

# Analysis of Traditional Italian Food Recipes: Experiments and Results

---

MARIA TERESA ARTESE\*, GIANLUIGI CIOCCA§, AND ISABELLA GAGLIARDI\*

\*IMATI - CNR (NATIONAL RESEARCH COUNCIL), MILAN, ITALY  
{TERESA,GAGLIARDI}@MI.IMATI.CNR.IT

§UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA, MILAN, ITALY  
GIANLUIGI.CIOCCA@UNIMIB.IT



# CookIT Portal

---

- ❖ CookIT, a web portal related to Italian Traditional Recipes,
- ❖ with the aim to spread the knowledge of Italian food recipes and the Mediterranean diet

## Why a new web site and a dataset have been created?

- ❖ A complete collection of recipes related to Italian culinary tradition, of all Italian regions, with the aim of preserving and disseminating them
- ❖ A controlled dataset, to be able to test new visualization tools and search algorithms for content based retrieval on text and images

# CookIT: Characteristics

---

- ❖ Multimedia Information system, oriented towards traditional Italian recipes, in Italian language
- ❖ Designed as Collaborative tool, to integrate data coming from web on the basis of people real traditions.
- ❖ Innovative navigation, retrieval, and visualization tools
- ❖ Testbed for tools and algorithms for content based retrieval (text and images).

## Ingredienti

**Formaggio (14)** Frattaglie di vitello (2) Friarielli (2) Profita candida (2) Fritti di mare (1) **Funghi (4)** Gallette (3) Galina (1) **Gamberi (4)** Gattuccio (1) Genepi (1) Gherghi di nocca (2) **Ginepro (4)** Glucosio (2) **Grana Padano (9)** Granchio (1) **Grano (3)** **Grappa (6)** Grappe di prugne (1) Guanciale (2) Insalata (2)

Lampredotto (1) **Lardo (8)** **Latte (35)** Latte di cocco (1) Latterino (1) **Lenticchie (3)** **Lievito di birra (21)** **Lievito in polvere (9)** **Lievito madre (3)** **Limoni (47)** Lucido (1) Luganega (1) Lumache (2) Lumachine di Mare (1) **Maggiorana (5)**

**Maiale (12)** Mastrano da saba (1) Mello (2) Mandarini (1)

**Mandorle (20)** Mandorle amare (2) **Manzo (18)** Maraschino (1) Marmellata (1) **Marsala (3)** Marzapane (1) Mascarpone (1) **Melanzane (10)** **Mele (5)** Mendo (1) **Menta (4)** **Miele (11)** Mirio (1) **Mortadella (3)** **Moscadella (2)** MosGatte di forno (1) **Mostarda (1)** **Mozzarella (9)** Mozzarella di bufala (1) Nocciolo (2) **Noce moscata (15)** **Noci (10)**

**Olio di arachidi (15)** Olio di girasole (4) Olio di mais (1) Olio di mandorle (3) Olio di roo (1) **Olio di oliva (29)**

Tag

Categorie

Regioni

Stagionalità

Livello di difficoltà

Tipo di cottura

## Ricette trovate: 297

Icose Grafico 2D **Grafo** Rose dei venti

PER REGIONE PER CATEGORIA PER TAG TUTTI

Use la rotellina del mouse per lo zoom  
Premi CTRL per aprire il dettaglio della ricetta selezionata

● Regioni

● Tag

● Antipasti

● Pane/Lievitati

● Dolci

● Contorni/Inselate

● Piatti Unici

● Primi Piatti

● Secondi Piatti

● Torte salate

● Street Food

● Salse/Sughi

● Marmellata/Conserve

● Bevande



Ingredienti

- TUTTI GLI INGREDIENTI **Abbacchio (2)** **Acciughe (19)**  
**Aceto (28)** **Acqua (65)** **Acqua di fiori di arancio (6)** **Aglio (79)** **Agnello (5)** **Albume (13)**  
**Alchermes (2)** **Alcol (3)** **Alici (4)** **Alloro (24)** **Amarene (1)**  
**Amaretti (6)** **Anguille (1)** **Anice (1)** **Anice (3)** **Arancia (11)**  
**Asparagi (4)** **Astice (1)** **Baccalà (2)** **Barbabietola (1)** **Basilico (25)** **Besciamella (2)** **Bianchetto (3)** **Bicarbonato di sodio (2)**  
**Biete (6)** **Birra (2)** **Biscotti (1)** **Borragine (1)** **Bottarga (1)** **Brodo di carne (12)** **Brodo di pesce (3)** **Brodo vegetale (9)**  
**Burro (59)** **Cacao (11)** **Cacao amaro (1)** **Caciocavallo (7)** **Caffè (2)** **Calamari (2)** **Canditi (4)** **Cannella (19)**  
**Canocchie (1)** **Capone o Lampuga (1)** **Capperi (9)** **Cappone (4)**  
**Capretto (1)** **Capriolo (1)** **Carciofi (4)** **Cardi (2)** **Carne (10)**  
**Carote (28)** **Castagne (2)** **Cavolfiore (3)** **Cavolo cappuccio (1)** **Cavolo verza (1)** **Ceci (4)** **Cedro (5)** **Cernia (1)** **Cetrioli (3)**  
**Chiodi di garofano (10)** **Cicoria (1)** **Cioccolato fondente (9)** **Cipolle (74)** **Cipolle rosse (5)**

Tag

Categorie

Regioni

Stagionalità

Livello di difficoltà

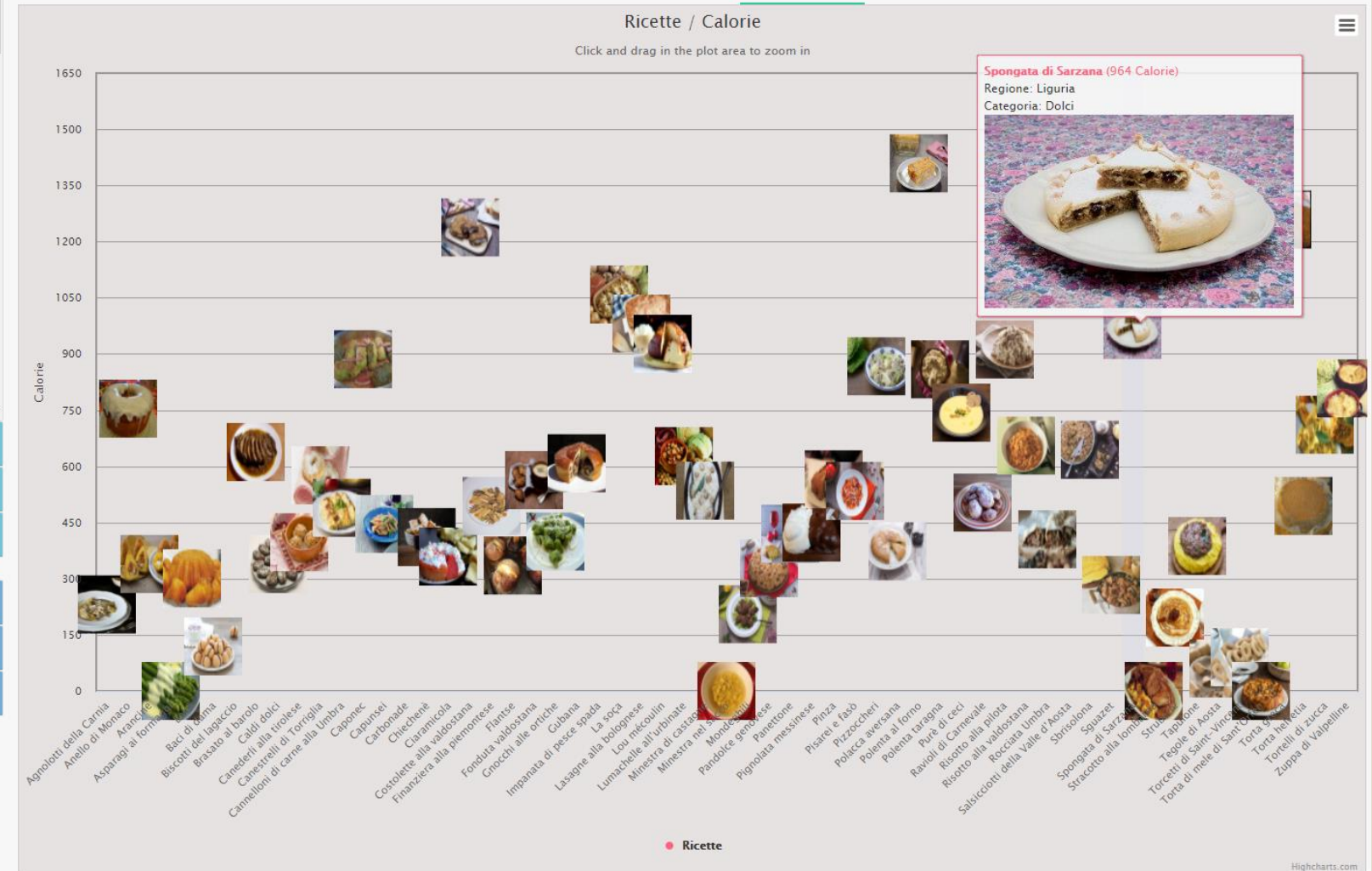
Tipo di cottura

Ricette trovate: 59

Burro

Icone Grafico 2D Grafo Rosa dei venti

RICETTE/CALORIE REGIONI/CALORIE CATEGORIE/CALORIE



# Recipe analysis

---

Aim: To identify the elements of the recipe -> recipe structured document (according to a defined DTD)

Constituent elements of food making process:

- ingredients (flour, oil, almonds, eggplants, ...) and intermediate products
- cooking utensils (bowl, knife, ...)
- actions (cut, mix, join, ...)
  
- duration of actions (let it rest for 30 minutes, ...)
- conditional actions (if the dough is too hard , ... otherwise, ...)
- order in which the actions are executed
- whether the actions are (almost) simultaneously performed

# Recipe analysis: steps

---

For each recipe:

- for each sentence:
  - tokens are identified: text (sentence) is decomposed in a set of terms or words, called tokens
  - tokens are pos tagged: to each word, the corresponding grammatical category (noun, adjective, verb, ...) is associated
  - tokens are dependency parsed: relationship between the head words and other words are defined and graphically represented
  - for each token they are identified:
    - - compound nouns (mainly nouns and adjectives)
    - - compound verbs
    - - expressions of time
    - - conditional expressions

Post processing:

- harmonize compound words/verbs/expression of time
- classify terms (compound words/verbs) as belong/not belonging to the food domain

# Recipe analysis: NLP

---

A problem solved with NLP tools

NLP Steps:

- Tokenization
- Annotation with pos tagging
- Normalization: Lemmatization and stemming
- Dependency parsing.

NLP Tools used (and integrated) for Italian language:

- ❖ SpaCy: used to identify tokens and find rule-based components
- ❖ TreeTagger: used for lemmatizer and POS tagging
- ❖ Stanza: used for dependency parsing



# Recipe analysis: classification

---

Aim: classify a term as belonging/non belonging to food domain

Classifiers used: Logistic Regression and Knn (K=2, 3, 5)

Dataset used for training / testing:

- Created ad hoc
- Wikipedia pages belonging to the 9 root categories + food domain
- Same number of labelled entries of food/non food classes
- Title and abstract
- Single terms / compound terms (n-grams)
- Language: Italian

The methods used here to classify terms are based on word embeddings

Word embedding models:

- Transform terms in vectors, using deep learning techniques, trained on huge corpus
- Semantic representation of terms, able to capture the semantics of words and their context
- Each abstract/title/n-grams is calculated as the average of the vectors of all terms present, except stop-words, and returns a vector
- Pre-trained (word2vec/GloVe) for Italian language

# Recipe analysis: evaluation

Training Set	Class	Precision	Recall	F1 Score	Accuracy
Logistic Regression	Food	0.92	0.93	0.92	0.914
	Non Food	0.91	0.90	0.91	
KNN (k=2)	Food	0.89	0.96	0.92	0.914
	Non Food	0.95	0.86	0.90	
KNN (k=3)	Food	0.92	0.94	0.93	0.921
	Non Food	0.93	0.90	0.91	
KNN (k=5)	Food	0.89	0.94	0.91	0.900
	Non Food	0.92	0.85	0.89	

Test Set	Class	Precision	Recall	F1 Score	Accuracy
Logistic Regression	Food	0.91	0.91	0.91	0.896
	Non Food	0.88	0.89	0.88	
KNN (k=2)	Food	0.86	0.95	0.90	0.881
	Non Food	0.92	0.80	0.86	
KNN (k=3)	Food	0.88	0.92	0.90	0.884
	Non Food	0.89	0.84	0.86	
KNN (k=5)	Food	0.86	0.93	0.89	0.876
	Non Food	0.90	0.81	0.85	

# Recipe analysis: evaluation

Human-created lists of n-gram terms containing ingredients, tools, and actions, taken from the web and manually integrated with missing terms.

	Class	Precision	Recall	F1 Score	Accuracy
Logistic Regression	Food	0.81	0.97	0.88	0.820
	Non Food	0.87	0.45	0.60	
KNN (k=2)	Food	0.75	0.99	0.85	0.757
	Non Food	0.85	0.20	0.33	
KNN (k=3)	Food	0.75	0.97	0.85	0.754
	Non Food	0.76	0.23	0.35	
KNN (k=5)	Food	0.75	0.98	0.85	0.750
	Non Food	0.79	0.19	0.31	

# Recipe analysis: results on Caponata di Melanzane

Per preparare la caponata cominciate lavando le melanzane e privandole delle due estremità. Quindi tagliatele a cubetti di circa 1,5 cm per lato, tagliandole prima in senso verticale e poi in orizzontale. Su un vassoio sistemate uno scolapasta e versatevi i cubetti di melanzana poi cospargete con sale grosso e adagiate sulla superficie un peso, come per esempio delle sfere di ceramica riposte in un recipiente, e lasciate così per un'ora. Questa operazione è necessaria per far fuoriuscire l'acqua di vegetazione delle melanzane e farle perdere un po' di amaro. Intanto, dopo aver lavato i gambi di sedano, trasferiteli su un tagliere e tagliateli a tocchetti di circa 1 centimetro dopodichè versateli in una pentola con acqua bollente e salata e, dopo 3 - 4 minuti di cottura, scolateli e lasciateli asciugare su un canovaccio pulito. Mondate le

Ingredients are highlighted in green, utensils in cyan, and actions in yellow. Light gray terms are those discarded by the classifier

Structured recipe

```
<?xml version="1.0" ?>
<RECIPE>
  <INGREDIENT INGR_ID="00-02i">
    < INGR_NAME>melanzane di due estremita</INGR_NAME> </INGREDIENT>
  <INGREDIENT INGR_ID="01-03i"><INGR_NAME>cubetti</INGR_NAME> </INGREDIENT>
  <INGREDIENT INGR_ID="02-07i"><INGR_NAME>scolapasta</INGR_NAME> </INGREDIENT>
  <INGREDIENT INGR_ID="02-08i"><INGR_NAME>cubetti di melanzana</INGR_NAME> </INGREDIENT>
  <INGREDIENT INGR_ID="02-09i"><INGR_NAME>sale grosso</INGR_NAME> </INGREDIENT>
  <INGREDIENT INGR_ID="02-10i"><INGR_NAME>superficie</INGR_NAME>
  <INGREDIENT INGR_ID="03-13i"><INGR_NAME>acqua di vegetazione di melanzane </INGR_NAME></INGREDIENT>
  <INGREDIENT INGR_ID="04-14i"><INGR_NAME>tagliere</INGR_NAME> </INGREDIENT>
  ...
  <ACTION ACTION_ID="00-03a"><ACTION_NAME>lavando</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="00-02i"><INGR_NAME>melanzane di due
    estremita</INGR_NAME></INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="00-04a"><ACTION_NAME>privando</ACTION_NAME></ACTION>
  <ACTION ACTION_ID="01-05a"><ACTION_NAME>tagliatele</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="01-03i"><INGR_NAME>cubetti</INGR_NAME> </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="02-08a"><ACTION_NAME>sistemate</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="02-07i"><INGR_NAME>scolapasta</INGR_NAME> </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="02-09a"><ACTION_NAME>versatevi</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="02-08i"><INGR_NAME>cubetti di melanzana</INGR_NAME>
    </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="02-10a"><ACTION_NAME>cospargete</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="02-09i"><INGR_NAME>sale grosso</INGR_NAME>
    </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="02-11a"><ACTION_NAME>adagiate</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="02-10i"><INGR_NAME>superficie</INGR_NAME> </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="03-13a"><ACTION_NAME>far fuoriuscire</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="03-13i"><INGR_NAME>acqua di vegetazione di melanzane</INGR_NAME>
    </INGREDIENT></PRE></ACTION>
  <ACTION ACTION_ID="04-14a"><ACTION_NAME>trasferiteli</ACTION_NAME>
    <PRE><INGREDIENT INGR_ID="04-14i"><INGR_NAME>tagliere</INGR_NAME> </INGREDIENT></PRE></ACTION>
  ...
  <ORDER_RELATION ACTION_IDPrec="0" ACTION_ID Succ="00-01a"> </ORDER_RELATION>
  <ORDER_RELATION ACTION_IDPrec="00-01a" ACTION_ID Succ="00-02a"> </ORDER_RELATION>
  <ORDER_RELATION ACTION_IDPrec="00-02a" ACTION_ID Succ="00-03a"> </ORDER_RELATION>
  <ORDER_RELATION ACTION_IDPrec="00-03a" ACTION_ID Succ="00-04a"> </ORDER_RELATION>
  ...
</RECIPE>
```

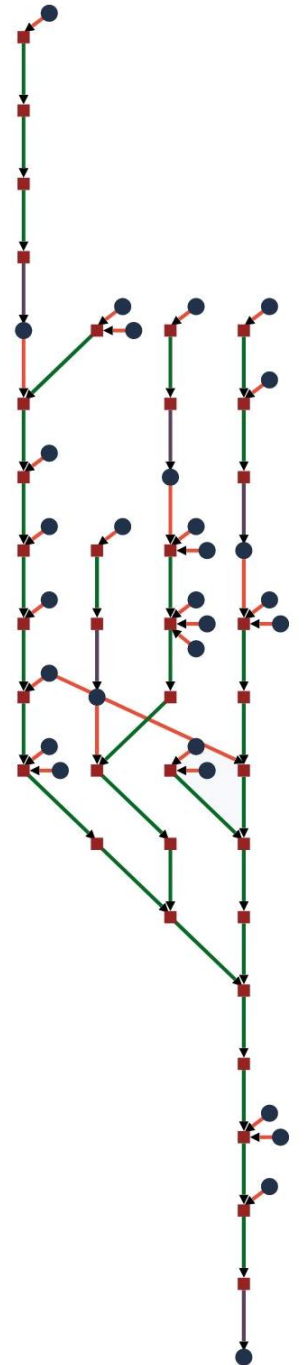
# Recipe visualization application

Results of Recipe analysis are:

- Ingredients, tools, and intermediate products (nodes)
- Actions (arcs)
- Connected to form a graph -> Recipe graph

Application

- To visualize and browse the recipe
- To navigate and edit the recipe -> GT creation
- As a recipe analysis tool
  - Concurrent actions analysis
  - Recipe complexity analysis
  - Recipes comparison
  - ...



# Recipe visualization application

**RECIPE CONTROL PANEL**

Grafico

MENÙ

- Tabelle Locali >
- Liste Globali >
- Processate Prev

**Legenda**

- :ingrediente -> azione
- :azione -> ingrediente
- :azione -> azione
- - - :azione -> condizione
- :ingrediente
- :azione
- :azione con condizione

## caponatal

Download File Updated   XML File Updated

Refresh

**Id : 1a**  
Nome : lava le melanzane  
Durata :

**Id : 2a**  
Nome : priva delle due estremità  
Durata :

**Id : 3a**  
Nome : taglia in senso verticale  
Durata :

**Id : 4a**  
Nome : taglia in senso orizzontale  
Durata :

**Id : 1i**  
Nome : melanzana  
Quantità : 1Kg

**Testo della ricetta**

Per preparare la caponata cominciate lavando le melanzane e privandole delle due estremità. Quindi tagliatele a cubetti di circa 1,5 cm per lato, tagliandole prima in senso verticale (1) e poi in orizzontale (2). Su un vassoio sistemate uno scolapasta e versatevi i cubetti di melanzana (3) poi cosparrete con sale grosso (4) e adagiate sulla superficie un peso, come per esempio delle sfere di ceramica rioste in un recipiente. (5) e lasciate

**Modifica Testo**

# Conclusions and Future Works

---

Tools for automatically analyzing and modeling cooking procedure

Automatic identification of ingredients, kitchen utensils, and actions:

- NLP methods and tools
- ML Classification as belonging/not belonging to the food domain
- Preliminary results: satisfactory for ingredients, tools and duration
  - Actions and their order require further refinement
  - Classification does not generalize well on n-gram terms

Recipe visualization tool

We plan to:

- Integrate these tools to query CookIT and visualize results by recipe pattern, methods, number of steps,...
- study LSTM and RCNN deep learning frameworks
- develop more analysis tools to compare recipes, ...
- automatically gain information on the features of recipes by analyzing their graphs ...

# Thank you

---

Maria Teresa Artese            teresa@mi.imati.cnr.it  
Isabella Gagliardi            gagliardi@mi.imati.cnr.it  
IMATI - CNR (National Research Council), Milan, Italy

Gianluigi Ciocca            gianluigi.ciocca@unimib.it  
Università degli Studi di Milano-Bicocca, Milan, Italy

**CookIT || CookIT Online Archive**

Ricette e immagini della cucina tradizionale italiana

<http://arm.mi.imati.cnr.it/cookIT>

