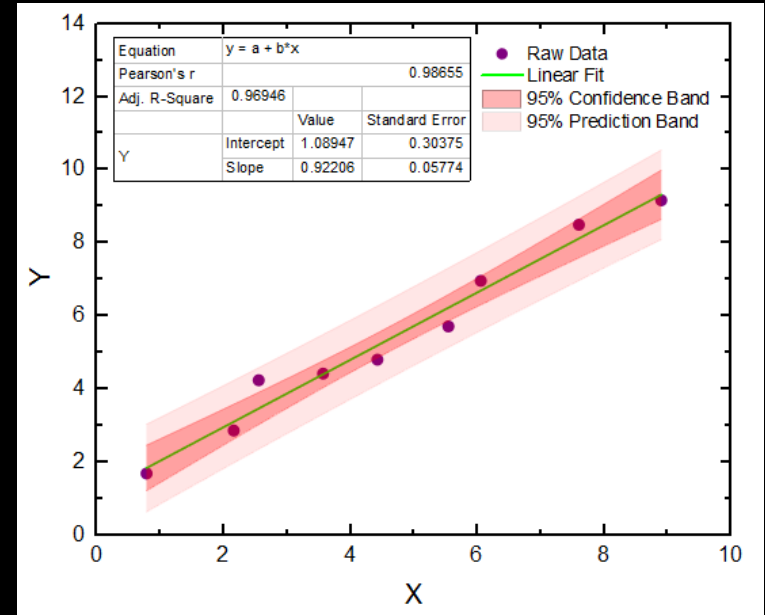
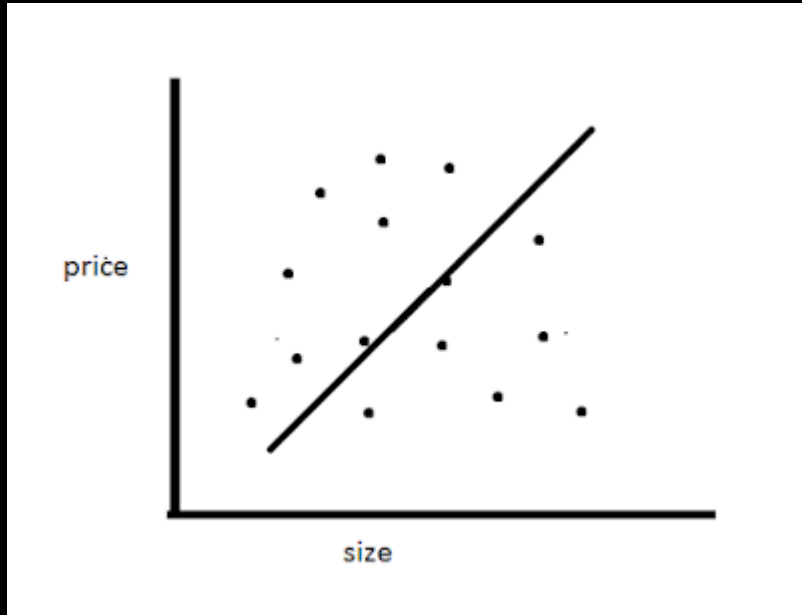


Hvad kan data science gøre for bibliotekerne?

**Christian Boesgaard
cbo@dbc.dk**

- **data science**
- **recommendere**
- **prototyper**





Data science

Artificial Intelligence

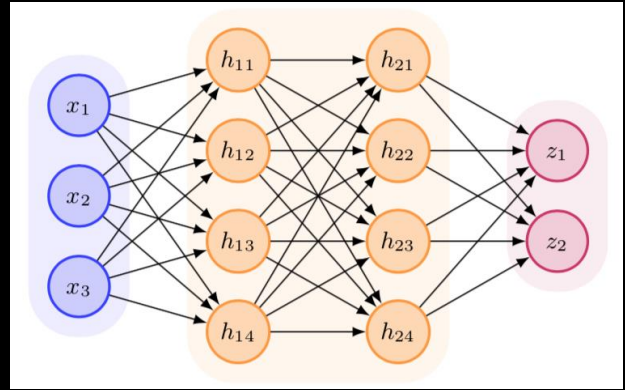
Machine Learning

Deep Learning

Data Science



IN CS, IT CAN BE HARD TO EXPLAIN
THE DIFFERENCE BETWEEN THE EASY
AND THE VIRTUALLY IMPOSSIBLE.



Matrix Factorization

m = number of users, n = number of items
choose d , the number of features

$$\begin{matrix}
 \begin{matrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{matrix} & \times & \begin{matrix} ? & ? \\ ? & ? \end{matrix} & = & \begin{matrix} & & 1 \\ 5 & & 2 \\ 2 & 3 & \\ & 1 & \end{matrix} \\
 m \times d & & d \times n & & m \times n
 \end{matrix}$$

$\hat{r}_{ui}^{d=2} = q_i^T p_u$

Teknik

Google!

Search the web using Google!

10 results ▾

Google Search

I'm feeling lucky

Index contains ~25 million pages (soon to be much bigger)

About Google!

[Stanford Search](#) [Linux Search](#)

Get Google! updates monthly!

your e-mail

Subscribe

[Archive](#)

Copyright ©1997-8 Stanford University



Wayfinding

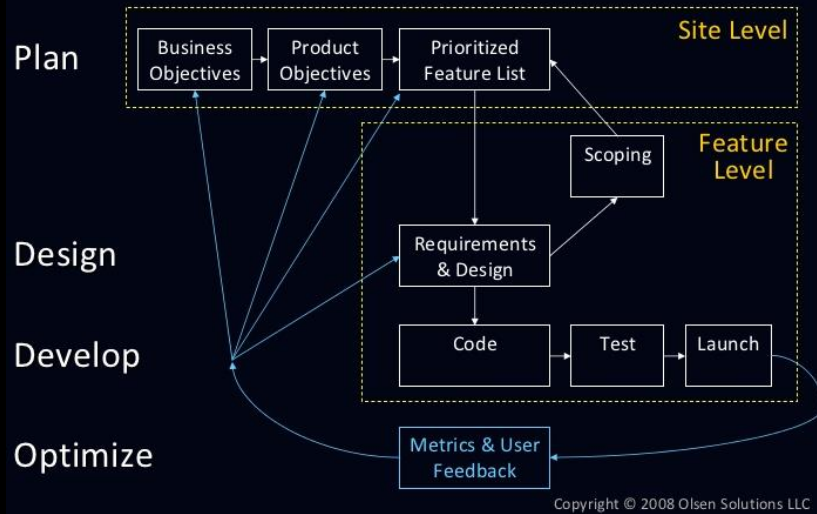


Data



Domæneviden

Adding Metrics and Optimization to your Product Process



Metriker og målinger



Etik

UMBERTO ECO

DRONNING LOANAS
MYSTISKE FLAMME



A

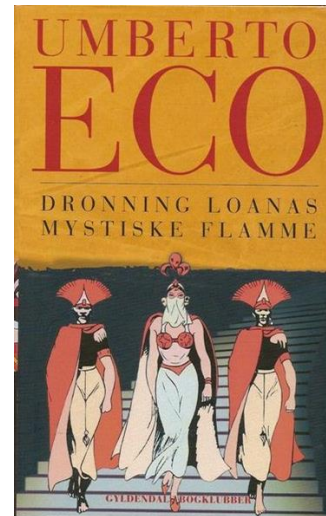
Milan Kundera: Tilværelsens ulidelige lethed

Paul Auster: Opfindelsen af ensomhed

Alain De Botton: Om kærlighed

Andrew Wilson: Talent for mord : den mystiske mrs. Christie

Siegfried Lenz: Dykkeren



B

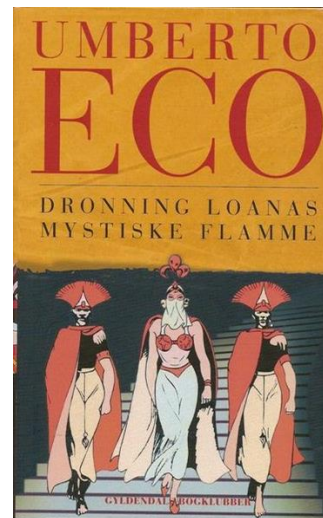
Vindens skygge af Carlos Ruiz Zafón

Stalins køer af Sofi Oksanen

Fiskerne af Hans Kirk

Svøm med dem som drukner af Lars Mytting

Fattigmandsære af Heðin Brú



C

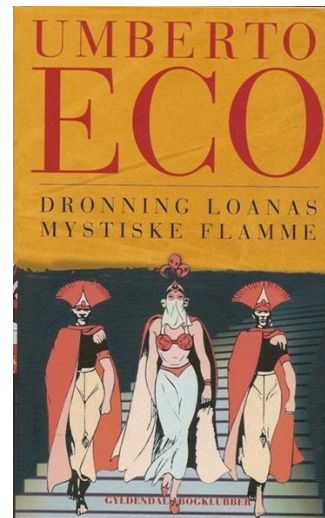
Rita Monaldi: Imprimatur

Ildefonso Falcones: Sigøjnerinden fra Sevilla

Enrique Moriel: Den tidløse by

Jon Fasman: Korttegnerens bibliotek

Chufu Lloréns: Jeg vil give dig verden



D

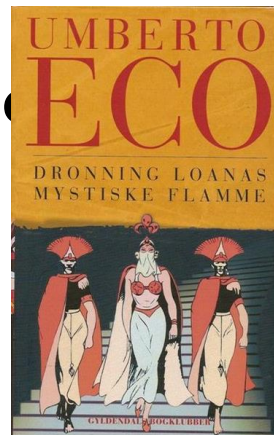
Råchokteksterne af Steven Hall

**Dumas-Klubben: eller I Richelieus skygge af Arturo Pa
Reverte**

2666 af Roberto Bolaño

Gravity's Rainbow af Thomas Pynchon

Fiktioner og andre fiktioner af Jorge Luis Borges



Hvad er en recommender?

Funktionalitet der kan svare på hvor meget to bøger ligner hinanden:

- 100 % de er identiske
- 0 % de er helt forskellige

Hvor meget ligner A og B hinanden: 7 %

Hvor meget ligner A og C hinanden: 15 %



(dette er en simpel forklaring af en type recommender)

Hvad kan vi med sådan en recommender

Finde de bøger der ligner en konkret bog mest:

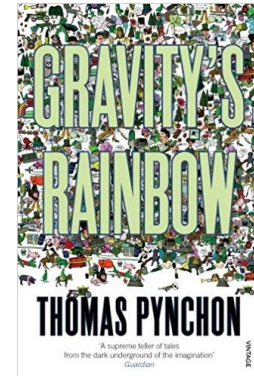
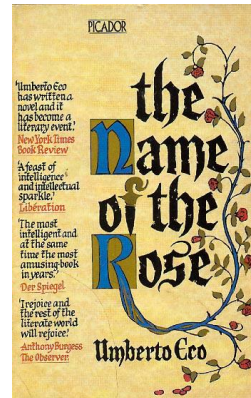
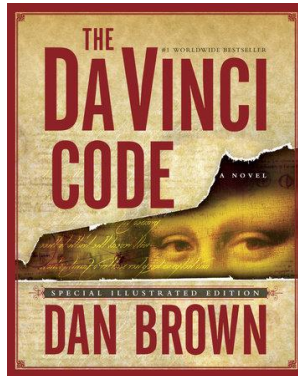
- fx ved visning, søgning, aflevering...
- understøtte “uendelig” browsing

Forklarer forslag: du ser X fordi den ligner Y 🐱



Komplekse anbefalinger

- foreslå bøger der hjemme
- foreslå bøger der ikke er ret udlånt
- foreslå mere udfordrende bøger
- ...



Personalisering og gruppering

En brugers profil kan være et antal bøger, fx:

- udlån
- bøger der er kigget på

Vi kan bruge dem til at finde bøger der ligner

- og gruppere dem efter hvad de handler om (metadata)
- eller så dem der ligner hinanden mest sættes sammen (clustering)



Data

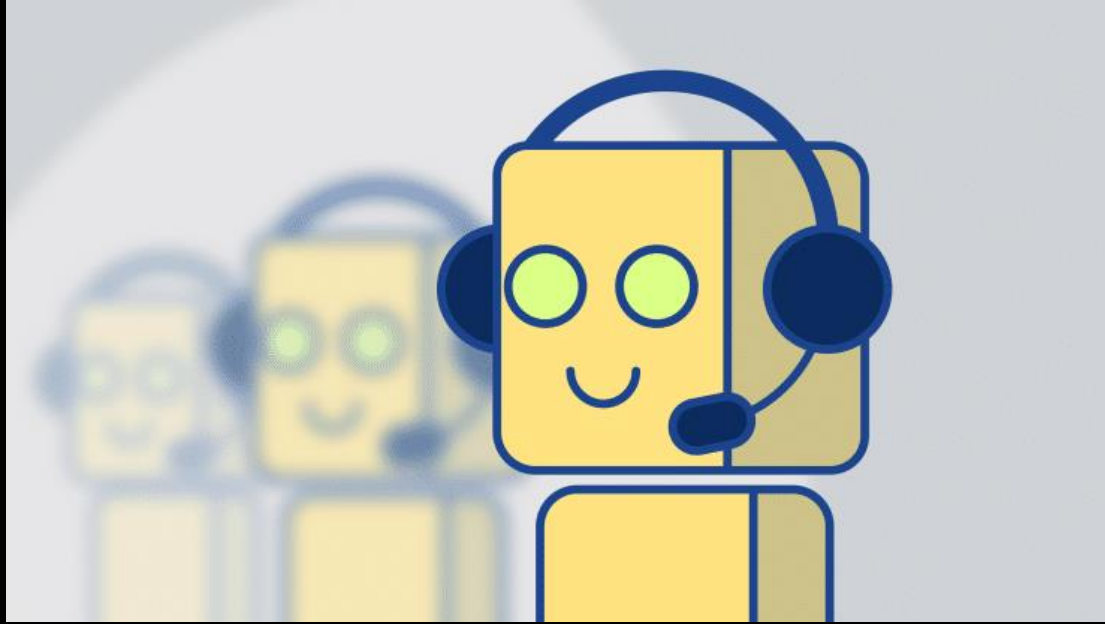
- **“klassisk”- eller “læsekompas”-metadata**
- **brugeradfærd: udlån, visninger, feedback på anbefalinger**
- **indhold: forsider, fuldtekst, sideopsætning**



Brugeradfærd

Anonymiseres før brug

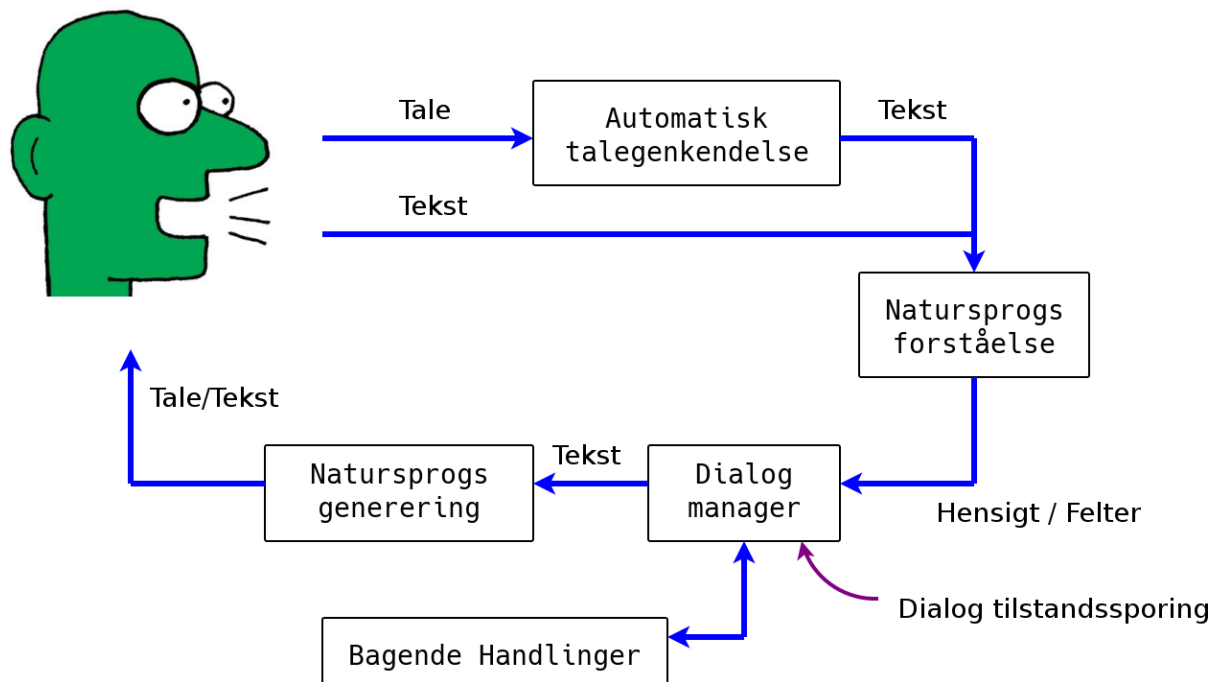
userid	pid
=====	=====
1f98a7a492	870970-basis:51326938
1f98a7a492	870970-basis:29965110
1f98a7a492	870970-basis:27175953
9367ff4d98	870970-basis:29934088



Chatbots

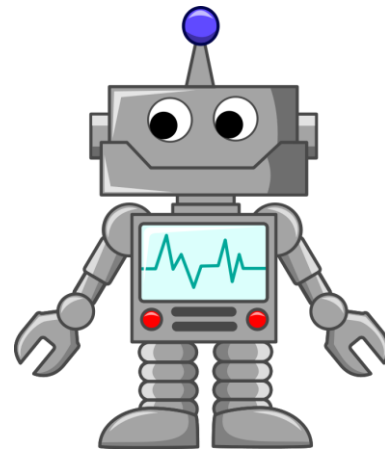
prototyper

Hvad er en Chatbot?



Hvor kan den bruges?

- Dedikerede grænseflader
- Plugin i eksisterende grænseflader
- Førstelinjes support
- ++



Demo

- Jytte
Integration af søgeservices i chatbot
- Talekompasset
Dedikeret grænseflade til finde inspiration gennem samtale