive.org/>
150 miliardów stron

- The Internet Archive – cyfrowa biblioteka internetowych stron i innych kluturowych wytworów w postaci cyfrowej.

Informacja – fundament komercji w Sieci

- towarem jest informacja:
 - każda (*any*)
 - wszędzie (*anywhere*)
 - zawsze (*anytime*)
- Internet zdominował sprzedaż detaliczną informacji

Zasoby informacyjne Sieci

- zarysowuje się brak równowagi między narzędziami umożliwiającymi korzystanie z Internetu, a oferowanymi tam zasobami informacyjnymi i usługami (rafinacja informacji sieciowych).
- Sieciowe zasoby zazwyczaj są szczególnie cenne ze względu na ich aktualność i możliwość szybkiego i powszechnego do nich dostępu.
- Wiedza o tym potencjale ma kapitalne znaczenie w obszarach społeczno-ekonomicznym oraz nauki i edukacji.
- W odniesieniu do mediów zasoby informacyjne są wartościowym źródłem informacji backgroundowych i bieżących oraz kontroli poczynań konkurencji.

Wiarygodność informacji sieciowych

- Nawet niewielka część nieprawdziwych danych w wyróżnionym zasobie czyni jego całość mało wiarygodną.
- Stwarza to ryzyko, często bezkrytycznego, wykorzystywania mało wartościowego informacyjnego potencjału Internetu.
- Mamy do czynienia z paradoksem, iż „nieograniczonemu dostępowi do informacji towarzyszy ryzyko uzyskiwania mniejszej wiedzy”.
- Współczesne technologie informacyjne istnieją bez odpowiednio wykształconych ram, w postaci struktur wartości i wzorców mających wpływ na sieciowe zasoby.
- Ramy te nie mają stosownych związków z regulacjami prawnymi oraz instytucjonalnymi strukturami.

Godne zaufania źródła informacji, USA, skala 1-10

	1997 University of Massachusetts survey	2007 Bridge Ratings survey
Friends, family and acquaintances	8.8	8.6
Strangers with experience	4.2	7.9
Teachers	9.2	7.3
Religious leaders	9.0	6.9
Newspapers and magazines	8.1	6.1
Favorite radio personality	6.8	5.5
TV news reporters	7.5	5.2
Bloggers	*	2.8
Advertising	3.3	2.2
Telemarketers	2.1	1.8

Wiarygodne anglojęzyczne źródła informacji

- [Bureau of the Census](#) — dane demograficzne i ekonomiczne
- [Rand Corporation](#) — edukacja, ubóstwo, przestępczość, środowisko naturalne i bezpieczeństwo narodowe
- [Pew Charitable Trusts](#) — Internet w USA, opinia publiczna, klimat globalny, oceany, dziennikarstwo, Hiszpański i inne
- [MacArthur Foundation](#) — globalne bezpieczeństwo i stabilizacja, humanitaryzm, problematyka społeczna
- [JupiterResearch](#) — biznesowe wykorzystanie Internetu i pokrewnych technologii
- [The Gallop Organization](#) — zachowania konsumentów, problematyka samochodowa, usługi biznesowe, edukacja, usługi finansowe, służba zdrowia, sprzedaż i produkcja
- [BizMiner](#) — marketing i finansowe profile w 16 000 biznesowych kategoriach w USA
- [First Research](#) — przemysłowe profile i ekonomiczne profile państw
- **CIA** – (<http://www.cia.gov/cia/publications/factbook/>) - dane o wszystkich krajach świata, od lokalizacji po średnią długość życia.

Eliminowane są szczeble pośrednie wędrowki informacji od twórcy do odbiorcy

- w odniesieniu do tradycyjnych – papierowych publikacji, z jednej strony spowalniały przepływ informacji, z drugiej natomiast czyniły je bardziej wyważonymi, ze znacznie mniejszym ryzykiem upowszechniania informacji niesprawdzonych, niewiarygodnych.
- W efekcie dużego kredytu zaufania, jakim obdarzono zasoby informacyjne online, nastąpiło redukowanie lub wręcz wyeliminowanie pośredników na drodze informacji od „producenta” do odbiorcy.
- Konsekwencja - problem ogromnej liczebności i wiarygodności źródeł online - nazwany przez R. Tadeusiewicza „smogiem informacyjnym”.

Ograniczenia „smogu”

- Korzystanie ze stron internetowych o powszechnie akceptowanej wiarygodności, o wysokich ocenach nadawanych przez uznane autorytety. Ich właścicielami są zazwyczaj instytucje o dużym prestiżu. Innym sposobem uwiarygodniania są:
 - społeczna, środowiskowa akceptacja źródeł, których wyraz stanowi **liczba i częstotliwość powoływania** się na te źródła przez uznane autorytety, np. w świecie nauki.
 - **porównywanie informacji** z różnych niezależnych źródeł sieciowych,
 - korzystanie z **komercyjnych zasobów informacyjnych** oraz **stron urzędów** biorących udział w tworzeniu prawa i generujących raporty na potrzeby państwa.

Wiarygodne źródła informacji w Sieci

- strony ministerstw, rządu i innych agencji biorących udział w tworzeniu prawa i generujących raporty na potrzeby państwa
- Komercyjne źródła informacji tematycznych (np. D&B)
- źródła informacji eksperckiej
- firmy zajmujące się badaniem rynku serwisy biblioteczne
- agencje informacyjne
- dystrybutorzy specjalistycznych materiałów informacyjnych (w tym biznesowych)

GIODO

- Wyszukiwanie

Funkcja "Wyszukiwanie" umożliwia wyszukiwanie zbiorów w prowadzonym przez **GIODO** ogólnokrajowym rejestrze zbiorów danych osobowych, według kryteriów podstawowych do których należą: nazwa administratora danych, nazwa miejscowości, nazwa zgłoszonego zbioru, numer księgi rejestrowej oraz numer zgłoszenia.

- Wyszukiwanie zaawansowane

Funkcja "Wyszukiwanie zaawansowane" umożliwia przeszukiwanie rejestru według kryteriów rozszerzonych, do których należą: nazwa administratora danych, siedziba administratora (nazwa miejscowości, nazwa ulicy, kod pocztowy, REGON, nazwa zgłoszonego zbioru, numer księgi rejestrowej oraz numer zgłoszenia).

Ustawy, dzienniki urzędowe

- Od 1 stycznia 2012 tylko w wersji elektronicznej
- Nieodpłatnie z Rządowego Centrum Legislacji

Wiarygodne źródła informacji

- Informacje o prawie unijnym [Eur-lex.europa.eu/pl/index.htm](http://eur-lex.europa.eu/pl/index.htm)
- Informacje o zamówieniach publicznych
- Wyszukiwarka firm wpisanych do KRS – pdi.cors.gov.pl/KRSSED
- Rejestr zlikwidowanych firm:
www.zus.pl/default.asp?p=4&id=582

Rankingi, nagrody - wiedza o zasobach informacyjnych

- Są to na przykład witryny: CheckPageRank, Alexa i Page Rank Google.
- **nagrody**, na przykład nagroda Pulitzera nadawana materiałom publikowanym online

Handle System - inicjatywa Wspólnoty Europejskiej

- umożliwia wyszukiwanie aktualnych i wiarygodnych danych dotyczących UE.
- Wykorzystywany w nim *Digital Object Identifiers* (DOI) jest standardową metodą identyfikowania publikowanych cyfrowo informacji.
- *Official Publications of the European Communities* (OPOEC) zawiera 60 serwisów z 1700 tytułami. Informacje o tych zasobach ukazują się dwa razy dziennie w specjalnym piśmie, w dwudziestu językach Wspólnoty.
- *CrossRef* (www.crossref.org) – strona wskazująca linki publikacji naukowych do cytowanych dzieł, co stanowi formę gwarancji wiarygodności przywoływanego źródła.
- *Content Directions* (CDI) – stosowne referencje, w tym adresy stron dotyczące określonego tematu.



LexisNexis™

- Popularne archiwum treści gazet, magazynów, aktów prawnych i innych źródeł drukowanych
- Księgarnia internetowa



Lexis + Nexis

LexisNexis™

Dwie osobne strony internetowe:

- **Lexis.com** – informacje prawne
- Bezpieczne przechowywanie danych
- Dostęp do zasobów sądowych
- **Nexis.com** – biznes i artykuły prasowe
- Teksty od 1970 r.

Sposoby korzystania z serwisów Lexis

- Wykupienie abonamentu dla firmy
- Uaktywnienie hasła przez SMS
- Korzystanie z księgarni internetowej
- Specjalne oferty dla studentów

LexisNexis w Polsce

- LexPolonica – największy serwis prawniczy na rynku
- Serwis kadrowy
- podatkowy
- nieruchomości
- biznesowy
- akademicki

TeleAdreson

- zbiór danych o niemal milionie firm w Polsce.
- Poszukiwany podmiot można wybrać zgodnie z katalogami kategorii aktywności biznesowej (NACE (EKD) lub/i SIC).
- TeleAdreson umożliwia rozwinięcie owych kategorii i wybranie tej, która obejmuje zakres aktywności poszukiwanej firmy.

Bezpłatne informacje teleadresowe

- Panorama Firm (www.pf.pl),
- Polskie Książki Telefoniczne (www.pkt.pl),
- www.pf.pl, www.bazafirm.pl.

Interia.pl

- Oferuje swoją własną bazę, zasoby TeleAdresonu, Kompassa i Katalog firm Getin.
- Interia umożliwia wyszukiwanie danych o firmach także według informacji dokładniejszych (zatrudnienie, obroty, rok założenia).

Kompass

- Globalny serwis informacyjny, informacje o 1,6 mln firm z 70 krajów (2004 rok).
- Przeszukiwanie setek tysięcy kluczowych produktów i usług oraz ponad trzy miliony nazwisk osób pełniących ważne funkcje w firmach.
- Kompass umożliwia, w ograniczonym zakresie, bezpłatne korzystanie z następujących informacji: produkty/usługi, nazwy firm/handlowe, przegląd geograficzny (cały świat, regiony lub kraje), listy firm, profile firm (niepełne profile zawierają tylko dane kontaktowe).
- Komercyjna oferta: listy firm (z numerami telefonów i faksów), pełne profile firm, sortowanie listy firm (kryteria sortowania to m.in. liczba zatrudnionych osób, wielkość obrotu/przychodu ze sprzedaży, dane finansowe).

Dun & Bradstreet

- Informacje z niemal 100 mln firm, z ponad 200 państw i 95 stref językowych (2005 rok).
- Każdego dnia wprowadza się około milion aktualizacji – przede wszystkim i najdokładniej z miejsca powstania firmy – z USA.
- Ponad 160-letnie doświadczenia stanowią o profesjonalizmie oferty tej firmy.
- Ma ona także swój oddział w Polsce (www.dnb.com.pl).

Raport – podstawowa forma udostępnianej przez Dun & Bradstreet informacji o firmach

- dane identyfikacyjne (nazwa, adres itp.);
- ocena wykonana przez Dun & Bradstreet, a w tym opinia o kondycji firmy (np. „raczej słaba” zdolność finansowa)
- podstawa wydanej oceny; dane rejestrowe (forma prawna, data założenia, NIP itp.); kierownictwo; struktura własności; działalność według klasyfikacji branżowej; partnerzy handlowi; nieruchomości; zatrudnienie (łącznie z ilustracją graficzną w ostatnich dziesięciu latach);
- banki obsługujące;
- powiązania kapitałowe; dane finansowe – sprzedaż w ostatnich latach,
- zysk/strata, sprawozdania finansowe; historia; informacje prasowe.
- użytkownicy chętnie korzystają z możliwości wybierania tylko niektórych informacji z raportu o ocenianej firmie – koszt uszczuplonego raportu jest wówczas odpowiednio mniejszy.

Podobne do Kompassa i Dun & Bradstreet serwisy

- Francja - Delphes,
- Niemcy – Hoppenstedt,
- Anglia – ICC.

Dialog od 1972

- zawiera setki baz danych, takich jak na przykład Kompass czy Dun & Bradstreet.
- niemal 1000 baz danych.

Narzędzia Dialogu

- Wzór wyszukiwarki
- wstępna obróbka odszukanych informacji (np. sortowanie według specyficznych kryteriów) oraz wyboru najbardziej odpowiadającego poszukiwaniom formatu wyświetlania danych.
- Formatem może być dowolna część danych określonego dokumentu (np. nazwa firmy, adres, obroty, tytuł).
- różne formaty mają inne ceny, a zatem można zmierzać do kompromisu między ceną a zakresem uzyskanej informacji.

Na wielkość opłaty za uzyskane informacje składają się dwa elementy

- czas, w jakim jest przeglądana określona baza, i ilość pobranej informacji.
- Ta zasada jest uniwersalna i dotyczy wszystkich baz.

Zasoby Dialogu

- tworzą między innymi: 5 mld dokumentów i 6 tys. pełnych tekstów bieżących i archiwalnych pism.
- Na wstępie wędrówki po zasobach Dialogu pomagają tzw. kartki – przewodniki po bazach.
- Są ich trzy rodzaje, najważniejsze z nich to Bluesheets (BS) – dotyczą one baz zbudowanych samodzielnie przez Dialog.
- BS-y są precyzyjnymi informatorami o formacie i zawartości każdej bazy. Zawierają dokładne wytyczne o sposobie poszukiwania informacji, opis bazy, wyróżnione w niej pola, częstotliwość aktualizowania danych, źródła danych oraz przykład postaci rekordu w danej bazie.
- Najważniejszą częścią BS-ów jest jednak instrukcja dotycząca możliwości i sposobu wyszukiwania informacji.

Bogactwo informacyjne Dialogu

- zasoby Dun & Bradstreet - informacje o ponad 20 mln firm z 200 krajów (2004 rok).
- informacje o 16 mln europejskich firm zlokalizowanych w 57 krajach – baza D&B – *European Dun's Market Identifiers*. Obie bazy zawierają takie informacje, jak np.: nazwa firmy, adres, profil działalności

Investext

- Największa światowa baza danych zawierająca ekspertyzy, wyniki badań firm i branż na całym świecie.
- Pełne teksty zawierają raporty o 1,6 mln firm, o branżach (np. przemysł wydawniczy w Polsce), które zostały wypracowane przez ekspertów z 600 uznanych banków, brokerów i firm konsultingowych na całym świecie (2004 rok).
- Raporty te są nadzwyczaj użyteczne w badaniach rynkowych, marketingowych, w opracowywaniu strategicznych planów, analizie konkurencji i finansowych prognozach. Format tej bazy jest zgodny z zasygnalizowanymi wcześniej standardami Dialogu.

Zakres informacji serwisu informacyjnego Dialog

- informacje o firmach na świecie – katalogi i dane finansowe
- informacje o gałęziach przemysłu – trendy, oceny, badania marketingowe i specjalistyczne wiadomości branżowe wraz z raportami
- Międzynarodowe, bieżące wiadomości z agencji informacyjnych i gazet od USA do Azji
- wiadomości administracji USA, łącznie z danymi o problematyce publicznej, prawie i innych ważnych regulacjach
- patenty i znaki towarowe w skali świata
- chemia, badania podstawowe, nauka i technologia
- informacje ogólne: ludzie, książki, wiadomości dla konsumentów i podróżujących
- informacje farmaceutyczne, medyczne i dotyczące opieki zdrowotnej
- informacje o firmach europejskich (łącznie z Polską) – katalogi i dane finansowe
- wiadomości lokalne ekonomiczne, polityczne i biznesowe z całej Europy, bieżąca i archiwalna legislacja w ramach UE
- raporty – pełne teksty informacji o państwach, firmach i rynkach, informacje opisujące przemysł samochodowy i spożywczy – biznes i technika

Bezpłatne biznesowe zasoby informacyjne świata

- **World Email Directory (WED)** – adresowe informacje o 140 mln firm na świecie,
- **The European Business Directory** – oferty kupna (600 tys. firm) i sprzedaży (500 tys. firm) w 33 krajach Europy,
- **BigYellow** – katalog z danymi USA: miliony firm, tysiące kategorii.

Dane o osobach

- dostępne we wszelkiego rodzaju katalogach typu Who is Who (np. obecny w Dialogu – *Marquis Who's Who*®),
- Social networking - SoFlow.com – miejsce w wirtualnym świecie, które pomaga odszukiwać osoby mające wspólne zainteresowania, usiłujące nawiązać współpracę na płaszczyźnie prowadzenia interesów (partnerzy, inwestorzy, kontakty). Jest to rodzaj aukcji skrzyżowanej z blogiem, gdzie przedmiotem są osoby.

Jupitermedia Corporation

(www.jupitermedia.com)

- przykładowy globalny serwis informacyjny w zakresie technologii informacyjnych. Tworzą go firmy: internet.com i EarthWeb.com w zakresie IT; DevX.com dla developerów; ClickZ.com dla interaktywnego marketingu i Graphics.com dla artystów grafików.
- JupiterWeb zawiera ponad 150 stron i generuje ponad 150 newsletterów dostarczanych via mail do około 20 mln użytkowników.
- O bogactwie zasobów Jupitermedia świadczy baza ponad 5 mln clipartów, animowanych obrazów, fotografii, znaków i dźwięków.
- Wyróżniającą się częścią Jupitermedia jest The ClickZ Network
- (<http://www.clickz.com/about/>) – rzetelne źródło wiadomości rynkowych świata.

Clickz.com

- ClickZ News: informacje z ostatniej chwili, aktualne wiadomości i analizy,
- ClickZ Experts: komentarze, opinie, rady,
- ClickZ Stats: fakty, wyniki badań i dane dotyczące IT na świecie,
- ClickZ Features: dane o firmach, branżach, analizy, trendy i przegląd literatury przedmiotu,
- Search Engine Watch: bieżące wiadomości, analizy dotyczące problematyki wyszukiwania informacji w Internecie,
- ClickZ Resources: informacje o zasobach Sieci,
- E-Mail Reference: bogate archiwum dotyczące marketingu mailowego i bieżących doniesień w tej tematyce,
- SEM Archives: zasoby dotyczące marketingu z wykorzystaniem szperaczy,
- Events: komunikaty o konferencjach, seminariach organizowanych przez clickZ i inne firmy,
- ClickZ Marketing Excellence Awards: wyniki własnych rankingów dotyczących technologii, firm związanych z IT.

Wikipedia

- Wirtualną encyklopedię powołali do życia 15 stycznia 2001 r. Jimmy Wales i Larry Saneer.
- Obsługiwana przez 378 serwery w 2 centrach
- 17 mln artykułów – 3,5 – UK; 0,8 – PL
- 273 edycje językowe

Wiki PL

- Trzon polskiej Wikipedii tworzy około 2000 najaktywniejszych autorów.
- Nad działaniem portalu czuwa 168 administratorów, posiadających szczególne uprawnienia i mogących na przykład całkowicie usuwać cudze artykuły.

Wikipedia, wiki

- Dynamicznie rozwijający się ogólnoswiatowy zasób informacyjny w postaci encyklopedii (dwanaście razy większa od *Encyklopedia Britannica*, zawiera trzy miliony artykułów w 200 językach).
- Cel stworzenie największej wielojęzycznej, dostępnej dla wszystkich encyklopedii.
- Autorem jej haseł może być każdy. Stanowi to o coraz bardziej podkreślanej, dyskusyjnej wartości merytorycznej tego typu zasobu informacyjnego.
- Eksperti nauk przyrodniczych stwierdzili porównywalną liczbę błędów w encyklopedii brytyjskiej (123) i Wiki (162).

Język polski jest szóstym, po angielskim, niemieckim, japońskim, francuskim, włoskim, językiem Wikipedii, *The Wiki Principle. Among the Sudence, A Survey of New Media*, „The Economist” 2006, April 22.

Wiarygodność Wiki

- Dziennikarze Nature porównali 42 tematy/obszary Wikipedii i encyklopedii Britannica Online
- Wiki: 4 poważne błędy + 4 drobne nieścisłości
- Britannica: 3 nieścisłości, 4 błędy
- Błędy faktograficzne: 162 i 123 odpowiednio w Wikipedii i Encyklopedii

Wiarygodność

„bardzo ograniczonych mechanizm weryfikacji”

- Przedwojenny działacz komunistyczny, więzień polityczny, uczestnik wojny domowej w Hiszpanii i wreszcie patron ulicy na warszawskim Służewcu - historia Henryka Batuty została dokładnie opisana w listopadzie 2006 r.
- Postać wykreowana, nieistniejąca.
- Mistyfikacja pozostała niewykryta przez ponad rok.

WikiScanner – śledzenie zmian

- narzędzie do śledzenia (dyskusyjnych) ingerencji w treści publikowane w wikipedii w celu promocji, kreowania wizerunku, poprawy wizerunku

flagged revisions

- Nowa funkcja Wikipedii – upoważniony wolontariusz musi podpisać każdą zmianę wprowadzoną przez czytelnika.
- Np. dotyczy to „wrażliwych” stron [Britney Spears](#) czy [President Obama](#),

Australia – przykład manipulacji

- Australijski rząd to kolejna instytucja, która poprawia informacje umieszczane w Wikipedii na swój temat.
- System Wikipedii wykrył, że australijskie ministerstwo obrony wprowadziło 5 tys. zmian w hasłach encyklopedii. Częstych zmian w Wiki dokonuje również CIA oraz BBC.

Przykłady wandalizmu w Wikipedii

- Usuwanie fragmentów z krytycznymi informacjami (politycy, korporacje)
- White-washing – zastępowanie negatywnych/neutralnych ocen pozytywnymi (szczególnie popularne wśród polityków)
- Dodawanie negatywnych informacji konkurencji (głównie firmy)

KNOL vs Wikipedia

- Serwis Knol, skrót od knowledge, pozwoli tworzyć strony każdemu, na dowolny temat. Zawiera usługę komentarzy, ocen i sugestii zmian.
- Zmiany może wprowadzić tylko autor.
- Wikipedia jest not-for-profit i bez reklam. Knol może zamieszczać reklamy Google'a.

PLATFORMY NIESKRĘPOWANEJ WYMIANY INFORMACJI I OPINII

- W grudniu 2006 r., powstała nowa platforma swobodnej dystrybucji informacji ukrywanych przed społeczeństwami różnych krajów: WikiLeaks.
- Upowszechnienie technologii informacyjnych prowadzi do większej transparentności władzy politycznej. Głównym motorem tego stanu, wyznacznikiem nowej ery w komunikacji, Internet.

WikiLeaks

- Opublikował materiały dotyczące sekty scjentologów,
- Podał utrzymywaną w tajemnicy listę nazwisk wraz z adresami członków brytyjskich partii neofaszystowskiej.
- W czerwcu 2009 roku WikiLeaks zostało uhonorowane przyznawaną przez brytyjskich dziennikarzy nagrodą za opracowanie raportu w sprawie morderstwa dokonanego przez policję tego kraju. Nagrodę odebrał australijski programista i pisarz, a jednocześnie założyciel i twarz WikiLeaks - Julian Assange.

WikiLeaks, ale

- wiceprezydent Joe Biden o Sarah Palin nazwała J. Assange terrorystą,
- doradca premiera Kanady Tom Flanagan, który powiedział, że Assange'a powinno się zniszczyć za pomocą bezzałogowego samolotu.
- portal oraz jego główny animator znaleźli się pod ogromną presją z wielu stron - WikiLeaks zostało przez polityków okrzyknięte wrogiem publicznym numer jeden.

- Usługa ułatwiająca rozumienie znaczeń
- Definicje z różnych źródeł
- Ilustracja kontekstami

Define Relate List Discuss See Hear Love

Examples

"Your sample picture shows white clipped at ~108IRE (**digital** 256) but black level around 15IRE (~**digital** 32)."
[VideoHelp.com Forum](#)

"We can learn a great deal about these new communications paradigms by the various definitions of the term **digital** media found in the syllabi of **digital** media courses and programs emerging in high schools, colleges, and universities across this country and abroad."
[Netvouz - new bookmarks](#)

"The term **digital** curation is ... the actions needed to maintain **digital** research data and other **digital** materials over their entire life-cycle and over time for current and future generations of users."
[ArchivesBlogs](#)

"GREENFIELD: In fact, some conservatives have been very angry at Al Gore for using the term **digital** brownshirts to describe some of the more conservative media folk."
[CNN Transcript Jul 26, 2004](#)

"I was disappointed in this video for the definition given to the term **digital** story."
[Digital Storytelling In Plain English](#)

Wordnik

digital

Definitions

American Heritage® Dictionary of the English Language, Fourth Edition

adj. Of, relating to, or resembling a digit, especially a finger.

adj. Operated or done with the fingers: *a digital switch.*

adj. Having digits.

adj. Expressed in numerical form, especially for use by a computer.

adj. Computer Science Of or relating to a device that can read, write, or store information that is represented in numerical form. See Usage Note at [virtual](#).

adj. Using or giving a reading in digits: *a digital clock.*

n. A key played with the finger, as on a piano.

Century Dictionary and Cyclopaedia

n. One of those primaries, or quill-feathers, most commonly four in number, which are attached to the phalanges of the second and third digits of a bird's wing.

Of or pertaining to a digit or digits: as, the *digital* phalanges.

Resembling digits; [digitate](#).

n. A digit; a finger or toe.

n. The fifth and last joint of the pedipalp of a spider. It is generally larger than the preceding joints, sometimes much swollen, and in the males modified to form the complicated sexual or palpal organs.

n. One of the keys or finger-levers of instruments of the organ or piano class.

Dane statystyczne o Sieci

- Netcraft. news.netcraft.com,
- ranking.pl,
- www.InternetWorldStats.com,
- Alexa. <http://www.alexa.com/>.

Dane statystyczne w Sieci

- Gemius S.A
- Megapanel (PBI/Gemius)
- Internet Standard / Sklepy 24.pl
- CBOS
- GUS, UOKiK

World Clock

Poodwaddle.com



Thursday 2012-2-9
PM 5:00:58

World Population 6,960,430,533
 Births 14,581,746
 Population Growth 8,376,627
 Prison Growth 11,608
 US Illegal Immigration 54,360
 US Divorces 100,029
 Abortions 5,001,104
 Mothers Die/Abortions 7,610
 HIV Infections 467,495
 Cancer Incidence 1,195,916

Earth Temp (C) 14.6510802789
 Species Extinct 2,935
 Forest Lost (ha) 1,413,356
 Oil Produced (bo) 3,293,662,040
 Cars Produced 3,913,908
 Bicycles Produced 11,855,879
 Computers Produced 9,154,195



Deaths 6,205,119

from TB 170,123
 STD/HIV/AIDS 336,658
 Diarrhoeal diseases 203,122
 Childhood Cluster Diseases 121,795
 Meningitis 18,805
 Hepatitis 17,003
 Malaria 99,012
 Tropical Cluster Diseases 17,161
 Leprosy 670
 Dengue 2,017
 Other Infectuous Diseases 2,812
 Respiratory infections 436,839
 Maternal conditions 55,426
 Perinatal conditions 267,288
 Nutritional deficiencies 52,636

Cancer 789,013
 Diabetes 106,878
 Endocrine disorders 26,231
 Psychiatric conditions 120,552
 Cardiovascular diseases 1,816,106
 Respiratory diseases 436,839
 Digestive diseases 213,564
 Other noncommunicable 164,875
 Traffic Accidents 129,321
 Poisonings 38,042
 Falls 42,543
 Fires 33,852
 Drownings 41,492
 Other Accidents 100,136
 Suicide 94,876
 Violence 60,656
 War 18,593

www.InternetWorldStats.com

- **Internet Usage Stats and Population Statistics**

[Index](#) | [World Stats](#) | [Africa Stats](#) | [America Stats](#) | [Asia Stats](#) | [Europe Stats](#) | [EU Stats](#) | [Middle East Stats](#) | [Oceania Stats](#)

[Spanish Internet Stats](#) | [Top Internet Usage](#) | [Top Internet Penetration](#) | [Top Internet Languages](#) | [Top Ten](#) | [DSL](#) | [Search Site](#) **Country Links and International**

Directory

[Africa](#) | [North America](#) | [Asia](#) | [Central America](#) | [Europe](#) | [European Union](#) | [EU](#)

[Enlargement](#) | [Middle East](#) | [South America](#)

[South Pacific and Oceania](#) | [The Caribbean](#) | [World Population](#) | [Country List](#) | [World](#)

[Internet Users Associations](#) **Internet Marketing Pages**

[Broadband Usage](#) | [B2B Trade Exchanges](#) | [Industry Stats](#) | [Internet Usage](#) | [Internet Divide](#) | [Internet Growth](#) | [Internet Library](#)

[iBurst](#) | [Market Reports](#) | [Market Research](#) | [Newsletter](#) | [PR Coaching](#) | [SEM](#) | [SEO](#) |

[Search Engines & Directories](#) | [Web Stats](#) **Site Resources and Services**

[About Us](#) | [Blog](#) | [Dictionaries](#) | [Tools](#) | [Internet Telephone Calls](#) | [Conference Calls](#) |

[Online Profits](#) | [NFL Tickets](#) | [Privacy Policy](#)

[Archives](#) | [World News](#) | [Software](#) | [e-Learning](#) | [Translation Tools](#) | [Web Services](#) |

[Time and Weather](#) | [Site Surfing Guide](#) **Country Internet Usage and**

Telecommunications Reports

[Africa](#) | [America](#) | [Asia](#) | [Caribbean](#) | [Europe](#) | [EU27](#) | [Middle East](#) | [South America](#) |

[South Pacific](#) **All About**

Market Research

Zasoby fotograficzne

- Fotografia stokowa: Licencja Royalty Free
- Flickr.com – przechowywanie, wyszukiwanie, sortowanie i współdzielenie zdjęć

Agencje microstockowe

- iStockphoto
- Shutterstock
- Dreamstime
- BigStockPhoto
- Fotolia
- 123RF
- inne

Co oferują trzy największe agencje

- iStockphoto: fotografie, ilustracje, wideo
- Shutterstock: fotografie, ilustracje, wideo
- Dreamstime: fotografie, ilustracje

iStockphoto - opłaty

Small or One-Off Projects			
6 CREDITS	12 CREDITS	26 CREDITS	50 CREDITS
\$9.99 (\$1.67/credit)	\$19.50 (\$1.63/credit)	\$41.50 (\$1.60/credit)	\$78.50 (\$1.57/credit)
Buy now	Buy now	Buy now	Buy now
Large or Ongoing Projects			
120 CREDITS	300 CREDITS	600 CREDITS	1,000 CREDITS
\$180.00 (\$1.50/credit)	\$440.00 (\$1.47/credit)	\$850.00 (\$1.42/credit)	\$1,380.00 (\$1.38/credit)
Buy now	Buy now	Buy now	Buy now

Ceny c.d. główna i ekskluzywna kolekcja

XSmall	Small	Medium	Large	XLarge	
1	3	7	10	15	
CREDITS	CREDITS	CREDITS	CREDITS	CREDITS	
Photo	Photo+	Illustration	Video	Audio	Flash

XSmall	Small	Medium	Large	XLarge
3	5	10	15	20
CREDITS	CREDITS	CREDITS	CREDITS	CREDITS
Photo	Exclusive+ photos	Illustration	Audio	

Najpopularniejsze słowa kluczowe - Shutterstock

Top 100 Keywords Ever

1	<u>christmas</u>
2	<u>medical</u>
3	<u>family</u>
4	<u>business</u>
5	<u>flowers</u>
6	<u>people</u>
7	<u>beach</u>
8	<u>wedding</u>
9	<u>computer</u>
10	<u>woman</u>

Top 100 keywords - Past Month

1	<u>christmas</u>
2	<u>medical</u>
3	<u>halloween</u>
4	<u>family</u>
5	<u>flower</u>
6	<u>flowers</u>
7	<u>business</u>
8	<u>wedding</u>
9	<u>music</u>
10	<u>vector</u>

Top 100 Keywords - Past Week

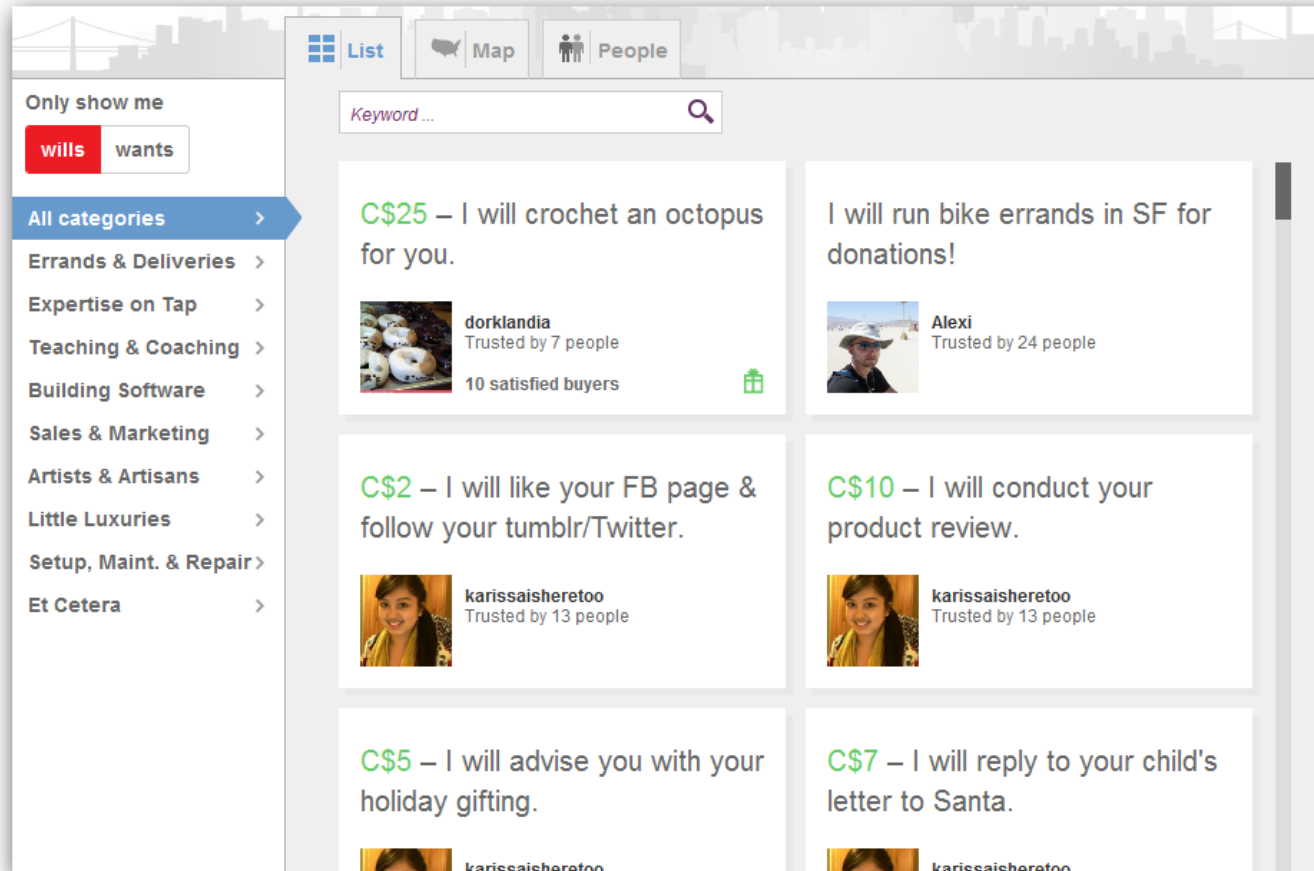
1	<u>christmas</u>
2	<u>halloween</u>
3	<u>medical</u>
4	<u>flower</u>
5	<u>music</u>
6	<u>family</u>
7	<u>vector</u>
8	<u>business</u>
9	<u>flowers</u>
10	<u>background</u>

RapidShare

- Globalny komercyjny serwis do dzielenia się plikami
- Rapid=5.99€ -30 dni, transfer 150 GB
- Internauci wrzucają do niego wszystko, co tylko da się przerobić na postać cyfrową: filmy, muzykę, programy, gry, zdjęcia, e-booki
- Nie ma własnej wyszukiwarki – pomocne www.filestube.com
- Wrzucasz na niego plik, do którego automatycznie jest generowany link; tym linkiem dzielisz się ze znajomymi

Buy & sell small jobs, 15% prowizji

<http://www.coffeeandpower.com>



Podobne serwisy: [Task Rabbit](#), [Freelancer.com](#) i Amazona [Mechanical Turk](#)

Medycyna

- Duża część zasobów medycznych w pełni spełnia kryteria tego, co można określić jako „media”.
- W połowie XIX wieku lekarz Ignaz Semmelweis dzięki prostemu zabiegowi higienicznemu – myciu rąk – radykalnie zmniejszył liczbę umierających pacjentów w swoim szpitalu. Korzystając z tych doświadczeń, usiłował przeforsować obowiązek mycia rąk przez swoich kolegów wykonujących pracę. Nie znalazł jednak zrozumienia.
- Za marnotrawienie czasu na „nikomu niepotrzebne mycie” wyrzucono go z pracy. Podobnie obecnie korzystanie z potencjału komputerów i udostępnionych przez nie zasobów często jest traktowane jako marnotrawienie czasu.

Medycyna online w USA

- Przeciętnie 55,3 miliona osób miesięcznie (31% internautów), odwiedziło Sieć w poszukiwaniu informacji medycznych (Q1 2007).

Medline

- Medyczna baza danych, zawierała w 2004 roku 3672 artykuły na temat zawału serca u dorosłych. Przeczytanie każdego z nich (każdy ok. 15 minut) wymagałoby 115 ośmiogodzinnych dni pracy. Większość osób robiących sprawunki powyżej pięciu rzeczy, sporządza sobie listę zakupów.
- Jak zatem mają sobie poradzić lekarze podczas szybkiego diagnozowania chorego na serce z analizowaniem informacji o 12 000 znanych chorobach tego organu? Komputery (ze specjalistycznym oprogramowaniem) i Sieć jawią się tutaj jako wspomniana troska o higienę.

Mapy, plany

- Google Earth
- Google Maps
- Zumi (mapy Techmeksu i Imagis)
- Panorama Firm
- W Polsce: Navteq i TeleAtlas
- Mapy, plany, zdjęcia + usługi, hybryda map i obrazów

Problem archiwizacji informacji

- Informacje w postaci cyfrowej nanoszone są na przeróżne nośniki (taśmy, dyski, dyskietki, CD, DVD). Jednak wszystkie one gwarantują stosunkowo krótki okres trwałości przechowywanych na nich informacji.
Maksymalnie do 100 lat
- Maleje proporcja informacji gromadzonych na tradycyjnych nośnikach informacji – papierowych – na korzyść zapisu cyfrowego

Cyfrowe nośniki vs papierowe

- Papier przestał być głównym nośnikiem informacji
- Przykład częstotliwości korzystania z książki telefonicznej
- Konsekwencje
 - zagrożenie trwałości zapisów
 - większa konsumpcja energii elektrycznej

Wzorce budowane na podobieństwo wzorców miar

- Niezmienne w czasie, o dużej dokładności, łatwe w odtwarzaniu i stosowaniu.
- Obejście problemu – wykorzystywanie wirtualnych komputerów – odpowiedników ich starszych generacji, np. *Universal Virtual Komputer* (UVC), na którym jest uruchamiane oprogramowanie dekodujące dokumenty zapisane we wcześniejszych formatach.
- Nadal pozostaje jednak nierozwiązany problem trwałości nośników i dostępności urządzeń do ich odczytania (już dzisiaj duża część sprzedawanych pecetów i laptopów nie ma napędu dyskietek).

Bezpieczne formaty

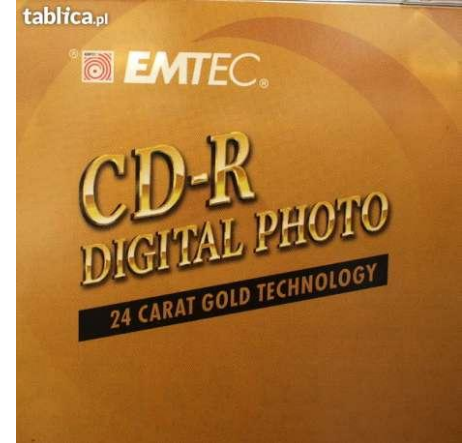
- PDF/A – dla długowiecznej archiwizacji
- Teksty - ODT (OpenOffice) – otwarta specyfika, nikt nie może w niej manipulować
- Obrazy i zdjęcia - TIFF i PNG jw.

Wytrzymałość nośników danych

- Dyskietki, CD/DVD, HD, pendrive (flash) – 5–10 lat
- VHS – 10–15 lat
- Taśma magnetyczna, DVD-RAM – 20–30 lat
- Srebrne CD, DVD do 100 lat, warunek – idealne warunki
- HD – do 5–10 lat
- Flash – ok. 5 lat
- DVD-R – 30 lat, warunek, j.w.
- Preservation Kodaka daje gwarancję do 300 lat
- Uszkodzenia – zadrapania po stronie nośnika i po stronie etykiety (niszczenie warstwy odbijającej, np. przez klej naklejek!)

Jakość nośników

- Kupuj markowe czyste płyty ze złotą, odporną na korozję warstwą odbijającą i twardą, odporną na zadrapanie powierzchnią,
- Np. Emtec CD-R Gold (5 zł sztuka), TDK ScratchProof DVD (8 zł)
- Najpewniej zapisywać na HD – cztery kopie na różnych dyskach



Groźba fizycznej utraty archiwizowanych informacji cyfrowych

- Dzięki tabliczkom z pismem klinowym (IV tysiąclecie p.n.e.), zapisom na papirusie (2500 lat p.n.e.), później na papierze, sznurach *kipu* (pismo węzełkowe Inków) informacje w dużej części przetrwały setki, nawet tysiące lat.
- Takiej gwarancji nie zapewniają nośniki informacji z zapisami w postaci cyfrowej

Ułomność komputerowych nośników danych

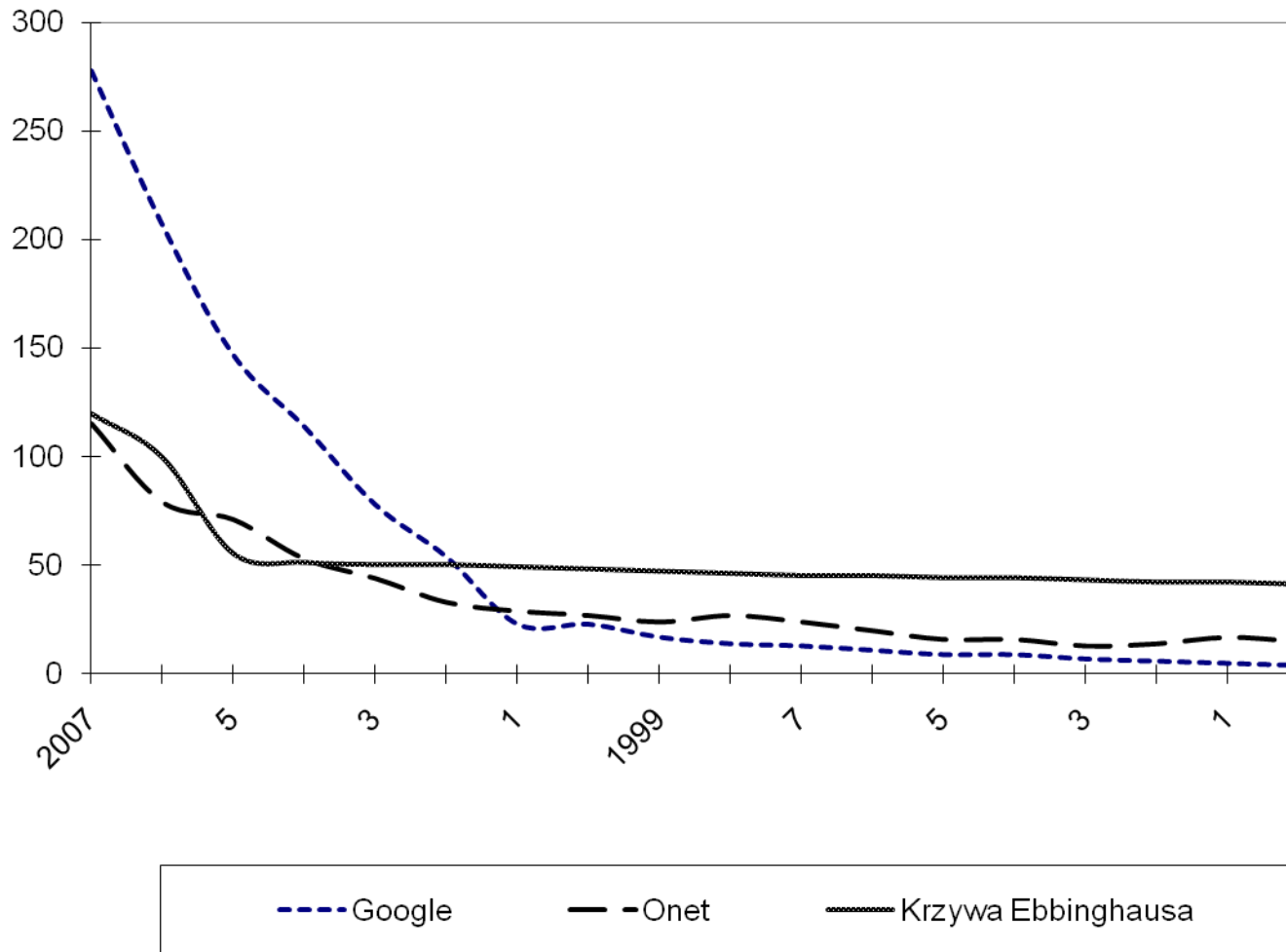
- Np. dyskietki trzyipółcalowe, zanikające CD-ROM-y i doskonalone DVD.
- Przykład: inicjatywa BBC z 1986 roku, której efektem było zapisanie wartościowych danych na temat Wielkiej Brytanii, łącznie z mapami, zapisami wideo i tekstami. Nośnikiem tych zapisów były specjalne laserowe dyski, bazujące na, współczesnych wówczas, komputerach BBC Micro. Nośniki te obecnie są nie do odczytania na jakimkolwiek urządzeniu.
- Po dwuipółletnich bardzo kosztownych pracach ostatecznie udało się jednak przenieść te zapisy na współczesne pecety (WWW.domesday1986.com.)
- Inny przykład: Kilkaset nośników CD z 4 milionów w zbiorach Niemieckiej Biblioteki Narodowej nie do odczytania.
- [NASA](#) scientists found in 1999 that they were unable to read digital data saved from a Viking space probe in 1975; the format had long been obsolete.

Analog/cyfra – koszty archiwizacji

- Przechowywanie cyfrowych obrazów i filmów nie jest tak trwałe jak analogowych
- Koszt rocznego przechowywania cyfrowego filmu \$12514 vs analogowa wersja - \$1059
- Współcześnie niemal wszystkie studia filmowe („Click” lub „Miami Vice”) używając cyfrowych zapisów, korzystają z trwalszej (>100 lat) wersji analogowej
- Cyfrowy zapis jest mniej trwały od analogowego
- Doświadczenia wskazują, że nieuruchamiany przez dwa lata dysk twardy utraci informacje. Podobnie DVD – po 15 latach będzie czytelne tylko połowa zbiorów.

**Internet rejestruje wszystko
niczego nie zapominając**

Zmiany wartości miar czasu magazynowania informacji w Sieci i człowieka (wg Krzywej Ebbinghausa)



Legenda: w celu uzyskania większej czytelności wykresu miara ilości informacji (procent przechowywanych sylab) Krzywej Ebbinghausa została podwojona. Oś X dla tej krzywej skalowana jest w dniach, dla Google'a i Onetu w latach.

Poleganie na Sieci

- Lepiej pamiętane są informacje, o których wiemy, że ich nie ma w Sieci
- Korzystanie z wyszukiwarki zmienia to, co zapamiętujemy. Badani, którzy mieli dostęp do Google'a, łatwiej zapamiętywali miejsce przechowania informacji niż właściwe dane.

<http://www.nytimes.com/2011/07/15/health/15memory.html?nl=todaysheadlines&emc=tha26>

Negatywne informacje

- Trudne do usunięcia. Koszt 500 – 1000 USD
- Blokada w wyszukiwarkach negatywnych informacji – tysiące USD + kolejne tysiące w danych o osobach: Intellius lub PeopleFinder
- Usunięcie oszczerczych (ktoś o firmie) informacji – dziesiątki tysięcy USD

Czas życia profilu w Sieci

- Prośba o kontakt (Facebook) od osoby, która zmarła.
- Problem narasta ze wzrostem wieku internautów
- Facebook umożliwia zgłoszenie śmierci (krewny, znajomy) stosownym dowodem, np. informacją z drukowanej gazety.

Sieciowa pamięć

- Nieprzemyślane informacje o sobie (tekst, obrazy) pozostają na lata w Sieci bez możliwości ich usunięcia. **Sieć nie daje drugiej szansy!**
- 75% profesjonalistów HR kontroluje kandydatów w Sieci (szperacze, społecznościowe, foto, wideo, osobiste strony, blogi, gry online. Efekt - 70% kandydatów odrzucona.

Zasoby dla HR

- 75% specjalistów HR ocenia kandydatów do pracy wg informacji z Sieci – wynik 70% odrzuceń
- Użytkownicy Facebooka zostawiają 25 miliardów wpisów każdego miesiąca

Edytory, przeglądarki

- Mimo że minęło niespełna dziesięć lat, nie zawsze wyświetlają one treści w tej samej formie jak czyniły to poprzednie wersje oprogramowania.
- Np. dokumenty napisane kilkanaście lat temu w edytorach polskich tekstów, takich jak np. Chi Writer czy QL Tekst.
- Sytuacja ta zwraca uwagę na konieczność stworzenia **wzorców formatów** zapisu wszystkich form multimedialnych informacji.

Komercjalizacja rynku informacyjnego

- W mediach dotyczy przede wszystkim informacji biznesowych.
- Już w 2001 roku czwarta część wpływów właścicieli stron, które uzyskiwali oni ze sprzedaży miejsc na reklamy online, pochodziła z wpływów za informacje.
- Eksperci szacują, że sprzedaż informacji via Internet będzie rosła każdego roku dwukrotnie.
- Wydatki za treści online w 2002 roku osiągnęły 50 mld USD, a szacunek na 2006 rok wynosi znacznie ponad 100 mld USD.

Podstawowe kryteria oceny towaru w postaci informacji

- wiarygodność źródeł,
- dokładność i kompletność,
- najszybciej rozwijającymi się kategoriami sprzedawanych informacji jest rozrywka (cyfrowa muzyka i strony multimedialne; jej wartość sprzedaży podwaja się każdego roku) oraz biznes/inwestycje, sport i gry.

Komercyjna wartość dóbr informacyjnych - zagrożenia

- Informacja zastępuje własność gruntów, energii i surowców.
- Następuje proces kanalizowania dystrybucji wszelkich treści i czerpania korzyści z tej operacji.
- Tylko w USA roczne wpływy z licencjonowanych technologii przynoszą 45 mld USD, a na świecie ponad 100 mld USD.
- Dzięki wątpliwej jakości regulacji rodzą się instytucjonalni rentierzy, którzy wyłącznie za fakt posiadania praw do oprogramowania, patentów, gazet, książek, muzyki, filmów pobierają opłaty możliwe do zaakceptowania tylko przez niektórych.
- W rezultacie tego pogłębiają się różnice pomiędzy krajami i grupami społecznymi będącymi na innych etapach drogi ku społeczeństwu informacyjnemu. Żarysowuje się między nimi przepaść, głównie spowodowana różnicami w dostępie do cyfrowych dóbr. Zagroza to społeczną i kulturową stagnacją.

Cyberbullying – sieciowa nagonka

- Podszywanie się pod kogoś w celu skompromitowania
- Tworzenie opinii w publicznych miejscach wymiany informacji online
- Tworzenie i rozsyłanie tendencyjnych maili
- Kojarzenie w szperaczach niepochlebnych słów z nazwiskiem ofiary

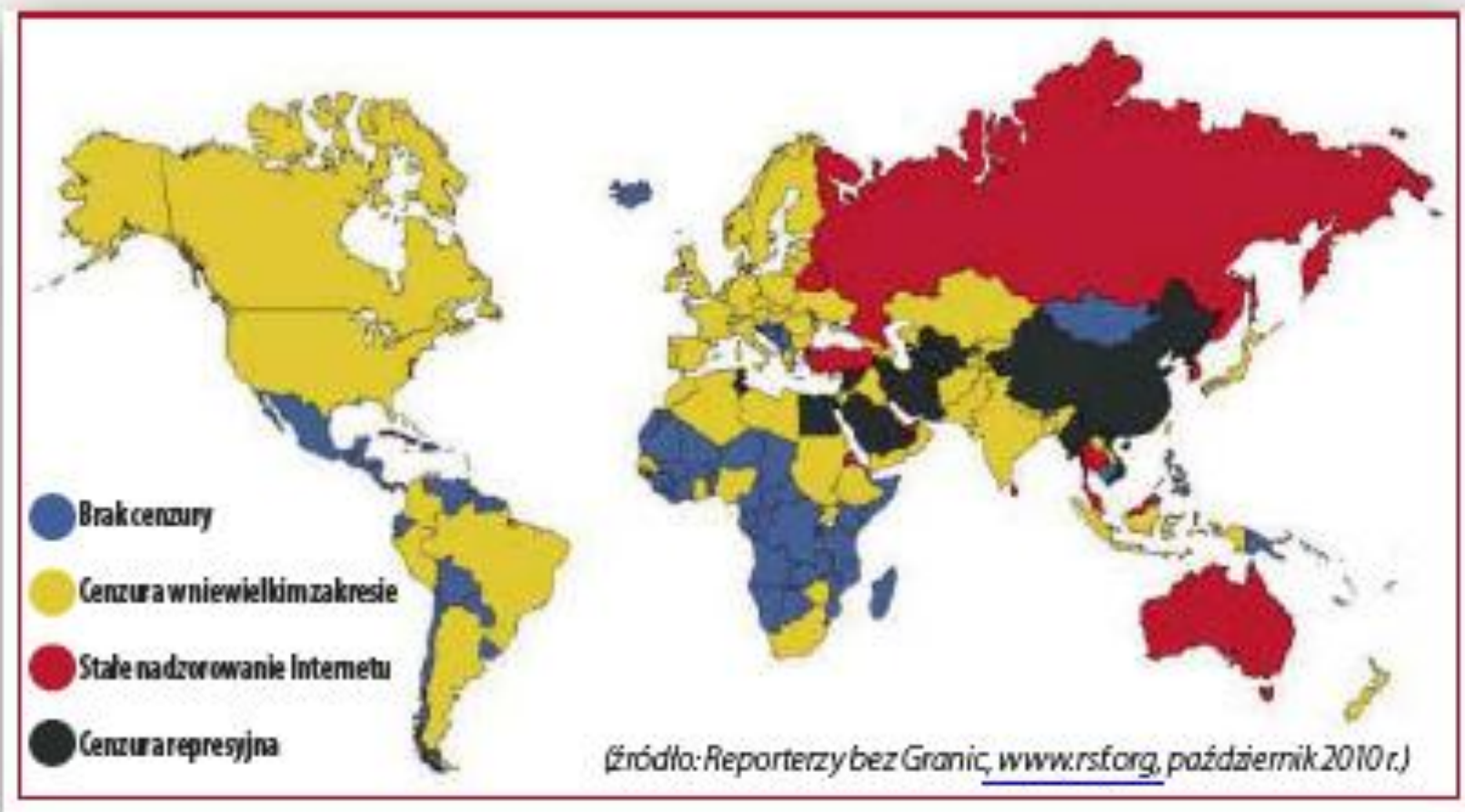
Wzmacniająca funkcja IT: „rewers”: Cyberbullying – negatywna siła wirtualnego tłumu

Cyberbullying obejmuje użycie informacji i technologii informacyjnych takich jak mail, komórki, IM, strony w celu wspomagania dyskusji, powtarzania, inicjowania, nieprzyjaznych zachowań osób lub grup z zamiarem uczynienia komuś krzywdy

Kontrola treści Sieci, cenzura

- 50 pracowników Telecommunication On Demand sprawdza ok. 20 milionów obrazów każdego tygodnia
- Outsourcingowa firma - Caleris – 4,5 mln obrazów codziennie
- YouTube kontroluje materiały wskazane przez internautów.
- Facebook – kontrola w Palo Alto i w Dublinie.

Cenzura Internetu na świecie



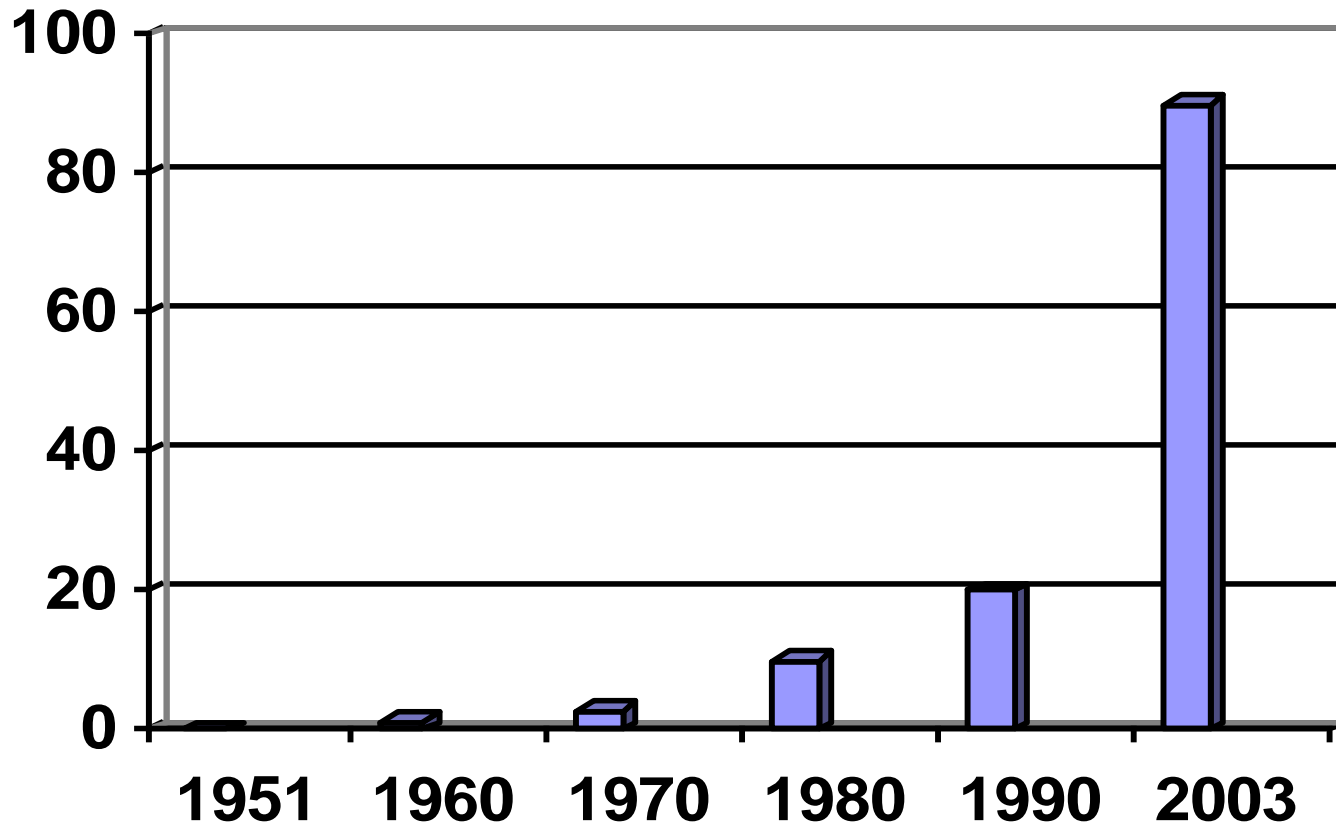
Granice obiektywizmu

- Na stronie Wikileaks opublikowano poufne, ukradzione dane o kontaktach klientów banku szwedzkiego.
- Strona ta głosi, iż umożliwia ludziom publikować dokumenty anonimowo, z każdego państwa, które piętnują nieetyczne zachowania władz i korporacji

„Cyfrowa przepaść” przeciwstawianie *copy-right (od 300 lat)– copy-left*

- Zwalnianie rozwoju nauki i rozprzestrzeniania się wartościowych idei.
- Kto zapanuje nad oprogramowaniem, ten zapanuje nad *technosferą*. Władza nad *technosferą* będzie zaś tożsama z władzą nad światem.
- Sztandarowe przykłady idei *copy-left*: standard *open source*, a w nim system operacyjny Linux, sieciowe oprogramowanie dla serwerów Apache i tworzona przez społeczność internautów encyklopedia – Wikipedia oraz inicjatywa Projektu Gutenberg (bezpłatne udostępnianie elektronicznych wersji książek).
- Celowość rewizji tego, co ma pozostać prywatne, państwowe, a co ogólnoludzkie – jak powietrze i woda.

Własność intelektualna – wpływy [mld USD]



Źródło: Athreye and Cantwell, 2005 za:

A market for ideas, A survey of patents and Technology, The Economist, October 22nd, 2005.

Street View (Google) Streetside (Microsoft)

- Microsoft. Pod koniec 2009 roku projekt Streetside będący częścią projektu Bing Maps.
- UltraCam mobile: 12 megapikselowy aparat z sześcioma obiektywami co 4 metry zapisuje 6 zdjęć o wielkości 16 MB i 16-bitowej głębi kolorów.



GPS

- Global Positioning System,
- Światowy system określania położenia
- baza: 24 satelity
- Dokładność do 10 m.

AGPS

- Ustalanie pozycji wspomagane BTSami – stacjami bazowymi GSM (anten) operatorów komórkowych
- Dokładniej i szybciej

Jak działa system GPS

1 Wszystkie satelity mają dokładnie zsynchronizowane zegary - każdy z nich pokazuje tę samą godzinę

2 Wszystkie satelity znają swoją precyzyjną pozycję dzięki danym przesłanym przez kontrolery systemowe

3 Każdy satelita wysyła dane informujące o swojej pozycji i czasie

4 Sygnał podróżuje do odbiorcy z opóźnieniem wywołanym przez dużą odległość

5 Różnica długości drogi przebytego sygnału powoduje, że sygnały nie docierają w tym samym czasie



Inne formy źródeł informacyjnych Sieci

Rafinacja informacji

Destylacja

**Informacyjne źródło
zasobów sieciowych
– biały wywiad**

Big data – baza rafinacji

- Big Data to zasoby w postaci cyfrowej, których istotną część tworzą wszelkie informacje dostępne w internecie: posty, blogi, strony, fora, profesjonalne wydawnictwa (dzienniki, periodyki, książki i inne podobne wydawnictwa, np. – 235 terabajtów danych zebranych w US Library of Congress do IV.2011).
- Zbiór 100 milionów wiadomości agencyjnych z całego świata, który obejmuje pół wieku ludzkiej cywilizacji, cyfrowa biblioteka 4% wszystkich książek wydrukowanych w okresie sześciu wieków, zasoby branżowe, np. Służba zdrowia.

Big data cd.

- Big data są także tworzone przez urządzenia komórkowe (5 mld w 2010 r.), komputery, cyfrowe kamery, aparaty fotograficzne, czytniki RFID, samochody, domy, inteligentne mierniki np. poboru energii elektrycznej i GPSy.

Inteligentne liczniki prądu – Smart Meter

- Źródło informacji o prywatnych zachowaniach użytkowników prądu:
 - Posiłki
 - Sen
 - Toaleta
 - Pranie
 - Typ TV (skok poboru mocy)

Rafinacja informacji

- Jest peryskopem pomiędzy dwoma poziomami informacji: poziomem bazowego i poziomem wyniku rafinacji.
- Umożliwia dostrzeganie informacji z otoczenia/poziomu bazowych informacji (czyste informacje z WWW) tych, które są na poziomie informacji będących wynikiem rafinacji.

Rafinacja - zasięg

- Analizując miliardy postów, blogów i artykułów dostępnych online możliwe jest uzyskanie nowych informacji o społecznych fenomenach, państwach, organizacjach i osobach (relacje, migracje itp.)
- Wynik: ocena emocji relacji (sympaia, antypatia, poczucie szczęścia, optymizm, pesymizm, strach).

Narzędzia rafinacji

- Wyszukiwarki
- Dane behawioralne (zachowania w Sieci)
- Personalizacja (np. kryteria opisujących osobę decydują o wynikach wyszukiwania)
- Zaawansowane narzędzia, np. ocena sentymentu (Attentio)

Diagnoza emocjonalnych odczuć adresata przekazu

- Tłumaczenie emocjonalnych odczuć na liczby
- analityczne narzędzia mierzące sentyment użytkownika Sieci (*sentiment analysis*) do określonych zdarzeń, informacji, marek towarowych oraz towarów, a nawet wskazywania dni, w których Internauci są najbardziej podatni na określone przekazy informacyjne.

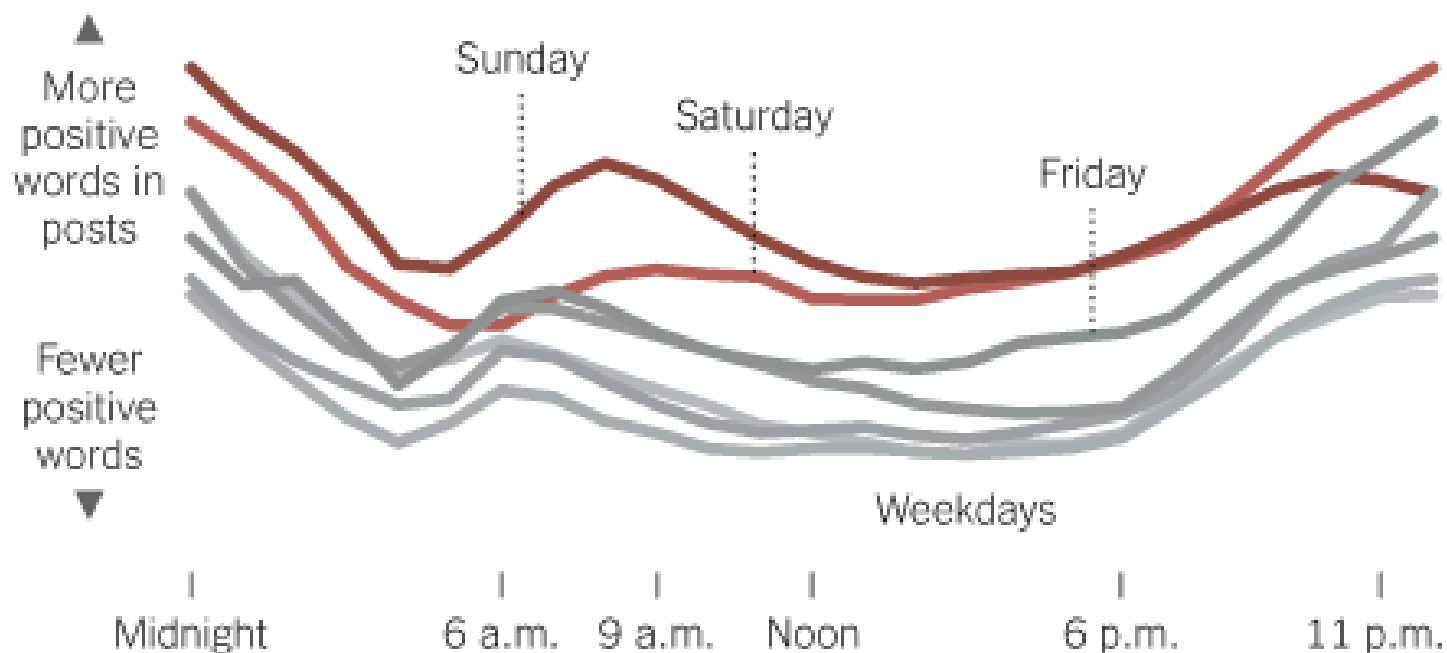
Ocena emocji z Twittera 1/2

- Analiza komunikatów (ang.) >2 mln osób z 84 krajów (II.2008 – I.2010)
- Emocjonalny ton (wybrane słowa i emotikony) komunikatów jest podobny w skali tygodnia, pół roku
- 7% - nocne marki, 16% - poranne skowronki
- Najlepsze samopoczucie 6-9 rano/koniec tygodnia/lato, kryzys po 15.00/początek tygodnia/jesień-zima

Ocena emocji z Twittera 2/2

Studying Moods Through Twitter

A textual analysis of more than 500 million Twitter messages found people around the world tend to express more positive emotions in the morning and evening, and are most positive on weekends. The recurring daily pattern suggests moods are influenced by sleep and circadian rhythms.

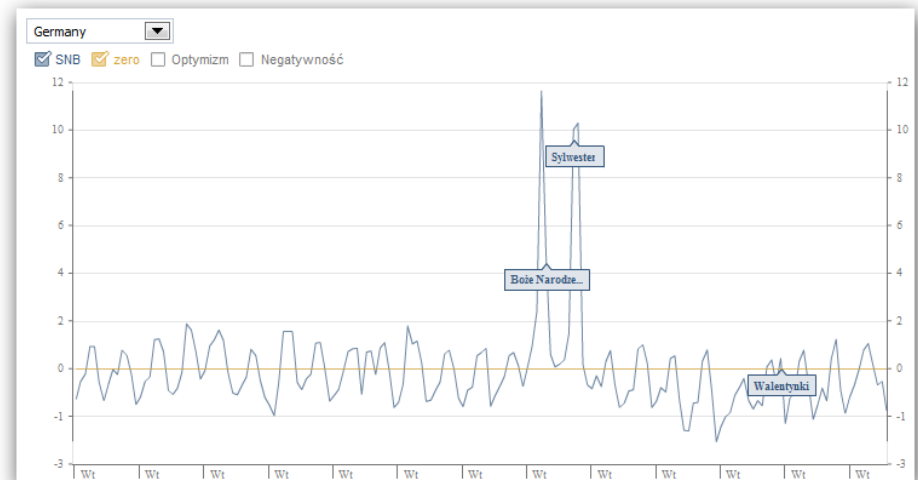
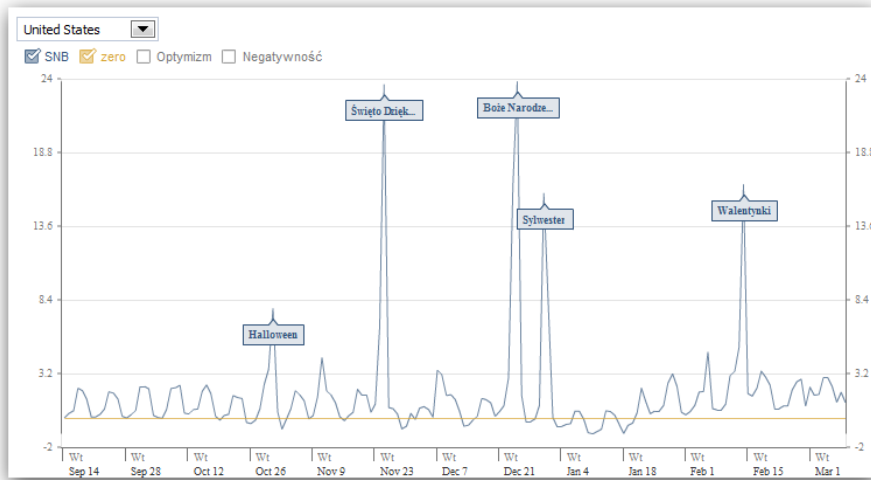


Source: Science

THE NEW YORK TIMES

Informacja o poczuciu szczęścia użytkowników Sieci

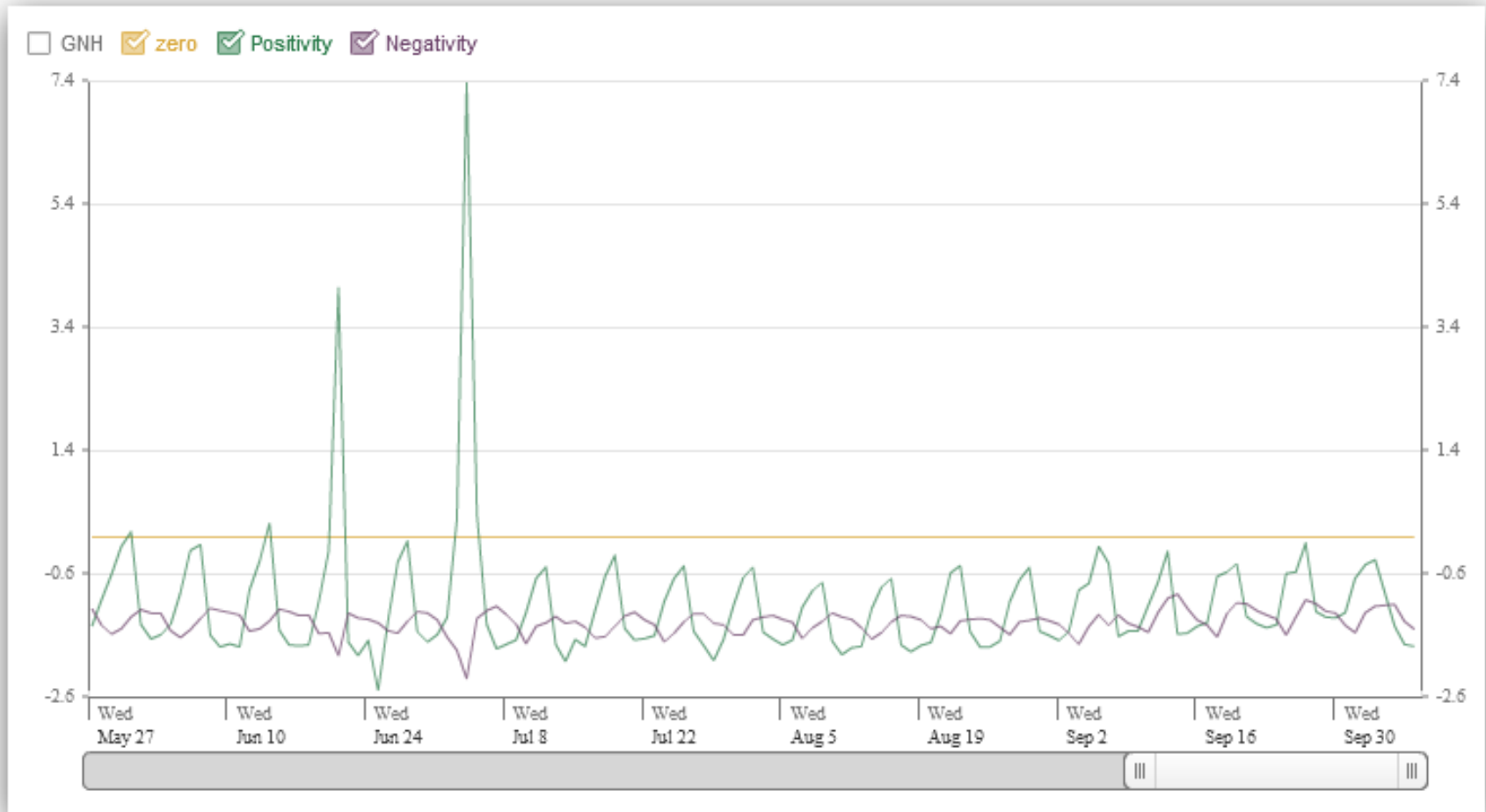
– wg. Liczby użytych na stronie słów:
szczęście, wspaniale!, smutek, wątpliwości i tragedia



http://apps.facebook.com/gnh_index/

October 12, 2009, NYT, Link by Link, **Is It a Day to Be Happy? Check the Index**, By [NOAM COHEN](#)

United States Gross National Happiness (2009)



WHERE PEOPLE GO (MOST POPULAR CHAINS ACROSS THE USA)



APPAREL

Old Navy
H&M
Victoria's Secret



BANKS

Bank of America
Chase
Wells Fargo



CONVENIENCE

7-Eleven
Circle K
Wawa



HOME

Home Depot
IKEA
Lowe's



RETAIL

Target
Walmart
Macy's

LIFE MOMENTS

WEDDINGS



225

Weddings at City Hall

804

Weddings in parks

5,321

Weddings in churches

BABIES



1,061

Birth announcements on hospital check-ins



1,750

Congratulations received in comments

MAYORSHIPS



78,387

Venue mayors ousted each day



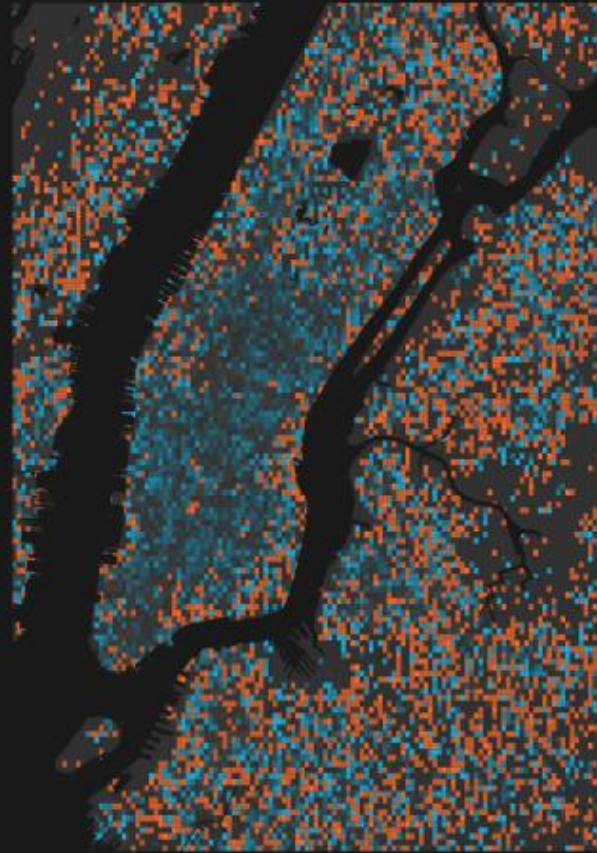
1

Real world mayors enthroned on foursquare (@MikeBloomberg, Mayor of City Hall, NYC)

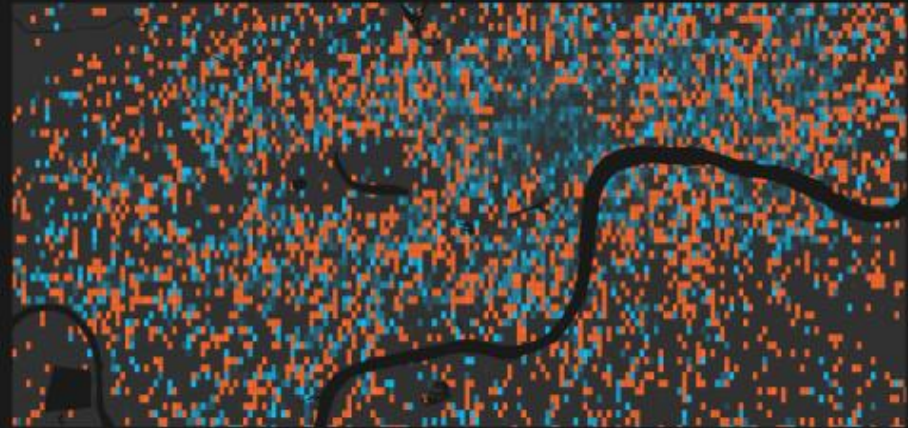
<https://foursquare.com/10million>

HAPPINESS ON FOURSQUARE

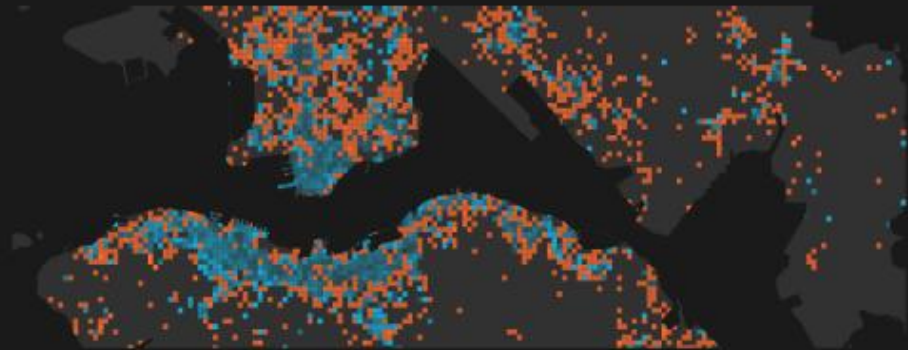
NEW YORK CITY



LONDON



HONG KONG



Based on a sentiment analysis of shouts on foursquare, looking for words indicating happiness*. Examples:

AWESOME

HAPPY

OKAY

BLAH

SUCKS

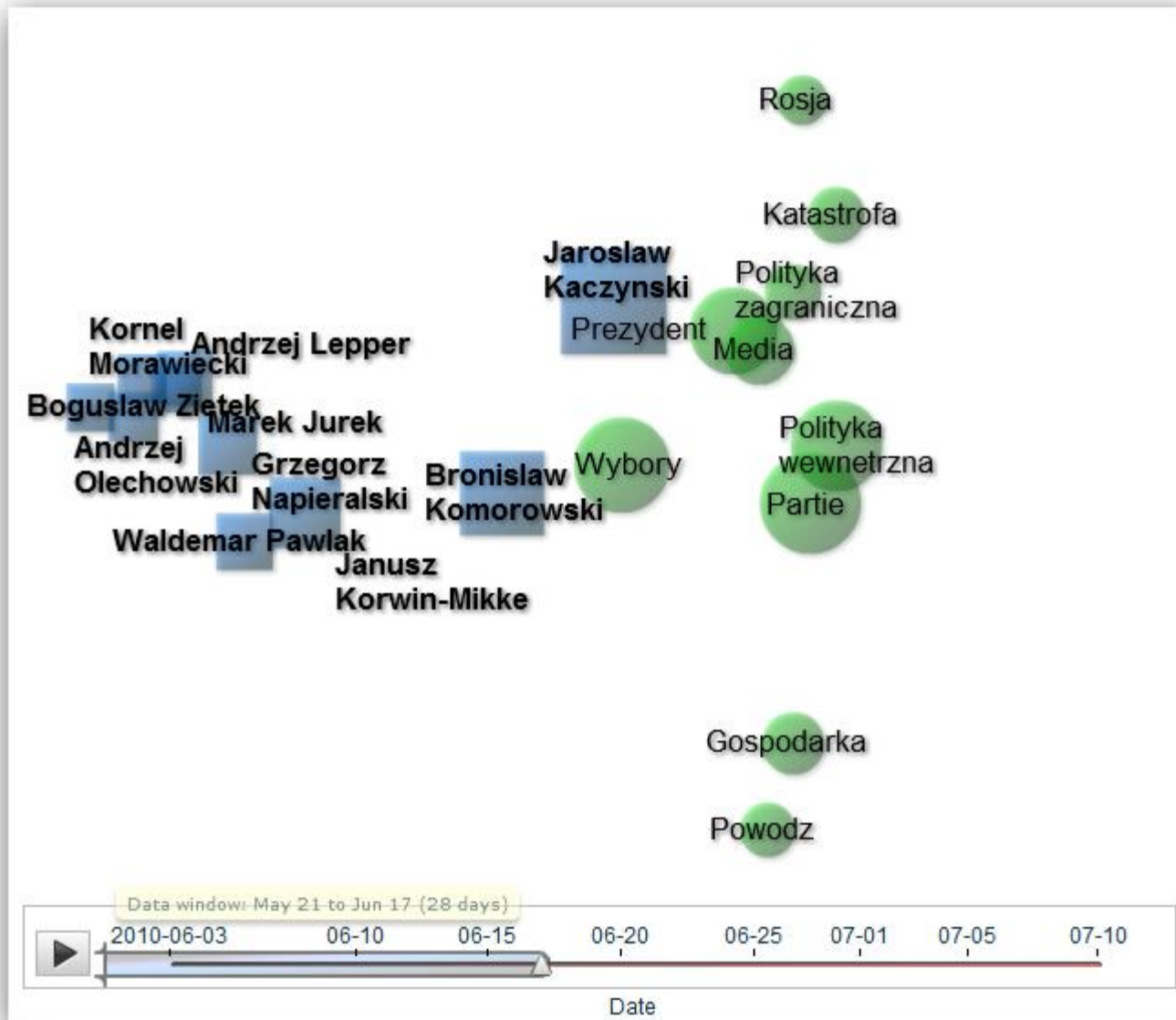
WTF

<https://foursquare.com/10million>

Dynamiczna analiza relacji informacji w Sieci

Attentio Brand Maps

Source: Forums Polish Daily

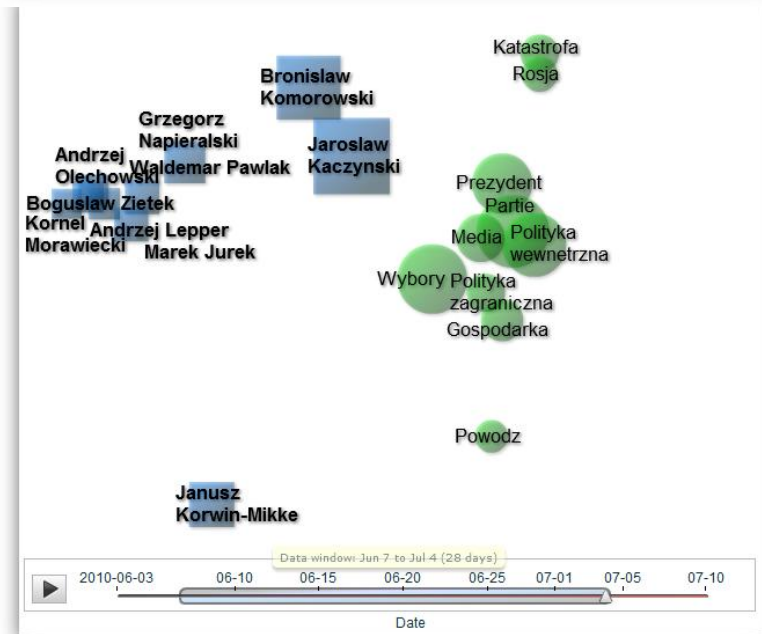
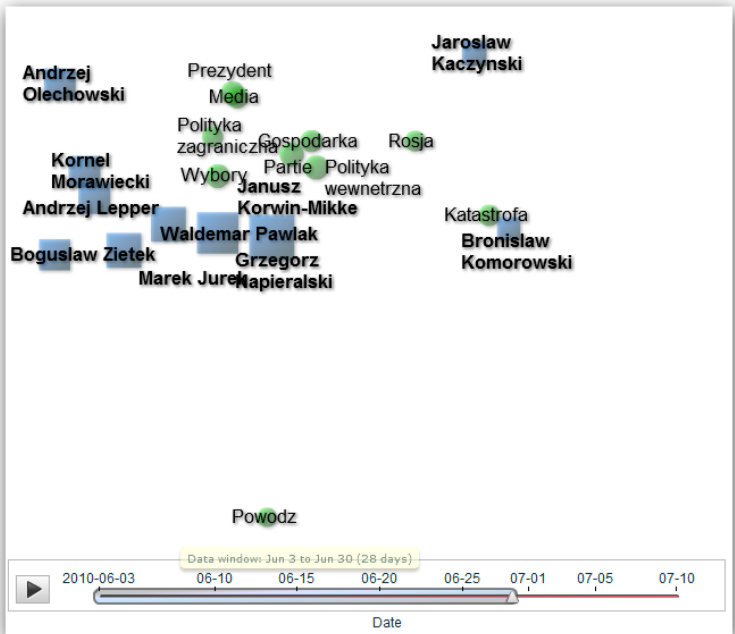
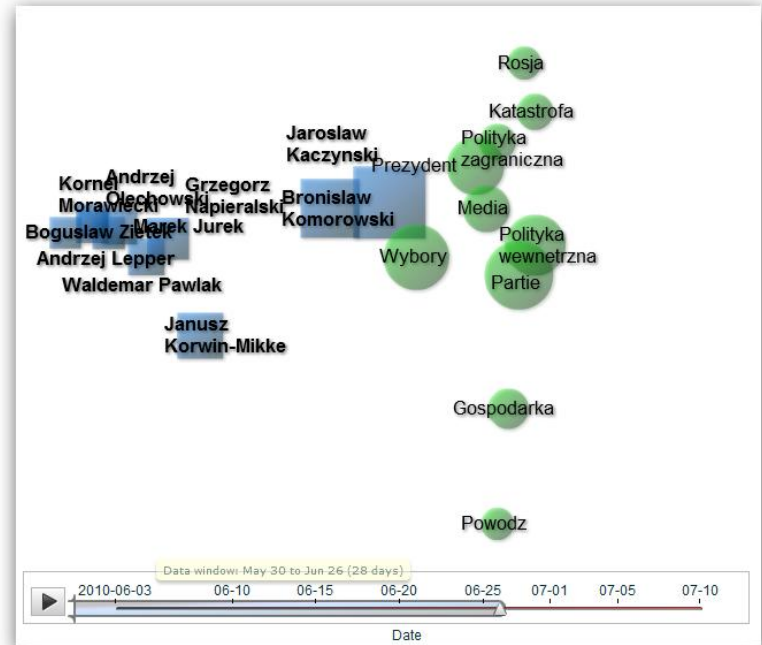
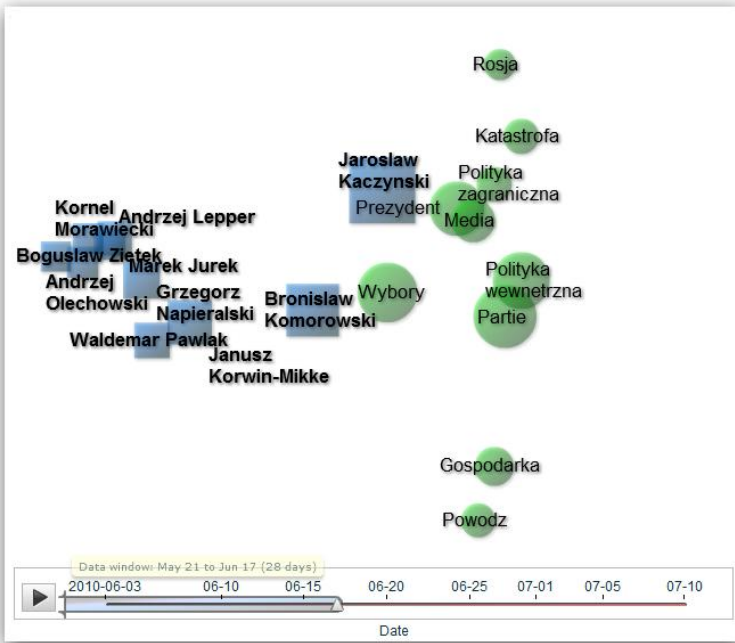


ntio Brand Dashboard™

czma | Inbox: () | Log out

Source: Blogs Polish Daily

- Blogs Polish Daily
- Facebook Polish
- Forums Polish Daily
- Twitter Polish Daily
- News Polish Daily
- Comments Daily
- Blip Daily



Wyniki rafinacji

- Ocena przeszłości, zmian na podstawie korpusu 5.2 milionów książek z lat 1500 - 2008 containing about 4% of all books ever printed.
- Np. badanie kulturalnych trendów, korzystanie z kolektywnej pamięci, ocena adaptacji nowych technologii, epidemiologia, cenzura itp.

Google wykresy i liczby

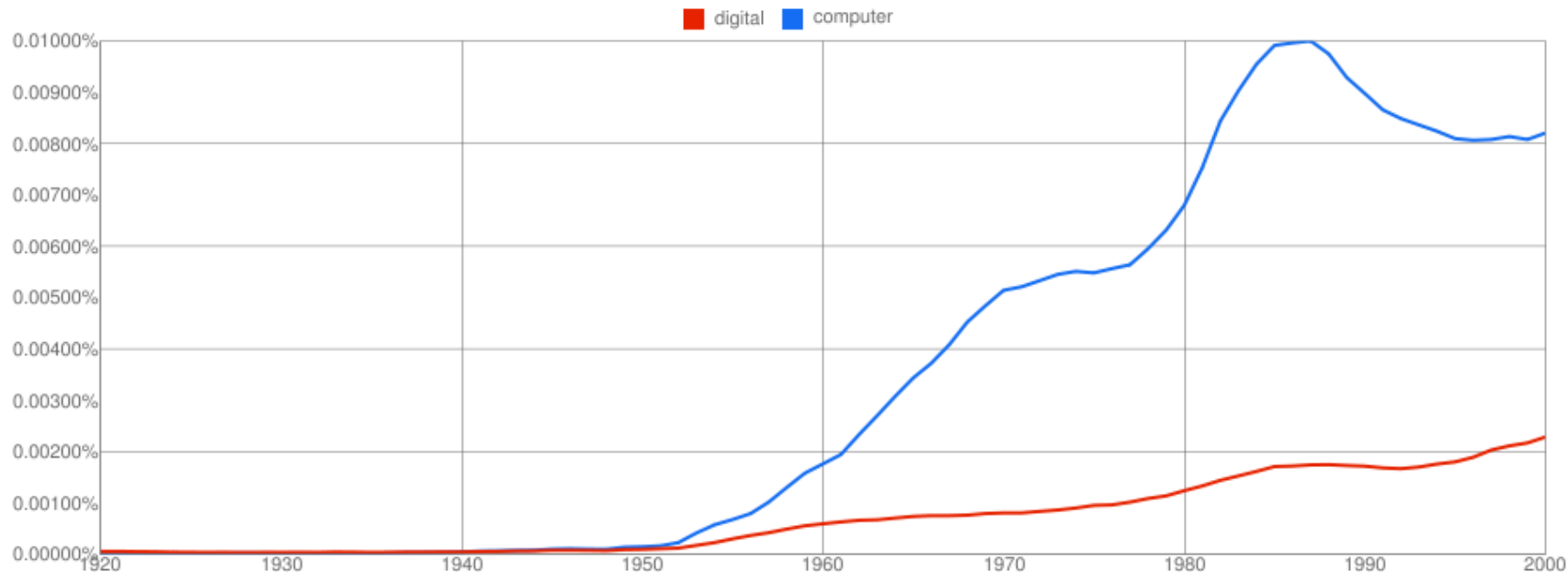
- Google udostępnił do analizy bazę 5.2 milionów książek dla onlinowych analiz – 4% wszystkich książek od 1800 roku.
- Baza zawiera 15 milionów książek – 12% od Biblii Gutenberga z 1450 r.

5,2 mln cyfrowych książek - Google

- 4% wszystkich wydrukowanych książek
- 1500 – 2008 rok
- 500 miliardów słów: English, French, Spanish, German, Chinese and Russian.
- Źródło oceny kulturowych trendów na świecie
- Inne rozumienie języka, kultury i przepływu idei
- Zjawisko: culturomics

Graph these **case-sensitive** comma-separated phrases:

between and from the corpus with smoothing of .



Search in Google Books:

1920 - 1966	1967 - 1990	1991 - 1993	1994 - 1996	1997 - 2000	computer
1920 - 1961	1962 - 1991	1992 - 1994	1995 - 1997	1998 - 2000	digital

Run your own experiment! Raw data is available for download [here](#).

© 2010 Google - [About Google](#) - [About Google Books](#) - [About Google Books NGram Viewer](#)

December 16, 2010, NYT.

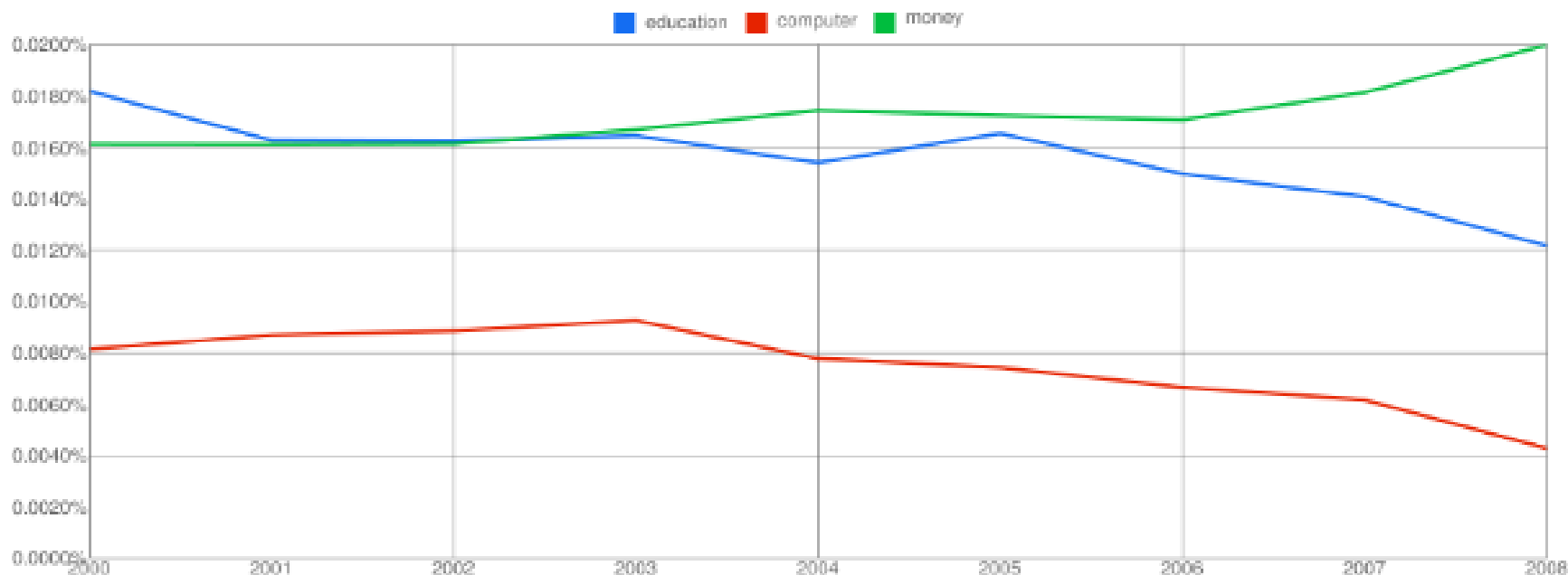
In 500 Billion Words, New Window on Culture

By [PATRICIA COHEN](#)

Google labs Books Ngram Viewer

Graph these case-sensitive comma-separated phrases:

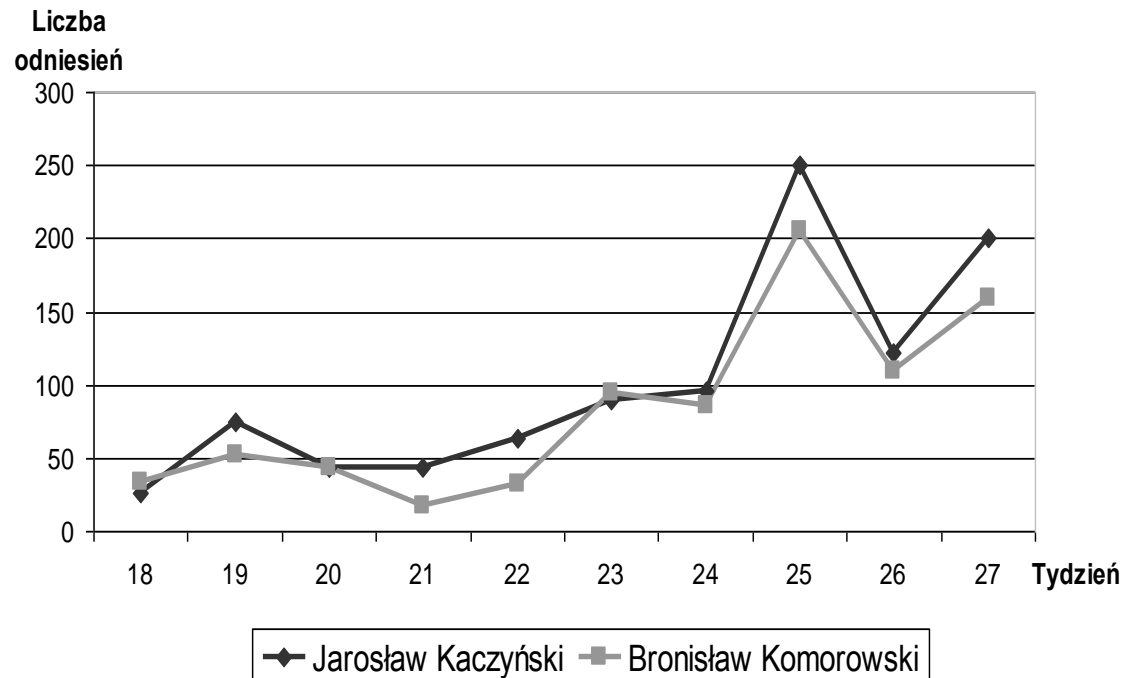
between and from the corpus with smoothing of .



http://ngrams.googlelabs.com/graph?content=education%2Ccomputer%2Cmoney&year_start=2000&year_end=2008&corpus=0&smoothing=0

Wyniki rafinacji – przewidywanie przyszłości

- Zachowania giełd, sprzedaż książek, wideo
- Choroby
- Zjawiska polityczne i społeczne
- Konflikty społeczne



Web data predict flu

NATURE|Vol 456|20 November 2008

Potencjał badawczy danych tworzonych podczas wyszukiwania informacji przez internautów w Sieci

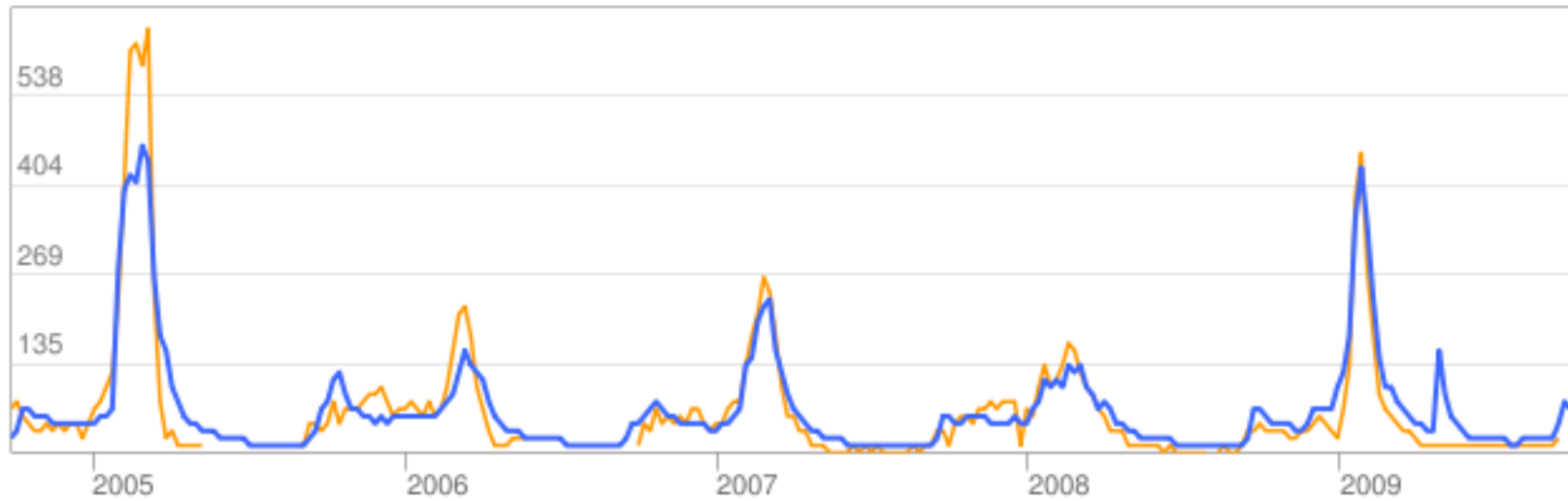
Oszacowania historyczne

Zobacz dane dla kraju:

Polska – aktywność grypy

Szacowane zachorowania na grypę

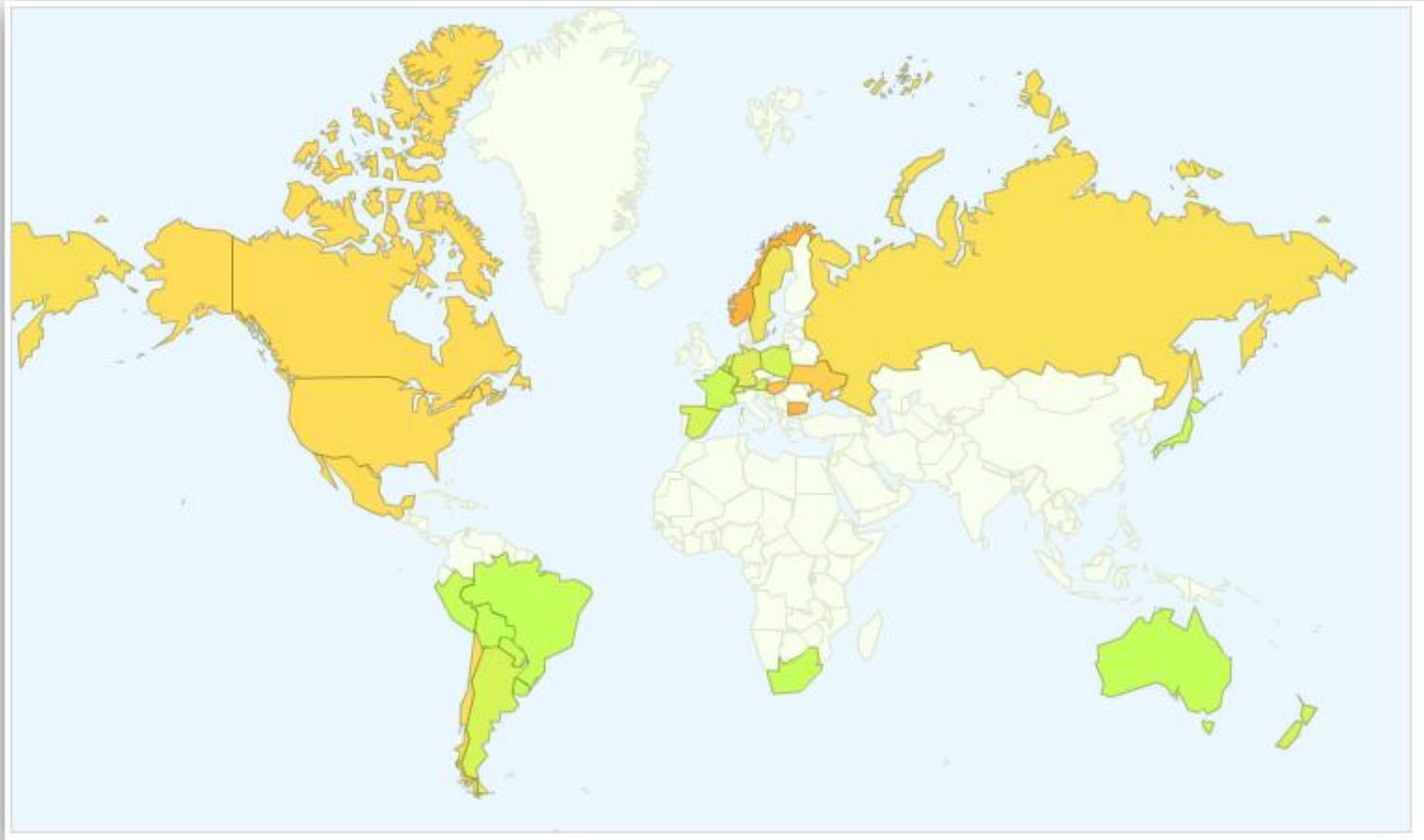
● Google Flu Trends - dane szacunkowe ● Polska – dane



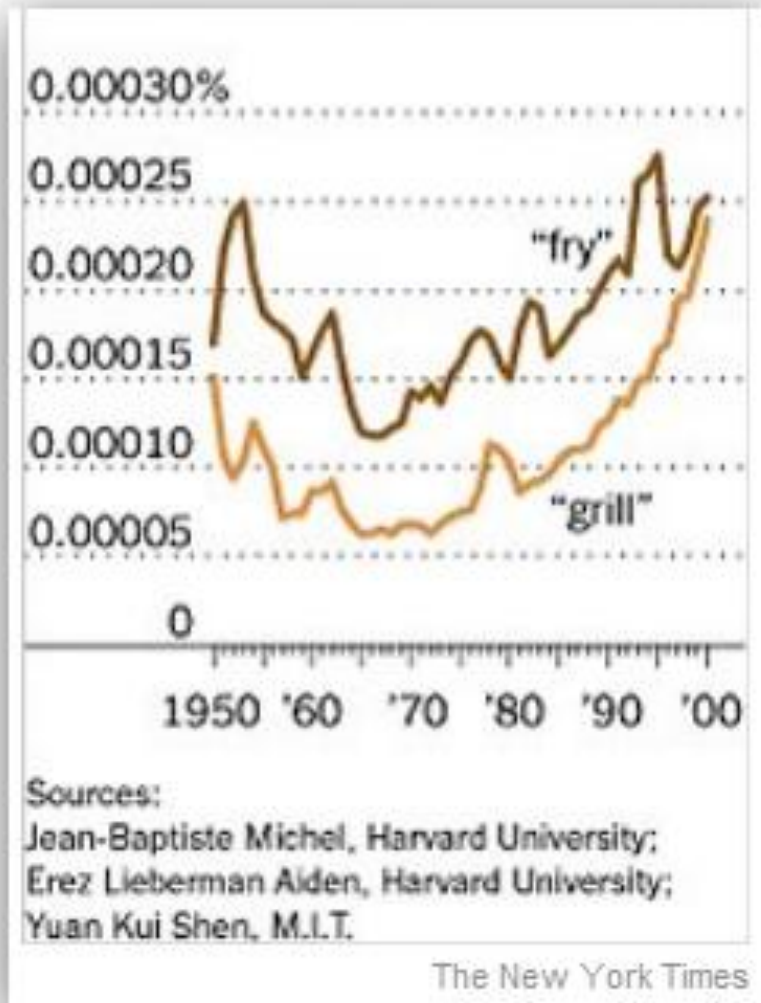
Polska: Dane o zespole objawów grypopodobnych opublikowane przez [Europejska Sieć Nadzoru nad Grypa](#) Europejskiego Centrum ds. Zapobiegania i Kontroli Chorób.

Przykład: tendencje aktywności wirusa grypy na świecie

Odkryliśmy, że niektóre z wyszukiwanych haseł są dobrymi wskaźnikami aktywności grypy. Google Flu Trends używa zagregowanych danych wyszukiwania Google do szacowania aktualnej aktywności grypy na całym świecie niemal w czasie rzeczywistym.

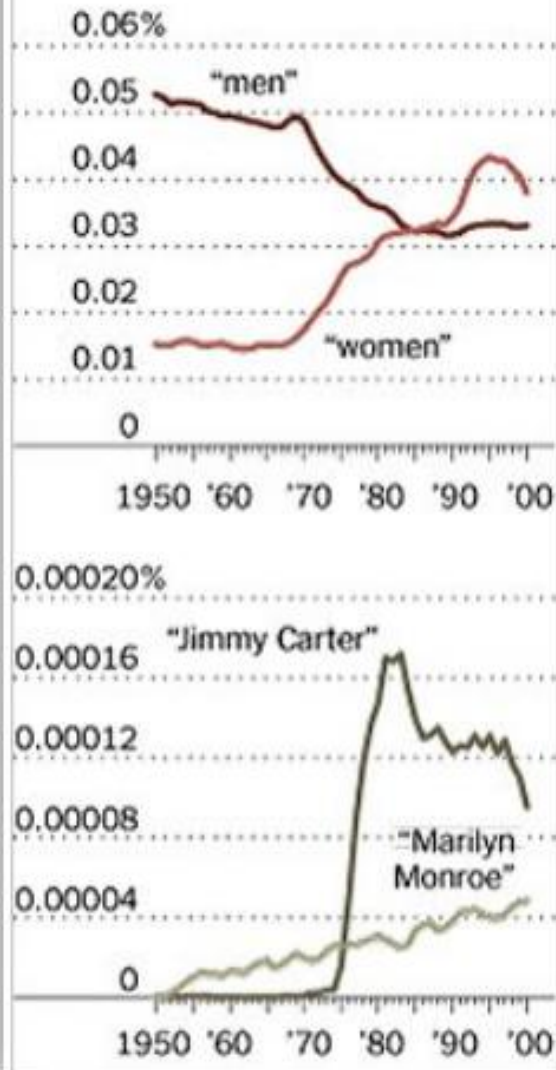


Częstotliwość słów w książkach/lata



Culturomics, Anyone?

The frequency of selected words in books published each year.



December 16, 2010, NYT.

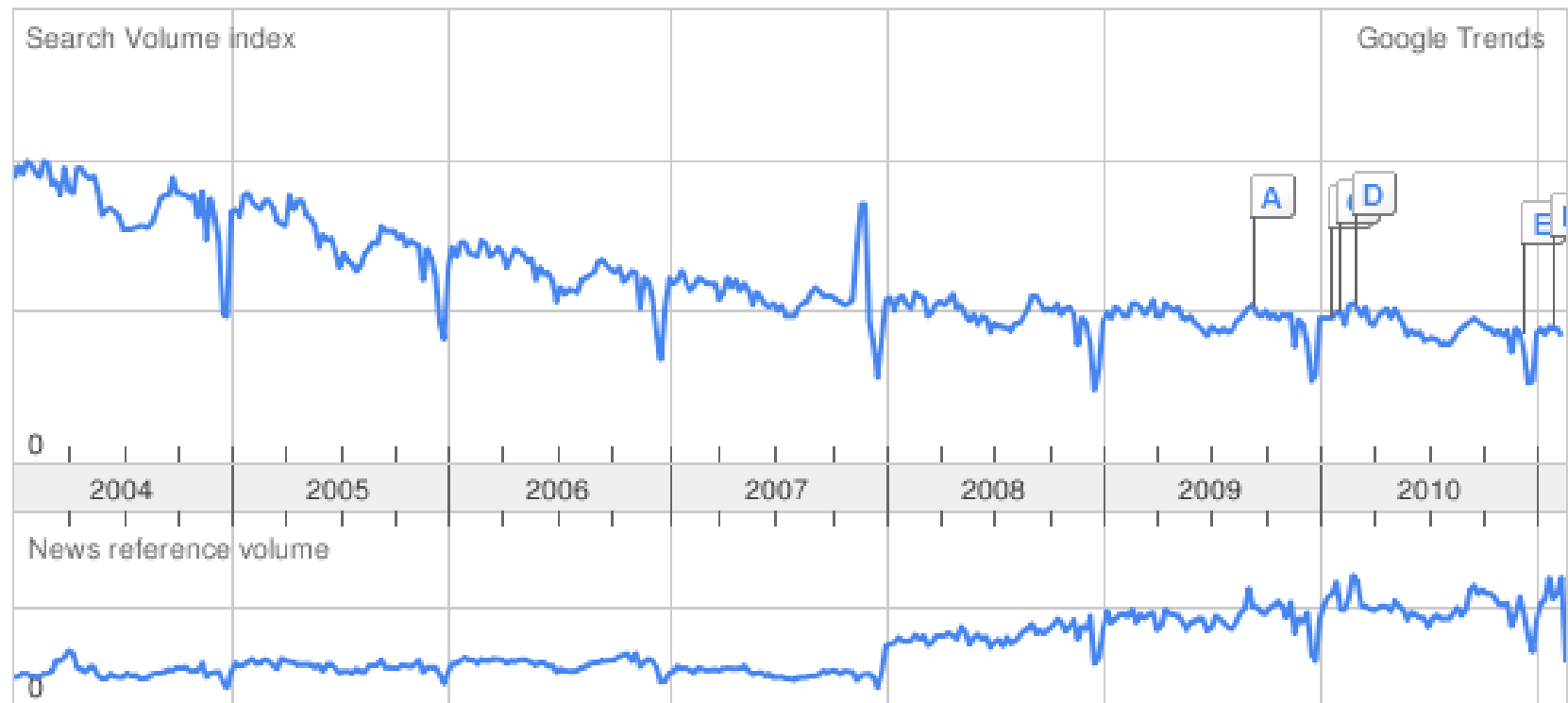
In 500 Billion Words, New Window on Culture

By [PATRICIA COHEN](#)

Tip: Use commas to compare multiple search terms.

Searches [Websites](#)

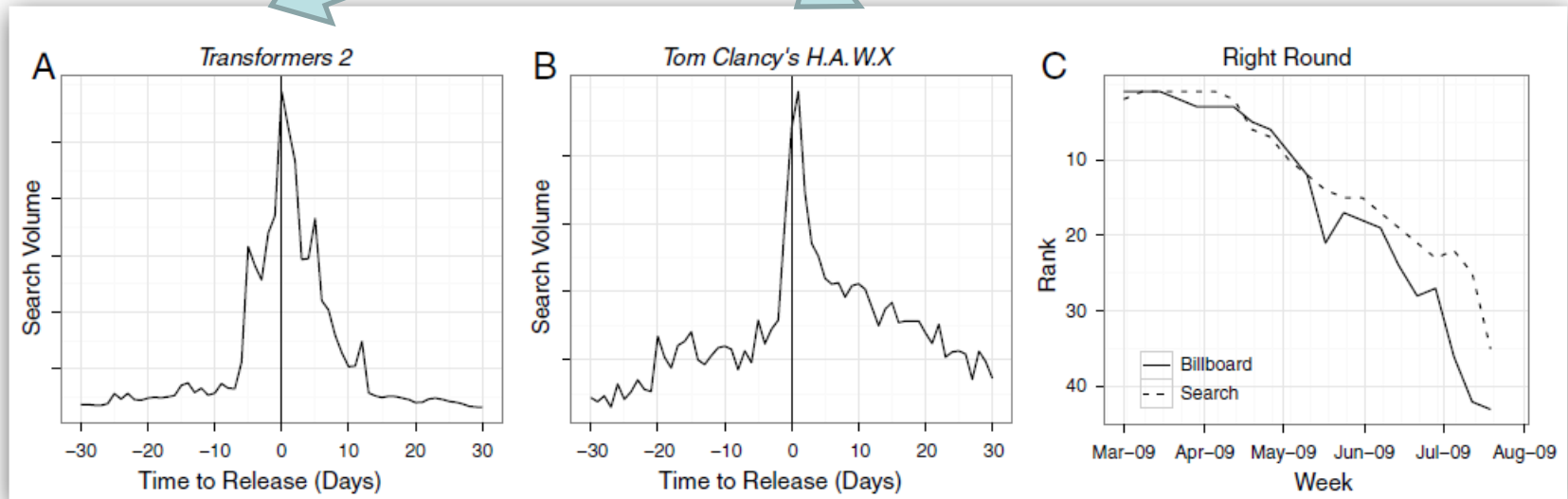
● education



Prognozowanie z informacji sieciowych

- Ludzie szukając dzisiaj wskazują co będą robili w przyszłości
- Np. bezrobocie, sprzedaż samochodów, choroby.
- Badania potwierdziły prognozy sprzedaży filmów i pozycję muzyki na liście przebojów
- Realne przewidywanie kolektywnych zachowań w zakresie ekonomii, kultury i polityki

Wartość wyszukiwania dla filmu, gry i piosenki



Zbieżność dat publikacji z wartością zapytań

Minerva – inicjatywa Pentagonu

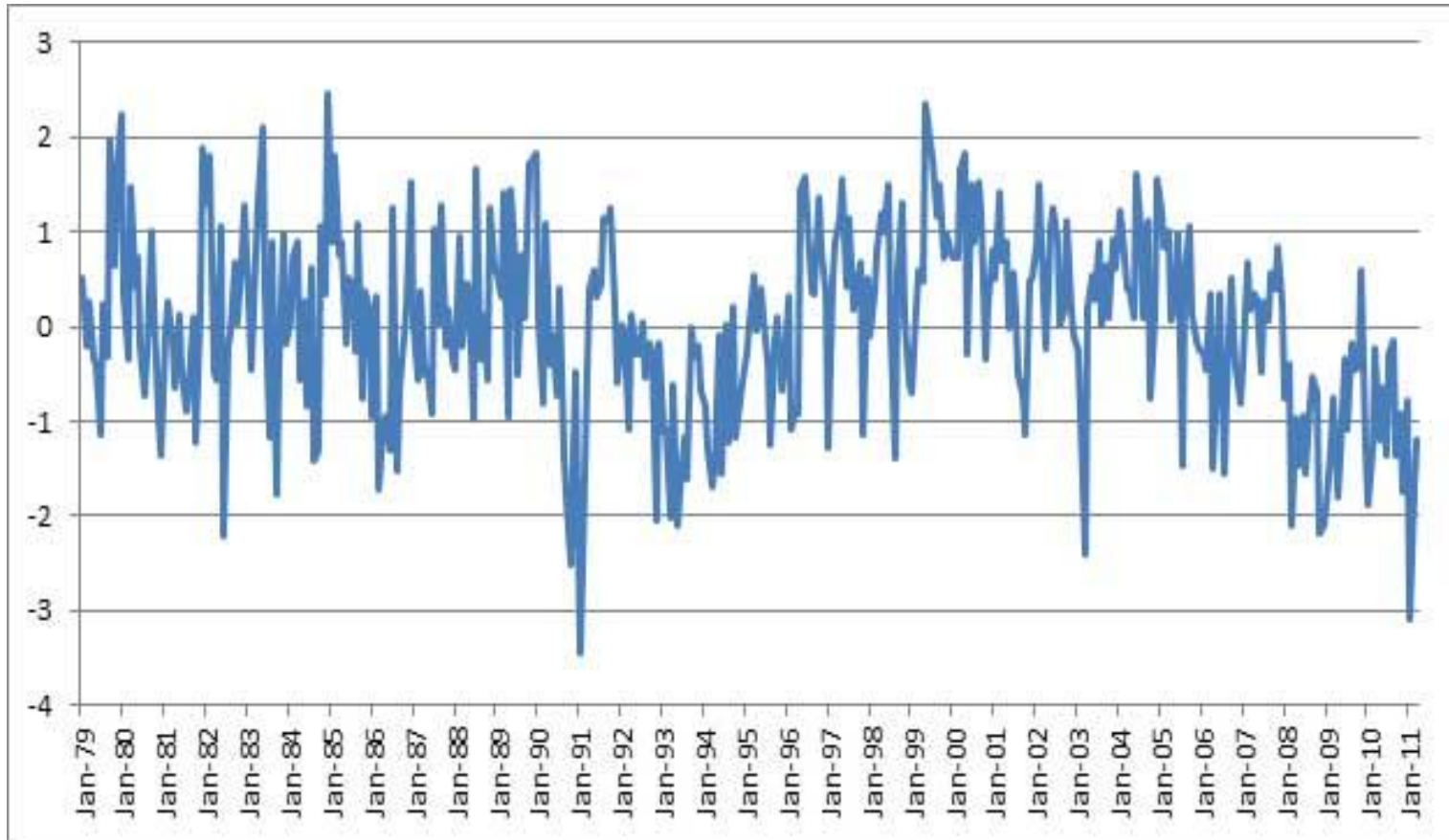
- Seek to **define and develop foundational knowledge about sources of present and future conflict** with an eye toward better understanding of the political trajectories of key regions of the world. Improve the ability of DoD to **develop cutting-edge social science research**, foreign area and interdisciplinary studies, that is developed and vetted by the best scholars in these fields.

Psychohistoria

Issac Asimov

- Kombinacja matematyki i psychologii do przewidywania przyszłości
- Obecnie bazą „**big data**”: RFID, POSy, Twitter, Facebook, blogi, Wiki, lokalizacja komórek, ruch w Sieci, zapytania, dane z rynków finansowych, webcamy z ulic. Efekt – przewidywanie politycznych kryzysów, rewolucji i innych form społecznej i ekonomicznej niestabilności, epidemii

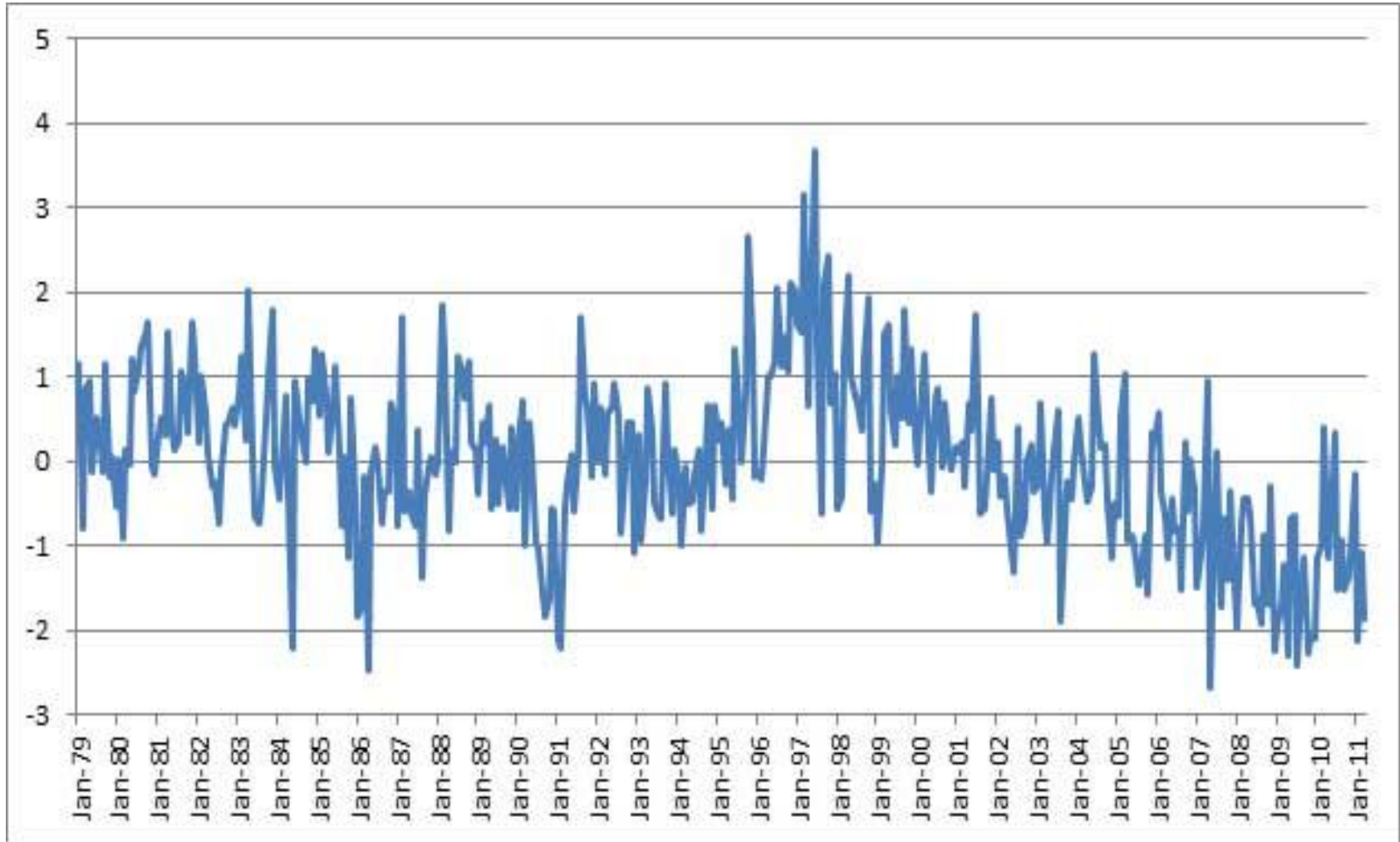
Ton dziennikarskich wiadomości z Egiptu



First Monday, Volume 16, Number 9 - 5 September 2011, in:

<http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3663/3040>. {January 2012}.

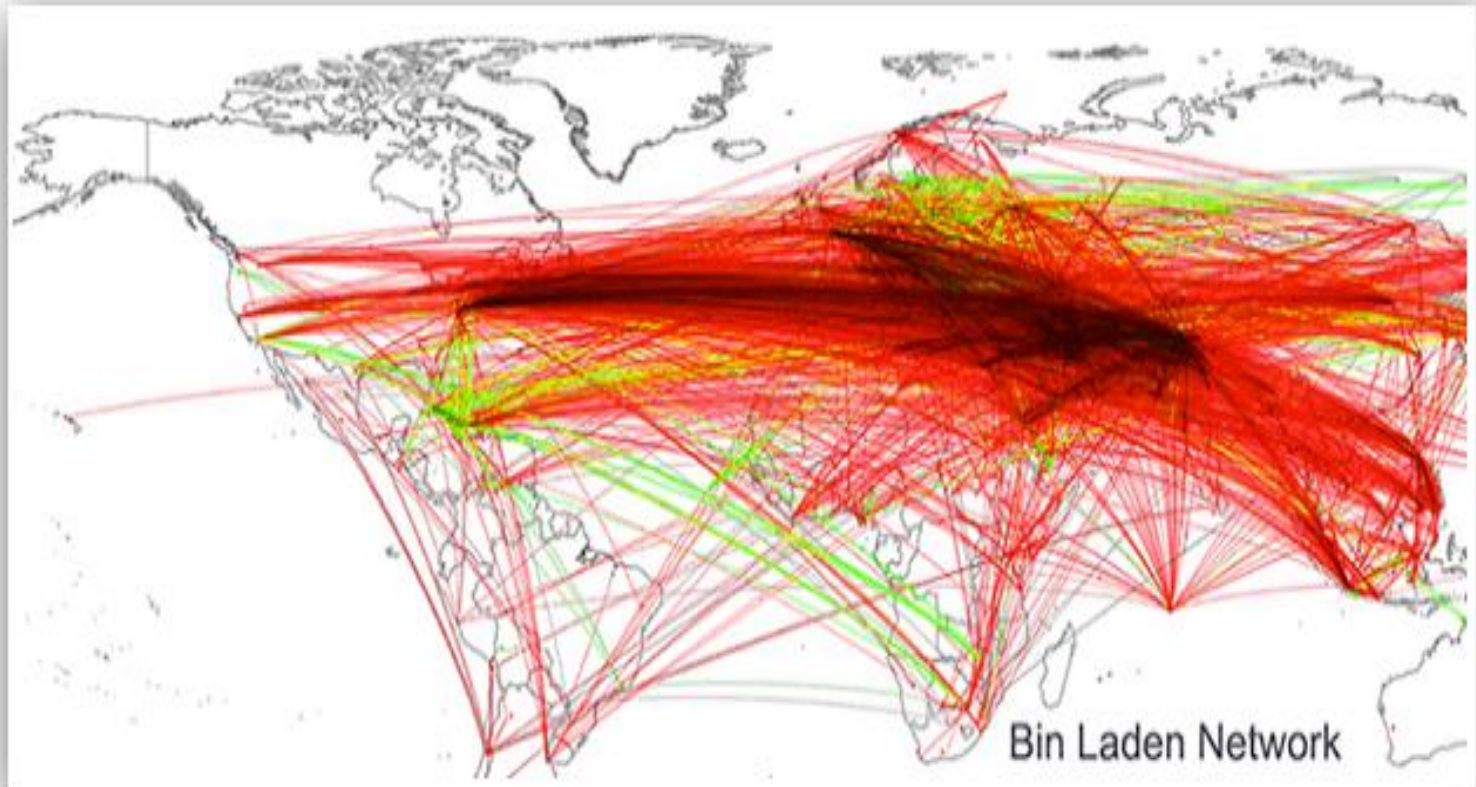
Tunezja



First Monday, Volume 16, Number 9 - 5 September 2011, in:

<http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3663/3040>. {January 2012}.

Nazwy miejsc blisko słów Bin Laden



By Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, Pickett JP, Hoiberg D, Clancy D, Norvig P, Orwant J, Pinker S, Nowak MA, Aiden EL, Quantitative Analysis of Culture Using Millions of Digitized Books, **Science**; 1/14/2011, Vol. 331 Issue 6014, p176-182, 7p.

Cyfrowy testament

- Facebook – kilka milionów kont osób nieżyjących. W tym roku przybędzie kolejne 1,5 mln kont należących do zmarłych.
- Usługa polegająca na spisaniu cyfrowego testamentu. Umożliwia to bliskim zmarłych użytkowników odzyskanie ich kont i zarządzanie nimi.
- Po zmarłych mogły pozostać pieniądze w PayPalu, zdjęcia w sieciowych galeriach, życiorysy i CV w serwisach rekrutacyjnych, poczta itp.

Przykłady miejsc i form testamentów

- Nagrania wideo i audio (legacylocker, greatgoodbye)
- Zamieszczanie komunikatów na blogu (mywebwill)
- Konieczność potwierdzenia przez dwie osoby lub cykliczne potwierdzanie życia.
- Testament sieciowy bywa łączony z tradycyjnym

Serwisy przechowujące cyfrowy testament

Legacy Locker

INFO: www.legacylocker.com

Serwis oferuje bezpłatne powiadomienie o testamencie trzech wybranych osób. W wariantcie płatnym umożliwia wgranie wiadomości w postaci nagrania wideo i przechowywanie kopii najważniejszych dokumentów.

The screenshot shows the Legacy Locker website homepage. At the top, there is a navigation bar with the logo and links for Features, Why Legacy Locker?, Pricing & Sign Up, and Support. The main content area features a large white cloud graphic with the text: "The safe and secure way to pass your online accounts to your friends and loved ones." Below this, a smaller text block states: "Legacy Locker is a safe, secure repository for your vital digital property that lets you grant access to online assets for friends and loved ones in the event of loss, death, or disability." To the right of the text is a photograph of a family. At the bottom, there is a footer with a navigation bar containing "The value of", "Protect your", "Privacy and", "Family and", "Financial and", and a prominent green "Sign Up Now" button.

My Webwill

INFO: www.mywebwill.com

Serwis wyróżnia się bezpłatną usługą automatycznego usuwania kont ze wskazanych serwisów po zgłoszeniu śmierci użytkownika. Za dopłatą otrzymasz możliwość przesyłania danych dostępowych wskazanym osobom i automatycznego opublikowania przygotowanej wcześniej ostatniej wiadomości, np. na tablicy w Facebooku.

The screenshot shows the My Webwill website homepage. At the top, there is a navigation bar with the logo and links for Home, About, FAQ, Plans & Pricing, Contact, and Log in. The main content area features a large yellow sun graphic with the text: "Your life online after death". To the right of the sun is a video player with a play button and a world map. Below the sun are social media icons for LinkedIn, Facebook, Twitter, and YouTube. At the bottom, there is a footer with a navigation bar containing "The value of", "Protect your", "Privacy and", "Family and", "Financial and", and a prominent green "Sign Up Now" button.