



## Vad är Business Intelligence?

Hercules Dalianis  
DSV-SU-KTH  
e-post:hercules@kth.se  
070-568 13 59 / 08-674 75 47

Hercules Dalianis 1



## Business Intelligence

- *Intelligence* är ett tvetydigt ord
- Både *förmåga* och *underrättelse*

Hercules Dalianis 2



## Andra termer

- Specifika termer för BI är:
  - konkurrenten - competitor intelligence
  - marknaden - market intelligence,
  - teknologin - technical intelligence,
  - kunden - customer intelligence



Hercules Dalianis 3



## Områden

- Marknadsanalys
- Statistik
- Affärsutveckling
- Strategisk planering
- Kommunikation och information

Hercules Dalianis 4





## Övergripande faktorer

Bevakning av

- makroekonomiska,
- politiska
- teknologiska
- sociala
- demografiska

faktorer



Hercules Dalianis 5



## Aktörer

- informatör
- branschanalytiker
- affärsenhetschef
- försäljningschef



Hercules Dalianis 6

## Omvärldsanalys

- Att på ett strukturerat sätt kontinuerligt bevaka och analysera de faktorer som påverkar företagets utveckling och framgång
- Att använda informationen för beslutsfattande



Hercules Dalianis7

## Tillgången på information

- Antalet öppna informationskällor har ökat dramatiskt
- Kräver urval för att värdesättas och analyseras.
- Allt detta bidrar till ett behov att utveckla ämnet Business Intelligence.



Hercules Dalianis8

## Se samband

- Se samband, tänka utanför de vanliga mönstren, ifrågasätta det som andra tar för givet och ställa frågor om det som observerats.
- Allt är följaktligen inte vetenskap eller möjligt att lära - det finns en hel del konst också!



Hercules Dalianis9

## Språk och kanaler

- 1000-tals språk i världen
- Tio stora språk och svenskan
- Skriven information: pressreleaser, tidningar, rapporter, vetenskapliga rapporter, böcker.
- En del digitalt en del bara på papper i bibliotek
- Talad information: Radio, TV



Hercules Dalianis10

## Observer

- Läsa och klippa i tidningar
- 100-tals personer
- 1000-tals tidningar per dag
- Ringa kunderna om en intressant nyhet dyker upp
- Posta eller faxa pressklipp till kunder



Hercules Dalianis11

## Nyhetsanalys

- Analysera nyheter utifrån kundens behov
- Vilka företag förekommer tillsammans i nyheter?
- Årsomsättningen på dessa företag

Hercules Dalianis12






## Nya tiden

- Nyhetsflöde
- Nyhetsarkiv - bibliotek
  - Kungliga biblioteket-sparar hela Sveriges bok och www-produktion
  - InternetArchive - WaybackMachine  
<http://www.archive.org/web/web.php>
  - SVT och SR arkiv
- Internet ett arkiv men stort och ostrukturerat
  - Semantic webb

---

Hercules Dalanis13






## Kanaler

- Tidningar
- Radio/TV
- Företagsnyheter
- Myndighetsinformation

---

Hercules Dalanis14






## Format

- HTML
- TXT
- PDF
- PS - .ps postscript
- MS Office
  - Word .doc
  - Excel .xls
  - Powerpoint .ppt

---

Hercules Dalanis15






## (forts)

- XML-formatet RSS
- NewsML
- RTF- Rich Text Format
- Ljud- wav, au, mp3, wma, real-audio
- Bilder - jpg, gif, tiff
- WML-Wireless Markup Language WAP
- RDF- Rich Description Format -Semantic webb

---



Hercules Dalanis16

- Spindlingsfrekvens
  - En gång per månad en gång per dag, en gång per 5 minuter
- Åtkomlighet
  - Javascript
  - Robots.txt
  - Osynliga webben och databaser
  - Intranät och extranät

---

Hercules Dalanis17

## Java-script



- Javaskriptlänkar hindrar sökspindlar att hitta sidorna javaskriptlänkarna pekar på
- Vissa webbrowsers t.ex. för synskadade, kan inte evaluera eller följa javaskriptlänkar.
- Den som tillhandahåller informaion kompletterar med klartextlänkar
 

```

      <script language="JavaScript1.2" src="menyer.js"></script>
      <noscript>
      <a href="menyal-1.html">Val 1</a>
      <a href="menyal-2.html">Val 2</a>
      ...
      <a href="menyal-n.html">Val n</a>
      </noscript>
      
```

---

Hercules Dalanis18





## Robots.txt

- Robots.txt i rotkatalogen på webbplatsen hindrar sökmotorsspindlar att spindla och indexera webbplatsen

```
User-agent:*  
Disallow: /search  
Disallow: /cgi-bin/  
Disallow: /intern/  
Disallow: /about/chris.html
```

Det går också att utesluta enstaka webbsidor genom att använda en metatagg i själva html-koden.



Hercules Dalianis 19



## Databaser

- Finns ingen enkel ingång in till
- Måste specialbehandlas manuellt /automatiskt
- Ibland har den som tillhandahåller information löst detta genom ett index med adresser



Hercules Dalianis 20



## Intranät Extranät

- Intranät innanför brandväggar
- Extranät innanför brandvägg men inloggningsbart för kunder- t.ex. nyheter, banktjänster, säljportaler



Hercules Dalianis 21



## Problemställning

- Veta allt lite före alla andra
- Få en överblick



Hercules Dalianis 22



## Hur gör jag?

- Bevakar några favoritsajter Computer Sweden, Aftonbladet, portaler Internetbrus.com
- Läser tidningar Ny Teknik, DN,
- Kollar några företag
- Kollar på offentliga webbplatser
- Allt detta på svenska



Hercules Dalianis 23



## Idealt?

- Bevaka nyhetsflödet automatiskt på flera språk, format, överallt och automatiskt
- TDT-Topic Detection Tracking (2000)



Hercules Dalianis 24



## TDT-Topic Detection Tracking

- TDT-begreppet kommer från DARPA programmet TIDES Translingual Information Detection, Extraction and Summarization,
- Med datorer automatiskt detektera nyheter, extrahera och översätta dessa.
- Engelska och kinesiska nyheter på radio, TV och vanliga textmedia bevakas med taligenkänning och statistiska och lingvistiska metoder.

Hercules Dalanis 25



## Fem olika uppgifter i TDT.

- Segmentation - att kunna urskilja olika nyheter i nyhetsfloden,
- Tracking - att kunna följa nyheter med samma ämne/topic.
- Detection - att kunna gruppera nyheter efter ämne.
- Link detection- har två nyheter ngn koppling ämnesmässigt och slutligen det svåraste och intressantaste
- First Story Detection kunna upptäcka en helt ny nyhet i nyhetsfloden.
- Man lyckas hyfsat bra med de fyra första uppgifterna, man gör 33-25 % fel där, men med First Story Detection hamnar man på 70% fel.



Hercules Dalanis 26



## Söktips med Google

- Site / domän
- Språk
- Format

Hercules Dalanis 27



## Vivisimo

- Sök Business Intelligence
- Sök Business Intelligence och

Hercules Dalanis 28