

# iCookWare

## Abschlussvortrag

Universität Heidelberg, Lehrstuhl für Computerlinguistik

Studienprojekt 2004-2005

Nicola Kaiser, Ana Kovatcheva, Olga Mordvinova, Stephanie Schuldes

### ☼ Warum das Ganze?

- ☼ Motivation

- ☼ Benutzerszenario

- ☼ Ziel

### ☼ Und was haben wir gemacht?

- ☼ Systemübersicht

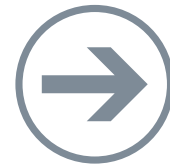
- ☼ Module & Techniken

### ☼ Diskussion

3

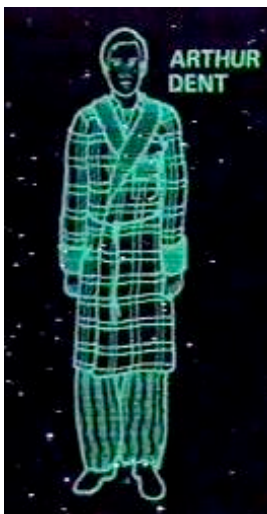


# Einführung



4

## Szenario I



```
Text — tcsh (tty1)
ArthurDent@marvin:~$ whoami
Profil:
  männlich
  Student
  Kocherfahrung: ***
  Vegetarier
  Allergie:
    Walnüse
    Karotten
  Küche:
    1 Pfanne
    2 Töpf
    2 Kochplatten
ArthurDent@marvin:~$
```

## iCookWare - intelligent Cooking Software

### ☼ Was SOLLTE es werden?

#### ☼ flexible Rezeptgenerierung

☼ Zutaten-Anpassung ✓

☼ Benutzer-Profil ✗

☼ Generierung der Rezepte (Templates oder Textplan) ✗

#### ☼ Phantasieanreger - kreative Kochvorschläge

### ☼ Was SOLLTE es nicht werden?

☼ keine einfache Datenbank

☼ keine Kochschule

☼ kein Kochroboter



☼ webserverbasiert

☼ frei

☼ zugänglich für alle

☼ erweiterbar



iCookWare - intelligentes Kochen

Anfrage My iCookWare Forum Super User das Team Impressum

**Anfrage**

Hallo Arthur!

Was hast du im Kühlschrank?

Gouda, Parmesan, 42 Eier, Bier, Milch, Butter, Nutella, Marmelade, Ketchup

Auf was hast du Lust?

warme italienische Hauptspeise

Für wieviele Leute willst du kochen?

2

Wie aufwendig soll das Rezept sein?

idiotischer  
 einfach  
 aufwendig  
 sehr aufwendig  
 ich will 5 Stunden in der Küche verbringen

Das Rezept soll sein:

für Suppenkasper  Date darf nicht schief gehen  Omas Küche  exotisch  experimentell

Wieviele Rezeptvorschläge sollen generiert werden?

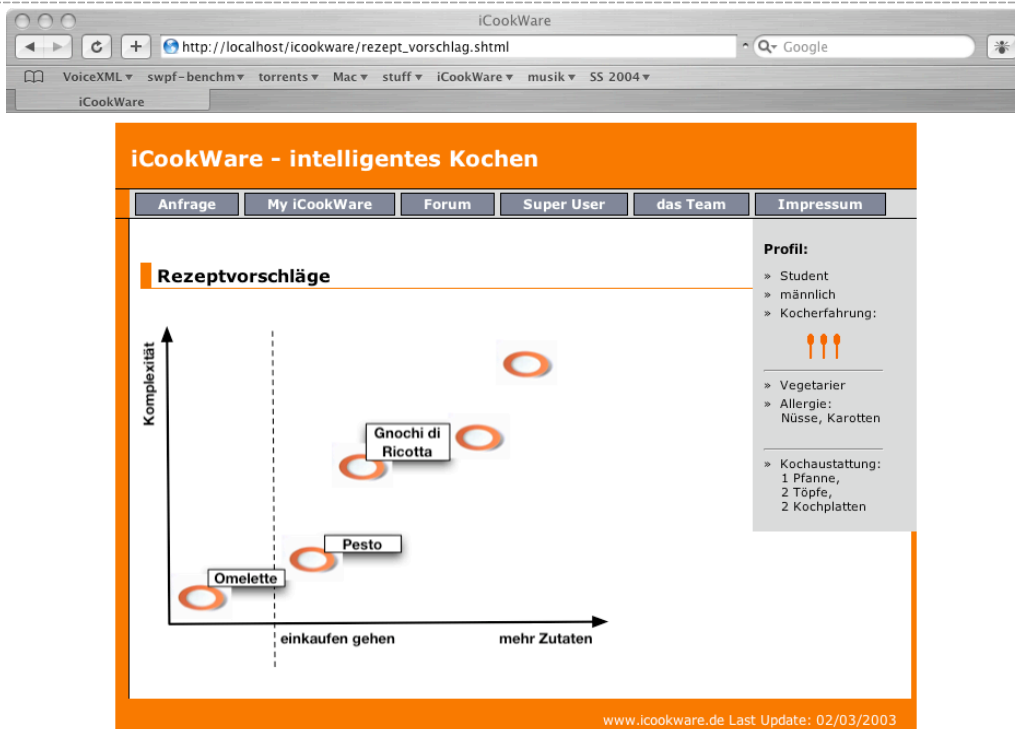
1  3  5  10  alle möglichen

Anfrage starten Zurücksetzen

www.icookware.de Last Update: 02/03/2003

**Profil:**

- » Student
- » männlich
- » Kocherfahrung:
  - ↑↑↑
- » Vegetarier
- » Allergie:
  - Nüsse, Karotten
- » Kochaustattung:
  - 1 Pfanne,
  - 2 Töpfe,
  - 2 Kochplatten



The screenshot shows a web browser window displaying the iCookWare website. The page title is "iCookWare - intelligentes Kochen". The main content area features a recipe for "Gnocchi di ricotta mit Gouda". The recipe is structured with sections for "Zutaten" (Ingredients), "Zum Servieren" (Serving), and "Zubereitung" (Preparation). The ingredients list includes 1 egg, 1/2 egg yolk, 150g Gouda, 37g Parmesan, 1/4 tsp salt, 1/2 tsp nutmeg, pepper, 25g flour, and 25g semolina. The serving instructions mention 1 1/2 tablespoons of butter and sage leaves. The preparation section describes the process of making the gnocchi, from mixing the dough to cooking and serving with sage butter. A sidebar on the right provides user profile information, including "Profil", "Vegetarier", "Allergie", and "Kochausstattung".

### ☘ verworfene Ideen:

#### ☘ Generierung

- ☘ schon Parsing zu schwer (Rezepte sind doch recht anspruchsvoll!)

#### ☘ XML

- ☘ keine iData mit zusätzlichen Informationen (Zeit etc.)

#### ☘ Zählen

- ☘ Schwierigkeitsgrad des Rezeptes (Anzahl der Verben im Algorithmus)
- ☘ Zeitaufwand

#### ☘ "semantischer" Tagger

- ☘ Rezepte bzw. Kochalgorithmen auseinandernehmen
- ☘ Erkennung von Hilfsmitteln und Methoden

#### ☘ Taxonomie für Zutatenbezeichnungen

- ☘ "Dose/n Tomaten", "Tomaten aus der Dose", "Geschälte Tomaten a. d. D." etc.

☘ umgesetzte Ideen:

☘ Webinterface

☘ Clustering (wenn auch anders als geplant)

☘ Taxonomie

☘ neue Ideen:

☘ Zutatenvektoren

☘ von Hand codiert

☘ mit Hilfe von Lebensmittel- und Nährwertinformationsseiten

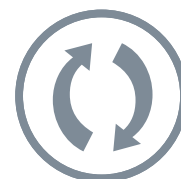
☘ Lemmatisierung statt Stemming

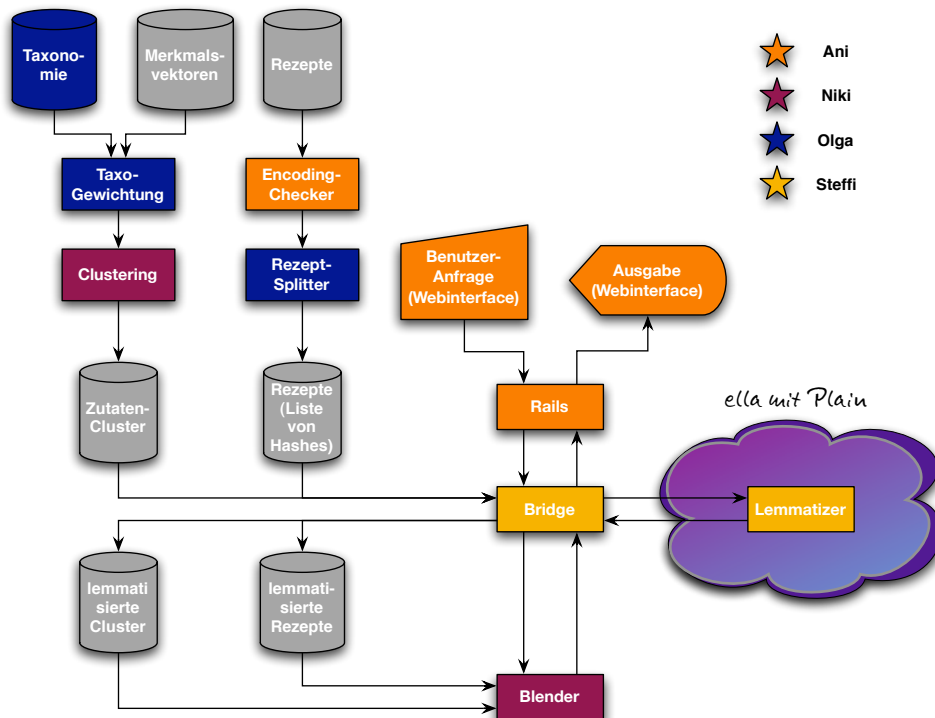
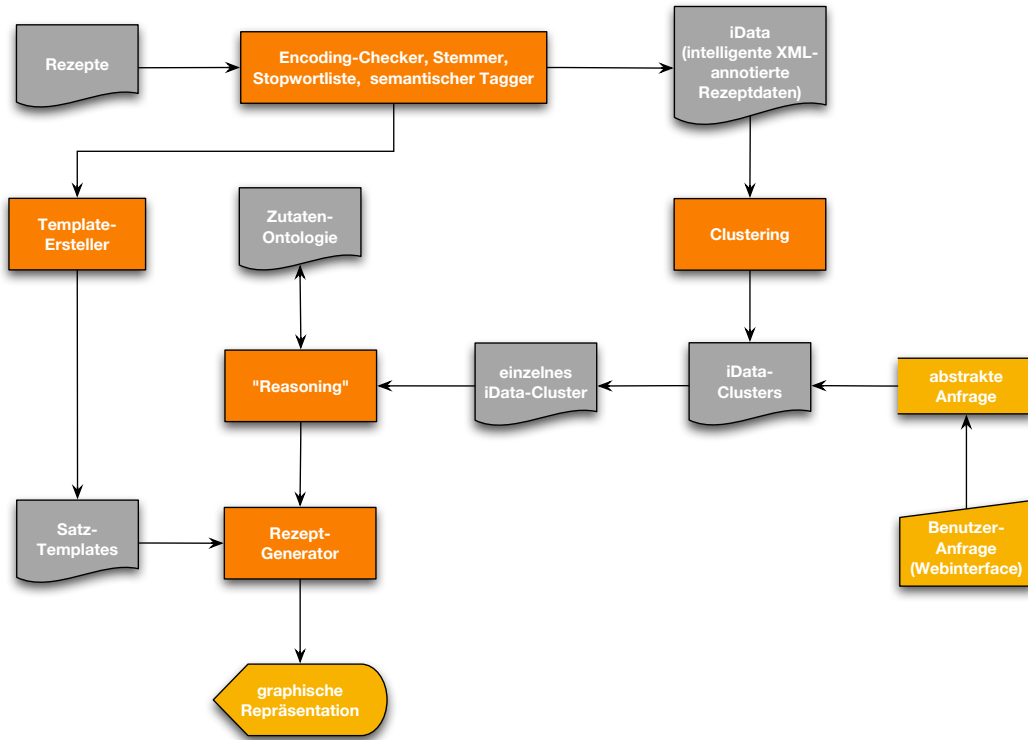
☘ Wiederverwendung von Plain

☘ Lemmatisieren ist besser als stemmen ;-)

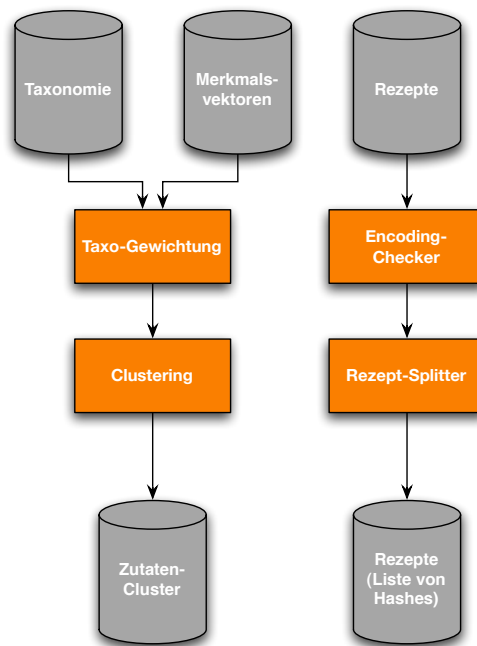


# Systemüberblick





- ★ Ani
- ★ Niki
- ★ Olga
- ★ Steffi



### MealMaster-Format

- sehr verbreitet

- aber durchaus nicht immer einheitlich!!!

- viele Pakete, die wir verwenden können

- <http://kochbuch.unix-ag.uni-kl.de/>

### Rezept-Corpus

- 100 Rezepte aus der italienischen Küche

- Analyse von Hand (Machbarkeit unserer Ideen?) und maschinell (Extraktion der Nomen etc.)



MMMMM----- 'Kalorio Vl.22' (reg.) nach Meal-Master

Title: Gnocchi di ricotta  
 Categories: Kaesegericht  
 Servings: 4 Portionen

2 Eier  
 1 Eigelb  
 300 g Ricotta  
 75 g Parmesan  
 1/2 ts Salz  
 1 Ms Muskat  
 Pfeffer  
 50 g Mehl  
 50 g Griess

MMMMM-----Zum Servieren-----

3 tb Butter  
 Salbeiblaetter

MMMMM-----QUELLE-----

- Betty Bossy Kochbuch  
 - Erfasst von Ralph Liebhaus

Die Eier und das Eigelb in eine Schuessel geben und verklopfen, dann stueckweise den Ricotta beigeben.  
 Parmesan dazugeben und mit Salz, Muskat und Pfeffer wuerzen. Mehl und Griess vermischen zu den  
 anderen zutaten geben und alles zu einer Masse verkneten, eine halbe stunde ruhen lassen.

Nun einen Topf mit Salzwasser zum kochen bringen. Die Ricotta-Masse in einen kalt ausgespuelten Spritzsack mit glatter Tuelle von ca. 1 1/2 cm Durchmesser fuellen, dann ca. 2 cm lange Stuecke heraus druecken und mit den Messerruecken abstreifen, direkt ins leise kochende Wasser fallen lassen. Die Gnocchi portionenweise ziehen lassen bis sie an die Oberflaeche steigen, herausnehmen und abgetropft in einer vorgewaermten Schuessel zugedeckt warm halten.

Etwas Parmesan zwischen die Gnocchiportionen streuen. Fuer den Service Butter in einer Pfanne auslassen und die in Streifen geschnittenen Salbeiblaetter dazugeben. Kurz anduensten, die Gnocchi in vorbereitete Teller geben und die Salbei-Butter daruebergiessen.

:Stichwort : Kaese  
 :Stichwort : Italien  
 :Erfasser : Ralph Liebhauser  
 :Erfasst am : 18.02.2001  
 :Letzte Aender. : 28.02.2001  
 :Quelle : Betty Bossy Kochbuch

MMMMM

- ❁ hat sich über die Zeit dynamisch entwickelt, deswegen mehrere Formate
- ❁ Inkonsistenz:
  - ❁ Formatmuster nicht immer durchgehalten
  - ❁ teilweise Mailadressen, Namen, Erfassungsdaten etc. in Rezepttext oder Zutatenliste
  - ❁ Zutaten enthalten Verarbeitungsschritte
    - ❁ 2 kleine Zwiebeln; fein gehackt
  - ❁ subjektive Anmerkungen
    - ❁ 2 kleine Zwiebeln (am besten rote)
  - ❁ Zutaten in verschiedenen Sprachen
  - ❁ Tippfehler



- ❁ Encoding der Rezept-Dateien überprüfen bzw. angleichen
  - ❁ UTF-8
- ❁ Splitten der Rezepte
  - ❁ title, category, yields, ingreds, data, keywords
- ❁ Datenformat der Rezepte
  - ❁ Array von Hashes von Arrays (von Arrays)
  - ❁ liegen in Datei vor (marshalled)



- ☘ KEIN Clustering von Rezepten zur Vorauswahl aus großem Corpus
- ☘ SONDERN Clustering von Zutaten
  - ☘ ermöglicht Ersetzung von Zutaten (alles in einem Cluster ist austauschbar)
  - ☘ basierend auf dem k-means Algorithmus
    - ☘ Cluster werden durch ein ausgewähltes Element (Zentroid) bestimmt
    - ☘ Initialisiert mit zufällig ausgewählten Zutaten, die einen Mindestabstand haben
    - ☘ Ordne jede Zutaten dem nächsten Zentroiden zu  
(mit Hilfe des Abstandmaßes Dice)  $\frac{v_1 * v_2}{v_1^2 + v_2^2}$
    - ☘ Berechne die Zentroiden neu
    - ☘ Iteriere, solange sich die Cluster verändern



- ☘ Beispiele für Cluster
  - ☘ Rahmfrischkäse, Quark, Ricotta, Mascarpone, Hüttenkäse
  - ☘ Steinpilze, Pfifferlinge, Austernpilze, Shiitake
  - ☘ Mais, Bohnen, Erbsen, Kichererbsen, Linsen, Sojabohnen



### besteht aus

258 Zutaten extrahiert aus MealMaster-Corpus

94 Klassen (42 im Projekt modelliert)

z.B. Gemüse, Fleisch, Nudeln, Fisch, Käse

nicht modelliert: Alkohol, Farbstoffe, Früchte, Gewürze und Kräuter

```
#GEMUESE
:ZWIEBELN
Zwiebel
Porree
Schalotte
Schnittlauch
Perlzwiebel
Gemüsezwiebel
Frühlingszwiebel
Lauch
:TOMATEN
Tomate
Kirschtomate
Eiertomate
Fleischtomate
```

### Gewichtung:

mit Hilfe einer manuell erstellten Gewichtungstabelle

33 Merkmale für jede Zutat (meist binär)

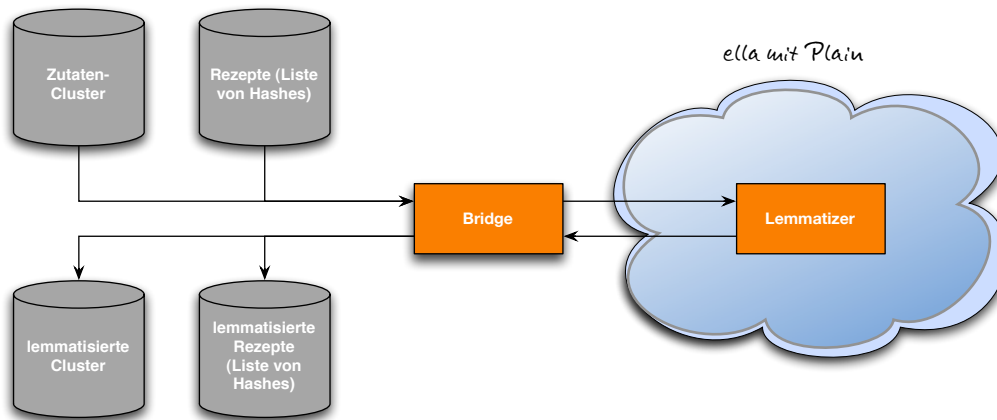
kochbar, bratbar, schneidbar, hackbar, reibbar, pürierbar, passierbar, schälbar, roh\_essbar, vegetarisch, faserig, pulver, mehlig, blättrig, füllbar, süß, salzig, sauer, scharf, bitter, fett, wasser, eiweiß, kohlenhydrate, dichte, zustand, prozent\_vom\_hauptgericht, rot, weiß, kuh, schaf, rotschmiere, blauschimmel, weißschimmel

distinktive Merkmale für einzelne Klassen extrahiert

```
class HARTKAESE < KAESE
  def initialize()
    elem=KAESE.new()
    @@reibbar=1.0
    @@schneidbar=1.0
  end
end

class RIND < FLEISCH
  def initialize()
    elem=FLEISCH.new()
    @@rot=1.0
    @@hackbar=1.0
  end
end
```

Rückgabe: normalisierter Vektor mit Gewichtung (für jede Zutat)



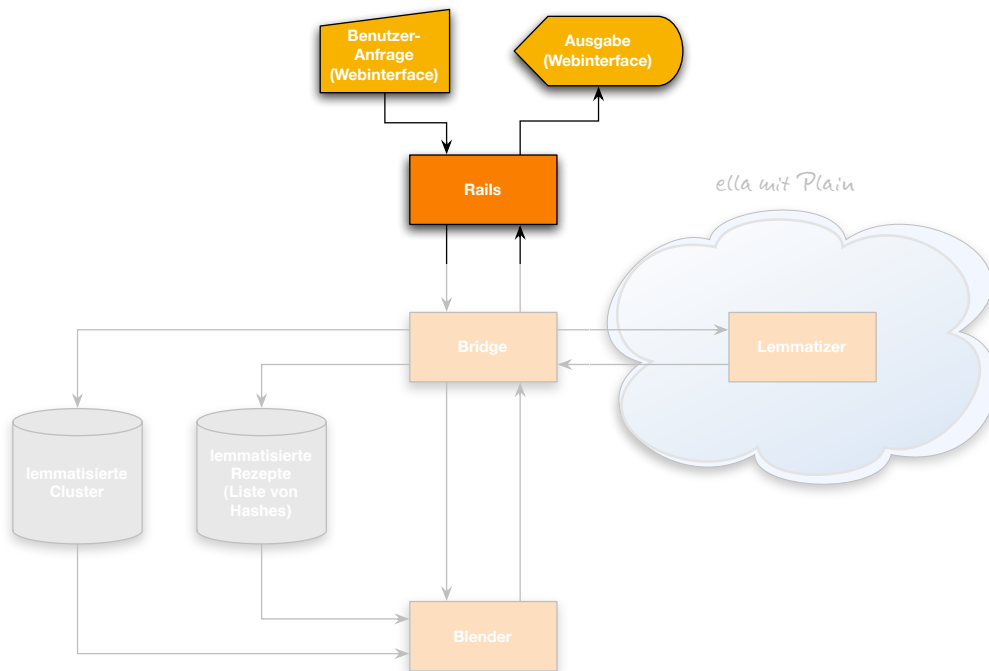
### Offline

#### Erweiterung der Vorverarbeitung

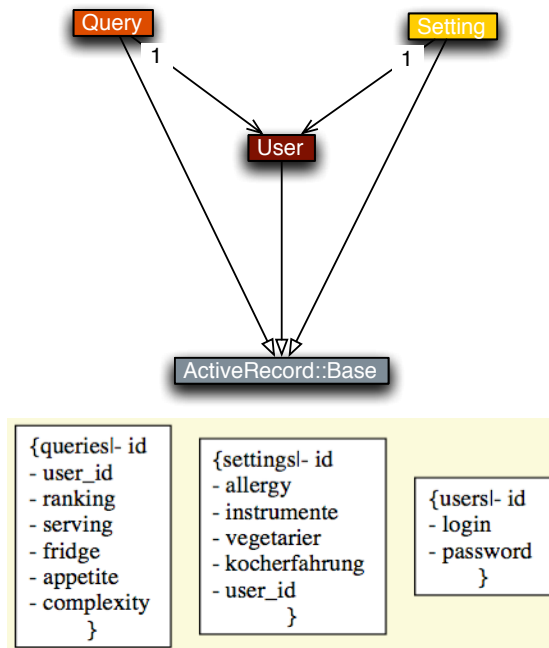
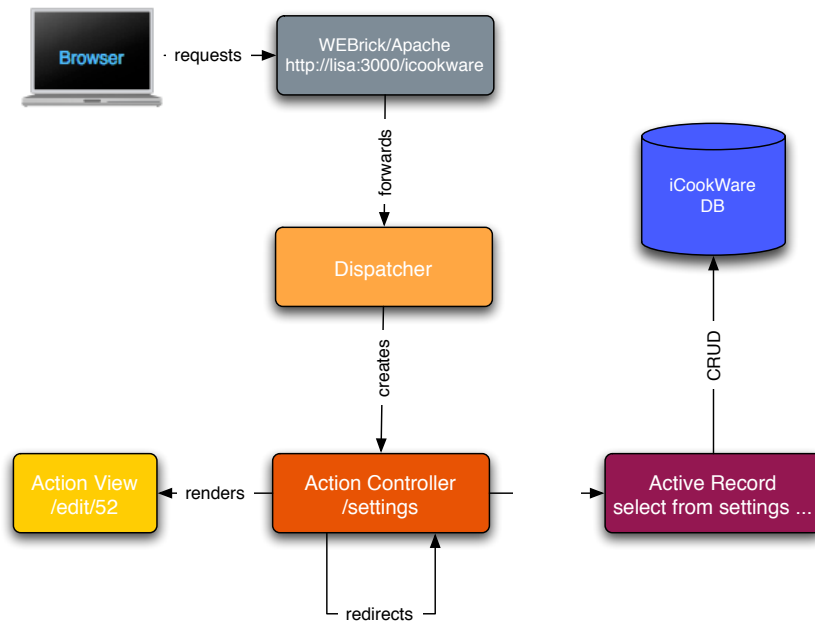
- Lemmatisierung der Rezepte und Zutaten-Cluster
- jeweils gespeichert in Datei (marshalled)

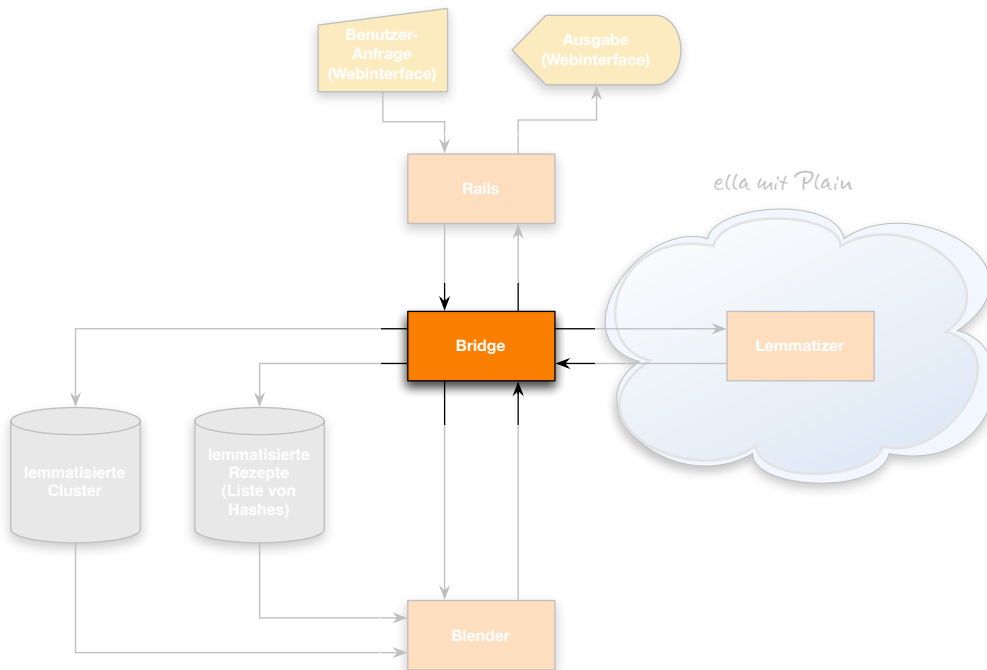
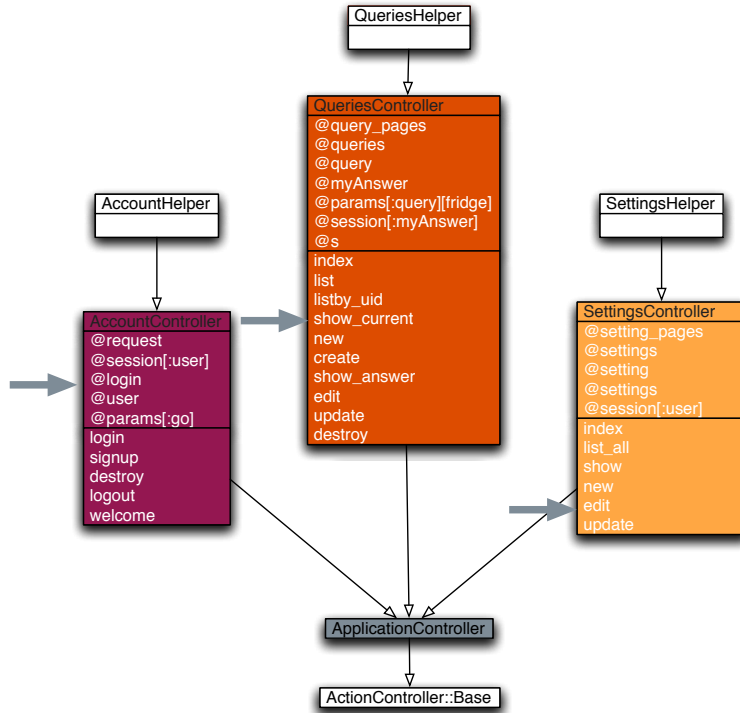
### Online

#### Vermittlung zwischen Webinterface (Rails), Lemmatizer und Blender



- 🍷 “Rails is a full-stack, open-source web framework in Ruby for writing real-world applications with joy and less code ...”
- 🍷 DRY - Don't Repeat Yourself
- 🍷 MVC - Model View Controller

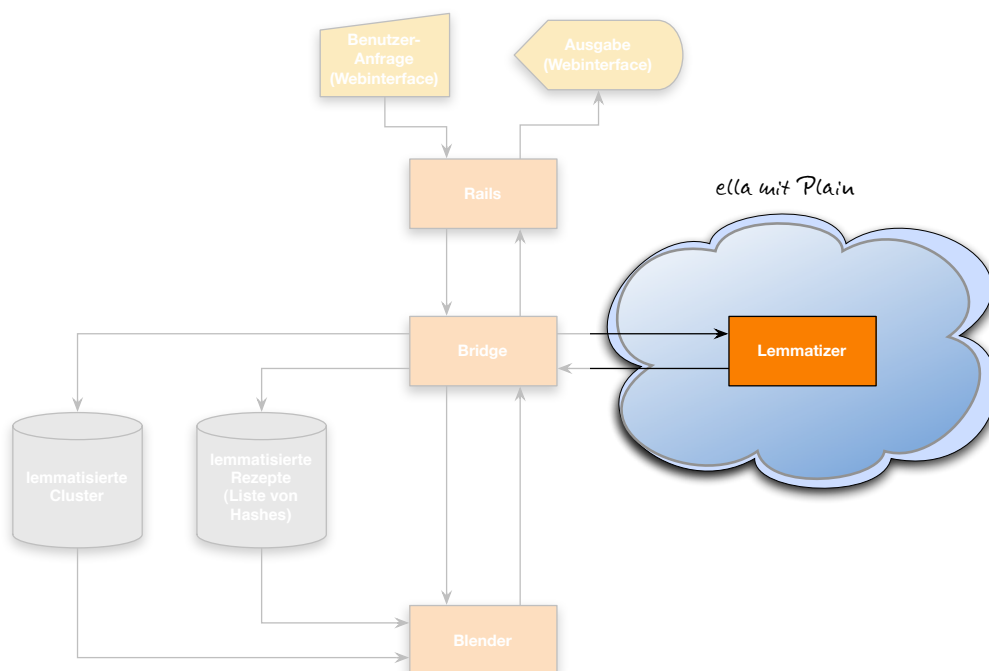






### 🍷 Online

- 🍷 Benutzeranfrage in Rails -> Zutaten an Bridge
- 🍷 Bridge schickt Zutaten an Lemmatizer und bekommt Lemmata zurück
- 🍷 schickt lemmatisierte Anfrage an Blender (zusammen mit den lemmatisierten Rezepten und Clustern)
- 🍷 Blender gibt Ergebnisse an Bridge und diese an Rails zurück -> Webseite

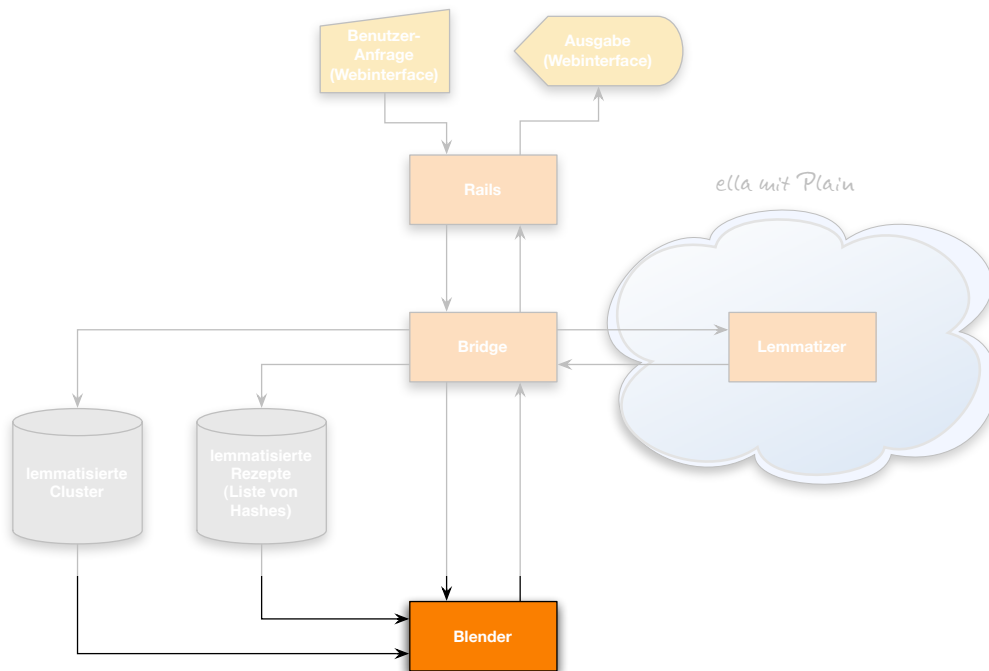


- Gegenstück der Bridge auf ella
  - startet TCP Server
  - wenn Anfrage von Client (Bridge) kommt,
    - Aufruf des Plain Lemmatizer
    - Ergebnis an Bridge zurückschicken



- Plain-Version von Heinz-Detlev Koch
  - geschrieben in C
  - Plain API
  - GUI für die Erstellung von Lexikon-Dateien, die Plain zur Laufzeit benutzt
- Lexikon
  - extrahierte Zutaten aus Rezepten + zusätzliche Liste von uns
  - handcodiert für Plain (CARDL-Forms)
- Problem: Plain läuft nicht ohne weiteres unter Mac OS X wie der Rest des Projektes
  - Lösung: Externes Modul und Kommunikation über Bridge (TCP Sockets)





- ❁ **Blender: "mixt" alles zusammen:**
  - ❁ Anfrage (Zutaten)
  - ❁ vorhandene Rezepte
  - ❁ ersetzbare Zutaten aus Cluster
- ❁ **Gibt es ein Rezept, das zu den Zutaten in der Anfrage passt?**
- ❁ **Wenn nicht: ersetze!**
- ❁ **Einstellbare Parameter:**
  - ❁ Anzahl der Zutaten, die ersetzt werden dürfen
  - ❁ Prozentsatz der Zutaten eines Rezepts, die mit der Anfrage übereinstimmen müssen
- ❁ **Gibt Rezeptvorschläge zurück**

# Zusammenfassung



### 🍷 Änderung des ursprünglichen Konzeptes

- 🍷 Webinterface (Rails)
- 🍷 Taxonomie für Gewichtung
- 🍷 Clustering der Zutaten
- 🍷 Lemmatizing mit Plain
- 🍷 Zutatenersetzung

### 🍷 Sprachen

- 🍷 Ruby
- 🍷 C



**Danke für Eure  
Aufmerksamkeit !**

