



# Text mining, metodik og forretningsmæssig anvendelse

Sune Birch  
Seniorkonsulent, SAS Institute

**sas** forum  
DENMARK 2008 copenhagen

# Agenda

- Motivation
- Forretningsmæssige anvendelser
- Metodik
- Case og demo af text mining
  - Klassifikation af servicereporter

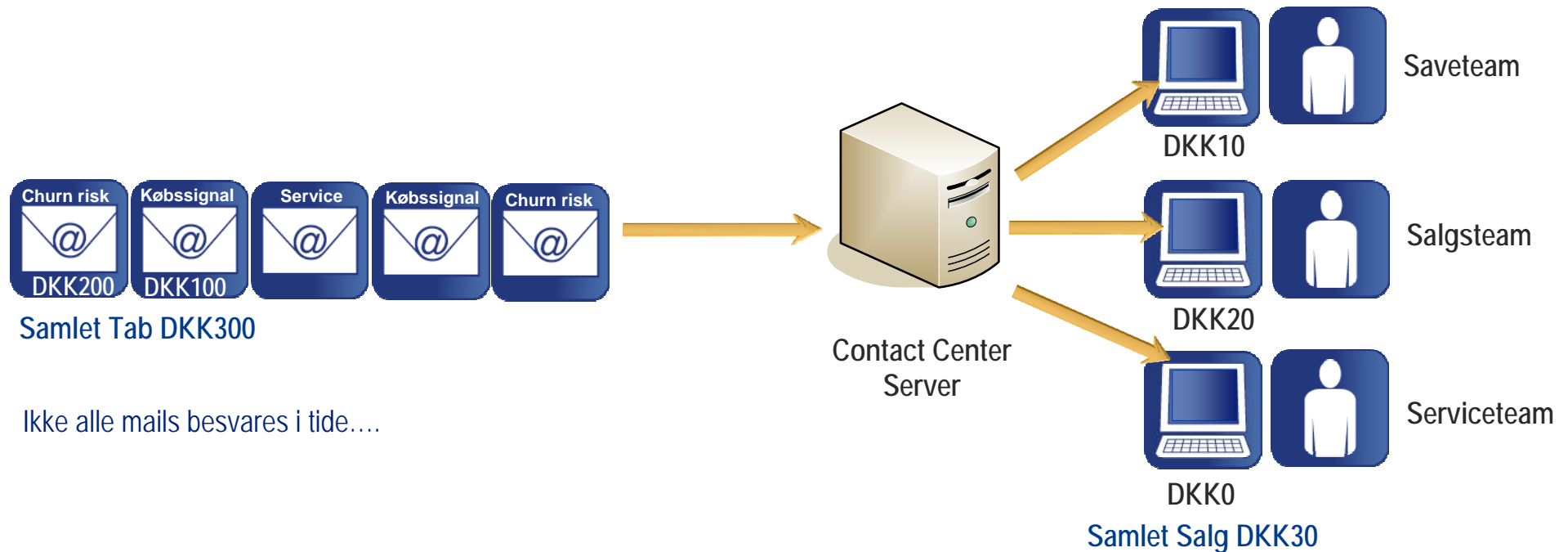
# Motivation

- Store mængder tekst lagret, men ikke anvendt
  - E-mails
  - Dokumenter
  - Websider
  - Service requests
- It-systemer gode til struktureret information
  - Gør tekst struktureret
  - Anvend "traditionelle" teknikker

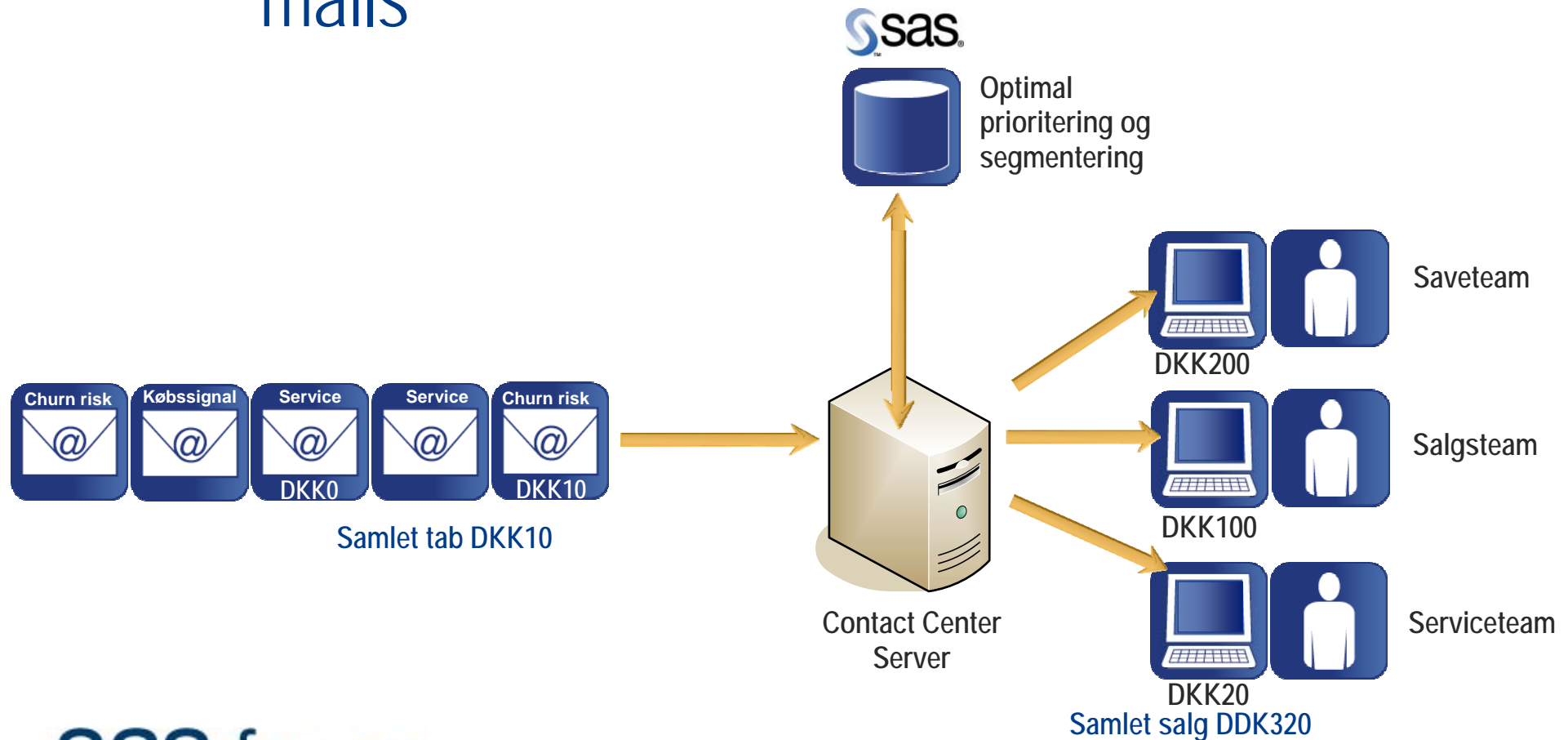
# Motivation - stigning i brugergenereret information

- Anvendelse af bruger til bruger-servicering (Web 2.0)
  - Diskussionsfora
  - Chat
  - Blogs
  - Wikis
  - Anmeldelser

# Kundesupport eksempel: Routing af e-mails baseret på text mining



# Kundesupport eksempel: Optimal prioritering og routing af e-mails



# Forretningsmæssige anvendelser

- Højne servicekvalitet
  - Routing af forespørgsler
  - Kundeindblik og prioritering af budskaber
  - Berigelse af kundeoplysninger
- Identificer salgsmuligheder og risici
  - Købssignaler
  - Opsigelsesrisiko

# Forretningsmæssige anvendelser

- Effektivisering af processer
  - Prioritering mellem kunder
  - Garantiforpligtelser
    - SAS<sup>®</sup> Warranty Solution
    - Early warning-systemer
      - Identificer kritiske problemer tidligere
      - Reducer omkostninger
  - Hvad anvendes resurser på?

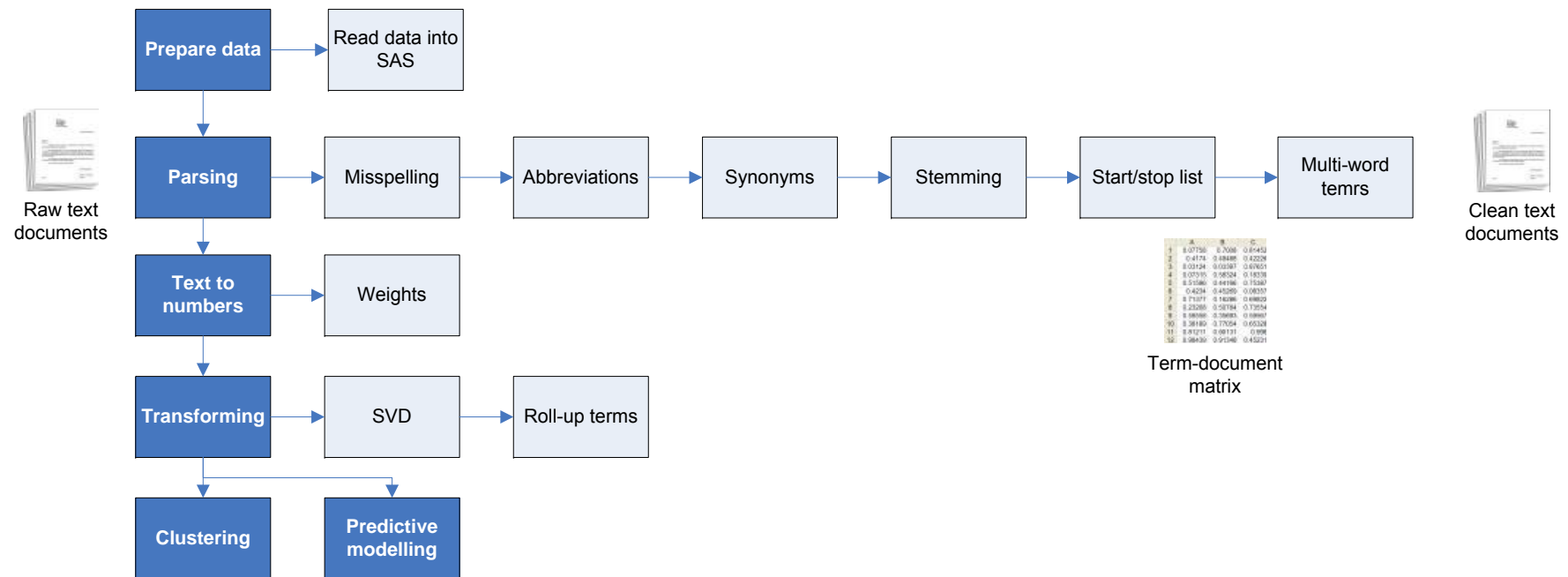
# Case og demo: Hvad anvendes resurser på?

- Case
  - Servicetekniker udfører arbejde
  - Servicetekniker skriver servicereport
- Optimering
  - Kan arbejdet tilrettelægges mere hensigtsmæssigt?
- Planlægning
  - Hvilke typer resurser har vi brug for?
  - Hvilke reservedele har vi brug for?

# Case og demo: Hvad anvendes resurser på?

- Løsning: Klassifikation af servicereporter
  - Hvilken type arbejde er udført? Eks.
    - *1 year maintenance*
    - *Software*
    - *Generator*

# Metodik - text mining steps





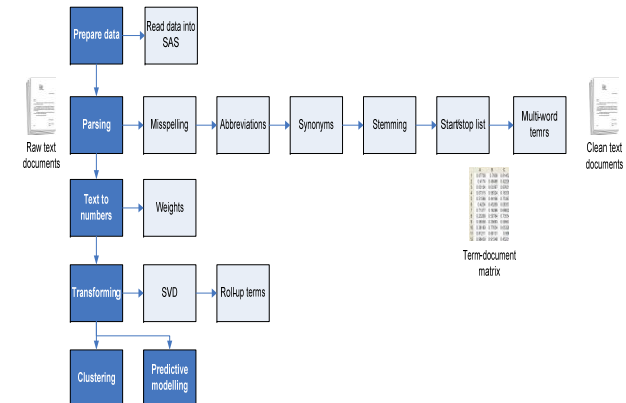
# Case og demo: SAS<sup>®</sup> Text Miner

# Case og demo - resultater

- Prædiktiv modellering
  - Mis-klassifikationsrate
    - Top 5 kategorier: 5%
    - Top 10 kategorier: 7%
- Clustering
  - Naturlige kategorier opdaget

# Text mining-værktøjskasse – hvad skal der til?

- Text processing
  - Gør ustruktureret information struktureret
- Domæneviden
  - Forkortelser, synonymer, fagtermer, multi-word termer
- Clustering
  - Udforsk nye kategorier af dokumenter
- Prædiktiv modellering
  - Prædikter køb, opsigelse, kreditrisiko...
  - Anvend struktureret information



## Til sidst...

- Spørgsmål?
- SAS<sup>®</sup> Enterprise Miner 5.3 indlæg  
fredag kl. 11.30

THANK YOU FOR ATTENDING

**sas**® forum  
DENMARK 2008 copenhagen