

# The multifactor recommender system @bol.com



Barrie Kersbergen  
expert software engineer  
bkersbergen@bol.com

**bol.com**

# Topics

- develop an operating recommender system
- crucial customer behavioral factors involved
- impact of visual presentation of item recommended.

# Me



- Computer Science
- 12 Yrs University Utrecht, CS/Mathematics
- @bol since 2010

# bol.com





This is no hype

Thursday 27 march 2014



**Bol.com awarded  
'Best webshop in the Netherlands'**

"Bol.com has put major steps forward in the past year, regarding personal recommendations and the use of big data.



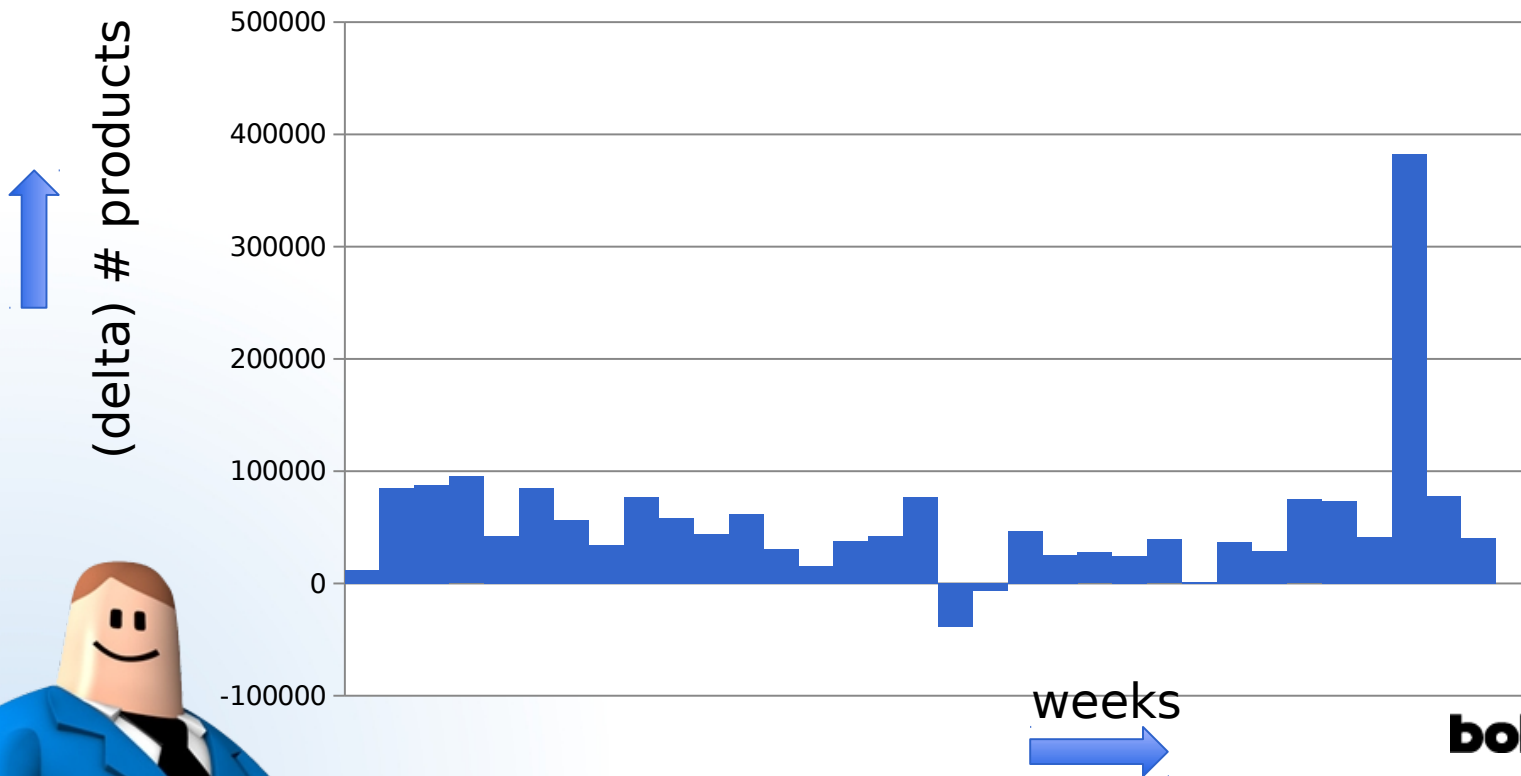
# Unique products per store

- Average supermarket
- XL supermarket

# # added products per week

+ 1 XL supermarket  
every 4 days

## Catalog delta per week





> 8.000.000 products online





> 20.000.000 products  
total





> 15.000.000 click events each



> 5.500.000.000 click events  
per year





# Multiple factors give us clues about our visitor

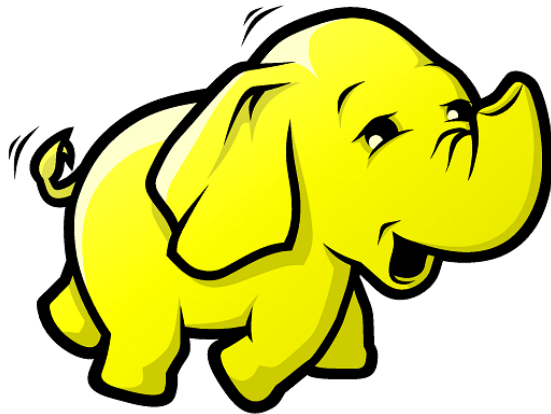


- How long does he look at the product?
- Does he read the product reviews?
- Does he share the product on social media?
- Add item to wishlist
- What products does the visitor click?
- Was the clicked product a recommendation?



# hadoop @bol.com

- Development in 2009
- First applications in production (2010)



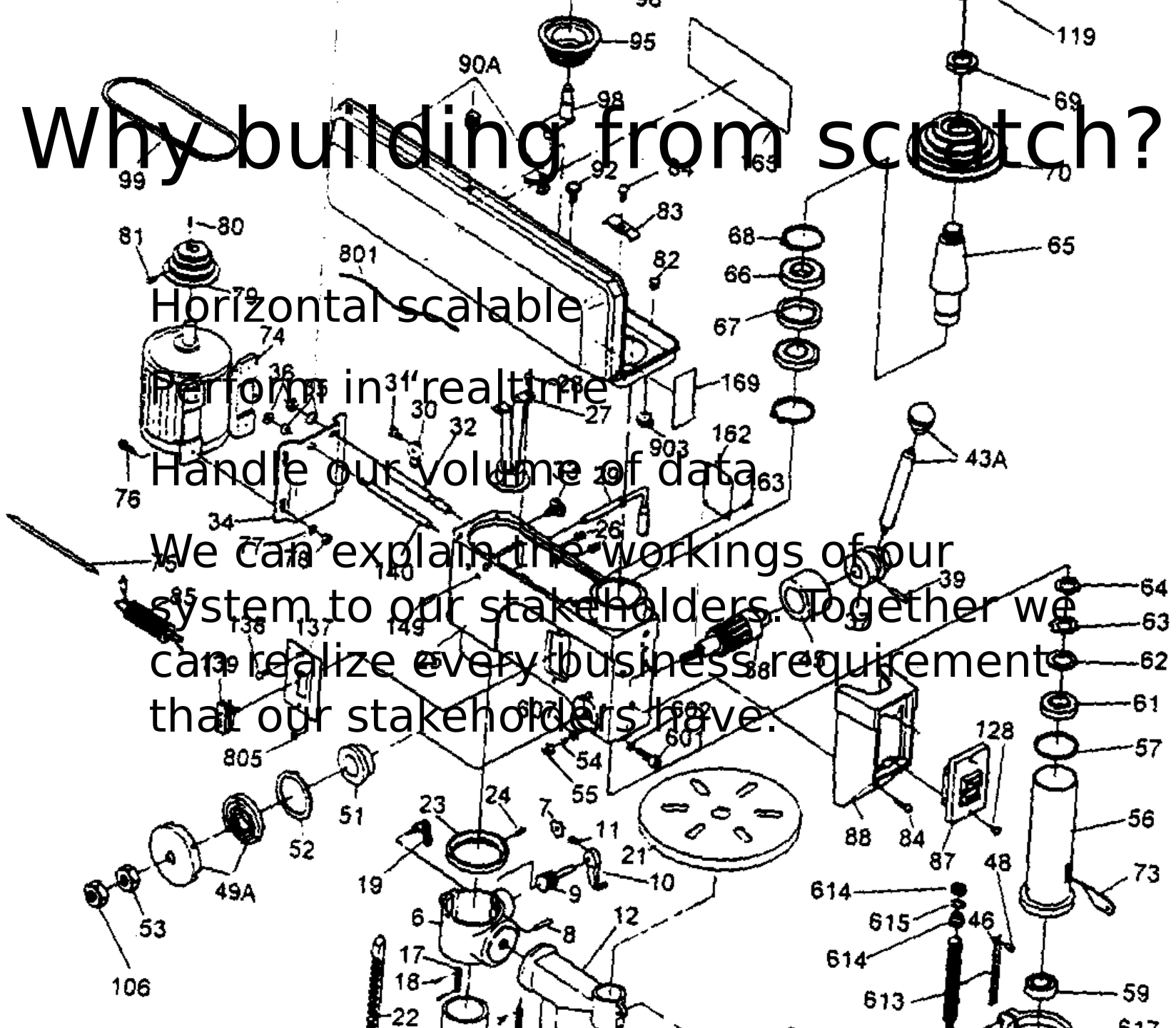
# Why building from scratch?

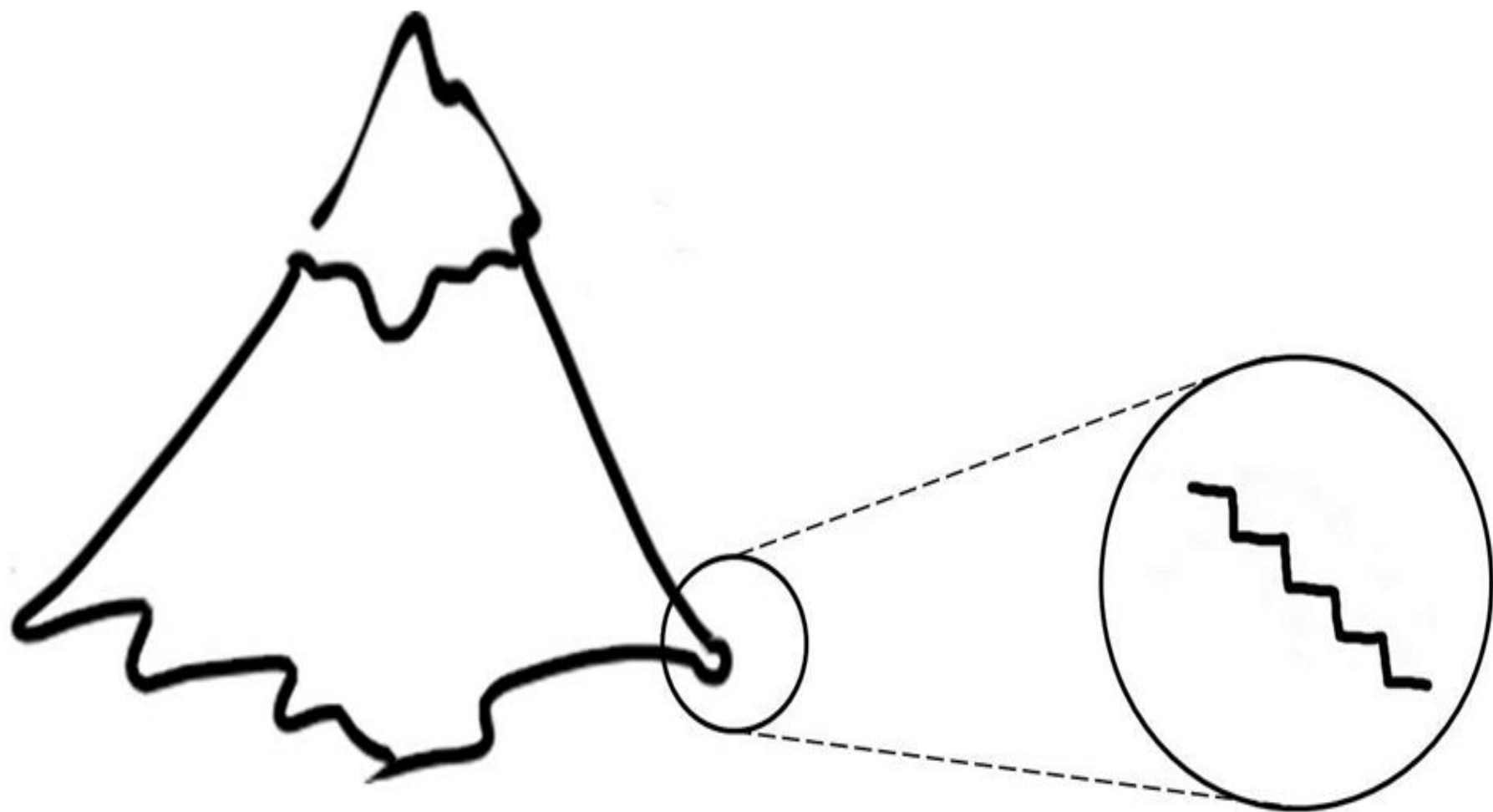
Horizontal scalable

Perform in real time

Handle our volume of data

We can explain the workings of our system to our stakeholders. Together we can realize every business requirement that our stakeholders have.



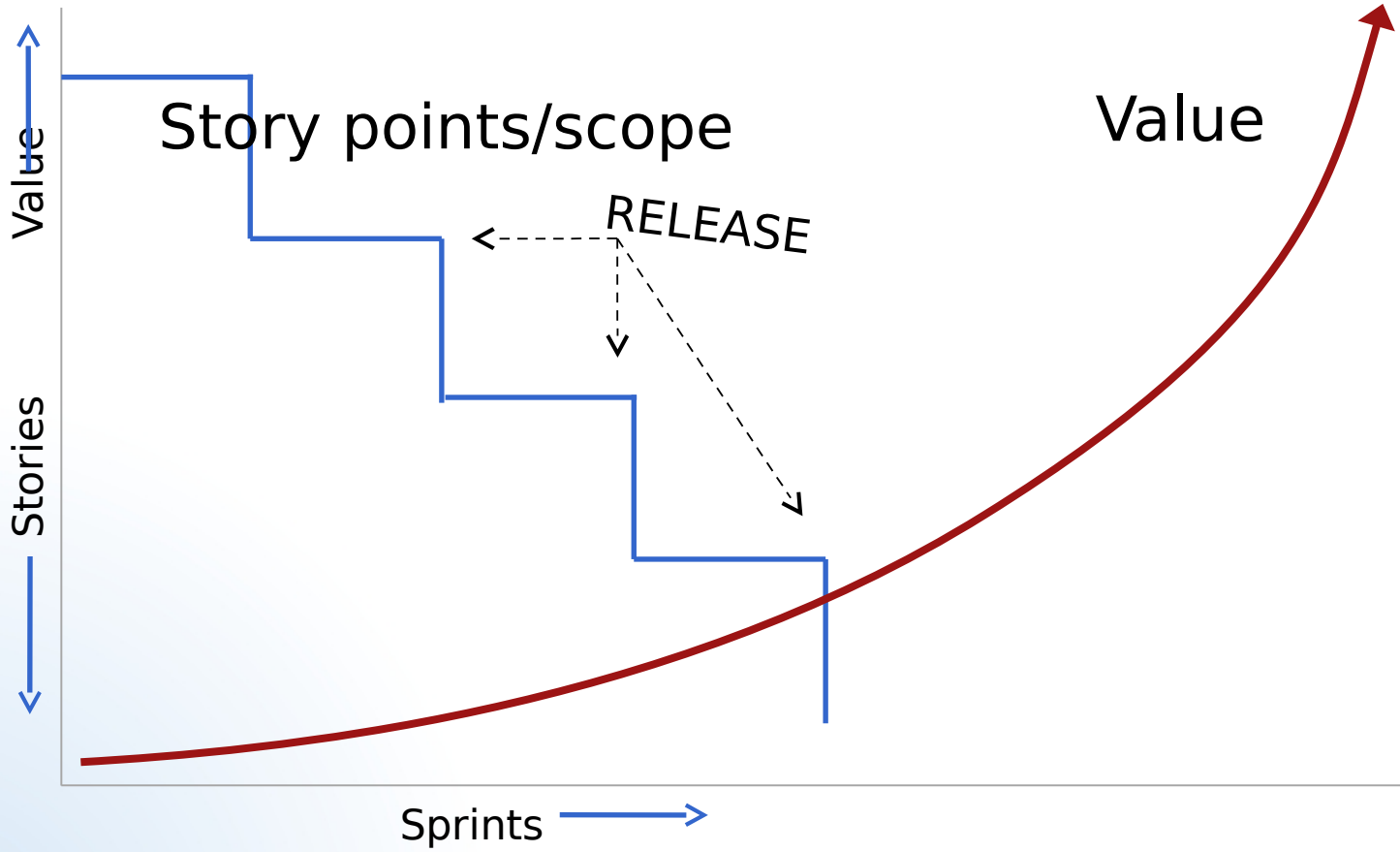


scary stuff :(



definitely do-able!



# As stories burn down, value increases





# Recommendations

Gratis verzending vanaf 20 euro & 30 dagen bedenktijd\*

Welkom B. | Mijn account | Bestelstatus | Klantenservice | Verlanglijstjes |  

**bol.com**

Cadeaubon Verkoppen Zakelijk bestellen Actiefolder Fotoservice

Alle Engelse boeken **ZOEKEN**



Home > **BOEKEN** MUZIEK, FILM & GAMES SPEELGOED BABY MOOI & GEZOND SPORT & VRUJE TIJD COMPUTER & ELEKTRONICA WONEN TUIN & KLUSSEN DIER AANBIEDINGEN

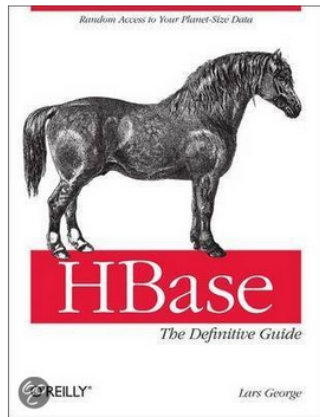
GRATIS RETOURNEREN 30 DAGEN BEDENKIJD DAG & NACHT BEREIKBAAR **Gratis verzending vanaf 20 euro\*** VEILIG BETALEN ACHTERAF BETALEN VANDAAG BESTELD MORGEN IN HUIS\*

Home > Boeken > Engelse boeken > ... > Data mining > HBase

## HBase

The Definitive Guide

Auteur: [Lars George](#) | ★★★★★ [Schrijf een review](#)



Bindwijze: **Paperback**

### Samenvatting

If your organization is looking for a storage solution to accommodate a virtually endless amount of data, this book will show you how Apache HBase can fulfill your needs. As the open source implementation of Google's BigTable architecture, HBase scales to billions of rows and millions of columns, while ensuring that write and read performance remain constant. HBase: The Definitive Guide provides the details you require, whether you simply want to evaluate this high-performance, non-relational database, or put it into practice right away. HBase's adoption rate is

Auteur: [Lars George](#)

- Engels
  - 554 pagina's
  - O'Reilly Media, Inc, USA
  - oktober 2011
- [Alle productspecificaties](#)

Favoriet  0  0  Like 0

**Bol.com** Nieuw vanaf € 28,08   
€ 33,99 Via andere verkopers (1)

**GRATIS VERZENDING!**

€ 37,99 **€ 33,99**

**Je bespaart 11%**

3-4 werkdagen

**+ In winkelwagentje**

Levering door **bol.com**

- Gratis verzending vanaf € 20,00
- 30 dagen bedenktijd
- Gratis retourneren

[Zet op verlanglijstje](#) | [Deel dit artikel met anderen](#)

Other people also looked at:

### Anderen bekeken ook:

 <b>Hadoop</b> Tom White € 41,99	 <b>Cassandra</b> Eben Hewitt € 48,99
 <b>Mahout in Action</b> Sean Owen € 35,99	 <b>HBase Administration Cookbook</b> Y Jiang € 39,99

[Bekijk de hele lijst](#)

**Gratis verzending**  **Achteraf betalen met acceptgiro** 

 **30 dagen bedenktijd**  **Veilig betalen**

 **Dag en nacht**

**bol.com**

# How can we tell the success of these recommendations?

- Visualize output data
- Offline recommender evaluation
  - Run algorithms that express the numerical errors of the recommender output.

## Root Mean Squared Error (RMSE)

The square root of the mean/average of the square of all of the error.

The use of RMSE is very common and it makes an excellent general purpose error metric for numerical predictions.

Compared to the similar Mean Absolute Error, RMSE amplifies and severely punishes large errors.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$





# How can we tell the success of these recommendations?

- Is that all we can do?



# The gasoline recommender system





# The gasoline recommender system

## Requirements:

- Support anonymous customers
- Realtime: response 100ms
- Low numerical error value
- Recommendation is successful if customer buys the item after seeing it first as a recommendation.



# The gasoline recommender system

- Customer: new customer
- Cartype: Tow Truck
- Recommendation: **Diesel**
- Does customer buy item after seeing this as recommendation? : yes





# The gasoline recommender system

- Customer: new customer
- Cartype: sportscar
- Recommendation: **Petrol (gasoline)**
- Does customer buy item after seeing this as recommendation? : yes





# The gasoline recommender system

- Offline evaluation:
  - Visual data inspection ✓
  - Low error value ✓
- Online evaluation: ✓
  - Accurately predict the recommendations ✓
  - Almost everything we recommend gets bought

# The gasoline recommender system

Is this a good recommender system?



# What is a good recommendation?

## Common error metrics

The accuracy to predict

## Commerce

Add value to customer experience





# How can we tell the success of this recommender?

- Visualize output data
- Offline recommender evaluation
  - Run algorithms that express the error of the recommender output in a numerical form.
- Long Term analyses
  - Analyze the performance of the recommender over time.
  - Do visitors return after buying a recommendation?
  - Analyze behavior to learn more from our visitors.
- Online recommender evaluation
  - Live user experiment 's



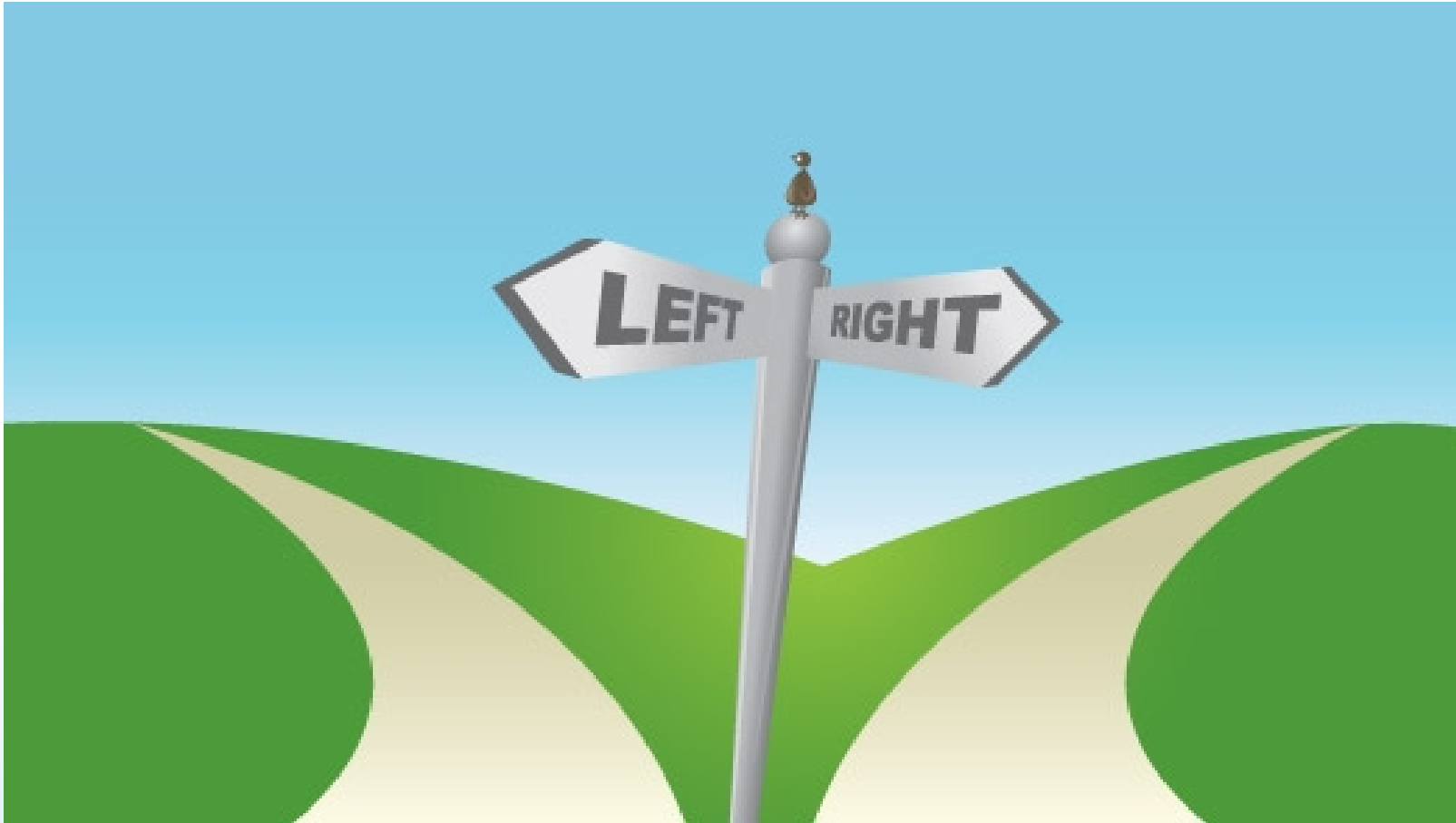
# Live user experiments?

Recommendations = Data-driven decisions

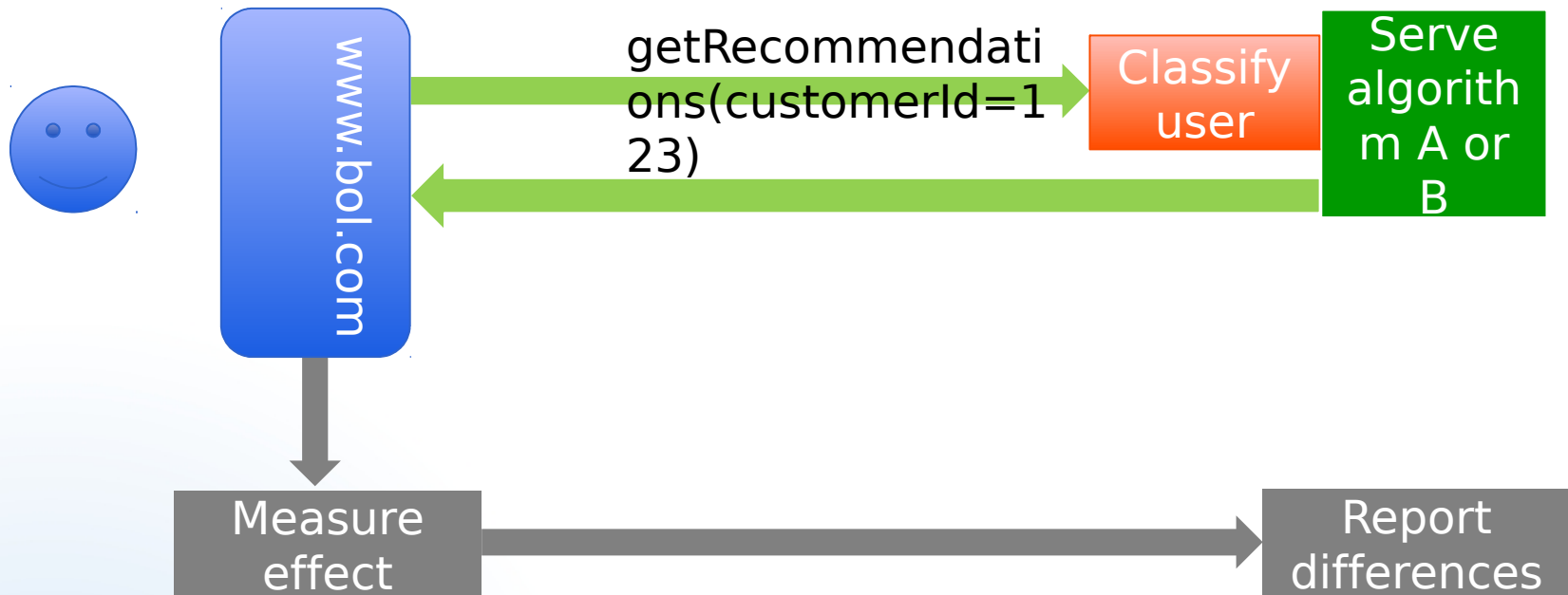
- Optimize combinations of recommendation algorithms
- User Interface
- User Interaction Flows



# Live experiments: What do visitors really prefer?



# Live experiments: our experimentation framework



# Experimentation results



# Impact of presenting the recommendations to the customer



Ø 20 cm  
Ø 18 cm  
Ø 24 cm  
Ø 14 cm  
Ø 16 cm

69,99  
**39,99**

Imperial  
houseware

69,99  
**39,99**

Imperial  
houseware

Scherp geprijsde 5-delige  
**Imperial Kitchen pannenset**

★★★★★

Volgens klanten een mooie pannenset die goed te reinigen is, heel gemakkelijk in het gebruik en met een prima prijs-kwaliteitverhouding.

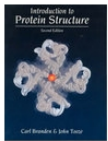
> Bekijk nu

# Impact of presenting the recommendations to the customer

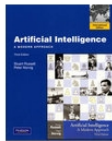
## Recommendation interaction

Especially selected for you

because you like 'data mining'



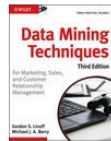
**Introduction to Protein Structure**  
Carl-Ivar Branden  
€-63,99 **€ 61,99**



**Artificial Intelligence**  
Stuart Russell  
**€ 60,03**



**Introductory Statistics with R**  
Peter Dalgaard  
€-53,99 **€ 48,99**



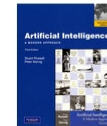
**Data Mining Techniques**  
Third Edition  
Michael J. Berry  
★★★★★ (1)  
€-40,99 **€ 38,99**

Especially selected for you

because you like 'data mining'



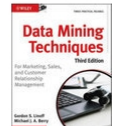
**Introduction to Protein Structure**  
Carl-Ivar Branden  
€-63,99 **€ 61,99**



**Artificial Intelligence**  
Stuart Russell  
**€ 60,03**



**Introductory Statistics with R**  
Peter Dalgaard  
€-53,99 **€ 48,99**



**Data Mining Techniques**  
Third Edition  
Michael J. Berry  
★★★★★ (1)  
€-40,99 **€ 38,99**



General personal reco's

# Recommendations

Recommend a deal

bol.com Cadeaubon Verkoop Zakelijk bestellen Actiefolder Fotoservice

Alle artikelen ZOEKEN

BOEKEN MUZIEK, FILM & GAMES SPEELGOED BABY MOOI & GEZOND SPORT & VRIJE TIJD COMPUTER & ELEKTRONICA WONEN TUIN & KLUSSEN DIER AANBIEDINGEN

GRATIS RETOURNEREN 30 DAGEN BEDENKTUJD DAG & NACHT BEREIKBAAR Gratis verzending vanaf 20 euro\* VEILIG BETALEN ACHTERAF BETALEN VANDAAG BESTELD MORGEN IN HUIS

NOG **5** DAGEN

**BULK** 10-DAAGSE tot 60% korting

GRIJP MIJN KANS

Naar de BulkDeals

tot 30% korting

In de luiers vanaf 15 cent per luier

2+1 gratis

Pick & Mix op Dove, Toni & Guy, Rexona, Vaseline en Neutral

40% korting

Rambo Tuin Beits op alle 5-literemmers

Recommend wishlist items

Op je verlanglijstje

Ook interessant?

Tegenstellingen, mijn eerste gruffalo € 8,95

de Gruffalo De Gruffalo € 13,95

Het Kind Van De Gruffalo € 8,95

DRAAK DRIES Draak Dries € 13,95

Recommend price change

Tip voor jou!

Salvador Dali, 1904-1989 Robert Descharnes Nederlands € 14,99

Bekijk het product

Recommend search queries

Recommend a category to the visitor

Speciaal voor jou geselecteerd:

R Graphics Cookbook Winston Chang € 37,99 € 33,99

The Art of R Programming Norman Matloff ★★★★★ (1) € 33,99

Python for Data Analysis Wes McKinney € 29,99

Discovering Statistics Using R John Fox, John Fox € 59,99 € 53,99

DonderDagdeal!

JE HEBT NOG 8 UUR, 11 MINUTEN, 27 SECONDEN

PLANES

9,99

Planes (Disney) Vlieg mee met de nieuwe



# What did we just recommend?

- Algorithms with outcome
- Personalized content
  - Products
  - Authors
  - Artists
  - Deals
  - Categories
- Personalized User Interface
- Determine the priority of algorithm outcome on a page
- Over channels (webshop, mobile, email)



# Level of personalization?

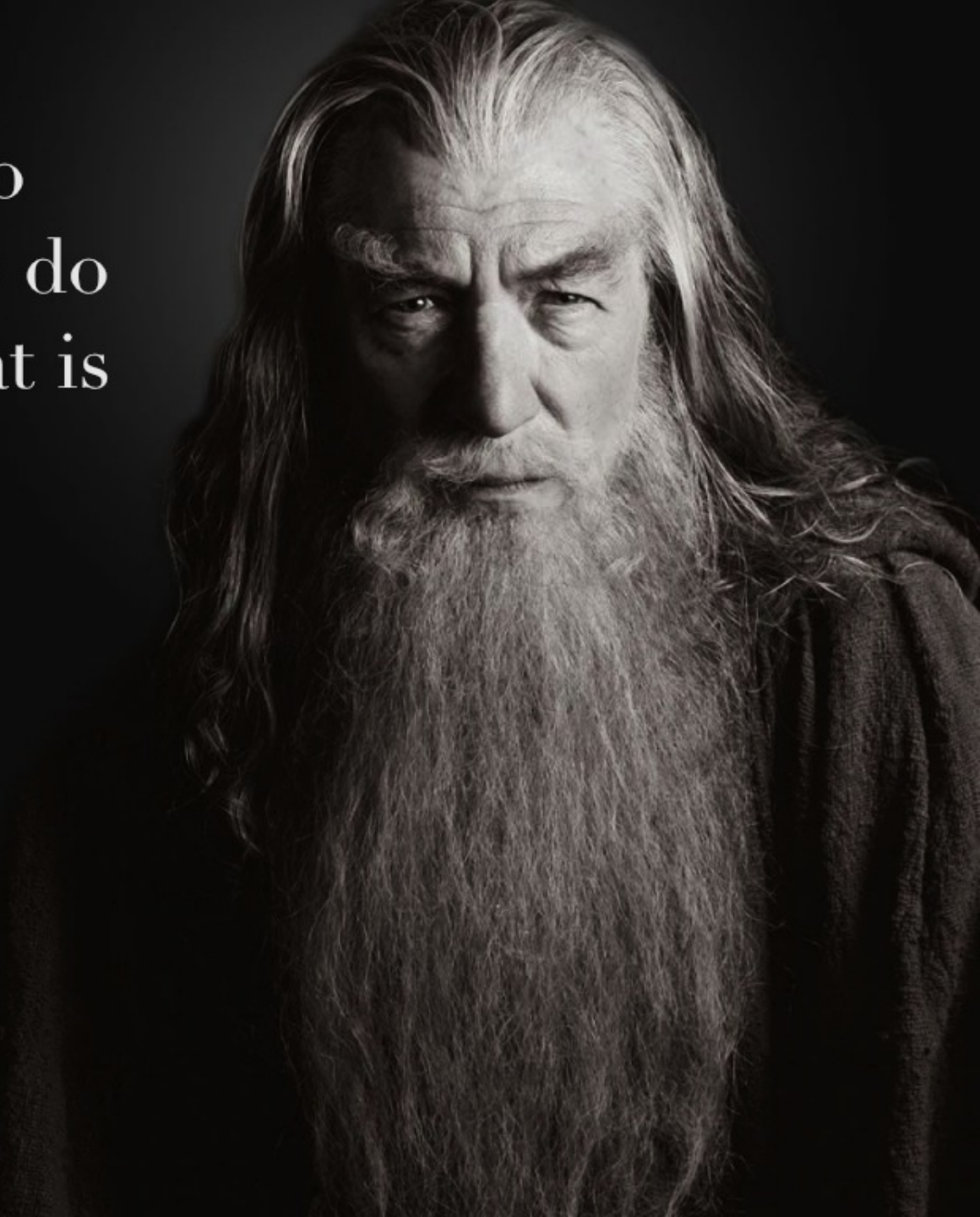


# How does the real time behavior of the visitor constrain the recommender system?

- What would add value to the customer given his current context?
- What do we already know about the visitor? (previous behavior)
- Determine the level of personalization



"All we have to  
decide is what to do  
with the time that is  
given to us."



# What to do with given time

- Cache 'expensive' data
- Tune/(try different) jdbc connectivity components
- Profile your code (Java VisualVM)



**WE'RE  
HIRING!**

Contact me **[bkersbergen@bol.com](mailto:bkersbergen@bol.com)**





The end

