

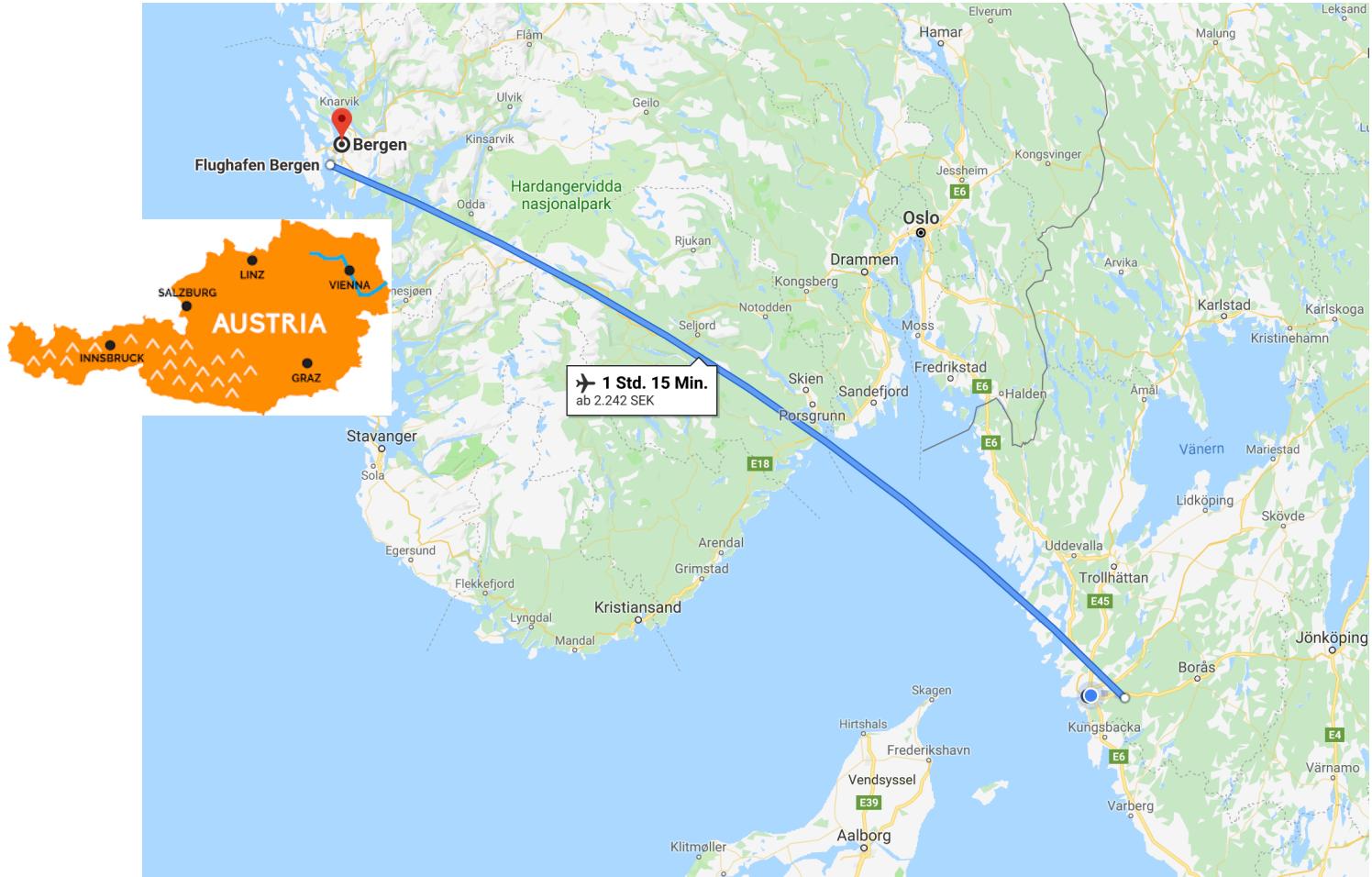


An Evaluation of Recommendation Algorithms for Online Recipe Portals

Christoph Trattner
University of Bergen

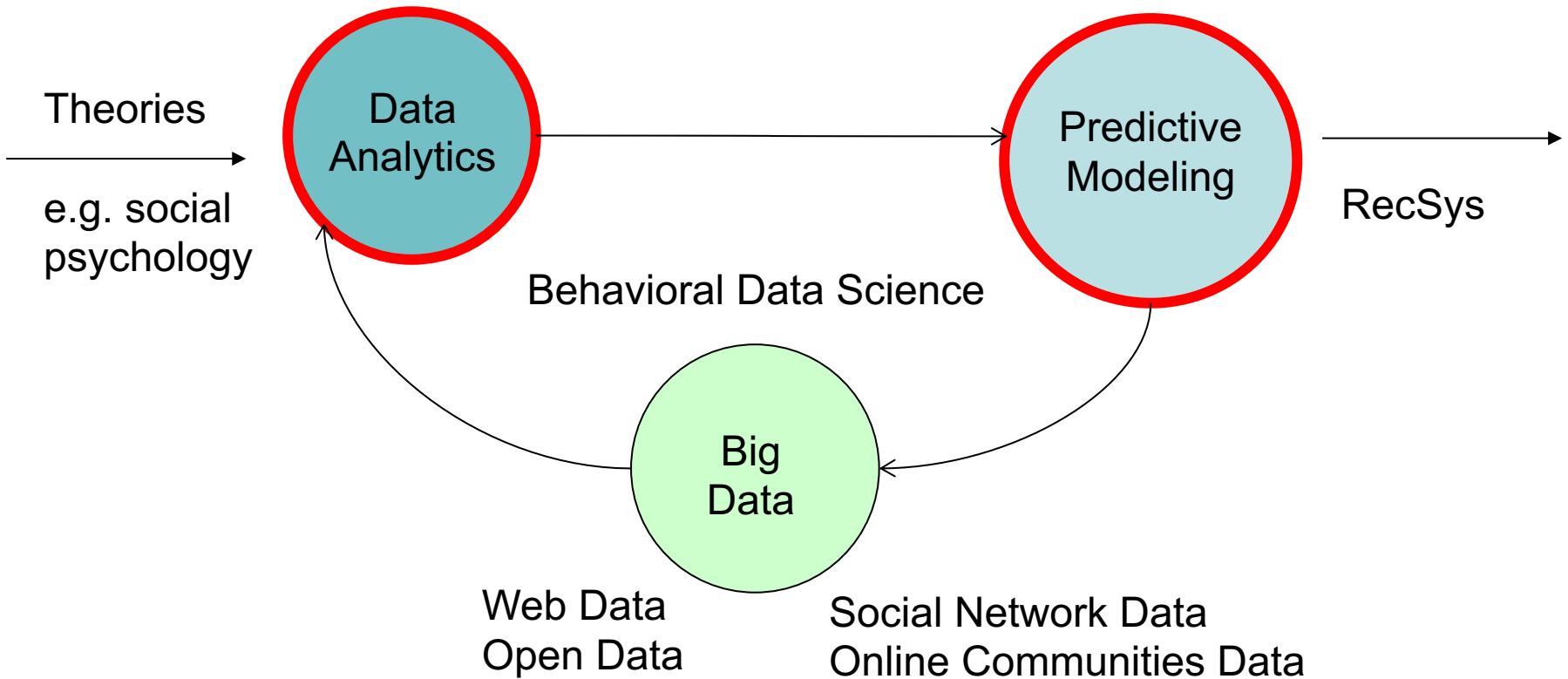
David Elsweiler
University of Regensburg

Where do I come from?



Research Focus

Understand how people behave



Agenda

- Motivation
- Current state-of-the-art in Food RecSys
- Content-based RecSys & Sim. Functions
- Protocol & Dataset
- Results
- Conclusions

Part 1: Motivation

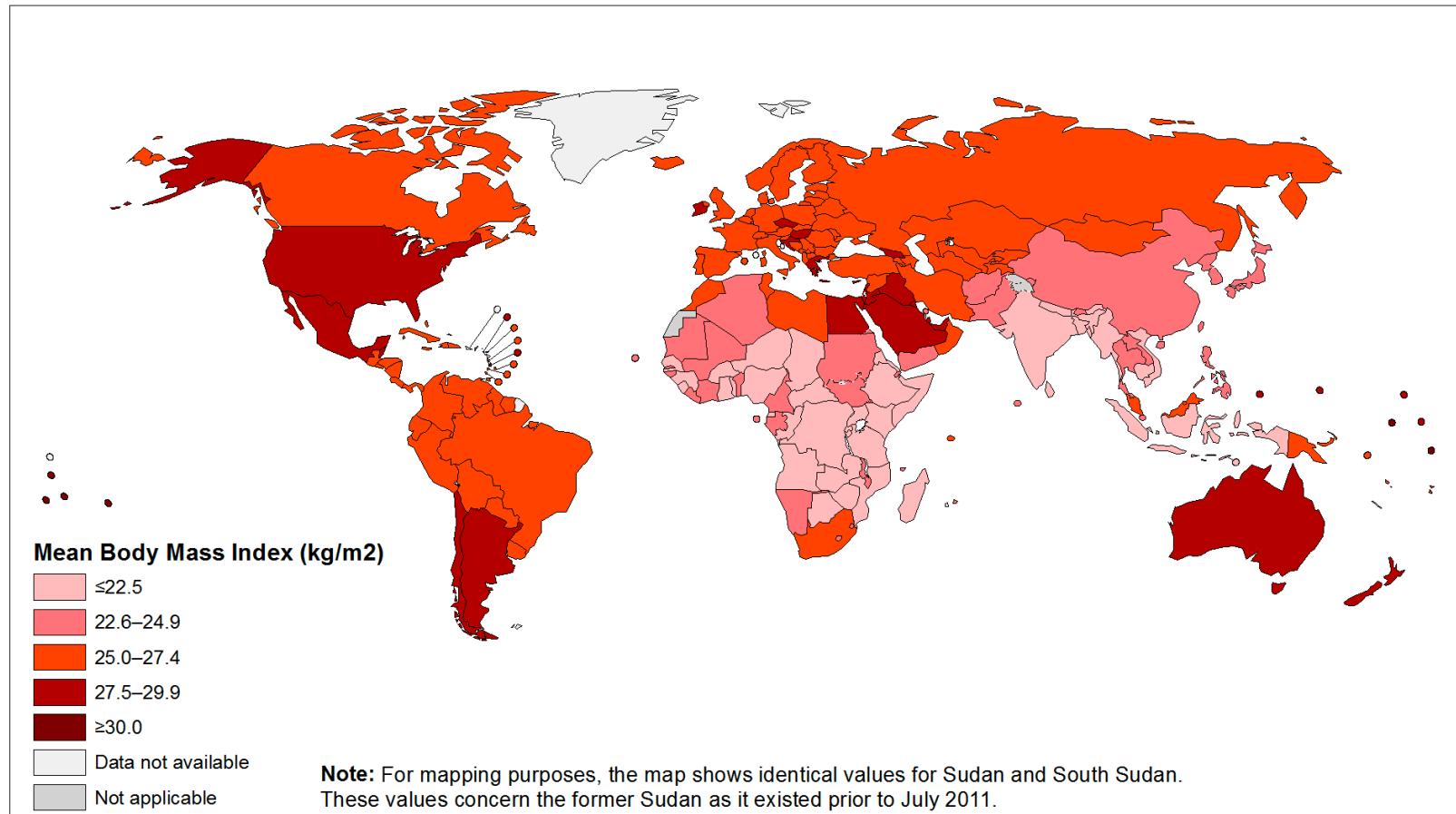
Why is research into Food Recsys Important?

Why is that important?

- Food is one the main concepts that **shapes how good we feel and how healthy we are**
- According to the WHO, if common lifestyle risk factors, among others diet-related ones, were eliminated, **around 80% of cases of heart disease, strokes and type 2 diabetes, and 40% of cancers, could be avoided** (European Comission Recommendation C(2010) 2587 final, 2010).

BMI is increasing World Wide

Mean Body Mass Index (kg/m²), ages 18+, 2016 (age standardized estimate)
Male



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: World Health Organization
Map Production: Information Evidence and Research (IER)
World Health Organization



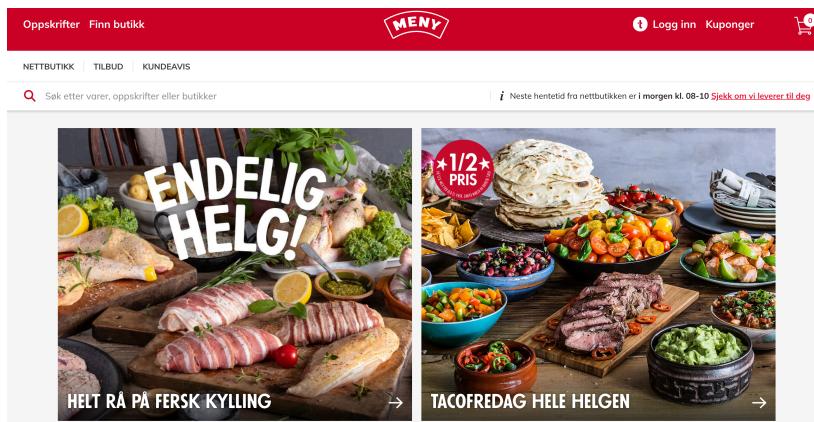
© WHO 2017. All rights reserved.

The approaches I am discussing today
are all online food recommender
approaches!

Why Online?

Most food interactions nowadays online

According to recent market research over 50%



The screenshot shows the homepage of the allrecipes website. At the top, there is a red navigation bar with links for "Oppskrifter", "Finn butikk", "MENY" (highlighted in white), "Logg inn", "Kuponger", and a shopping cart icon. Below the navigation bar, there are links for "NETTBUTIKK", "TILBUD", and "KUNDEAVIS". A search bar is followed by a message: "Neste hentetid fra nettbutikken er i morgen kl. 08-10 Sjekk om vi leverer til deg". The main content area features two large meal ideas: "HELT RA PÅ FERSK KYLLING" (left) and "TACOFREDAG HELE HELGEN" (right). Each idea includes a photograph of the meal and a "1/2 PRIS" (half price) badge.

Amazon

amazon

All ▾ 

Shop Today's Deals

 Deliver to Austria  Your Amazon.com Today's Deals EN  Hello, Sign in Account & Lists Orders 

Grocery Deals Snacks ▾ Breakfast ▾ Warm Beverages Cold Beverages Cooking Staples ▾ Baby Food ▾ Candy & Chocolate ▾ Prime Pantry ▾

Featured Shops

- New Year, New You
- Grocery Sales & Deals
- Subscribe & Save
- Prime Pantry
- Amazon Family
- Grocery Dash Buttons
- International Food Market

Show results for

Grocery & Gourmet Food

- Baby Foods
- Alcoholic Beverages
- Beverages
- Breads & Bakery
- Breakfast Foods
- Candy & Chocolate
- Canned, Jarred & Packaged Foods
- Condiments & Salad Dressings

Groceries & Gourmet Food

Shop groceries online for delivery of [coffee](#), [snacks](#), [chocolate](#), and everyday food.



NEW YEAR NEW YOU
Start 2019 with healthy food & beverages

[Shop now ▶](#)

Part 2: The Current State-of-the-art in Food RecSys

Food Recommender Systems: Important Contributions, Challenges and Future Research Directions. Trattner, C. and Elsweiler, D. Collaborative Recommendations: Algorithms, Practical Challenges and Applications, World Scientific Publishing Co. Pte. Ltd., 2018

Author(s)	Algorithm(s)	Person- alized	RecSys Type(s)	Feedback	Context/ Content Feature(s)	Dietary Constraints	Target	Dataset	FAS	
									9	WUW
(Elsweiler, Trattner & Harvey, 2017)	Logistic Random Forrest Naive Bayes	no	Recipes	Ratings Binary	Title Image Ingredients Nutrition Pop. & Appr	no	Single User	Allrecipes	FAS	
	LDA WRMF AR SLIM BPR MostPop User- ItemKNN BPR MostPop		yes/no	Recipes Meal Plans	Bookmarks Ratings Comments	WHO-FSA health score			FAS	
(Trattner & Elsweiler, 2017)	AR SLIM BPR MostPop User- ItemKNN BPR MostPop Learning to Rank					no	Single User	Allrecipes	FAS	
	UserKNN MostPop MF CB								FAS	
(Cheng, Rokicki & Herder, 2017)	MF CB	yes	Recipes	Ratings Tags	City Size Image Embeddings	no	Single User	Kochbar	FAS	
	CB								FAS	
(Yang et al., 2017)	MF CB	yes	Recipes	Ratings Tags	Gender Tags	no	Single User	Yummly	FAS	
	CB								FAS	
(Rokicki, Herder, Kuśmierczyk & Trattner, 2016)	MF CB	yes	Recipes	Ratings Tags	Tags	no	Single User	Kochbar	FAS	
	CB								FAS	
(Ge, Elahi, Fernández-Tobías, Ricci & Massimo, 2015)	MF CB	yes	Recipes	Ratings Tags	Tags	no	Single User	Wellbeing Diet Book	FAS	
	CB								FAS	
(Elsweiler & Harvey, 2015)	SVD-Hybrid	yes	Meal Plans (Set of recipes)	Ratings	Ingredients	yes	Single User	Quizine	FAS	
	UserKNN SVD Hybrid NL-PCA								FAS	
(Sano, Machino, Yada & Suzuki, 2015)	CB	yes	Groceries	Purchases	Food Categories	no	Single User	Grocery store data	FAS	
	CB								FAS	
(Trevisiol, Chiarandini & Baeza-Yates, 2014)	UserKNN CB	yes	Menus (Set of dishes)	Binary	Text Sentiment	no	Single User	Yelp	FAS	
	CB								FAS	
(Elahi, Ge, Ricci, Massimo & Berkovsky, 2014)	MF	yes	Recipes	Ratings Tags	tags	no	Group of Users	Wellbeing Diet Book	FAS	
	CB, CF								FAS	
(Harvey et al., 2013)	Logistic Reg. SVD-Hybrid	yes	Recipes	Ratings	Ingredients etc.	no	Single User	Quizine	FAS	
	CB								FAS	
(Teng, Lin & Adamic, 2012)	SVM	no	Recipes	Ratings	Ingredients Nutrition Cook effort Cook methods	no	Single User	Allrecipes	FAS	
	Graph-based CB		Menus (Set of recipes)	Tags	Ingredients					FAS
(El-Dosuky, Rashad, Hamza & El-Bassiouny, 2012)	CB KB	yes	Food items	Query	tags	no	Single User	USDA	FAS	
	KB								FAS	
(Freyne, Berkovsky, Baghaei, Kimani & Smith, 2011)	CF	yes	Meal plans (Set of recipes)	Ratings	-	no	Single User	Wellbeing Diet Book	FAS	
	KB								FAS	
(Ueta, Iwakami & Ito, 2011)	KB	yes	Recipes	Query Cooked recipes	tags Recipe content features	no	Single User	Cookpad	FAS	
	CB								FAS	
(van Pinxteren, Geleijnse & Kamsteeg, 2011)	CB	yes	Recipes	Ratings	Ingredients	no	Single User	Smulweb	FAS	
	UserKNN CB Hybrid								FAS	
(Freyne & Berkovsky, 2010)	UserKNN CB Hybrid	yes	Recipes	Ratings	Ingredients	no	Single User	Wellbeing Diet Book	FAS	
	UserKNN								FAS	

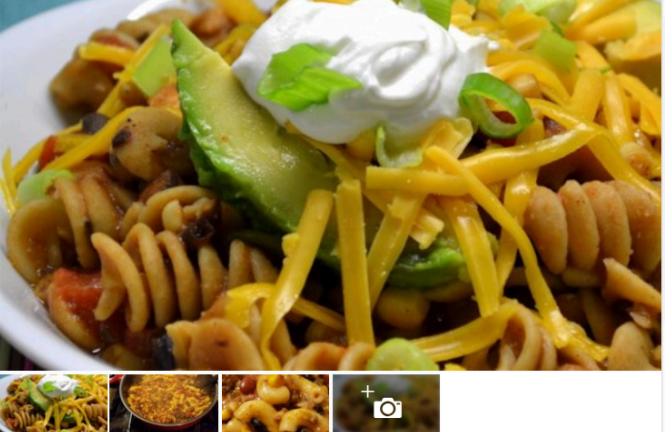
What types of features can we actually employ for CB in Food RecSys?

Problem

allrecipes

BROWSE ▾ Find a recipe Ingredient Search   

Home > Recipes > Soups, Stews and Chili > Chili > Vegetarian

Carole's Chili Mac 
Title  17 made it | 4 reviews | 3 photos
 Recipe by: Carole Moritz  3 "This is a vegetarian version with added pasta for those pasta lovers. I love the taste of the cumin in this recipe."

Image 

Prep 15 m Cook 50 m Ready In 1 h 5 m

- 1 Bring a large pot of lightly salted water to a boil. Cook elbow macaroni in the boiling water until cooked through but firm to the bite, 10 to 12 minutes. Drain.
- 2 Heat olive oil in a large skillet over medium heat; cook and stir onion, green bell peppers, and garlic in the hot oil until onion is softened, 5 to 10 minutes.
- 3 Mix tomatoes and their juices, tomato sauce, hot pepper sauce, chili powder, cumin, cayenne pepper, and black pepper into onion mixture; bring to a boil. Add kidney beans, black beans, and corn; reduce heat to low, cover skillet, and simmer chili for 30 minutes.
- 4 Stir macaroni into chili, cover skillet, and simmer until macaroni is heated, 5 minutes; top with Cheddar cheese and green onions to serve.

Ingredients

-  2 cups whole wheat elbow macaroni  1 teaspoon ground cumin
-  2 teaspoons olive oil  1/4 teaspoon cayenne pepper
-  1 small onion, chopped  1/4 teaspoon ground black pepper
-  2 green bell peppers, chopped  1 (15 ounce) can white kidney beans, drained and rinsed
-  2 cloves garlic, minced  1 (15 ounce) can black beans, drained and rinsed
-  1 (14.5 ounce) can diced tomatoes, undrained  1 (12 ounce) can whole kernel corn, drained
-  1 (8 ounce) can tomato sauce  3/4 cup shredded reduced-fat Cheddar cheese
-  1 tablespoon hot pepper sauce, or to taste  2 tablespoons chopped green onion, or to taste (optional)
-  2 teaspoons chili powder, or more to taste 

Directions

Part 3:

How do we calculate similarity between Recipes?

$sim(a,b)$

Linguine Pasta with Shrimp and Tomatoes



Ingredients

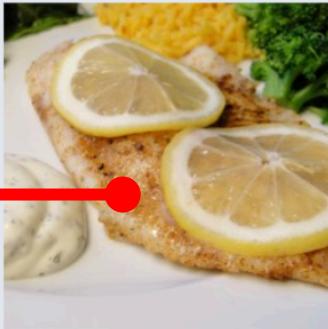
2 tablespoons olive oil
 3 cloves garlic, minced
 4 cups diced tomatoes
 1 cup dry white wine
 2 tablespoons butter
 salt and black pepper to taste
 1 (16 ounce) package linguine pasta
 1 pound peeled and deveined medium shrimp
 1 teaspoon Cajun seasoning
 2 tablespoons olive oil

Directions

Heat 2 tablespoons of olive oil in a large saucépan over medium heat. Stir in the garlic, cook 2 minutes. Add the tomatoes, and wine. Bring to a simmer and cook 30 minutes, stirring frequently. Once the tomatoes have simmered into a sauce, stir in the butter and season with salt and pepper. Fill a large pot with lightly-salted water, bring to a rolling boil, stir in the linguine and return to a boil. Cook the pasta uncovered, stirring occasionally, until the pasta has cooked through but is still firm to the bite,

 $sim(a,b)$

Hudson's Baked Tilapia with Dill Sauce



Ingredients

4 (4 ounce) fillets tilapia
 salt and pepper to taste
 1 tablespoon Cajun seasoning, or to taste
 1 lemon, thinly sliced
 1/4 cup mayonnaise
 1/2 cup sour cream
 1/8 teaspoon garlic powder
 1 teaspoon fresh lemon juice
 2 tablespoons chopped fresh dill

Directions

Preheat the oven to 350 degrees F (175 degrees C). Lightly grease a 9x13 inch baking dish.
 Season the tilapia fillets with salt, pepper and Cajun seasoning on both sides. Arrange the seasoned fillets in a single layer in the baking dish. Place a layer of lemon slices over the fish fillets. I usually use about 2 slices on each piece so that it covers most of the surface of the fish. Bake uncovered for 15 to 20 minutes in the preheated oven, or until fish flakes easily with a fork.

 $sim(a,b)$
 $sim(a,b)$

Paper

Trattner, C. and Jannach, D. **Learning to Recommend Similar Items from Human Judgements.** User Modeling and User-Adapted Interaction Journal. 2019.

Features for Similar Recipe Recommendations

Table 1: Similarity metrics computed based on recipe titles, images, ingredients and cooking directions.

Name	Metric	Explanation
Title:LV	$sim(r_i, r_j) = 1 - dist_{LEV}(r_i, r_j) $	Title Levenshtein distance-based similarity
Title:JW	$sim(r_i, r_j) = 1 - dist_{JW}(r_i, r_j) $	Title Jaro-Winkler distance-based similarity
Title:LCS	$sim(r_i, r_j) = 1 - dist_{LCS}(r_i, r_j) $	Title Least Common Subsequence distance-based similarity
Title:BI	$sim(r_i, r_j) = 1 - dist_{BI}(r_i, r_j) $	Title Bi-gram distance-based similarity
Title:LDA	$sim(r_i, r_j) = \frac{LDA(\text{Title}(r_i)) \cdot LDA(\text{Title}(r_j))}{\ LDA(\text{Title}(r_i))\ \ LDA(\text{Title}(r_j))\ }$	Title LDA cosine-based similarity (LDA = LDA vector)
Image:BR	$sim(r_i, r_j) = 1 - BR(r_i) - BR(r_j) $	Image Brightness distance-based similarity
Image:SH	$sim(r_i, r_j) = 1 - SH(r_i) - SH(r_j) $	Image Sharpness distance-based similarity
Image:CO	$sim(r_i, r_j) = 1 - CO(r_i) - CO(r_j) $	Image Contrast distance-based similarity
Image:COL	$sim(r_i, r_j) = 1 - COL(r_i) - COL(r_j) $	Image Colorfulness distance-based similarity
Image:EN	$sim(r_i, r_j) = 1 - EN(r_i) - EN(r_j) $	Image Entropy distance-based similarity
Image:EMB	$sim(r_i, r_j) = \frac{EMB(r_i) \cdot EMB(r_j)}{\ EMB(r_i)\ \ EMB(r_j)\ }$	Image Embedding cosine-based similarity (EMB= image embedding vector)
Ing:COS	$sim(r_i, r_j) = \frac{amount(\text{Ing}(r_i)) \cdot amount(\text{Ing}(r_j))}{\ amount(\text{Ing}(r_i))\ \ amount(\text{Ing}(r_j))\ }$	Ingredients Cosine similarity (amount-based weighting in grams per 100g of a meal)
Ing:JACC	$sim(r_i, r_j) = \frac{ \{Ing(r_i)\} \cap \{Ing(r_j)\} }{ \{Ing(r_i)\} \cup \{Ing(r_j)\} }$	Ingredients Jaccard similarity
Ing:TFIDF	$sim(r_i, r_j) = \frac{TFIDF(\text{Ing}(r_i)) \cdot TFIDF(\text{Ing}(r_j))}{\ TFIDF(\text{Ing}(r_i))\ \ TFIDF(\text{Ing}(r_j))\ }$	Ingredients Text-based cosine similarity (TFIDF = TF-IDF weighted vector)
Ing:LDA	$sim(r_i, r_j) = \frac{LDA(\text{Ing}(r_i)) \cdot LDA(\text{Ing}(r_j))}{\ LDA(\text{Ing}(r_i))\ \ LDA(\text{Ing}(r_j))\ }$	Ingredients LDA-based cosine similarity (LDA = LDA vector)
Dir:TFIDF	$sim(r_i, r_j) = \frac{TFIDF(\text{Dir}(r_i)) \cdot TFIDF(\text{Dir}(r_j))}{\ TFIDF(\text{Dir}(r_i))\ \ TFIDF(\text{Dir}(r_j))\ }$	Cooking Directions Text-based cosine similarity (TFIDF = TF-IDF weighted vector)
Dir:LDA	$sim(r_i, r_j) = \frac{LDA(\text{Dir}(r_i)) \cdot LDA(\text{Dir}(r_j))}{\ LDA(\text{Dir}(r_i))\ \ LDA(\text{Dir}(r_j))\ }$	Cooking Directions LDA cosine-based similarity (LDA = LDA vector)

Part 4: Experiment & The Dataset

Offline Evaluation with 10-fold Cross-Validation

LibRec/AUC

<http://allrecipes.com>

allrecipes BROWSE Find a recipe Ingredient Search Q 1 Heart

Home > Recipes > Fruits and Vegetables > Vegetables > Rhubarb

Strawberry Rhubarb Custard Pie

★★★★★ 80 made it | 62 reviews

Recipe by: Chef John  31K

"One of the most delicious and easiest pie recipes I know. I got this wonderful recipe from my mother Pauline, who I believe got it from my Aunt Angela. I love all their pies, but this might be my favorite."



Save I Made It Rate it Share Print

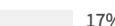
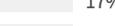
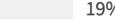
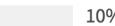
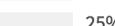
Ingredients

- + 1 (9 inch) unbaked pie crust (see footnote for recipe link)
- + 3 cups rhubarb, sliced 1/4-inch thick
- + 1 cup fresh strawberries, quartered
- + 3 large eggs
- + 1 1/2 cups white sugar
- + 3 tablespoons all-purpose flour
- + 1/4 teaspoon freshly grated nutmeg
- + 1 tablespoon butter, diced
- + 2 tablespoons strawberry jam
- + 1/4 teaspoon water

2 h 20 m 8 servings  342 cals 

Nutrition

Amount per serving (8 total)

Calories:	342 kcal		17%
Fat:	11.1 g		17%
Carbs:	57.4g		19%
Protein:	4.8 g		10%
Cholesterol:	74 mg		25%
Sodium:	159 mg		6%

Nutrition Facts
to compute FSA

Determining the healthiness of recipes

What the colours mean:



RED

means **HIGH**

indicating that the food is **high** in fat, sugars or salt

It's fine to eat this food occasionally or as a treat, but think about how often you choose it and how much of it you eat.



AMBER

means **MEDIUM**

making it an **OK** choice

Although going for green is even better!



GREEN

means it's **LOW**

*Which makes it a **healthier** choice.*



FSA food health criteria

Check how much fat, sugar and salt is in your food



Remember that the amount you eat of a particular food affects how much sugars, fat, saturates and salt you will get from it.

Food Shopping Card

	Sugars	Fat	Saturates	Salt
What is HIGH per100g	Over 15g	Over 20g	Over 5g	Over 1.5g
What is MEDIUM per100g	Between 5g and 15g	Between 3g and 20g	Between 1.5g and 5g	Between 0.3g and 1.5g
What is LOW per100g	5g and below	3g and below	1.5g and below	0.3g and below

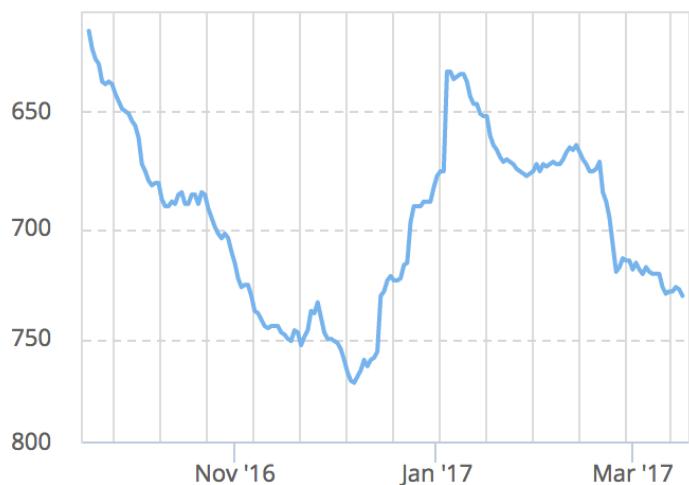
Trattner, C. Elsweiler, D. and Simon, H. Estimating the Healthiness of Internet Recipes: A Cross-Sectional Study. Frontiers in Public Health, 2017.

How popular is allrecipes.com?

Allrecipes.com popularity

Alexa Traffic Ranks

How is this site ranked relative to other sites?



According to Alexa.com

Global Rank ?

 730 ▼ 2

Rank in United States ?

 217

Country	Percent of Visitors	Rank in Country
 United States	69.6%	217
 Canada	8.0%	219
 United Kingdom	2.5%	1,375
 Germany	2.0%	2,017
 India	1.3%	4,349

Part 5: Results

Table 2: Results of the recommender experiment – collaborative (CF) vs content-based (CB) – in the dense data sample with all users. Best features in each set (CF and CB) are bolded. Top-5 (\uparrow) and Bottom-5 (\downarrow) single content features are also marked.

Method	Algorithm	AUC
CF	BPR	.7094
	WRMF	.6881
	UserKNN	.6962
	ItemKNN	.6909
	MostPopular	.6864
	LDA	.6863
CB	Title:Levenstein-Distance	.5468 (\uparrow)
	Title:Bigram-Distance	.5500 (\uparrow)
	Title:LCS-Distance	.5424
	Title:LDA-Text-Cosine	.5353
	Title:Jaro-Winkler-Distance	.5324
	Title>All	.5523
	Image:Cosine-Embeddings	.5322
	Image:Colorfulness-Distance	.5072 (\downarrow)
	Image:Contrast-Distance	.5175
	Image:Sharpness-Distance	.5109
CB	Image:Entropy-Distance	.5080 (\downarrow)
	Image:Brightness-Distance	.4991 (\downarrow)
	Image>All	.5425
	Ingredients:Cosine-Text	.5547
	Ingredients:Cosine-LDA-Text	.5653 (\uparrow)
	Ingredients:Jaccard	.5502
	Ingredients:Cosine	.5575
Directions	Ingredients>All	.5718
	Directions:Cosine-LDA-Text	.5606 (\uparrow)
	Directions:Cosine-Text	.5210
	Directions>All	.5731
Ratings	Ratings:Number-Distance	.4789 (\downarrow)
	Ratings:Average-Distance	.4832 (\downarrow)
	Ratings>All	.5249
Health	Health:FSA	.5775 (\uparrow)
	CB>All	.5883
Random		.4989

CF vs CB in Recipe RecSys

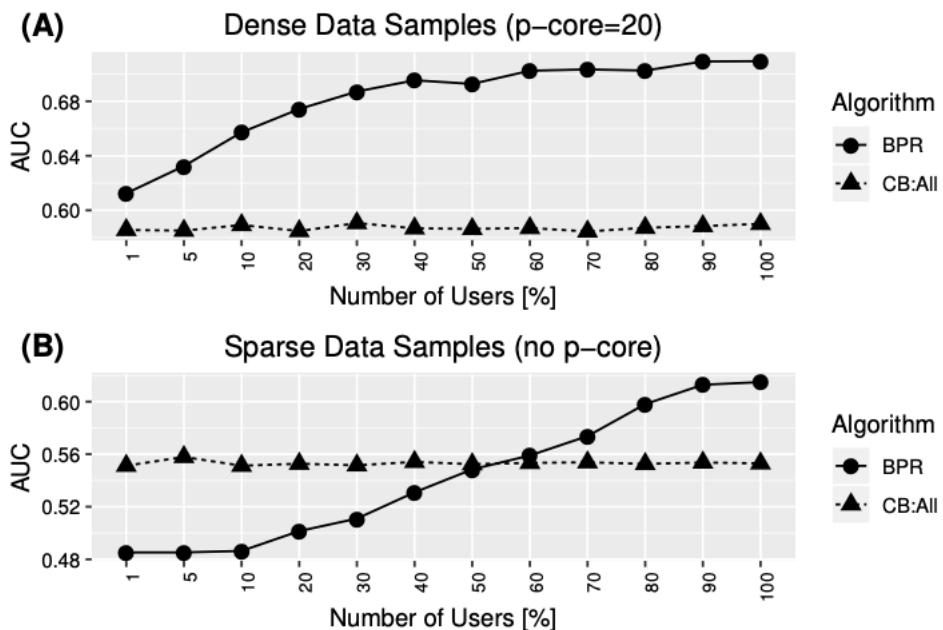


Figure 1: (A) shows the results in the dense data samples (= p-core filtered) where each user has at least 20 item interactions and each item is at least 20-times interacted with, (B) shows the results in the sparse data samples (=no p-core).

Summary

- **CF methods consistently outperform CB methods over the full dataset.**
- **CF requires either a small number of highly active users or over six hundred users, selected randomly to achieve competitive performance.**
- **There is a useful signal in the CB facets, which would be useful in cold-start situations.**
- **One of the most robust content features is the nutritional healthiness of the recipe as defined by a measure derived from the United Kingdom Food Standards Agency (FSA).**

What is the Future?

Sustainable Food Recommender Systems

What online data say about eating habits. Trattner, C. and Elsweiler, D. NATURE Sustainability, 2019

Thank you!



Prof. Christoph Trattner

Email: trattner.christoph@gmail.com
Web: christophtrattner.com

Twitter: [@ctrattner](https://twitter.com/ctrattner)



GRACIAS DANKSCHEEN
ARIGATO SPASSIBO
SHUKURIA NUHUN
JUSPAXAR TAVTAPUCHI
BAHNA
MEDAWAGSE
TASHAKKUR ATU CHALTU
SNACHALHYA
YAQHANYELAY WABEEJA
MAITEKA
DHANYABAAD
ANNA
ATTO
GRAZIE KOMAPSUMNIDA
MAAKE
MERASTAWHY
GAEJTHO
LAH
MEHRBANI PALDIES
DENKAUJA
UNAHLCHEESH
MEERI
TINGKI HUI
YUSPAGARATAM
THANK BİYAN
SHUKRIA
YOU GUI
NATUR
EKOJU
SIKOMO
MAKETBI
BOLZİN MINMONCHAR
MERCI