

# ReciPic: A Tool for Generating Infographic from Recipe Procedure Text

Lechang Zhang\*     Itoh Takayuki†

**Abstract.** In this paper, we propose a tool for generating infographic from recipe procedure text, to supplement the pure text information on the recipe websites. This tool accepts one step text of the cooking procedures, and then separate the text stream into appropriate units. After that, it extracts the food material, cooking tool, cooking action segmentations and label them with F, T, Ac (F: Food material, T: Cooking tool, Ac: Cooking action) respectively by users. Finally, it matches the words with prepared pictures and place them with appropriate positions in an infographic and represents it to users.

## 1 Introduction

In recent years, more and more people choose to use recipe websites to look for recipes, cooking advices or to share their own cooking ideas. The most popular recipe websites in Japan is Cookpad, which has over 50 million monthly users in Japan, and an additional 42 million monthly users across 74 countries, with 3 million recipes. In order to support such a large user community and improve their using satisfaction, it is necessary to analyze the recipes in recipe websites.

A recipe in recipe websites simply contains the recipe name, ingredients, and cooking procedures. Among these components, cooking procedure is the key part that users will follow. However, they are often displayed by pure text in the websites. It is difficult for users to grasp the valuable information and have a big picture about the food materials and cooking methods at a glance.

In this paper, we propose ReciPic, an interactive visualization tool to make the cooking procedure text more understandable. Our tool allows users to enter recipe procedure step text, and then transforms it into the corresponding infographic.

## 2 Method

### 2.1 Outline

Our research developed a tool called ReciPic, which is a web application that allows users to enter recipe procedure step text, then perform word segmentation on the text and return the useful information that we use to generate infographic. Next, we ask users to edit the information to improve the accuracy of generating infographic. Finally, we generate the infographic and present it to users.

### 2.2 Text analysis

In order to generate infographic from the input recipe procedure step text, we first analyze the input text and segment the text into useful parts. There are two popular Japanese text analysis tools. One is called kytea and the other is called mecab. We chose mecab for the following reasons.

- Mecab has a dictionary called mecab-ipadic-NEologd that is constantly updated, while kytea has not been updated for many years.
- Mecab can be used by python, while kytea can only be used in terminal.

A text analysis result is shown below:

**Input:** 鍋に水、本だし、料理酒、みりん、濃口醤油、上白糖を入れて、手羽元を入れて強火にかける。(Put water, hondashi, cooking sake, sweet sake, dark soy sauce, caster sugar and the chicken wings in the pot, and set the strong fire.)

**Output:** 鍋 (pot) 水 (water) 本だし (hondashi) 料理酒 (cooking sake) みりん (sweet sake) 濃口醤油 (dark soy sauce) 上白糖 (caster sugar) 入

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* Ochanomizu University

† Ochanomizu University

れ (put) 手羽元 (chicken wings) 入れ (put) 強火 (strong fire) かける (set)

### 2.3 User modification

After text analysis, we classify the segmented words into appropriate categories in order to make ReciPic have a rough understanding of them. The recipe terms definition is shown in Table 1. We ask the users to label the result with the

Table 1. Recipe terms definition.

Recipe term	Tag	Examples
Food	F	egg, rice
Tool	T	fry pan, rice cooker
Action of chef	Ac	put, cut

recipe tags.

After text analysis, we also lost the relationships between words. e.g. 水 (water), 本だし (hondashi), 料理酒 (cooking sake), みりん (sweet sake), 濃口醤油 (dark soy sauce), 上白糖 (caster sugar) and 手羽元 (chicken wings) is related to tool 鍋 (pot) and action 入れ (put). Therefore, we ask users to mark the word relationships, the detailed steps are as follows:

- Mark the tools and actions with their tag and occurrence order.
- Mark the food with its related tools and actions.

The result is shown below:

**User modification:** 鍋/T/T0 水/F/T0Ac0 本だし/F/T0Ac0 料理酒/F/T0Ac0 みりん/F/T0Ac0 濃口醤油/F/T0Ac0 上白糖/F/T0Ac0 入れ/Ac/Ac0 手羽元/F/T0Ac1 入れ/Ac/Ac1 強火/T/T1 かける/Ac/Ac2

Then we ask the user to submit the result.

### 2.4 Infographic generating

In order to generate the infographics, we prepared an image database which contains the illustrations of food and tools. The database consists of two collections, one is food, and the other is tool. In each collection, it contains different objects that have two properties — name and image address — to indicate the name and image of a certain food or tool.

ReciPic will choose the illustration for each food and cooking tool by their name, and then combine them with certain rules, which are shown in Figure 1.

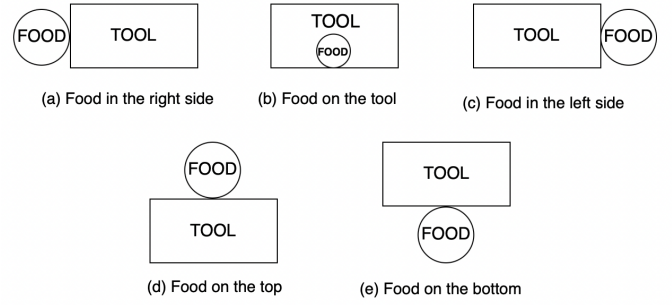


Figure 1. Rules for generating infographics.

## 3 Result

**Input:** 鍋に水、本だし、料理酒、みりん、濃口醤油、上白糖を入れて、手羽元を入れて強火にかける。(Put water, hondashi, cooking sake, sweet sake, dark soy sauce, caster sugar and the chicken wings in the pot, and set the strong fire.)

**a. Text analysis:** 鍋 (pot) 水 (water) 本だし (hondashi) 料理酒 (cooking sake) みりん (sweet sake) 濃口醤油 (dark soy sauce) 上白糖 (caster sugar) 入れ (put) 手羽元 (chicken wings) 入れ (put) 強火 (strong fire) かける (set)

**b. User modification:** 鍋/T/T0 水/F/T0Ac0 本だし/F/T0Ac0 料理酒/F/T0Ac0 みりん/F/T0Ac0 濃口醤油/F/T0Ac0 上白糖/F/T0Ac0 入れ/Ac/Ac0 手羽元/F/T0Ac1 入れ/Ac/Ac1 強火/T/T1 かける/Ac/Ac2

**c. Infographic:**

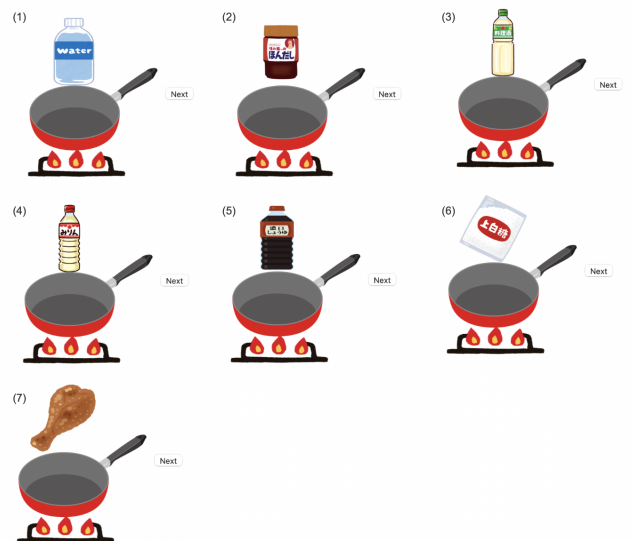


Figure 2. Infographic.