

*Användning av big data och hantering
av kvalitetsaspekter kring
databaser i näringslivet*

Jörgen Hasselgren, HNO3 Data Science AB

Jörgen Hasselgren

- 25+ år i branschen
 - Cognos
 - Informix
 - Resco
 - Itera
 - HNO3 Data Science
- Konsult Business Intelligence / Analytics
- Lärare
 - BI, Machine Learning m m.

- Arbetat med bl a

PWC

Swedish Match

Systembolaget

Scania

Aco

Roche

Pfizer

SEB

Edenred

RNB

Telenor

IL Recycling

Fujifilm

Ragnsells

Investor

Landsting

Sandvik

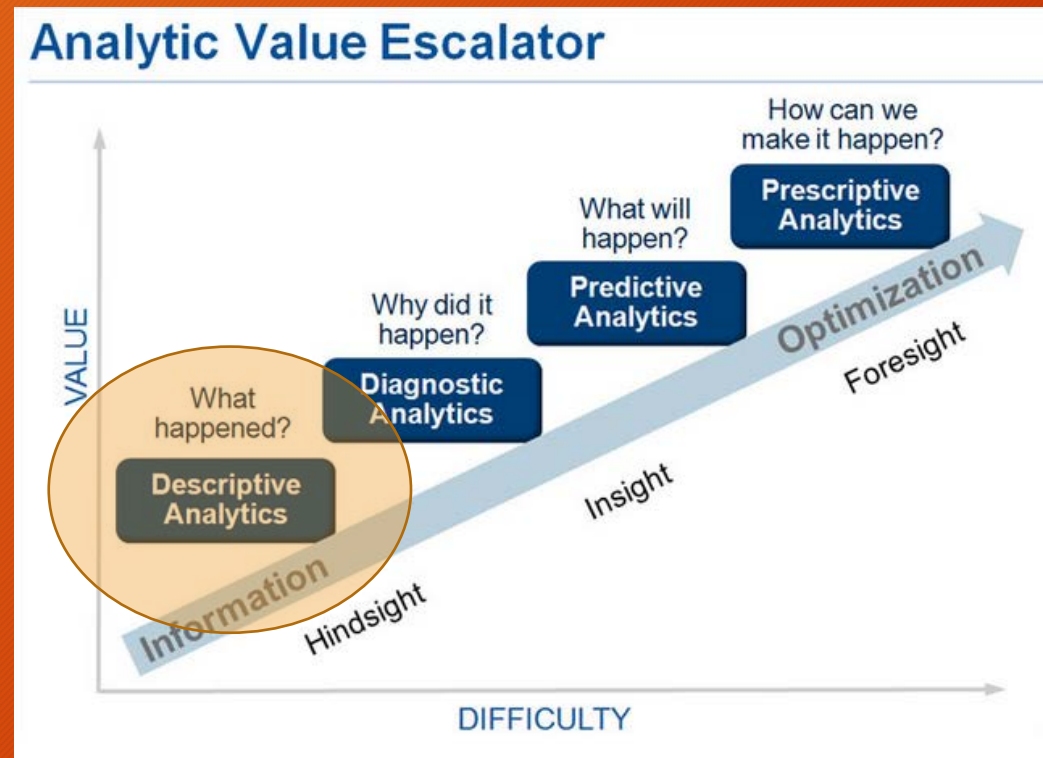
Preem

SPP

M fl



Var jobbar vi idag?



Är datakvalitet ett problem?

- Ta fram telefonen

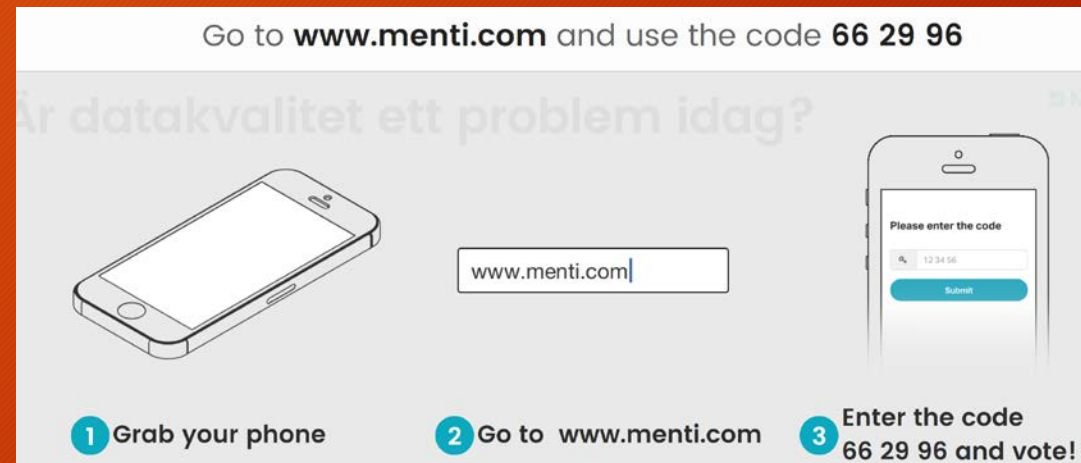


- Öppna favorit-browser
- <http://menti.com>

66 29 96

Go to www.menti.com and use the code **66 29 96**

Är datakvalitet ett problem idag?



- 1 Grab your phone
- 2 Go to www.menti.com
- 3 Enter the code 66 29 96 and vote!

Resultat

*Statistiskt
säkerställt?*

Bortfall?

Reproducerbart?

Bra Sample?

Användbart?

(good enough)

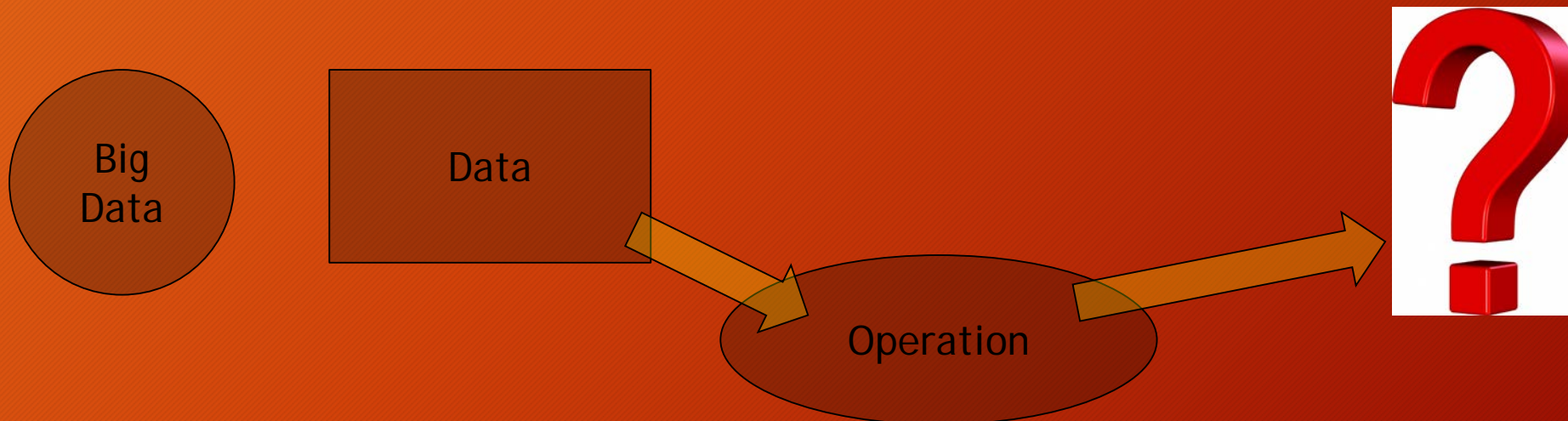
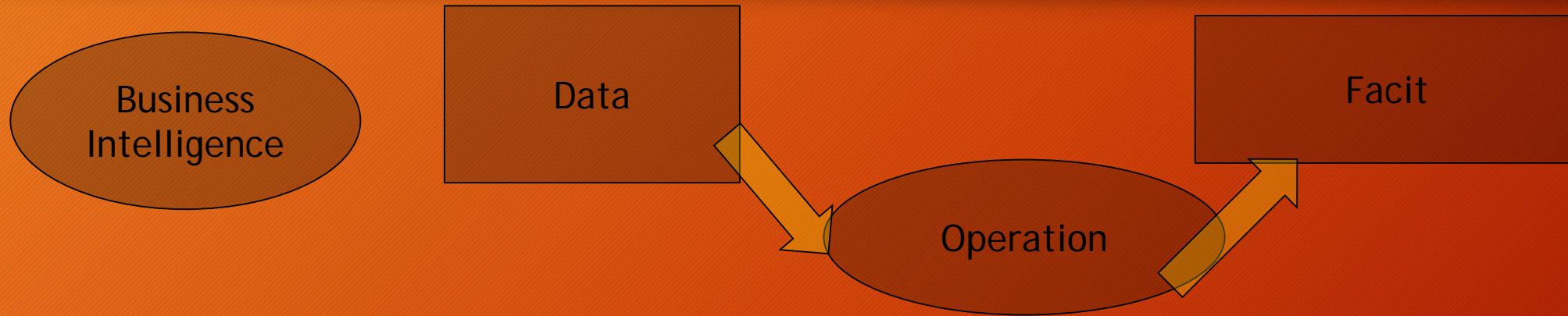
Vad är kvalitet?

Data är av hög kvalitet om användarna säger att det är så.

Kvalitet = Uppfyllandet av satta krav



Vad är hög kvalitet / korrekt?



Big Data / Business Intelligence

Data lake



Snabbt
Förändringsbart

R&D, Marknadsföring,
beteenden



Pålitligt
Säkert
Uthålligt

Rapportering
Kundinformation

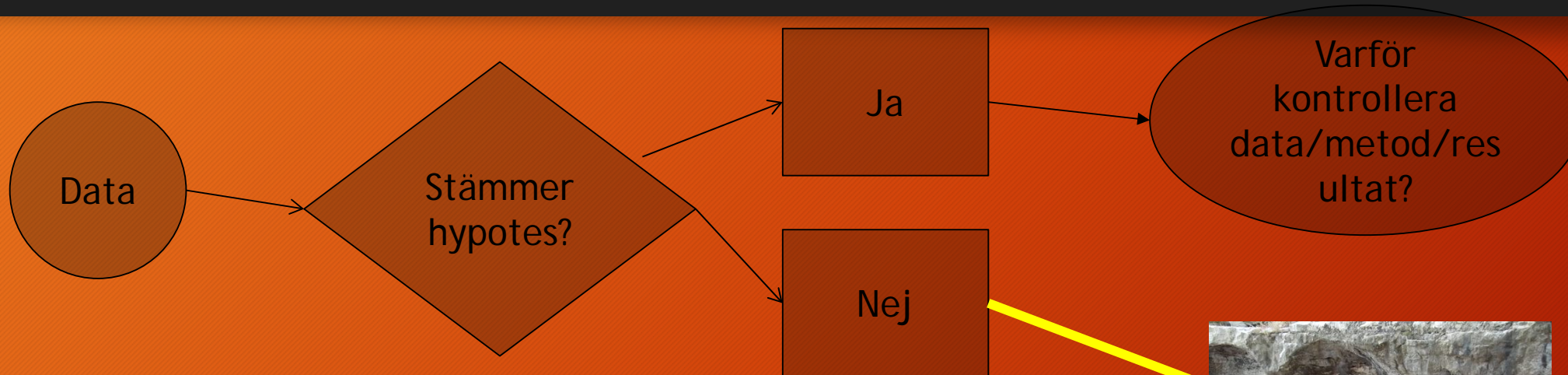
Vad används Big Data för i kommersiella sammanhang (ur min synvinkel)

- Bekräfta / förneka idéer
- Hitta nya grupper att bearbeta kommersiellt
- Förstå dina kunder
- Och förstås mycket mer 😊

De största hindren vid Big Data analyser

- Integration
- Data kvalitet

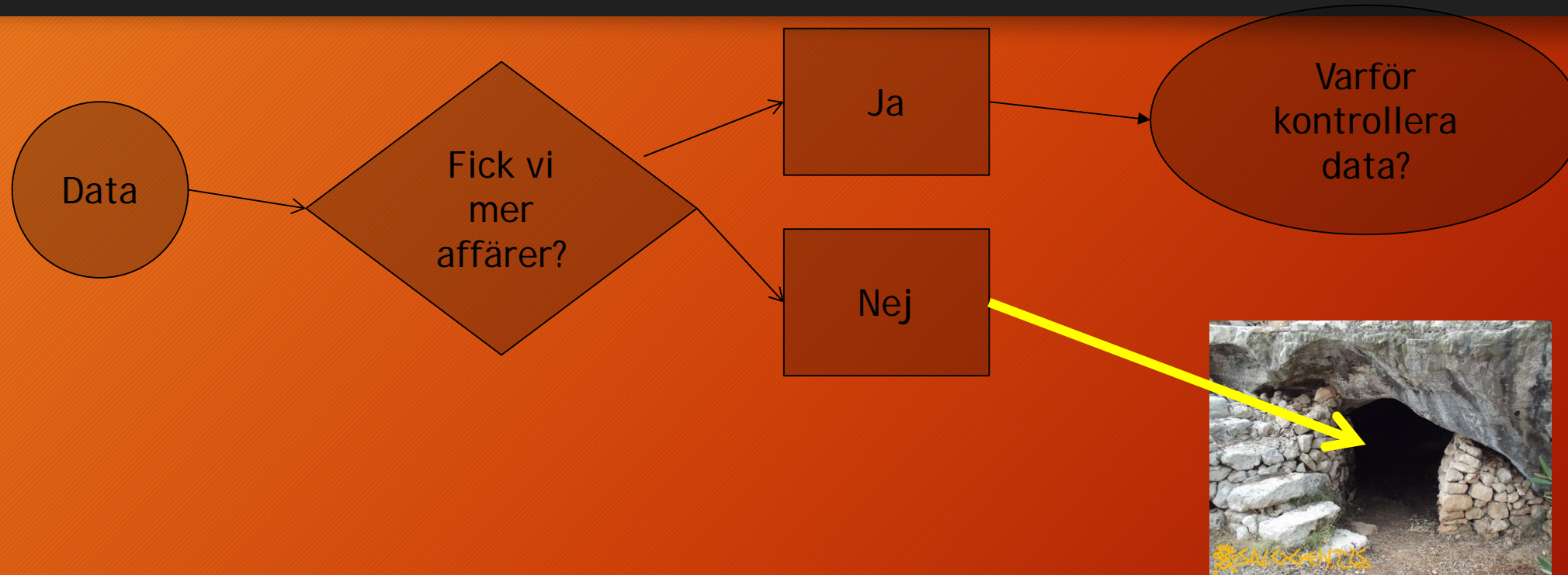
Fall 1 hypotestest



I detta fall hypotes om t ex beteende. Något som inte går att mäta direkt. "Vi tror att kunderna gör så eller så"



Fall 2 - Generera affärer



Bovar i dramat

P O C
Proof
Of
Concept



Kortsiktighet

Generera affärer

- Marknadskampanj
 - Lyckas - Varför förklara
 - Misslyckas - Ingen tid att få en förklaring. Gör ett nytt försök. Precis som vetenskapliga studier, man drar inga lärdomar av misslyckade projekt. Ingen vill förknippas med ett misslyckat projekt
- Black Box - Spelar det någon roll om den fungerar?

Saker kunder aldrig frågar om

- Algoritmer
- Urval (endast sporadiskt)
- Bortfall

Praktiskt exempel Internet of Things

- Hur många är i butiken?
- Rörelsemönster i butik?
- Hur länge är de i butiken?
- Mäta Wifi-signaler
- Tolka bilder (kön, ålder, ansiktsuttryck)

Plats och tid

IOT

- Hur mäta kvalitet?
- Kostar lika mycket att mäta som att göra grundprojektet
 - Rörelser över ytor t ex
- Visa slutresultat

Slutsats

- *“Good enough”*
 - *Data skall vara komplett, korrekt och ha samma betydelse överallt*
- *Intern data ALLT ska vara rätt*
- *Big Data - Good enough Quality*
- *Kvalitet tänker man på i efterhand*
- *Omoget tänk kring kvalitet och dålig skolning kring dessa ämnen bland ekonomer/marknadsfolk*
- *NI behövs för att förklara och definiera kvalitet (en dröm?- en Q%)😊*