Lucerne University of
Applied Sciences and Arts

HOCHSCHULE
LUZERN

Technik & Architektur

Recommender Systems

Mindblast May 2014

Prof. Dr. Marc Pouly

Office D303
marc.pouly@hslu.ch

Agenda

1. Definition and Purpose of a Recommender System
2. Recommender Systems Paradigms
3. Content-based Recommender Systems
4. Collaborative Recommender Systems
5. Hybrid Recommender Systems
6. Case Study: Amazon’s Recommender System
7. Demographic Recommender Systems
8. What a Proposal on Recommender Systems must not lack

Definition of a Recommender System

*A recommender system is any system that produces **individualized** recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options*

(Burke, 2002)

*A recommender system is a software capable of suggesting interesting things to its users after learning their **preferences over time***

(Jennach & al, 2010)

*Let C be the set of all users and let S be the set of all items. Let u be a utility function that measures the usefulness of item s to user c, i.e. $u : C \times S \rightarrow R$, where R is a totally ordered set. Then, for each user c we want to choose an item s that **maximizes** the user's **utility***

(Adomavicius & Tuzhilin, 2005)

Diametric Purposes of Recommender Systems

1. Customer Perspective:

   Help users cope with the information overload problem

2. Business Perspective:

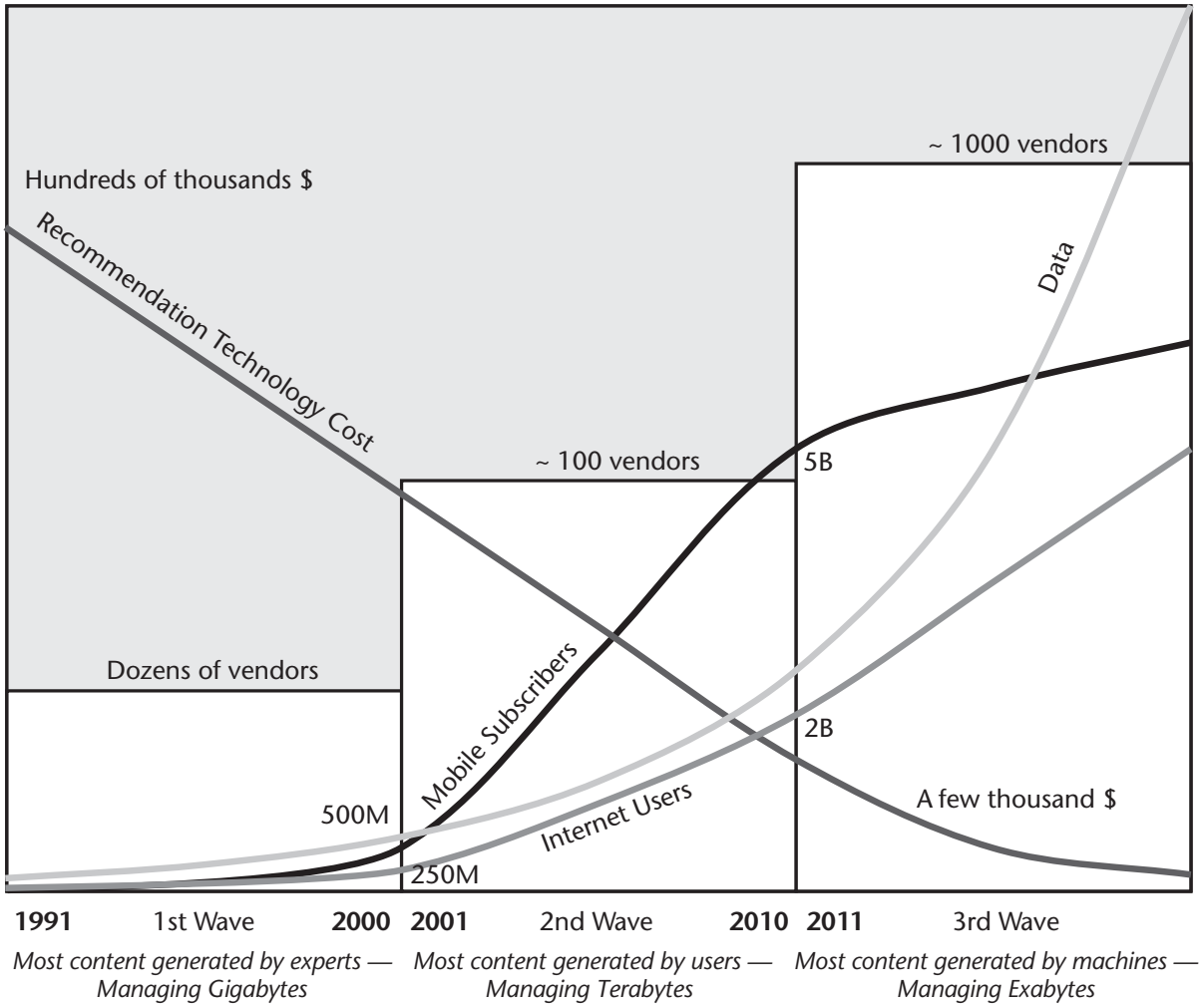
   Provide companies with a effective way to drive more sales or increase the level of engagement of the services they offer.

   Promises of recommendation engine vendors:

   - 300% revenue increase
   - 150% higher conversion rate
   - 50% higher average order value

Survey from 2010 consulted Avail Intelligence, Baynote, Certona, Adobe, RichRelevance (Aldrich, 2011)

Internet vs. Recommendation Systems Trends



Recommender System Paradigms

1. Content-Based Recommender Systems

   - Recommend items similar to those a user has liked in the past

   - Information structure: <u>item-to-item</u> association

2. Collaborative Recommender Systems

   - Identify users whose preferences are similar to those of the given user and recommend items they have liked.

   - Information structure: <u>user-to-user</u> and <u>user-to-item</u> association

3. Hybrid Recommender Systems

   - Combine advantages and reduce drawbacks of both worlds

Input for Content-Based Recommendations

Start from attributed data (relational structure) such as

ID	Name	Cuisine	Service	Cost
10001	Mike's Pizza	Italian	Counter	Low
10002	Chris's Cafe	French	Table	Medium
10003	Jacques Bistro	French	Table	High

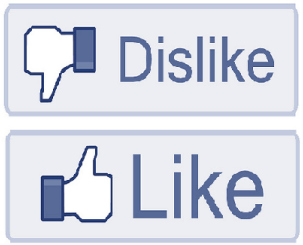
or from a Vector Space Model such as

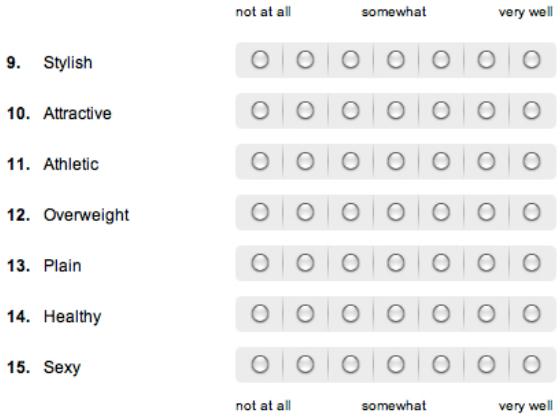
	Leisure	**Sport**	**Culture**
Guided City Tour	0.3	0.6	0.8
Early Bird Meal	0.8	0.1	0.1
Museum	0.6	0	1.0

User Profile for Content-Based Recommendations

A content-based recommender learns a user profile

- Users communicate explicit preferences
- Users rate product attribute values
- Systems create profiles from extracted (implicit) preferences





Weighted Similarity Approach

- User profiles are interpreted as idealized, virtual product specific to the current user

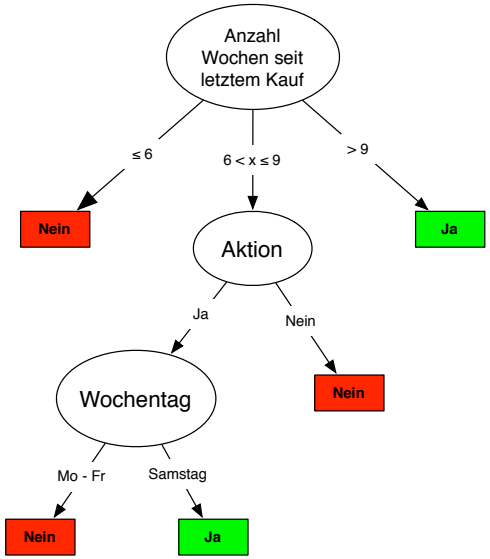
- We compare all products in the catalogue with the profile using a weighted similarity metric (e.g. cosine similarity for vector space model) and create a ranking

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

- In a relevance feedback loop user feedback on presented recommendations is used to enrich / improve user profiles

Other Content-Based Recommendation Approaches

- Other approaches exist for content-based recommendations such as decision trees or Bayesian classifiers
- But scalability is worse than with weighted similarity



MIGROS
Olten - Hammer

Menge	Rabatt		
1.000	0.00	ANDROS ORANGENSAFT 1L	4.95
1.000	0.00	CLEVERBAG 17L ROLLE	2.25
1.000	0.00	GURKEN TP ST S.	1.30
1.000	1.50	TR.WEISS KERNLOS AB 500G	2.10
1.000	0.75	ERDBEEREN AB 250G	1.95
1.000	0.00	SE OSTEREIER CH FL 50+4ER	3.25
1.000	0.00	DATTELTOMATEN AB 250G	2.50
1.000	0.80	THAIK KOKOS-MILCH 500ML	3.20
0.321	2.25	OP ADR P-SCHNITZEL 2STK	8.35
1.000	0.00	SOFT TORTILLAS 326G	4.80
2.000	0.00	HEI M-DRINK HOCHPAST 1L	3.30
0.179	0.00	BANANEN I LOSE	0.50
0.162	0.00	TS BRATSPECK MIDI	3.40
1.000	0.00	BUNDZWIEBEL BD S.	1.40
1.000	0.00	AB ENDIVIENSALAT 200G	2.40
1.000	0.00	BIO SULTANINEN 400G	2.80
0.156	0.00	WILDBACHKAESE FP 150G	3.75
1.000	0.00	TOMME LA FLEUR	5.50
0.214	0.00	APF.JAZZ I LOSE	1.05
1.000	0.00	EIER CH FL 53G+ 6ER	3.60
1.000	0.00	CUM +5x Pkt. ges Sort	0.00
2.000	0.00	GOUDA IN SCHEIBEN 150G	5.20
1.000	0.40	BIO JOGHURT HIMBEER 500G	1.50
0.096	0.00	SE VALENTINS-PLATTE MAXI	4.75
1.000	0.00	TS BUTTERZOPF 700G	4.50
1.000	0.50	MANGO ST S.	1.70
1.000	0.00	VALFLORA BUTTER 100G	1.40

Advantages of Content-Based Recommendation

1. User Independence:
   A user's profile is built only from her own ratings

2. Transparency:
   Computed recommendations can be explained (to clients or users) by unveiling the relevant attribute values

3. New items:
   Content-based recommender systems can recommend new items (not yet rated by any user) right from the beginning

Disadvantages of Content-Based Recommendation

1. Limited Content Analysis:
   Recommendations are computed based on attribute values, but the attributes do not necessarily reflect how users decide.

2. Overspecialization:
   Content-based recommender systems will never find anything unexpected or novel. Recommended products must score highly against the profile, and the latter has been created from the user's own shopping history.

3. New User:
   Lots of ratings must be collected before we can produce decent recommendations for a new user. In other words, the profile has to be learned first.

Collaborative Recommendations

- Recommendations are obtained from opinions of other users
  → computerized process of word to mouth

- **Hypothesis:** A set of users, who liked the same items in the past, probably share the same preferences

- Thus, picking a user from this set, we can suggest her all the unseen items the other members of this set liked in the past

- Profiles are represented as a user-item matrix where each cell (u,i) corresponds to a rating of user u for item i

- An algorithm identifies for each user the set of nearest neighbours

- Products are recommended that many nearest neighbours liked but the current user has not yet rated or purchased

Advantages of Collaborative Recommendation

1. Cross-Category Recommendations:
   Products from different categories can be recommended

Disadvantages of Collaborative Recommendation

1. Many people must participate:
   No nearest neighbours are found in too sparse matrices

2. New User Problem:
   System must first learn the preference of a user from her ratings

3. New Item Problem:
   Many users must rate new items before they can be recommended

4. Grey Sheep Problem:
   Users with unusual taste do not get accurate recommendations

5. Scalability Problem:
   Collaborative Recommendation Engines need a lot of resources

Hybrid Recommender Systems

Simple strategies for combining 2 or more recommender algorithms:

1. Weighted Approach
   Each recommendation method computes a score for every item.
   Combination of scores leads to final ranking.

2. Mixed Approach
   Results of different methods are presented to the user.

3. Cascade Approach
   One recommender methods refines the result of another method.

4. Switching Approach
   Online shops may use different methods at different places

A more sophisticated example is Amazon's recommender system …

The Amazon Recommender Algorithm 1

- The hybrid recommender algorithm used by Amazon is called item-to-item collaborative filtering

- This algorithm exploits that some items are often purchased together – it builds a similar-items table as follows

```
For each item I1 in product catalogue
  For each user U who purchased I1
    For each item I2 purchased by user U
      Record that a user purchased I1 and I2
    end
  end
  For each item I2
    Compute the similarity between I1 and I2
  end
end
```

The Amazon Recommender Algorithm 2

- The cosine similarity measure is used

- This offline computation of the similar-items table is extremely time-intensive with $O(N^2M)$ as worst case for N items and M users

- However, because most users only have a few purchases, complexity is in practice closer to $O(NM)$

- Based on the similar-items table, the algorithm finds items similar to each of the user's purchases and ratings, aggregates those items, and recommends the most popular or correlated.

- This online computation is very quick, depending only on the number of items the user purchased or rated

How to Counteract the Cold Start Problem

- Content-based recommendation systems suffer from the new user problem
- Collaborative recommender systems suffer from both the new user and the new item problem
- Not surprisingly, the new item problem can successfully be addressed with hybridization, but not the new user problem
- The new user problem also occurs with anonym website visitors
- Counteracting the new user problem is possible if
  - users are willing to state their preferences explicitly
  - we have access to additional (e.g. demographic) data sources

Demographic Recommender Systems

- Aim to categorize users starting from personal attributes

- Thanks to Facebook, Twitter, address dealers, website crawlers, etc. demographic data sources are abundant

- From such data sources demographic classes are built using machine learning techniques (see clustering)

- For a new user without rating or shopping history we determine her demographic class and provide recommendations of products that other people in this class liked.

Spiegel Online 14.05.14

SPIEGEL ONLINE WIRTSCHAFT

Login | Registrierung

Suche Kurse

Politik Wirtschaft Panorama Sport Kultur Netzwelt Wissenschaft Gesundheit einestages Karriere Uni Schule Reise Auto

Nachrichten > Wirtschaft > Unternehmen & Märkte > Twitter > #amazoncart: Twitter und Amazon kooperieren bei Einkaufen im Internet

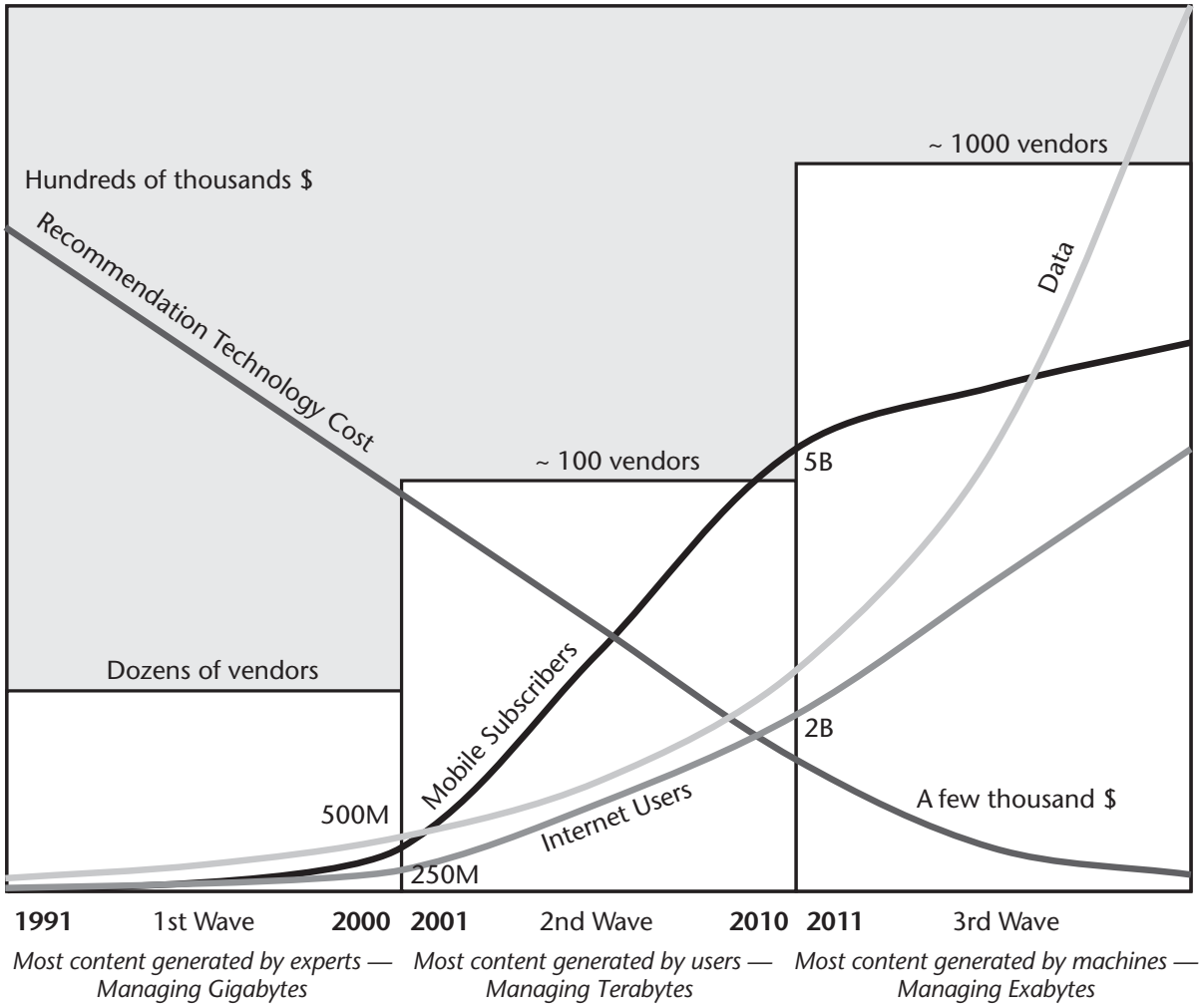
Geldanlage

Neue Vermarktungsstrategie: Amazon will Twitter-Nutzer ködern

Ein Tweet - und schon liegt der Einkauf im Warenkorb bei Amazon: So wollen der Internethändler und Twitter künftig beim Online-Shoppen zusammenarbeiten.

We are living in a Small World

Three Waves of Recommender Technology



Some Lessons Learned during the 1st Wave

1. Last Mile Problem: Integration is much harder than it looks

2. In-house competition fiercer than market competition
   - Recommendation as battle between IT and marketing
   - Return-on-invest not immediately measurable

3. Do not solve everything online !

4. Customer taste changes over time

5. Online shopping must be efficient → minimize explicit interactions

6. Cold Start Problem is a serious issue !

Some Lessons Learned during the 2nd Wave

1. Hybridization solves many issues and wins contests too
2. Cold start can be overcome using social web data
3. SaaS reduces the burden of integration …

   … but at the price of loosing control over the collected data
4. User experience is often more important than algorithms
5. Integration with CRM is a pending subject

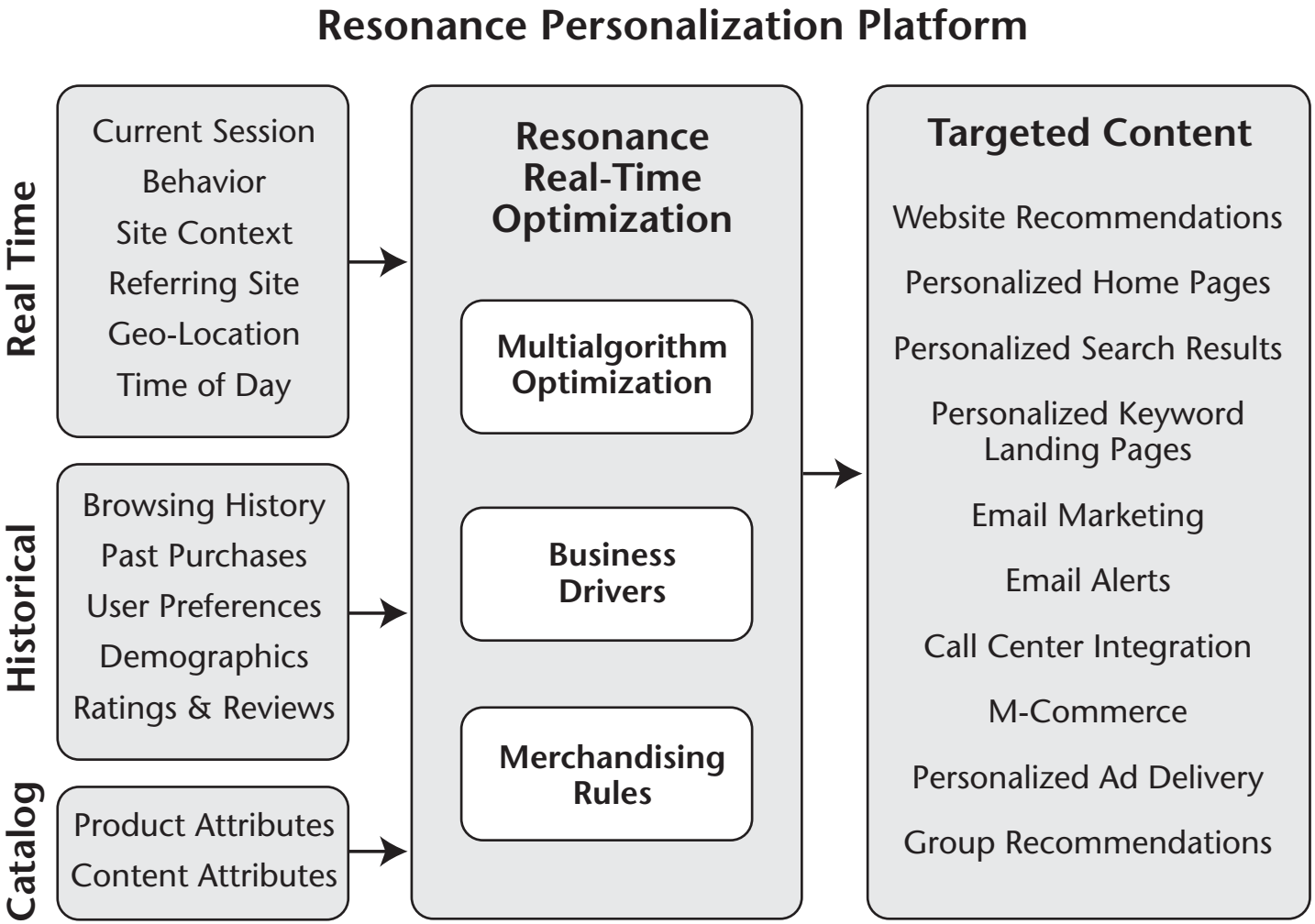
You can buy hundreds of products from Amazon and Apple getting better and better recommendations, but unbelievably their customer care services do not know if you are a good or bad customer.

(Martin et al, 2011)

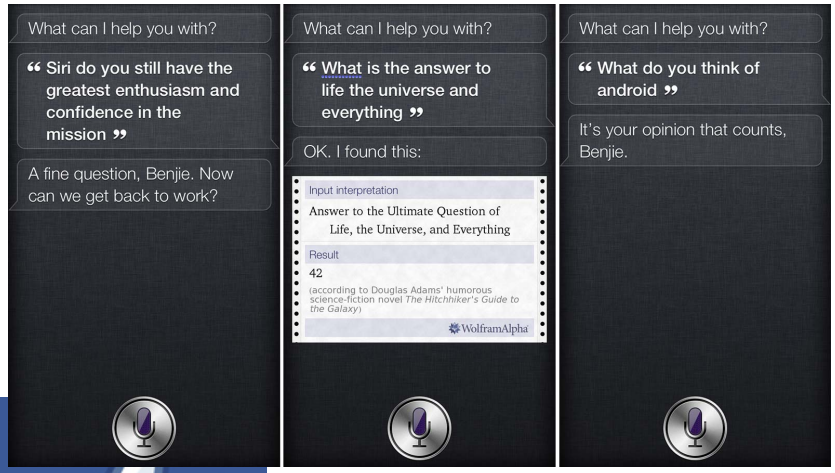
Catalysts for the 3rd Wave

- One in every 5 people in the world owns a smartphone; one in every 17 people owns a tablet (Dec. 2013)
- Mobile Commerce brings new levels of context awareness
  - Geo location, NFC tags, eye tracking, gesture detection, skin tension measurement
- New trends (e.g. live shopping) and services
  - For example, work out recommenders (e.g. Nike+), smart TV, ...
- Cloud and big data technology available:
  - Processing (e.g. Amazon EC2) and storing costs are decreasing rapidly
  - New computing paradigms (e.g. map reduce)
  - New libraries and frameworks (e.g. Apache Mahout, Google Prediction)

Generic Picture of a Commercial Recommender System



Genealogy of Recommender Systems



References

- Recommender Systems: An Overview
  Robin Burke, Alexander Felfernig, Mehmet H. Göker
- The Big Promise of Recommender Systems
  Francisco J. Martin, Justin Donaldson, Adam Ashenfelter, Marc Torrens, Rick Hangartner
- Recommender Systems in Commercial Use
  Susan E. Aldrich
- Comparative Rating of 5 Recommendations Solutions
  Susan E. Aldrich
- Preference Learning
  Johannes Fürnkranz and Eyke Hüllermeier
- Content-Based Recommendation Systems
  Michael J. Pazzani and Daniel Billsus
- Feature Weighting in Content Based Recommendation System Using Social Network Analysis
  Souvik Debnath and Niloy Ganguly and Pabitra Mitra
- Amazon.com Recommendations: Item-to-Item Collaborative Filtering, Industry Report
  Greg Linden and Brent Smith and Jeremy York
- A Framework for Collaborative, Content-Based and Demographic Filtering
  Michael Pazzani
- BI Intelligence Study
  https://intelligence.businessinsider.com