

# Personalisierung & User Profiling



WI/IM Seminar SS03

Gerald Haider

Christoph Radl

Josef Schachinger

# Einführung

---

- Motivation, Gründe
- Begriffsklärungen
- Methoden, (statistische) Hintergründe
- Beispielanwendungen
- Der Einzelne und die Community
- Mögliche Probleme
- Zusammenfassung

# Motivation für Personalisierung

---

## **Informationsflut:**

- In Suchmaschinen 3 Mrd. Seiten verfügbar
- Ansteigen der Newsgroup Postings
- Email Verkehr nimmt ständig zu

## **Anonymität:**

- User kommen sich verlassen vor
- Nicht jeder hat Mainstream Präferenzen
- Persönliche Note fehlt

# Was kann Personalisierung leisten?

---

- Herausfiltern von benutzerrelevantem Inhalt
- Maßgeschneiderter Zugang zu Portalen, Programmen usw.
- Zeitersparnis für den Benutzer
- Verwaltung von mehreren Benutzern

# Arten von Personalisierungssystemen

---

1. Customizing access to information services such as articles in newspapers or products
2. Filtering systems: - For email messages or newsgroup postings
3. Recommendation systems and services for assisting the browsing process
4. Tutoring systems
5. Search engines & systems

# Vorteile der Personalisierung & UP

---

## **Für den Benutzer:**

Wie bereits erwähnt, Eindämmung der Informationsflut, persönliche Note, usw.

## **Für den Information-Engineer:**

- ❑ Genauere Kenntnis des Kunden und seines Verhaltens.
- ❑ Gezieltere Angebote, bzw. Marketing ist möglich.
- ❑ Aufbau eines Vertrauensverhältnisses zum Kunden

# Methoden und Hintergründe

---

- Kollaboratives Filtern
- Content-based Filtern
- Kombinationen

# Kollaboratives Filtern

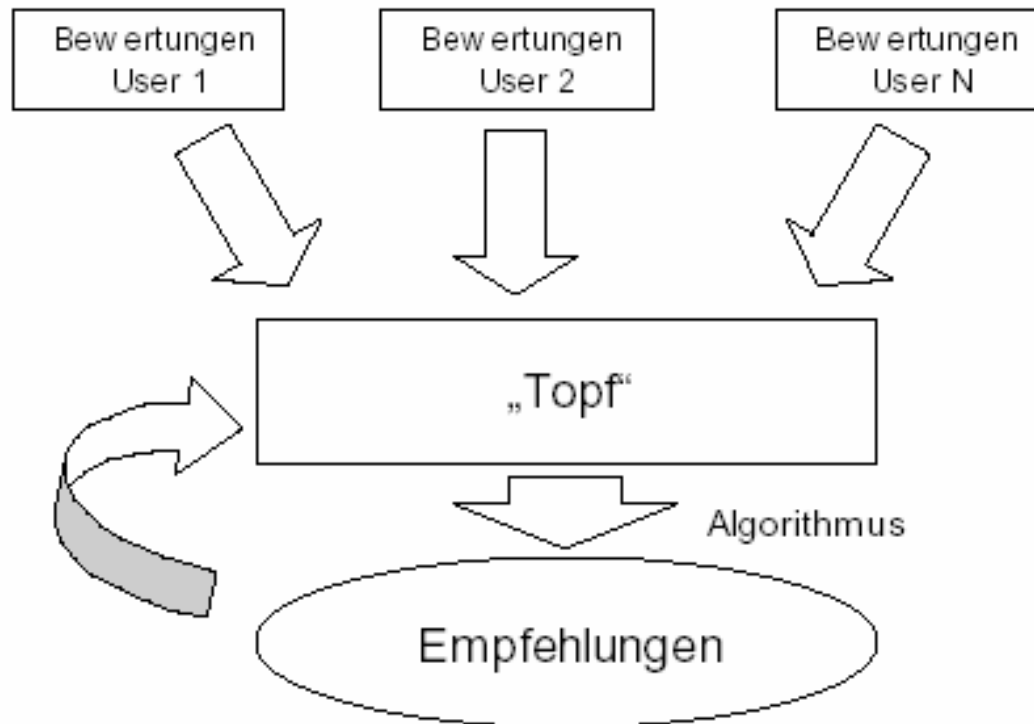
---

- Vergleichbar mit Mundpropaganda
- Information wird für eine große Gruppe von Benutzern gefiltert
- z.B. Regionalisierung
  
- Filtern von Bildern und Grafiken möglich

# Kollaboratives Filtern

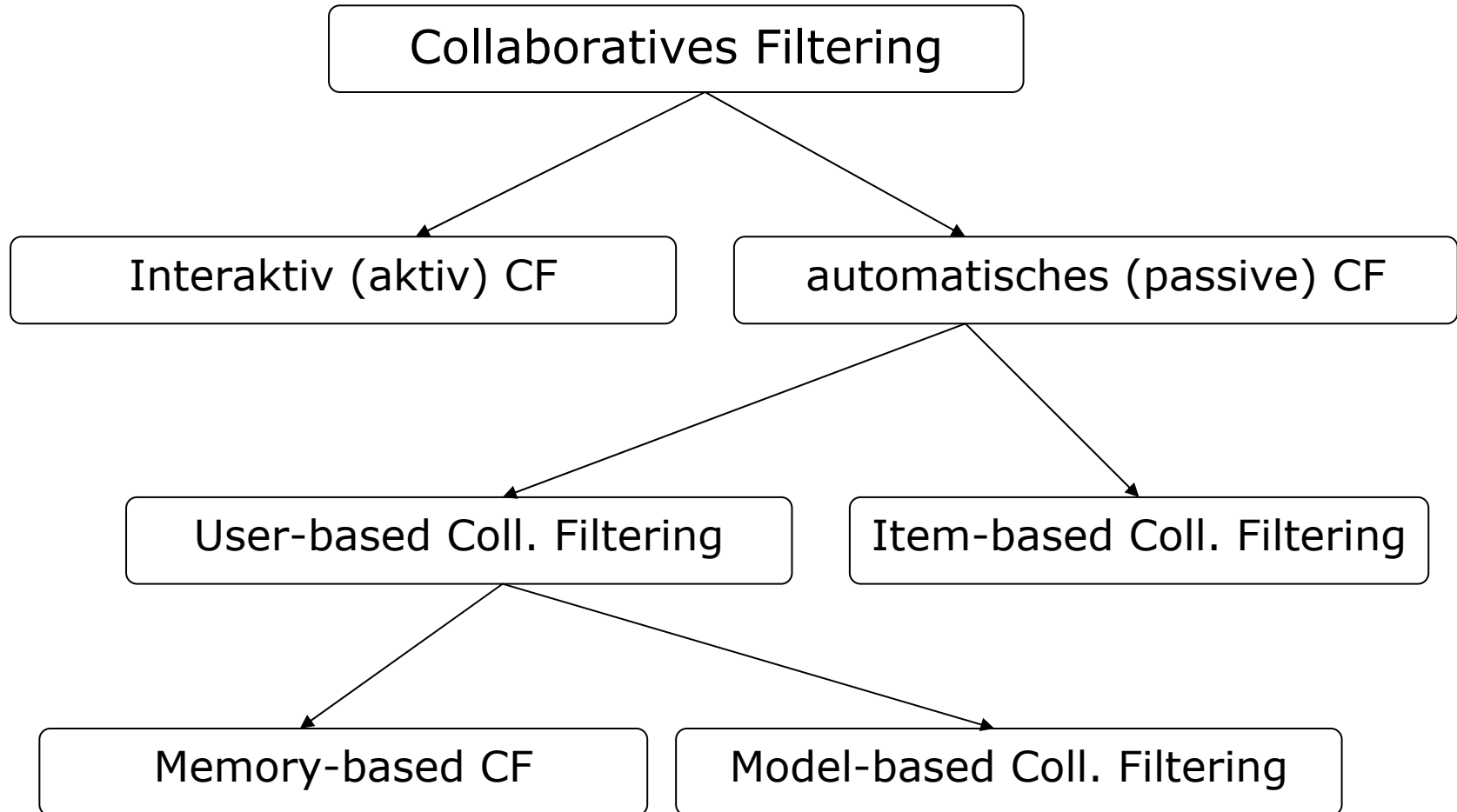
---

## □ Schema



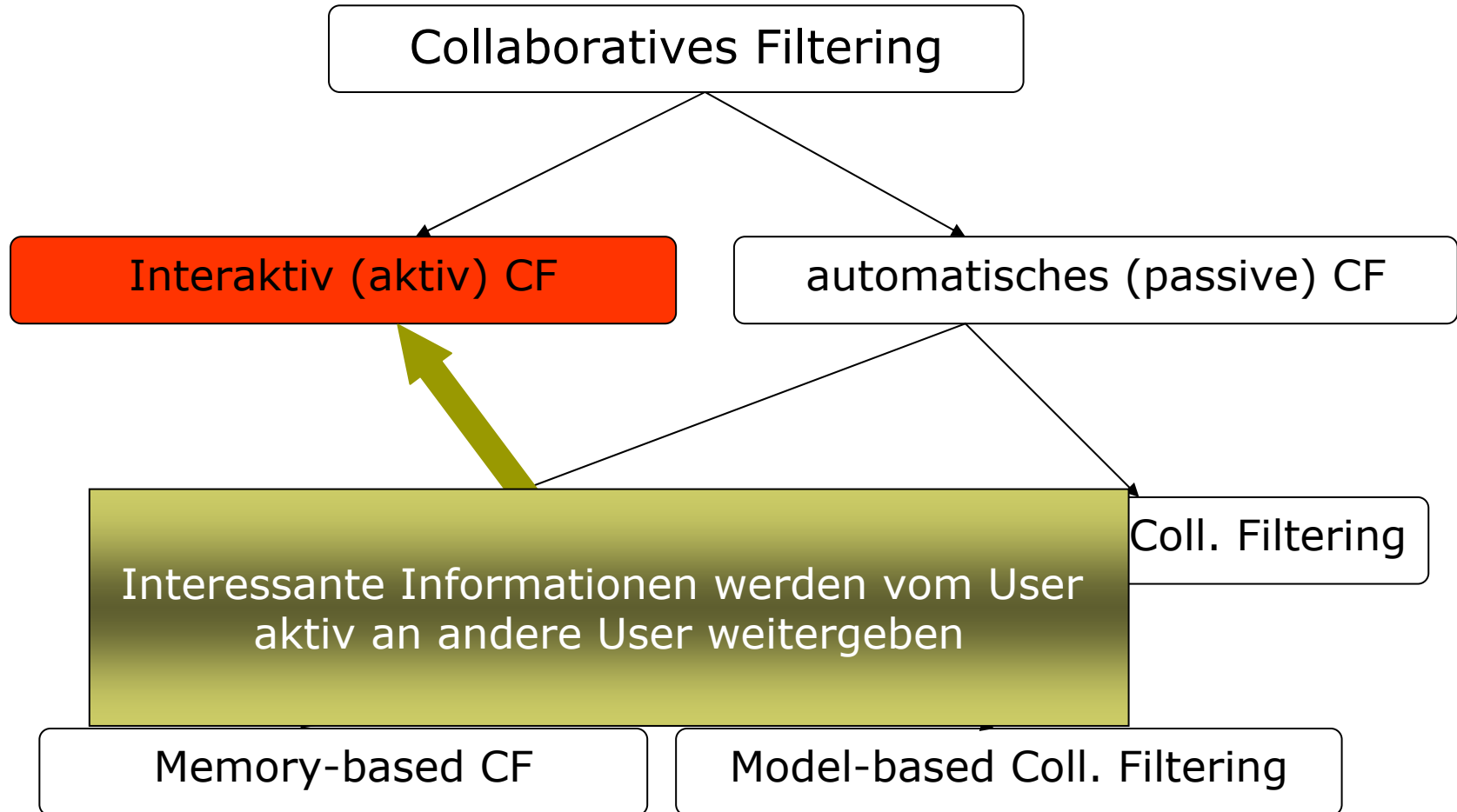
# Kollaboratives Filtern - Details

---



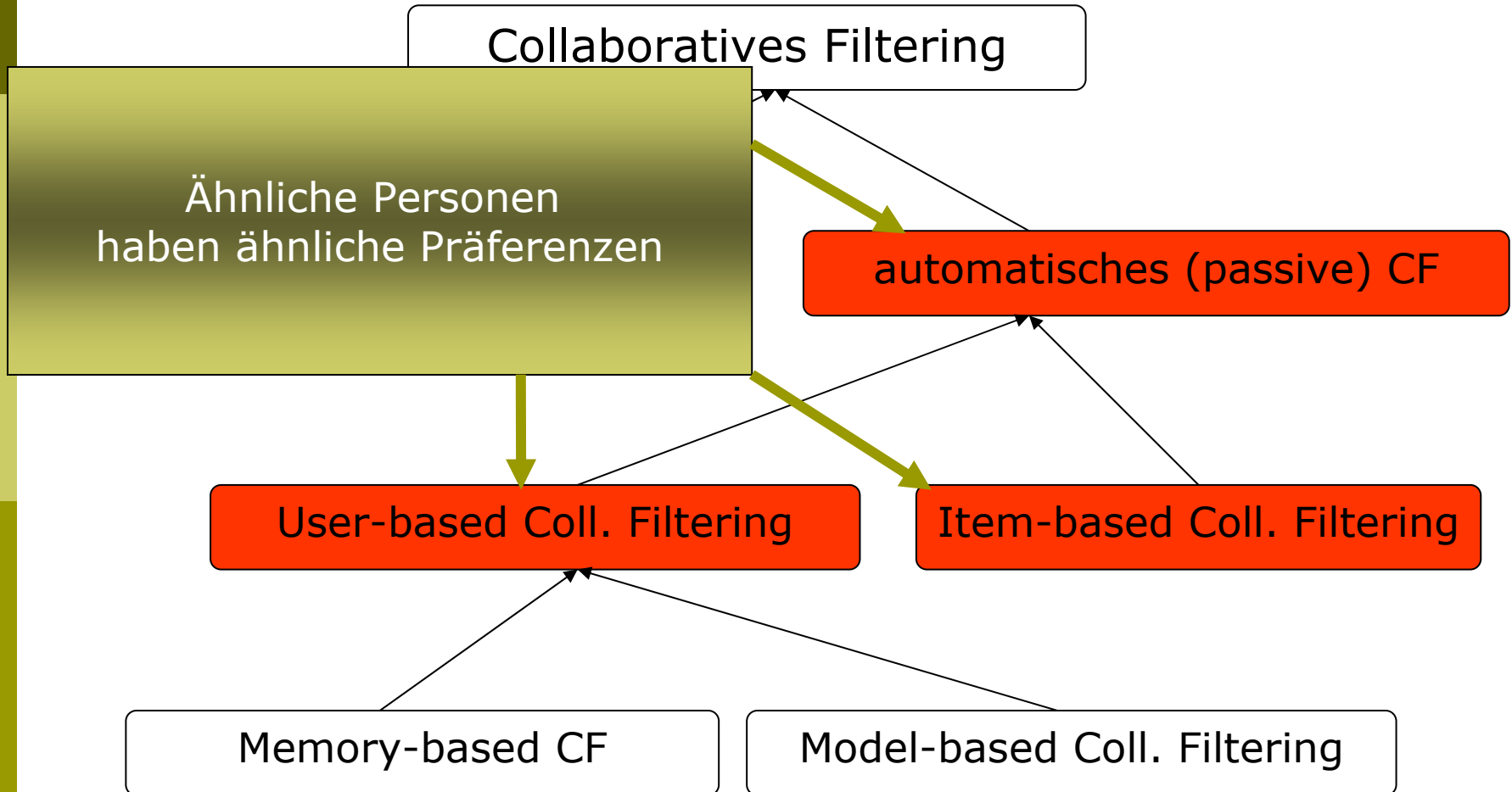
# Kollaboratives Filtern - Details

---



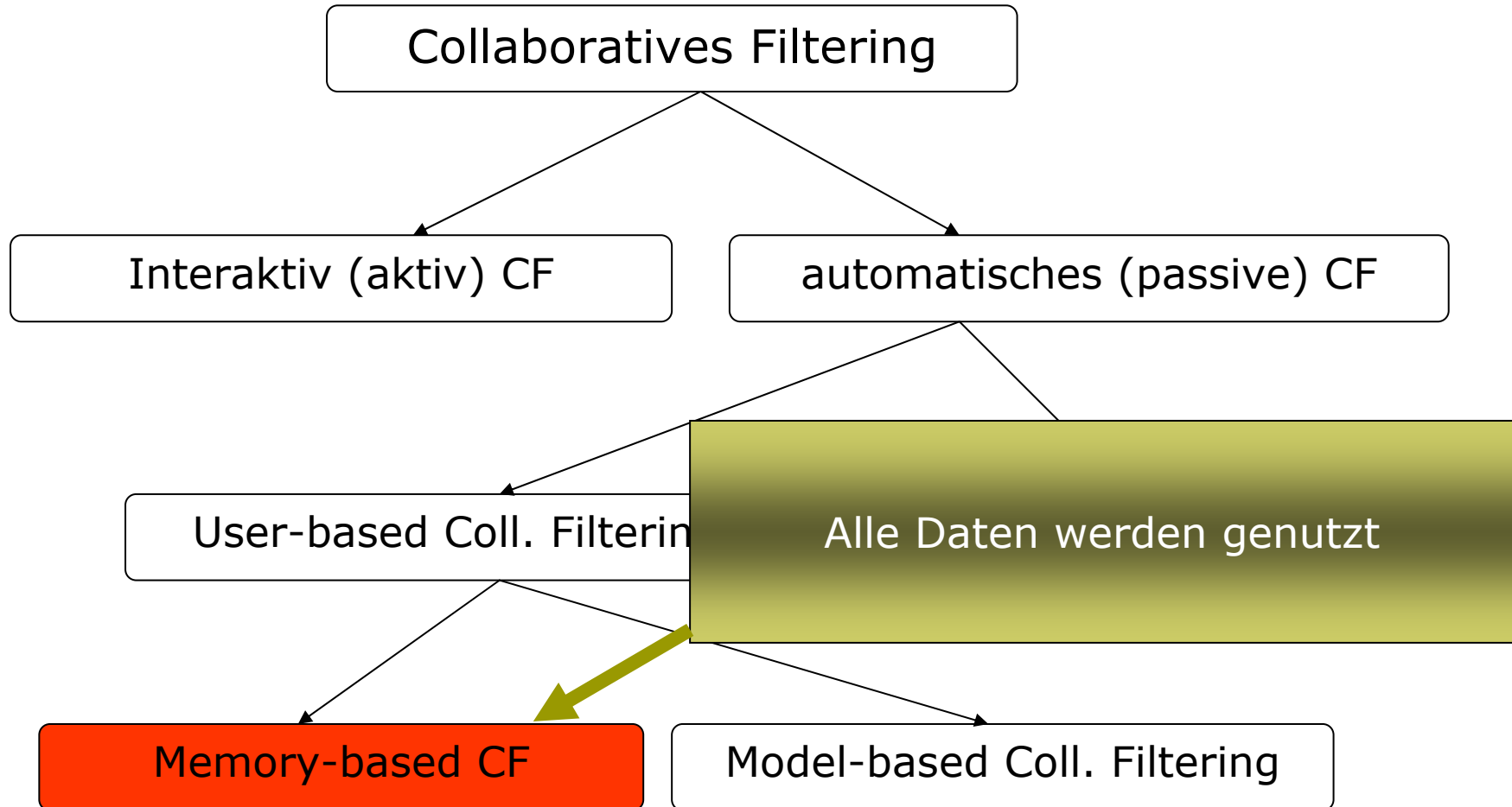
# Kollaboratives Filtern - Details

---



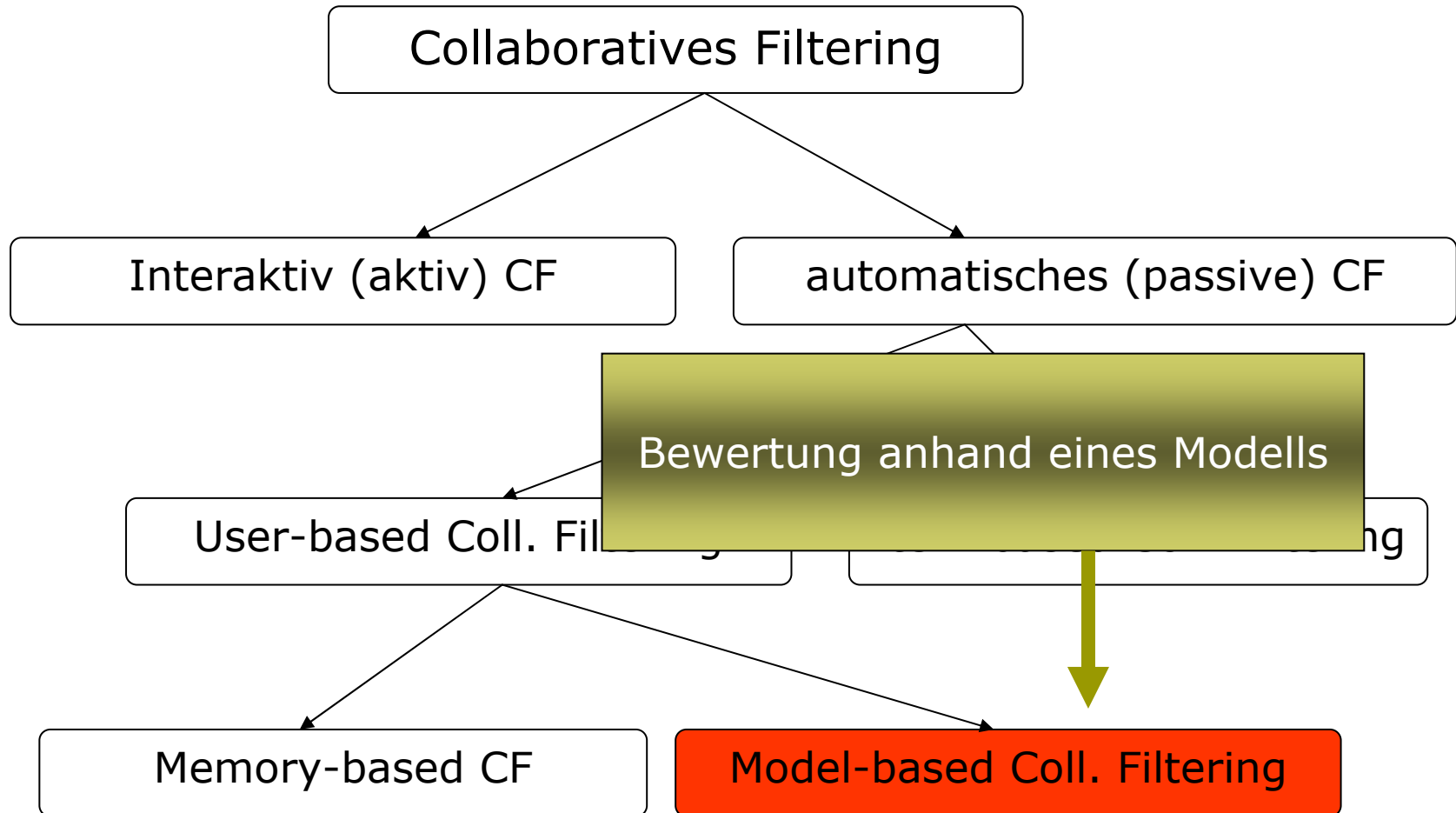
# Kollaboratives Filtern - Details

---



# Kollaboratives Filtern - Details

---



# statistische Hintergründe

---

## □ Korrelationskoeffizient

$$w(a, i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 (v_{i,j} - \bar{v}_i)^2}}$$

- Maß für den Zusammenhang zwischen zwei Datenreihen

# Content-Based Filtering

---

- Sammlung von „Keywords“ in einem Benutzerprofil
- Filtern von Bildern nur bedingt möglich
  
- 3 Phasen:
  1. Collection Stage
  2. Selection Stage
  3. Weitergabe der gefilterten Information an den Benutzer

# Content-Based Filtering (Details)

---

Differenz Algorithmus:

- Vektorfeld-Modell  $v_i = tf(i) \log \frac{n}{df(i)}$
- Bool'sche Funktion  
Gewichtung (0,1)  $\sum_{i=1}^I w_i x_i \geq \Gamma$
- Wahrscheinlichkeitsmethode

# User Tracking

---

- Log-Files
- Cookies
- Session-IDs
- Virtual Servers
- Login
- Panels and spyware-systems
- Benutzerbefragungen

# Probleme bei Personalisierung und User Modeling

---

## **Technische Probleme:**

### Cold-Start-Problem

- ❑ Anfangs oft unrealistische Wissensbasis.
- ❑ Durch einige extreme Eingaben am Anfang kann das System in die falsche Richtung driften.

## **Benutzbarkeitsprobleme:**

- ❑ Nicht jeder kann ein System verwenden das bereits auf einen User zugeschnitten ist.
- ❑ Was dem einen die Arbeit erleichtert kann sie dem anderen erschweren.

# Probleme bei Personalisierung und User Modeling

---

## **Datenschutzproblem:**

- ❑ Serverseitig gespeicherte Profile enthalten wertvolle Informationen.
- ❑ Verkauf dieser an z.B. Adressverlage.
- ❑ Verlinkung dieser Datensätze mit denen anderer Unternehmen (bereits Verboten, bzw. nur mit Zustimmung des Benutzers).
- ❑ Nutzung zu gezielten Angeboten oder auch SPAM.
- ❑ Gesetzlich geregelt, Wachsamkeit angebracht.
- ❑ Privacy Statements können vertrauensbildend wirken.

# Usenet News Filtering

---

## □ GroupLens

- Forschungsprojekt an der University of Minnesota 1992
- verkauft an NetPerception im Jahr 1996

## □ 3 Filtermethoden

- Personal recommendations (action-to-item)
- Anonymous recommendations (user-to-user)
- fast lookup recommendations (item-to-item)
  
- Berechnung erfolgt in Echtzeit

# Andere personalisierte Systeme I

---

## □ COGITO

- europäisches Forschungsprojekt 1999
- ist ein personalisiertes Vorschlagssystem
- Berechnung aufgrund von User-Profilen
- Spezialität: Chat-Robot zum Sammeln von User-Daten

# Anwendungen von Personalisierung I

---

## □ my.yahoo.com

- war einer der ersten am Markt
- keine Inhalte für Österreicher
  
- individuelles Layout
- personalisierter Inhalt
- Farbänderung möglich

# Anwendungen von Personalisierung II

---

## □ [www.amazon.at](http://www.amazon.at)

- vollständiges „recommender System“
- item-to-item, user-to-user und action-to-item

## □ [www.google.com](http://www.google.com)

- automatische Umleitung auf Landes-Seite
- basierend auf IP-Adresse

# Mail Filtering

---

- Cloudmark Spamnet
  - client-seitiges Filtersystem
  - Benutzer des Systems bewerten Mail
  - zentraler Server speichert Signatur von Spam-Mails
  - sehr hohe Trefferquote (ca. 95 %)
  - sehr niedrige Fehlerquote (false-positives)

# Andere personalisierte Systeme II

---

- Over the Air – Personalisierung am Handy
  - Klingeltöne
  - Logos
  - JAVA-Spiele
  - JAVA-Anwendungen
  - MMS (Multi Media SMS Message)
  - prominenter Vertreter: [www.sms.at](http://www.sms.at)

# Andere personalisierte Systeme III

---

- Lenkerpräferenzen im Autoschlüssel
  - Mehrere Personen teilen sich ein Fahrzeug
  - Jeder hat einen eigenen Schlüssel
  - In den Schlüsseln sind Sitz-  
Spiegeleinstellungen usw. für den jeweiligen  
Fahrer gespeichert

# Zusammenfassung

---

- immer mehr Informationen, daher immer mehr angepasste Informationen
- Datenschutzproblematik
- Verschiedene Arten Informationen zu Filtern (Kollaborative bzw. Content based Filterung)
- Usertracking
- verschiedene Anwendungen von Personalisierung