

SMB Roadshow Split 2013



Big Data: veliki podaci, velika prilika

Alen Gojčeta

Country Technical Sales Leader



Slijed predavanja

- Što je Big Data?
- Nastanak paradigme
- Primjene i ograničenja
- Tehnologija
- Prilike za Hrvatsku akademsku zajednicu i industrijsku politiku

Big Data i promjena paradigme



Prva promjena paradigme – od društvenog događaja do podatka

Stevan Dedijer: „Pojam Business Intelligence označava sposobnost da se snađeš u poslovnom svijetu, da postigneš svoje ciljeve.”



Druga promjena paradigme – od podatka do podatkovne masovnosti i sveprisutnosti (Volume, Velocity, Variety)

Volumen

12 terabajta
Tweetova dnevno

Sentimenti prema proizvodu

Brzina

5 milijuna
Burzovnih transakcija u sekundi

Potencijalna prijevara

Raznolikost

100-e video zapisa
S nadzornih kamera

Nadzor događaja od interesa

350 milijardi

otčitavanja električnih brojlara godišnje

Predviđanje potrošnje energije

500 milijuna

zapisa o telefonskim pozivima dnevno

Prevenција korisničkog napuštanja

80% rasta podataka

su slike, video i dokumenti...

Poboljšanje korisničkog zadovoljstva

Okruženje mijenja paradigmu

- U današnjem digitalnom svijetu postoji dostupno 1.8 trilijuna gigabajta podataka
- 2012 se dnevno proizvodi 2,5 eksabajta (milijarda gigabajta) podataka
- 90% svih svjetskih podataka generirano je tijekom posljednje dvije godine

Big Data – primjene u praksi



Big Data posvuda

5 Lessons Marketers Can Learn From Obama's Victory

November 2, 2012 by Todd Wasserman



1. It's the Big Data, Stupid

Big Data may have its flaws, but this election shows that it's indispensable. *New York Times* columnist Nate Silver showed how crunching numbers can render most pundits' gut instincts irrelevant. The Obama campaign proved the same for the marketer's gut. As *Time* chronicled, the Obama campaign relied on a team of dozens of number crunchers who made predictive calls on exactly the right type of pitch to right the right type of voter.

After consolidating its database into one megafile, the team relentlessly tested pitches based on the targeting and learned from its testing. As the article states:

A large portion of the cash raised online came through an intricate, metric-driven e-mail campaign in which dozens of fundraising appeals went out each day. Here again, data collection and analysis were paramount. Many of the e-mails sent to supporters were just tests, with different subject lines, senders and messages. Inside the campaign, there were office pools on which combination would raise the most money, and often the pools got it wrong.

- Izvor: Mashable.com

Big Data u nebrojenim scenarijima

- Izbori ☺
- Medicina (IBM Watson)
- Zrakoplovstvo
- Optimizacija gradskog prometa
- Lanac nabave
- Komunikacije
- Dinamika fluida
- Upravljanje klijentima / prevencija odlazaka
- Upravljanje klijentima / preciznije kampanje
- Sigurnost / prevencija prijevara
- Poljoprivreda
- Stočarstvo
- ...

Big Data @ Work

Organizations in all industries are under increasing pressure to capitalize on data.

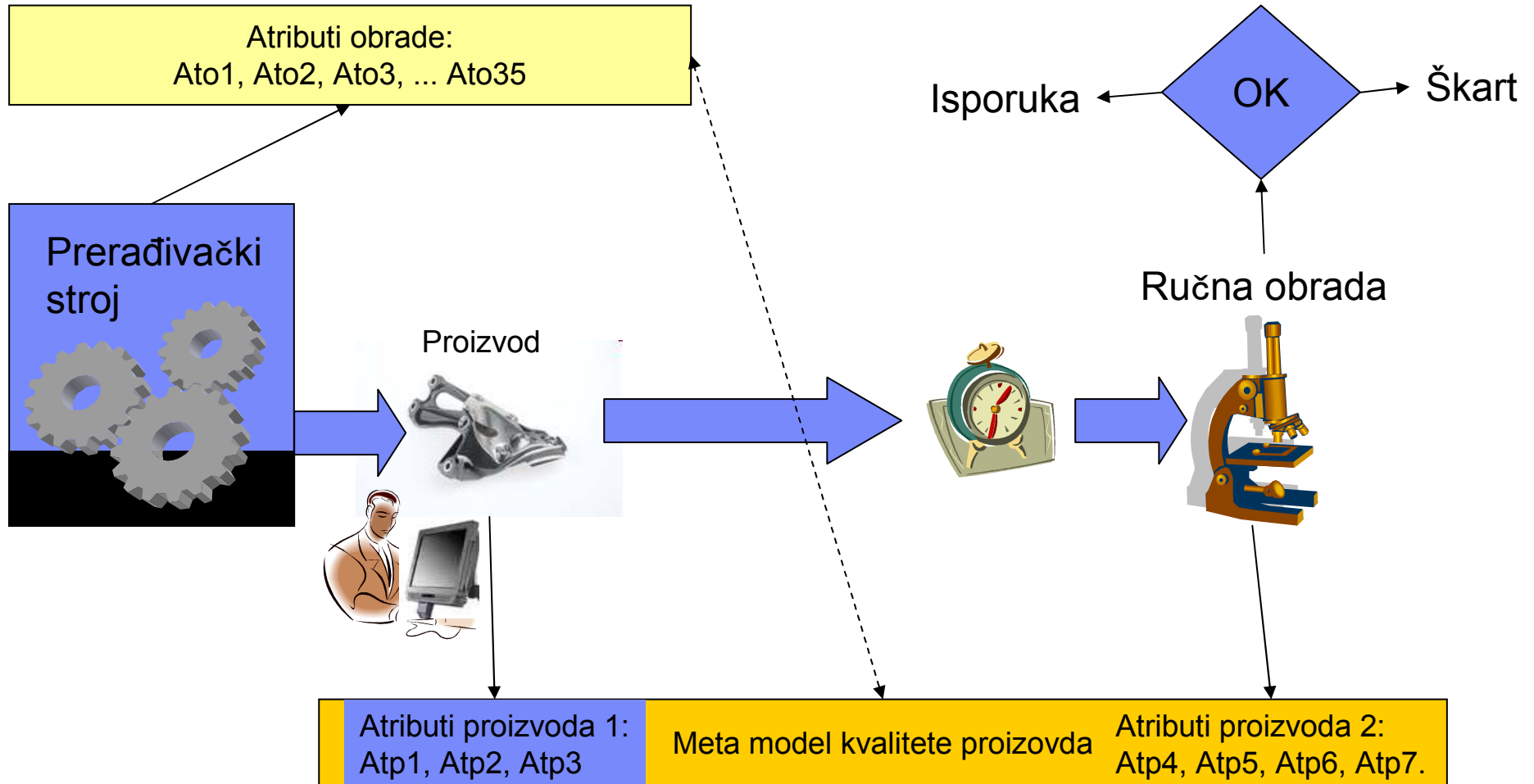
- Healthcare**
The average amount of data per hospital will increase from 147TB in 2011 to 445TB in 2015, driven by the enormous growth of medical images and electronic medical records.
With Big Data
Medical professionals can improve patient care and reduce costs by extracting relevant clinical information from vast amounts of data to better understand the past and predict future outcomes.
- Customer Service**
Today, 86% of consumers quit doing business with a company because of a bad customer experience, up from 59% four years ago.
With Big Data
Service representatives can use data to gain a more holistic view of their customers, understand top their likes and dislikes in real-time in order to resolve a problem or capitalize on happy clients faster.
- Insurance**
Insurance companies and government agencies each gather fraud data related to their own individual missions. But the kind, quality and volume of data compiled varies widely.
With Big Data
An insurance or citizen services provider can apply advanced analytics to data and detect fraud quickly, before funds are paid out.
- Financial Services**
Wall Street alone delivers 5 new research documents every minute. Dow Jones publishes upwards of 15,000 news items per day.
With Big Data
Financial services professionals can better understand market changes through improved business insight from data, helping to anticipate performance gaps and more accurately assess investment alternatives.
- Retail**
\$185 billion in total sales are missed each year because retailers don't have the right products in stock to meet customer demand.
With Big Data
Retailers can better understand their customers by analyzing sales trends and incorporating more accurate forecasting, ultimately increasing customer loyalty and revenue.
- Communications**
3 billion global subscribers in the telco industry are demanding unique and personalized offerings that match their individual lifestyles.
With Big Data
Communications providers can use data to create a more personalized customer experience and avoid losing customers to competitors.

*Information gathered by IBM
1. North American Health Care Provider Information by US Census & Forecast, Electronic Imaging Group
2. Healthcare Expenditures by Hospital Report, Health Economics
3. Statistics Applied Insurance Fraud
4. The Future of Retail in Trading, Ben Jones, June 2011
5. IHL Group, Dec 2011
6. IBM Survey Q1/Q2 2011
© Copyright IBM Corporation 2012. All rights reserved.

Big Data – ograničenja

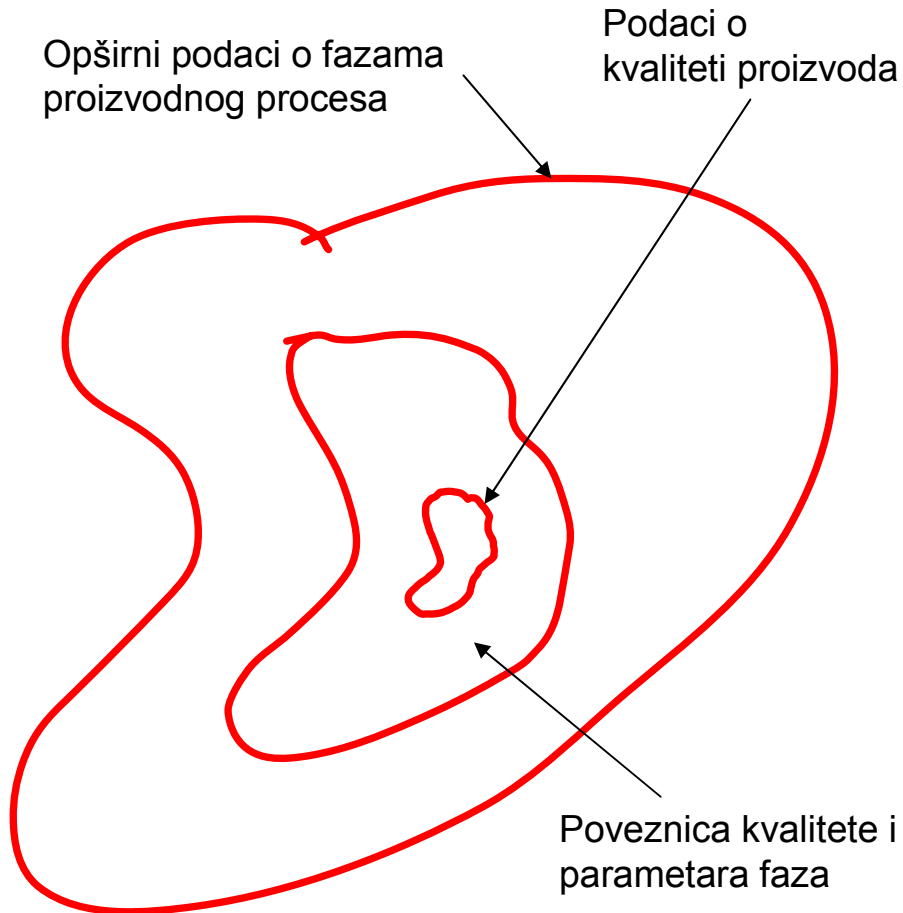


Slučaj jedne proizvodne tvrtke



ID ciklusa stroja kao poveznica između kvalitete proizvoda i parametara stroja

I Big Data podaci moraju sadržavati poželjan ishod – primjer 1



- Postojali su podaci o 35 parametara proizvodnje
- Data mining je davao uzorke devijacije parametara proizvodnje
- Nedostajala je spoznaja o tome kako je to utjecalo na kvalitetu proizvoda

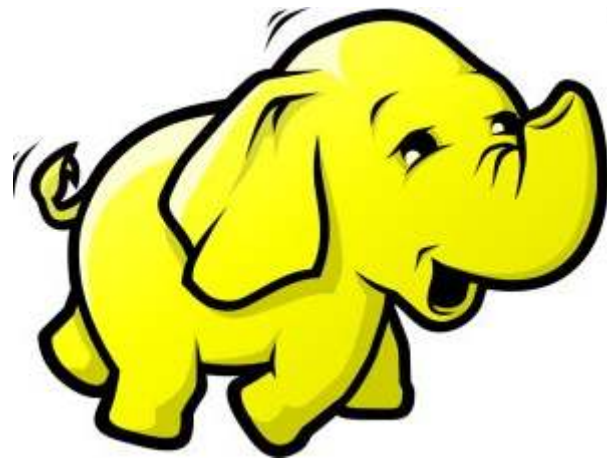
QR kôd kao sredstvo integracije?



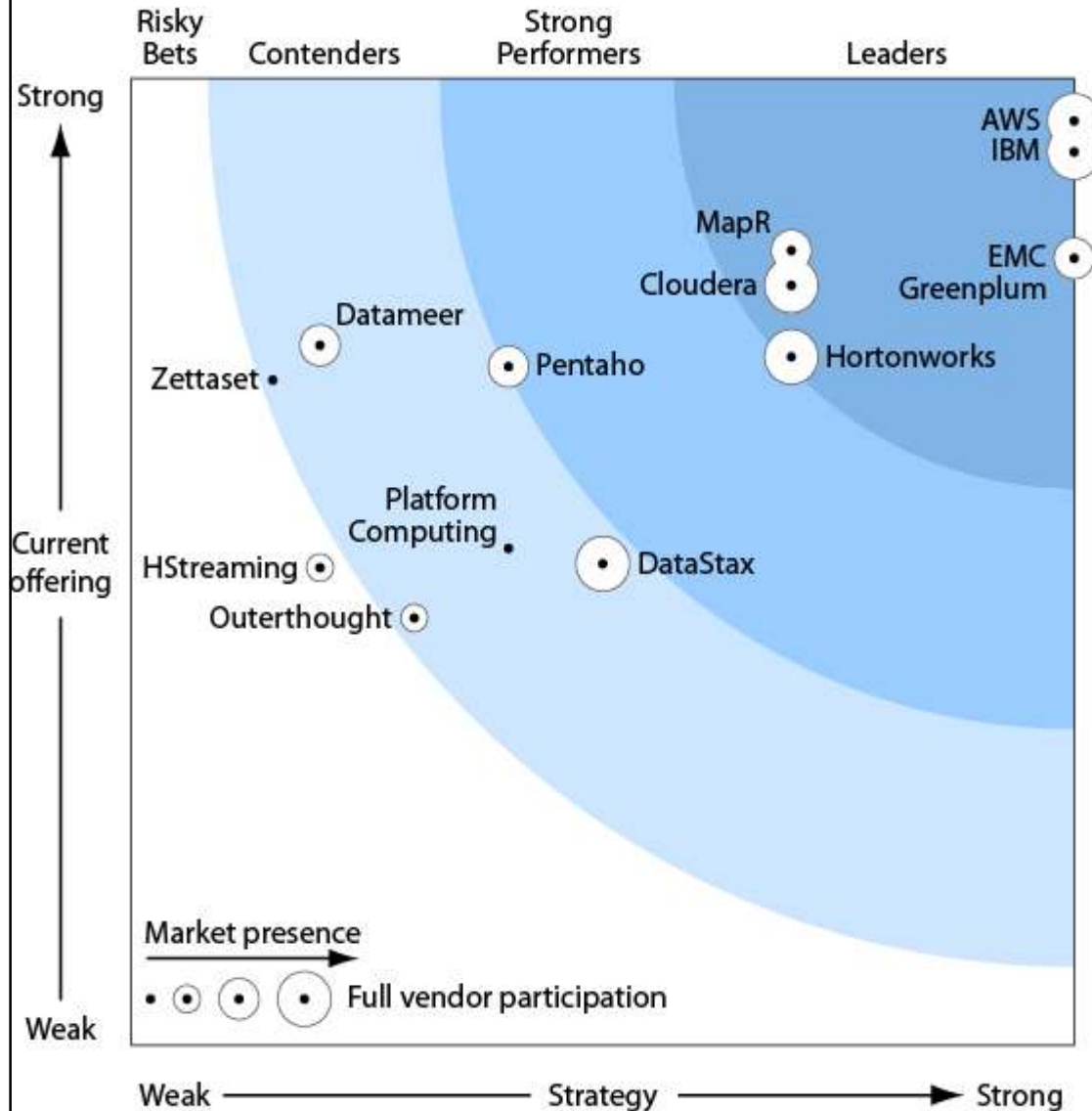
Big Data i tehnologija



Tko zna što je ovo?



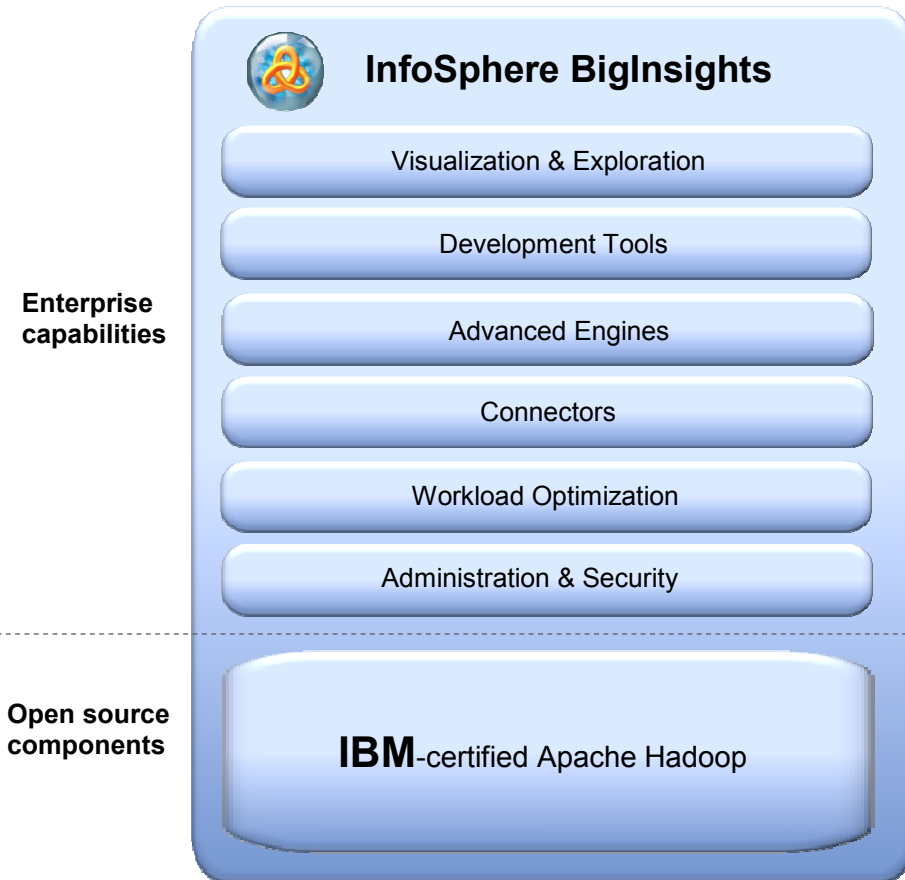
Forrester Wave na temu Hadoopa



“IBM has the deepest Hadoop platform and application portfolio. IBM, an established EDW vendor, has its own Hadoop distribution; an extensive professional services force working on Hadoop projects; extensive R&D programs developing Hadoop technologies; connections to Hadoop from its EDW.”

–The Forrester Wave™: Enterprise Hadoop Solutions, 1Q12

BigInsights – proširuje mogućnosti Open Source Hadoopa



Tehničke prednosti

- Ugrađena analitika
 - Text analytics engine, annotators, Eclipse alati
 - Interface prema project R (statistička platforma)
- Duboka integracija s IBM softwareom
- Spreadsheet-style analitički alat
- **Predefinirani biznis-proces akceleratori**
- Predinstalirane podržane open source i IBM komponente
- Web konzola za admin i aplikativni pristup
- Dodatna sigurnost, performansne značajke,...
- Standardno IBM licenciranje i vrhunska podrška
- Puna open source kompatibilnost
- Siguran daljnji razvoj novih mogućnosti

Hadoop Use Cases



Analyze a Variety of Information

Novel analytics on a broad set of mixed information that could not be analyzed before



Analyze Extreme Volumes of Information

Cost-efficiently process and analyze petabytes of information



Discovery and Experimentation

Quick and easy sandbox to explore data and determine its value

IBM u usporedbi s konkurencijom

- IBM BigInsights naširoko koristi Hadoop, uključujući analitike akceleratora (za tekst i vizualizaciju)
- IBM BigInsights nudi fleksibilnost i nižu cijenu rješenja jer je dostupan kao softver, u cloudu ili na fleksibilnoj IBM hardverskoj referentnoj arhitekturi
- IBM nudi kompletno Big Data rješenje, uključujući BigInsights, Streams, MPP Database, integraciju informacija
- IBM je prilagodio Apache Hadoop za enterprise, čvrsto povezan sa skladištem podataka
- IBM ima predefinirane opcije hardverske arhitekture na povoljnim System x (x86) ili PowerLinux (IBM Power) serverima



“Company XY (IBM competitor) underestimates Hadoop’s potential as an analytics platform in its own right” Ovum, February 2012

Prilike za Hrvatsku



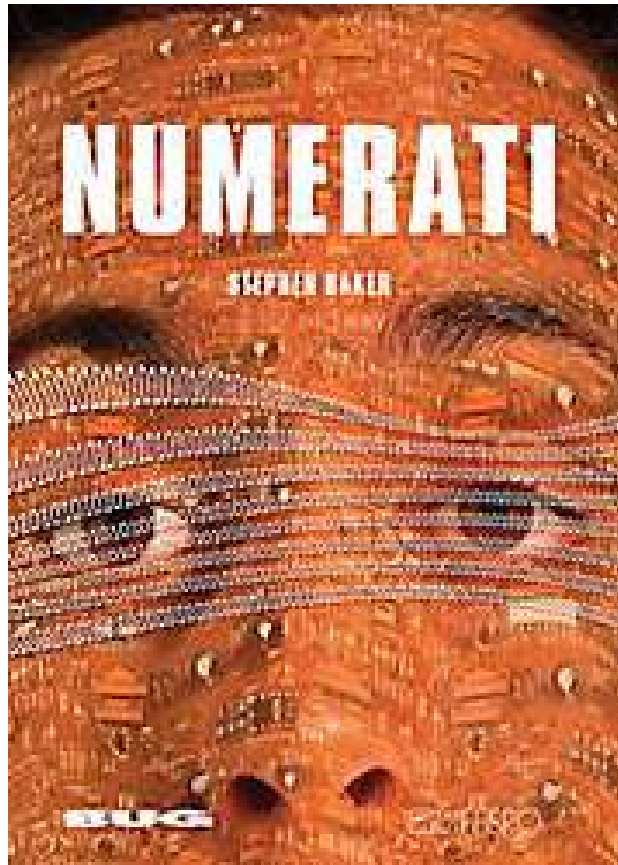
Prilike za hrvatsku akademsku zajednicu i industrijsku politiku

- Data scientist za profesionalce
- Matematički algoritmi za odlikaše
- Hadoop aplikacije za softveraše
- Industrija senzora za hrvatsku industrijsku politiku

Data scientist preko oglasa za profesionalce

The screenshot shows the Harvard Business Review website interface. At the top left is the Harvard Business Review logo. To its right is a search bar with a 'SEARCH' button. Below the logo is a navigation menu with links for 'THE MAGAZINE', 'BLOGS', 'AUDIO & VIDEO', 'BOOKS', 'WEBINARS', 'COURSES', and 'STORE'. A secondary navigation bar indicates 'Registered | limited access' and shows the user 'ALEN GOJCETA' with options to 'Subscribe' or view 'My Account'. The main content area features the article 'Data Scientist: The Sexiest Job of the 21st Century' by Thomas H. Davenport and D.J. Patil. Below the title are social media sharing icons and a 'Comments (0)' section. A 'RELATED' section is partially visible, showing 'Executive Summary' and 'ALSO AVAILABLE'. On the right side, there are promotional banners for 'Subscribe & Save Up to' and 'COMPREHENSIVE LEADERSHIP PROGRAM'.

Matematički algoritmi za odlikaše



Massive Paralell Processing (npr Hadoop) aplikacije za softveraše

The screenshot shows the ReversingLabs website. At the top left is the logo "REVERSING LABS". A navigation menu includes "HOME" (highlighted in red), "SOLUTIONS", "PRODUCTS", "TECHNOLOGY", "RESOURCES", and "ABOUT". The main banner features a large, stylized white geometric logo on a blue background. To the right of the logo, the text reads "TitaniumCore" in large white font, followed by "AUTOMATED STATIC DECOMPOSITION OF FILES" in a blue box. Below this, it says "File Analysis Platform is the next generation unpacking engine that takes any unknown binary object, removes all protection artifacts and then unwraps all formatting elements." Below the banner are three columns: "INTRODUCTION" with a paragraph about file analysis tools, "HIGHLIGHT" with a section for "TitaniumCore™ Automated Static Decomposition" and a link to download software, and "NEWS" with two news items dated "2012-10-12".

<http://www.reversinglabs.com/>

Što tu ima za hrvatsku industriju?

ENTRY

DESIGNER

MANUFACTURER / CLIENT



digitron, db 801-3
Taschenrechner

iF product design award 1974
Office Business

Not announced

digitron electronics equipment
factory
Buje,

Dobar dizajn Digitronove vage, no što je zaista bio uspjeh?



[HOME](#) [AWARDS](#) [DESIGN SPECIALS](#) [SUCCESS STORIES](#) [SEARCH](#) [JUROR INDEX](#) [iF HOME](#)

[Back to selection](#) | [Directlink to this entry](#)

[Browse through entries](#) < 1/1 >



Electronic Price Computing Scale "SKALA 02"

Electronic Price Computing Scala

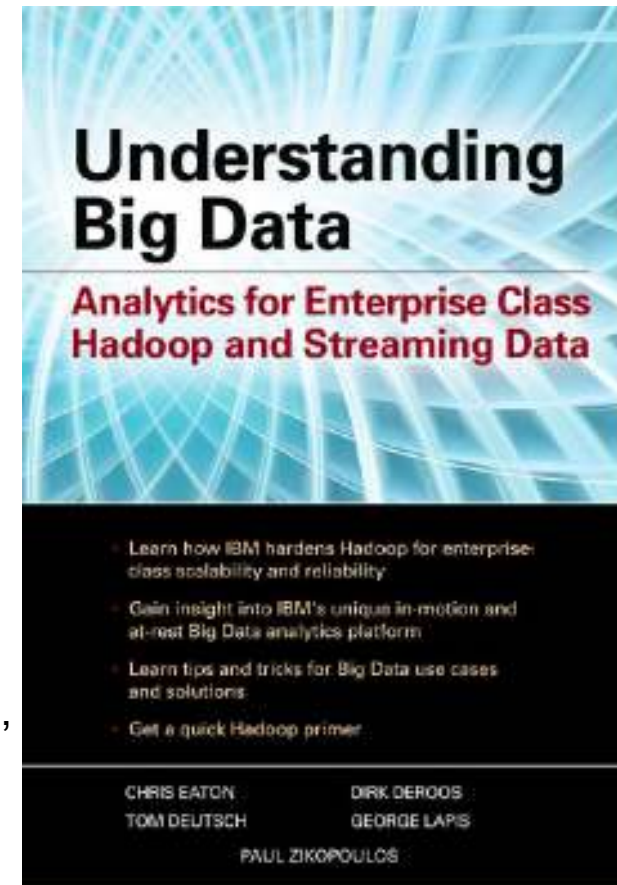
Awarded in iF product design award 1987

Digitron Scala je 1985 imala mjerni pretvornik sile, a što danas ima iPhone

- **Senzor blizine** – služi iskapčanju ekrana u svrhu izbjegavanja pokretanja komandi kad je telefon “na uhu”
- **Senzor pokreta / akcelerometar** – omogućuje promjenu landscape / portrait usmjerenja ekrana, gašenje / utišavanje kod okretanja prema dolje
- **Senzor svjetla** – detektira koliko je svjetlosti u okolini telefona, automatski prilagođava svjetlost ekrana te omogućava odluku o korištenju blica
- **Senzor vlage** – daje do znanja da je telefon smočen
- **Tro-osni žiroskop** – u kombinaciji s akcelerometrom omogućava šest osi korištenja pri čemu je telefon osjetljiviji i fleksibilniji za primjene poput igara
- **Senzor dodira ekrana, senzor oštine (kamera), ...**

Zaključno

- Big Data je sintagma koja označava analitička rješenja temeljena na masivnom paralelnom procesuiranju (MPP) strukturiranih i nestrukturiranih podataka.
- Primjene su bezbrojne
- Ograničenja su standardna
- Tehnologija se temelji na Apache Hadoopu, open source softveru kojeg je IBM širom prigrlio u svom Infosphere BigInsights rješenju
- Hrvatska mora profitirati od dobrobiti umreženog i digitaliziranog planeta kroz:
 - INDUSTRIJU: industriju senzora,
 - IT TVRTKE: razvoj Big Data rješenja temeljenih na MPP,
 - SVEUČILIŠTA: razvoj matematičkih algoritama,
 - POJEDINCI: stjecanje tehnika i vještina upravljanja velikim količinama raznolikih podataka



<http://www-01.ibm.com/software/data/bigdata/>

