

Extracting recipe ingredients from cookbooks

Der Beginn einer Master-Arbeit
von
Torsten Knauf

1 Hintergrund der Arbeit

- Was und wieso
- Digitalisierung
- Das Tagging

2 Ansätze

3 Literatursuche

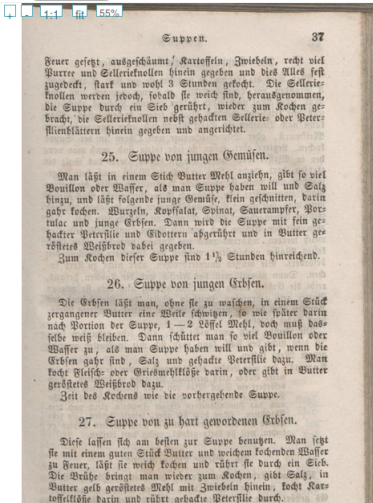
Was und wieso

- Was?
 - Zutaten der Rezepte automatisch extrahieren
- Wieso?
 - Ernährungswissenschaften
 - Soziologische Forschung wie z.B. Globalisierung, Wohlstand, ...

Digitalisierung

(Deutsches Textarchiv 2016)

Bild: 0075 : 37 << vorherige Seite



nächste Seite >>

Suppen.

Feuer gesetzt, ausgeschäumt, Kartoffeln, Zwiebeln, recht viel Purree und Sellerieknollen hinein gegeben und dies Alles fest zugedeckt, stark und wohl 3 Stunden gekocht. Die Sellerieknollen werden jedoch, sobald sie weich sind, herausgenommen, die Suppe durch ein Sieb gerührt, wieder zum Kochen gebracht, die Sellerieknollen nebst gehackten Sellerie- oder Petersilienblättern hinein gegeben und angerichtet.

25. Suppe von jungen Gemüsen.

Man läßt in einem Stück Butter Mehl anziehen, gibt so viel Bouillon oder Wasser, als man Suppe haben will und Salz hinzu, und läßt folgende junge Gemüse, klein geschnitten, darin gahr kochen. Wurzeln, Kopfsalat, Spinat, Sauerampfer, Portulac und junge Erbsen. Dann wird die Suppe mit fein gehackter Petersilie und Eidottern abgerührt und in Butter gebröstetes Weißbrod dabei gegeben.

Zum Kochen dieser Suppe sind 1½ Stunden hinreichend.

26. Suppe von jungen Erbsen.

Die Erbsen läßt man, ohne sie zu waschen, in einem Stück zergangener Butter eine Weile schmoren, so wie später darin nach Portion der Suppe, 1—2 Löffel Mehl, doch muß das selbe weiß bleiben. Dann schüttet man so viel Bouillon oder Wasser zu, als man Suppe haben will und gibt, wenn die Erbsen gahr sind, Salz und gehackte Petersilie dazu. Man kocht Fleisch- oder Griesmehlkloße darin, oder gibt in Butter gebröstetes Weißbrod dazu.

Zeit des Kochens wie die vorhergehende Suppe

```
<div n="3">
  <head>25. Suppe von jungen Gemü&#x017F;en.</head>
  <lb/>
  <p>Man läßt in einem Stuch Butter Mehl anzieh'n, gibt &#x017F;o viel<lb/>
    Bouillon oder Wa&#x017F;&#x017F;er, als man Suppe haben will und Salz<lb/>
    hinzu, und läßt folgende junge Gemü&#x017F;e, klein ge&#x017F;chnitten, darin<lb/>
    gahr kochen. Wurzeln, Kopf&#x017F;alat, Spinat, Sauerampfer, Por-<lb/>
    tulac und junge Erb&#x017F;en. Dann wird die Suppe mit fein ge-<lb/>
    hackter Peter&#x017F;ilie und Eidottern abgerührt und in Butter ge-<lb/>
    rö&#x017F;tetes Weißbrod dabei gegeben.</p> <lb/>
  <p>Zum Kochen die&#x017F;er Suppe &#x017F;ind 1½ Stunden hinreichend.</p>
</div>
```

```
<div n="3">
  <head>25. Suppe von jungen Gemü&#x017F;en.</head>
  <lb/>
  <p>Man läßt in einem Stich Butter Mehl anziehen, gibt &#x017F;o viel<lb/>
    Bouillon oder Wa&#x017F;&#x017F;er, als man Suppe haben will und Salz<lb/>
    hinzu, und läßt folgende junge Gemü&#x017F;e, klein ge&#x017F;chnitten, darin<lb/>
    gahr kochen. Wurzeln, Kopf&#x017F;alat, Spinat, Sauerampfer, Por-<lb/>
    tulac und junge Erb&#x017F;en. Dann wird die Suppe mit fein ge-<lb/>
    hackter Peter&#x017F;ilie und Eidottern abgerührt und in Butter ge-<lb/>
    rö&#x017F;tetes Weißbrod dabei gegeben.</p> <lb/>
  <p>Zum Kochen die&#x017F;er Suppe &#x017F;ind 1½ Stunden hinreichend.</p>
</div>
```

```
<cue:recipe type="Suppen." rcp-id="B-25">
  <head>Suppe von jungen Gemüsen.</head>

  <p>Man läßt in einem Stich Butter Mehl anziehen, gibt so viel Bouillon oder Wasser, als
    man Suppe haben will und Salz hinzu, und läßt folgende junge Gemüse, klein
    geschnitten, darin gahr kochen. Wurzeln, Kopfsalat, Spinat, Sauerampfer, Portulac und
    junge Erbsen. Dann wird die Suppe mit fein gehackter Petersilie und Eidottern
    abgerührt und in Butter geröstetes Weißbrod dabei gegeben. </p>

  <p>Zum Kochen dieser Suppe sind 1½ Stunden hinreichend.</p>

</cue:recipe>
```

Das Tagging

- **cueML** (culinary editions markup language)
 - Erweiterung von **TEI** (Text Encoding Initiative)
 - Erweitert mit **<https://schema.org/Recipe>**

...nehme man 1 Eßlöffel Mehl ...

...nehme man 1 Eßlöffel

<cue:recipeIngredient quantity="0.01" unit="kg">Mehl</cue:recipeIngredient>

...

Achtung

- *"[Man] legt ihn in einen eisernen Topf, darauf 1/4 Pfund in Scheiben geschnittenen rohen Schinken oder Sommerwurst"* → `<recipeIngredientAlt>`
- *"Wie die vorhergehende, nur mit weiß gebranntem Mehl und Bouillonzubereitet, die Nelken bleiben weg."*
- *"Fehlt ihm die gewünschte Süße, so wird zeitig ein Stück Zucker dazu gethan so wie beim Anrichten die Brühe mit etwas Kartoffelmehl gebunden gemacht."*
→ `<recipeOptionalIngredient>`
- Scorzoner-Wurzeln, Petersilienwurzel,
8 Pfennig Weißbrot verstehen, Loth, Quart

Ansätze

- Was muss die Lösung alles können?
 - Aus reinem Text alle Zutaten eindeutig erkennen
 - Zutaten erkennen (**Lemmatisierung**) und Zuordnen der Quantitäten und Einheiten zu den entsprechenden Zutaten (**Structure prediction for NLP**)
 - Unterscheidung ob eine Zutat verwendet werden soll, nicht verwendet werden soll, optional oder eine Alternative ist (**Klassifizierungs-Probleme**)

Literatursuche



Deutsches Textarchiv, Berlin-Brandenburgische Akademie der Wissenschaften (2016). *Davidis, Henriette: Praktisches Kochbuch für die gewöhnliche und feinere Küche. 4. Aufl. Bielefeld, 1849.* URL:
http://www.deutschestextarchiv.de/book/show/davidis_kochbuch_1849.

NLP nicht selber machen!