Ontology based modeling of Food Recipe Concepts and Process

Hammad Malik

Matriculation number: 3169270

October 25, 2021

Master Thesis

Computer Science

Supervisors:

Prof. Dr. Jens Lehmann
Dr. Giulio Napolitano

Institut für Informatik III

Rheinische Friedrich-Wilhelms-Universität Bonn

Declaration of Authorship

I, Hammad Malik, declare that this thesis, titled "Ontology based modeling of Food Recipe Concepts and Process", and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work. I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 27.10.2021

Acknowledgements

I would first like to thank Dr. Hajira Jabeen for her continuous support from choosing a topic for my thesis to its completion. Your throughout guidance, insight feedback and supervision pushed me forward and steered me in the direction of the finish line.

I would like to acknowledge the efforts of Mr. Fahad Maqbool and Mr. Saad Razzaq from University of Sargodha. I want to thank you for sharing your domain knowledge, your patient support, and all the time you spent to further my research.

I must express my profound gratitude to my parents and to my spouse for providing me with unfailing support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without them.

In addition, I would like to thank the Protege 2000 and WebVowl community, as a huge part of this research is performed with the help of their tools. Thank you.

Contents

List of Figures

Abstract

The creation of new recipes can be thought of as an art that combines food, science, culinary experience, and expertise. While culinary experts frequently develop innovative recipes, it is known that we are using a tiny proportion of possible food combinations in our recipes. It is humanly and scientifically challenging to find the optimal food combinations that can meet nutritional needs while being appetizing.

Food-for-future desires innovative recipes exploiting novel food combinations while handling the ever-demanding dietary constraints; special food for the sick, for the environment-conscious, for the health-conscious, etc. In order to develop efficient intelligent systems that can understand the dynamics behind the cooking process, a comprehensive ontology is required. An ontology defines a common vocabulary of concepts and their relationships to model a particular domain while making it machine-understandable. In this work, we present the FoodRecipe, an ontology for the cooking process. This ontology makes the recipes machine-understandable and standardizes the representation method for recipes. Additionally, modeling of these concepts allows the possibility to create and generate new recipes computationally.

In the FoodRecipe Ontology, we have identified the core elements after observing the top five hundred recipe websites. We have developed the Ingredients, Instructions, Time, and Quantity, as the main concepts or classes, whereas Source, Nutrition, Rating, Classification, Difficulty, and Utensils are designed as secondary concepts. For better modeling of the concepts, we have identified the relationships among the concepts e.g. Action (is a) Instruction. From the original recipe data, we mapped these concepts to the instances. e.g. Fish (is a) Ingredient, Heating (is a) Action.

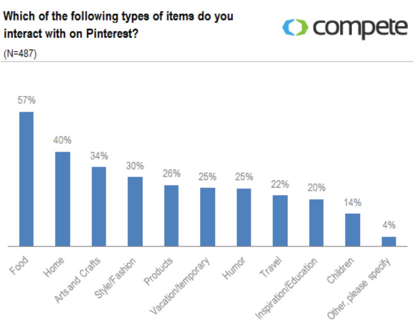
The FoodRecipe ontology allows representing recipes as a connected graph of concepts. An Intelligent system can use these identified concepts and learn the correlations among these concepts from the existing recipes (e.g. potatoes are cut, rice are boiled). This ontology will be further used in evolutionary algorithms like Genetic Programming to represent recipes and to create novel recipes while learning from the existing food combinations and substituting them with related combinations.

Chapter 1

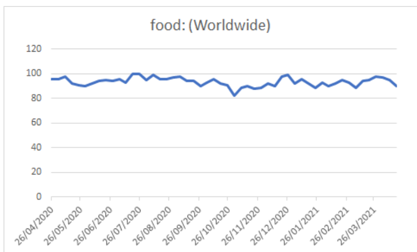
Introduction

The cooking of food is one of the key indicators of civilized life that sets humans apart from other species. Different cultures have developed different tastes, staple foods, and cuisines, based on their cultural and geographical differences. With the advancements in globalization, and availability of most of the ingredients, be it regional ones, across the globe, one can develop novel and specialized recipes to suit individual diet preferences.

In modern times, food researching is becoming more and more online, whether it's searching a website for a food recipe or using mobile services to find a place to eat in real-time. According to Google search trends in the U.S for 2020, the keyword "food near me" has been searched 16.6 Million times. The food interest over time worldwide has always remained at the peak. [3].



(a) Pinterest: Favourite item to interact



(b) Google trends world wide for keyword: food

Figure 1.1: Statistics about recipe

According to think with Google, 59 percent of youngsters (24 to 39 years old) use smartphones to take help during cooking, and the search for "best recipes" on YouTube has increased by 48 percent during the past year. [4]

In another study, by asking users about their favorite interactive content, it was found that food is the user's favorite interactive content, and 57 percent of users like food-related content interaction.

These statistics are a clear indication that a large number of people care about cooking at home and trying new recipes.

1.1 Motivation

Being a food lover, I also explore new recipe ideas but due to busy routine, and time limitations this hobby is for the weekends only. But the trend has changed now. During the past year of lockdown and quarantine, millions around the world were forced into the kitchen - many for the first time in years - to prepare food for survival.

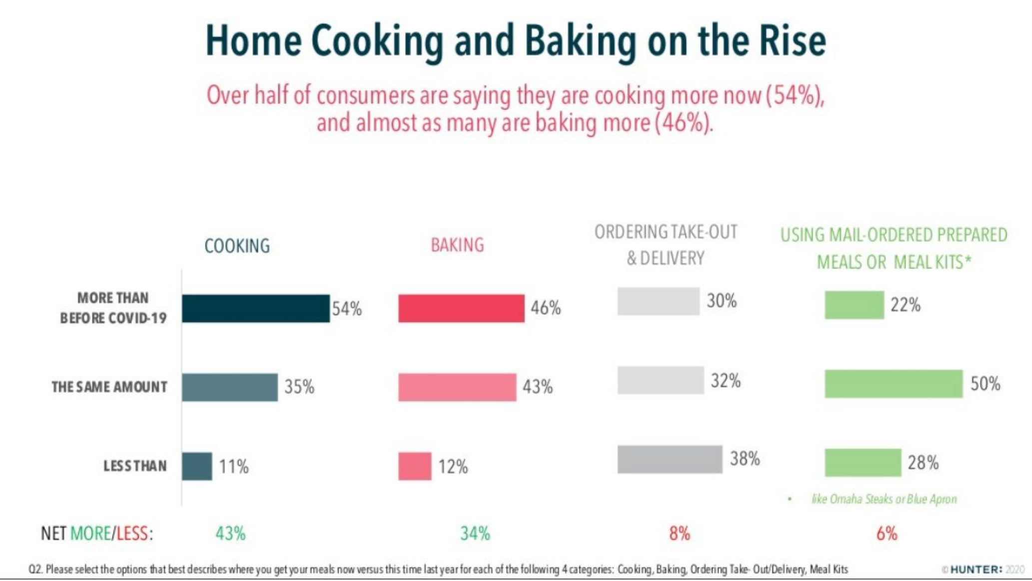


Figure 1.2: US Hunter survey year 2020

According to Hunter's consumer survey Figure:1.2, 54 percent of consumers are cooking more and 46 percent of people are baking more during the COVID-19 period, 51 percent of them reported they will continue this practice when the COVID crisis is over. [5]' As a result of this, cooking and recipe websites have seen a surge of traffic.

In light of the current situations, many people don't have the limited ingredients at hand or some don't have the liability of affording expensive ingredients, a cheaper and more commonly available substitute for those ingredients can allow them to use the recipes, and try the dishes.

The food we eat has an impact on our social lives but more importantly, it plays a huge role in our health. With limited exercise opportunities now, keeping your diet in check has become more important. The nutrition information and the flexibility to search on the basis of nutrition information can help the consumers keep a check of their diet while trying new and exquisite dishes with the resources at hand.

With little to no experience, ingredient availability, limited resources, and lack of cooking appliances, a simple recipe can also pose a challenge for new home cooks. But A rich and well-defined recipe ontology that takes all these factors into account, can propose a solution where the user can search recipes on the basis of the ingredients and resources at hand, evolve the recipes by substituting ingredients and their respective actions and also keep track of their diet by looking at the nutrition information.

1.2 Objectives

The problem to be addressed by this study is that of considering how the semantic technologies can be applied to the evolution of the recipes. This research attempts to create an ontology that expresses the domain and range of the recipe by (1) enabling the mapping of all kinds of recipes available over the internet to acts as a knowledge base for recipes, (2) re-usability of the knowledge base to evolve the recipes automatically and (3) selecting condition-appropriate meals. The focus will be specifically on gathering the recipe data from the internet and deriving a rich schema to map and highlight the key factors involved in any chosen recipe, whether it be (a) the required ingredients (b) cooking methods and instruction (c) the nutrients in the food product.

1.3 Thesis Structure

We start with the introduction to the issue under consideration, including a definition and explanation of the key terms such as recipe, ingredients, nutrition, and instructions. Afterwards, we discuss the currently available solutions and limitations with them. Later, we introduce basic concepts of the semantic web and suggests its potential for solving the problem and a proposed solution. Lastly, the areas where the proposed ontology can be used for Application development and improvements.

Chapter 2

Background

In this chapter, we will look at the background and discuss the problem statement mentioned briefly in the Introduction.

2.1 Background: Food and Recipe

2.1.1 Recipes

A set of instructions to put together a particular dish, including an inventory of the ingredients required [6]. It means that there are two essential parts of a recipe: the ingredient list and respective instructions for their usage.

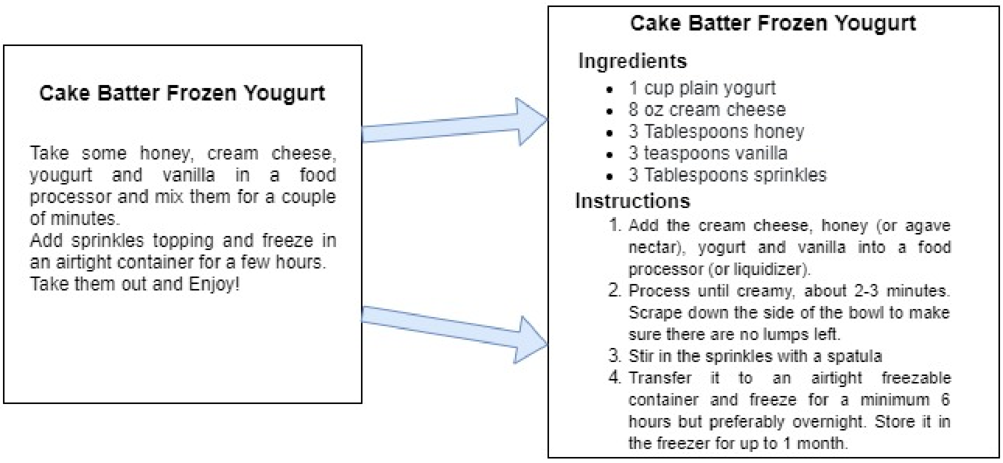
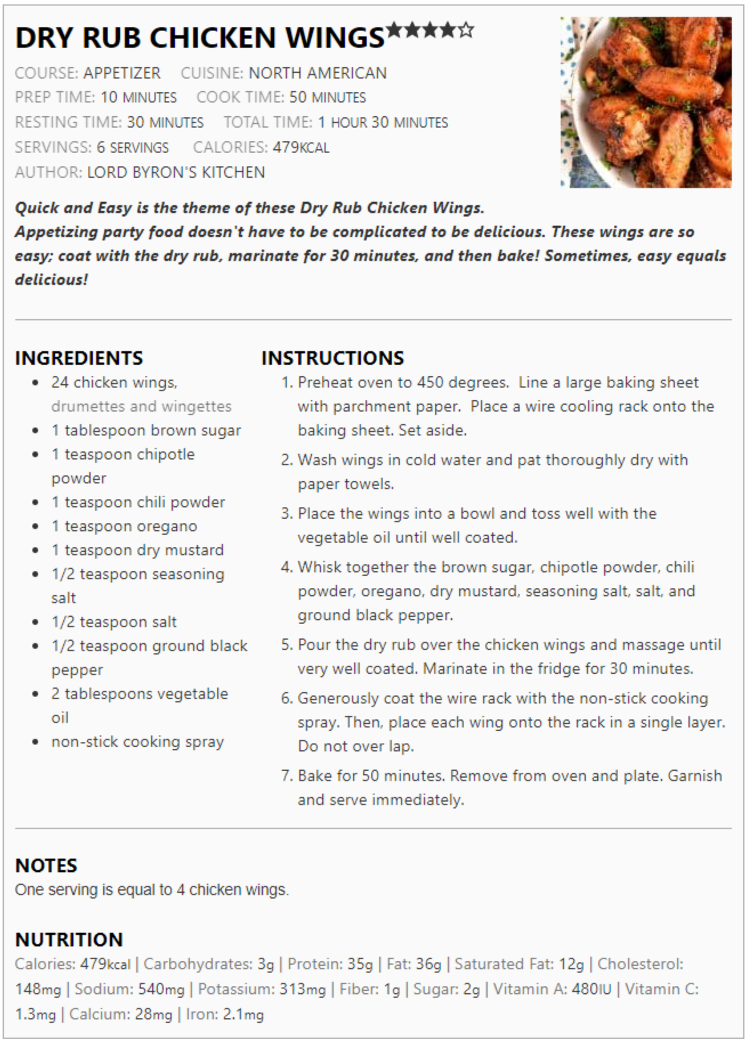


Figure 2.1: Basic components of recipes

In the figure 2.2, the text on the left doesn't qualify to be a recipe based on the definition because the list of ingredients and the clear instructions are missing. However, the section on the right fulfills the requirements thus a recipe.

Food recipes normally include more information than the ingredients and instructions, such as the quantity of the ingredients which allows the calculation of the nutritional information. Recipes also include the total time, cooking time, preparation time, and the instruments required for cooking as well. Based on the cooking method or the recipe type they are categorized. All of this information plays a dominant role in the selection of a distinct recipe by an individual.



DRY RUB CHICKEN WINGS★★★★☆

COURSE: APPETIZER CUISINE: NORTH AMERICAN
PREP TIME: 10 MINUTES COOK TIME: 50 MINUTES
RESTING TIME: 30 MINUTES TOTAL TIME: 1 HOUR 30 MINUTES
SERVINGS: 6 SERVINGS CALORIES: 479KCAL
AUTHOR: LORD BYRON'S KITCHEN

Quick and Easy is the theme of these Dry Rub Chicken Wings.
Appetizing party food doesn't have to be complicated to be delicious. These wings are so easy; coat with the dry rub, marinate for 30 minutes, and then bake! Sometimes, easy equals delicious!

INGREDIENTS

- 24 chicken wings, drumettes and wingettes
- 1 tablespoon brown sugar
- 1 teaspoon chipotle powder
- 1 teaspoon chili powder
- 1 teaspoon oregano
- 1 teaspoon dry mustard
- 1/2 teaspoon seasoning salt
- 1/2 teaspoon salt
- 1/2 teaspoon ground black pepper
- 2 tablespoons vegetable oil
- non-stick cooking spray

INSTRUCTIONS

1. Preheat oven to 450 degrees. Line a large baking sheet with parchment paper. Place a wire cooling rack onto the baking sheet. Set aside.
2. Wash wings in cold water and pat thoroughly dry with paper towels.
3. Place the wings into a bowl and toss well with the vegetable oil until well coated.
4. Whisk together the brown sugar, chipotle powder, chili powder, oregano, dry mustard, seasoning salt, salt, and ground black pepper.
5. Pour the dry rub over the chicken wings and massage until very well coated. Marinate in the fridge for 30 minutes.
6. Generously coat the wire rack with the non-stick cooking spray. Then, place each wing onto the rack in a single layer. Do not over lap.
7. Bake for 50 minutes. Remove from oven and plate. Garnish and serve immediately.

NOTES

One serving is equal to 4 chicken wings.

NUTRITION

Calories: 479kcal | Carbohydrates: 3g | Protein: 35g | Fat: 36g | Saturated Fat: 12g | Cholesterol: 148mg | Sodium: 540mg | Potassium: 313mg | Fiber: 1g | Sugar: 2g | Vitamin A: 480IU | Vitamin C: 1.3mg | Calcium: 28mg | Iron: 2.1mg

Figure 2.2: Example Recipe [1]

2.1.2 Searching for Healthy and Nutritious Recipes

The internet is full of data, When people query something on the browser, a large number of records are returned from all over the internet, containing information that may or may not be relevant or useful to them. Searching a recipe is no different and it takes human effort which can be more frustrating sometimes. There are platforms available that allow their users to search for available recipes bases on the ingredient name, or criteria-based or category-based. There are some platforms that focus more on nutrition information to target certain health conditions and provide food recommendations based on low calories, allergies, gluten-free diets, etcetera.

The available systems focus on retrieving recipe-based information via search queries based on ingredients or nutrition information, but a cook also decides on what to cook based on their personal preferences. This preference system can be based on culture, region, and ethnicity - varying the levels of chilly, sweetness, and other spices.

An ontology that takes all the key factors of a recipe into account and takes this wealth of data, by converting it into well-structured and usable information can be used as a base in developing a system that can provide food recommendations on what to eat and what to avoid by considering the personal preference, ingredient list at hand, or the health conditions of individuals.

2.2 The Linked Data and Web of Data potential

The vast majority of the web content available is for human readability, leaving machines with the inability to process the complete semantics of a web document. Although there are many recipe search websites on the internet, the schema and structure of the websites are not consistent, thus leaving machines unable to parse and process the information.

The "second generation: The Web" which we currently use is document centered, encoded using Markup languages, which enables the documents to connect with other documents creating a network of documents but the relationship between the documents or the nature of these links is not available. The information available on the web pages is retrievable using web crawlers but these documents are often unstructured or semi-structured.

The "third-generation: The web of Data" is an upgrade of the web of documents, as a decentralized database (knowledge base) of machine-accessible data also known as the Semantic Web. The semantic web removes the limitations of the web by adding meaning to the documents (semantics) and creating a common framework for reusability and sharing (vocabulary). The use of the semantic web will add meaning to the recipe information available on the web by merging the human-readable documents with machine-understandable data, thus creating the immense potential for improvement in this section.

2.2.1 The semantic Web

The Semantic Web, allows machines and humans to work and communicate better together through the power of linked data. Giving machines the ability to use logic on the information at hand and to choose the best course of action, by making choices to answer a question, is one of the most prominent advantages and uses of the semantic web.

The semantic web languages, such as Resource Description Framework (RDF) and Web Ontology Language (OWL) are powerful enough to define the entities, their relationship with other resources, and the logical rules applied to them in a structured way which is in a machine-understandable format but also compatible with human communication at the same time. In addition, these languages have the capability of well-defining and representing all the concepts and constraints of a particular domain in question.

Semantic Web technologies can be used in various applications like resource discovery and classification, data integration, cataloging, and content rating by intelligent software systems.

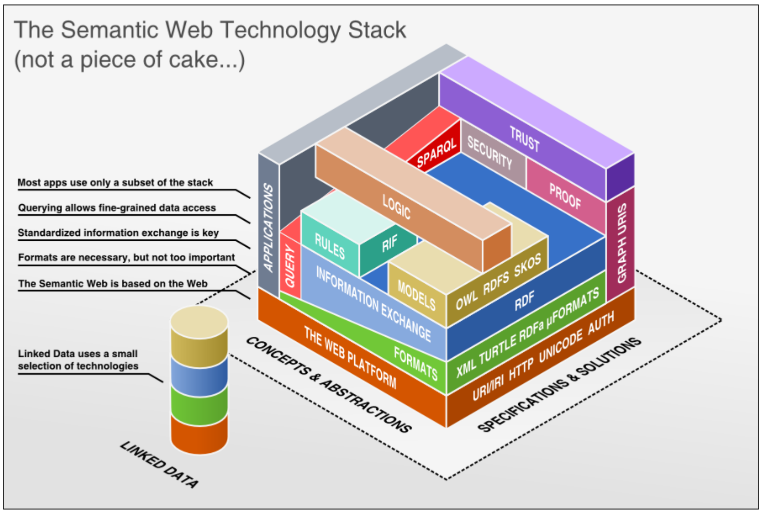


Figure 2.3: Semantic Web Technology Stack [2]

The Semantic Web Stack in the figure: 2.3 shows a hierarchy of different languages in layers on top of each other, and every layer uses and exploits the capabilities of the layers below it. It depicts how standardized Semantic web technologies are organized to make the Semantic Web possible. The illustration of the SWT was created by Tim Berners Lee [2].

2.2.2 Ontology - Why we need it?

An ontology is a formal description of knowledge as a set of machine-interpretable concepts and the relationships that hold between them for researchers who need to share information within a domain [7]. The most common uses of ontology development are sharing a common understanding of the structure of information amongst the domain experts or software architects, making domain assumptions explicit, enabling the reusability of domain knowledge, separating operational knowledge from the domain knowledge, and for the analysis of information in a particular domain [8].

The Web Ontology Language (OWL) is used to represent the machine-interpretable contents of the web through an explicit representation of terms in a vocabulary and relationships between those terms, OWL is better suited for expressing semantic information than previous languages such as XML and RDF variants. Although RDF provided a data model for objects and relations between them, OWL offers the ability to describe classes and properties including such things as disjoint relations, cardinality, equality, better property typing capabilities, enumerated classes, and property characteristics [9].

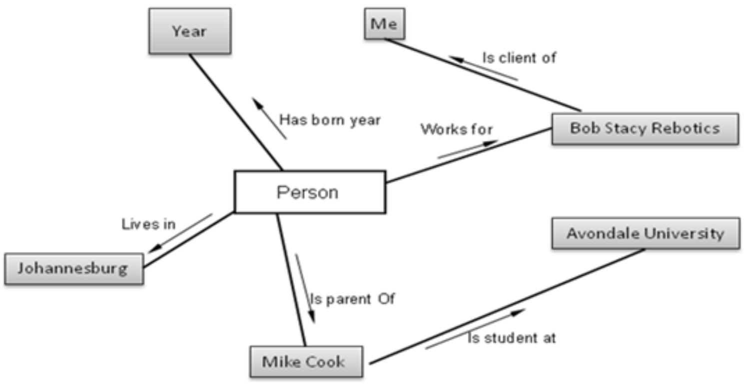


Figure 2.4: Person Ontology

Ontologies are used for various purposes, some of which include: providing a linked vocabulary for personalization and customization in a specific domain, defining an architecture that can be used for extracting information from a document, exemplifying word meaning, and for the semantic interpretation of a textual data. Ontologies can describe both concrete and abstract objects, the set of which, described by a set of attributes, is referred to as a class. Attributes are the features and characteristics of an object: for example, the class Person may have the attributes of age and name. Relationships show how objects are related to each other: for example, the classes of Person and City may be related through a relation such as wasBornIn.

2.2.3 Recipe Ontology

We have discussed the semantic web and its potential, now we need to get back to the following research questions:

1. Can Semantic Web technologies offer a solution for our problem statement?
2. How can an ontology provide a mechanism to improve the existing recipe system?
3. How to design a recipe ontology that provides a solution to our problem?

We have seen that semantic web technologies offer a mixture of human-readable and machine-understandable representation of data, The recipe ontology can act as the knowledge base for a recommendation system (mobile application or a website) which would take the user's search and run it against the knowledge available in the data set, therefore improving existing system. We do not simply wish to describe a recipe but also define and models all the underlying concepts related to it which will help us answer all of the questions we mentioned before, our ontology aims to:

- defines all the entities and concepts.
- specify and model the relationships between different domains of interest.
- design a well-defined and structured schema.
- represents recipe data from different sources.
- put constraints and rules on the entities.

Chapter 3

Related Work

In this chapter, we focus on previous work that is related to the problem stated in Background.

3.1 Food and Nutrition Related Databases and Vocabularies

The Information about food and nutrition is not only limited to the labels found on the back of the items. There are numerous existing systems, which were developed to hold relevant information. This information is available but not always accessible or understandable by non-technical people, however, this information can be made available to individuals by creating a system that extracts this information and presenting it in a formatted and easy-to-use platform. While designing our solution, we considered the existing available systems, which are relevant and can have an impact on defining the semantics of our ontology.

A formal ontology is a controlled vocabulary expressed in an ontology representation language. Controlled vocabularies provide a way to organize knowledge for subsequent retrieval. Its uses in making ontology not only decrease the redundant effort involved in creating an ontology from the start by using the available vocabulary but also initiate a system for allowing other vocabularies to be mapped onto the ontology. [10] The following research contributed towards the designing of our ontology structure.

- Linked Open Vocabularies (LOV) is a collection, catalog of high-quality reusable vocabularies for the description of data available on the Web. Its purpose is to support data sharing by gathering and making available the vocabulary's version, along with maintaining its editor history.

[11]

- The USDA Nutrient Database for Standard Reference is published by the United States Department of Agriculture and contains nutrient information for about 8,618 different foods. The database is searchable and provides full nutrient information for the foods stored within it, including calories, nitrogen-protein conversions, scientific names, and LanguaL codes. Users may run searches and download results as PDF or CSV files. [12]

- EuroFIR is a non-profit association which to develops, publishes, maintains, and exploits food composition. The EuroFIR maintains both databases and thesauri that draw information from compiler organizations across the world; data includes nutrients, bioactive, and food allergens. This data guides the key individuals in the food industry in the production of healthier foods and also provides information to nutrition experts, researchers, software developers, and consumers.[13]

- Agrovoc is a thesaurus or controlled vocabulary created by the Food and Agriculture Organization (FAO) of the United Nations (UN), which covers the interest areas of food and nutrition, as well as fisheries, agriculture, and environment. Currently, Agrovoc is a SKOSXL concept scheme and a linked open data set aligned with more than 10 other knowledge organization systems. Agrovoc may be downloaded in RDF core or LOD format for use in various applications. [14]

- The Factored Food Vocabulary (FFV) by the Food and Drug Administration describes food products characteristics that are important for food safety and nutritional quality by using a standardized language. Each food product is described using the following factors: type of product, source of food, plant or animal part, physical shape, form or state, preparation and cooking method, treatments applied, preservation method, container packaging or wrapping, contact surface, and user groups. [15]

- Recipe1M+ is a large-scale, structured data set containing more than 1 million recipes and thirteen million images. It is the biggest collection of data publically available in the recipe domain by training high-capacity models on aligned, multimodal data. [16]

- Yummly-28K is crawled from a recipe-sharing website called Yummly containing a total of 27,638 recipes. Each recipe contains an image, the list of ingredients, the cuisine, and the course information. There are 16 cuisines and 13 courses in this dataset. [17]

3.2 Food specific Ontologies

Ontology development is gaining popularity and is becoming a more preferred method in semantic web technologies to connect different types of information systems, though they suffer from the lack of a standard format that must be used for their development. The Ontology Search (OLS) contains a collection of ontologies and it provides a single point of access through the website as well as programmatically through the API for bioinformatic ontologies.

- FoodKG is based on a Knowledge Graph for Food Recommendation mapped on WhatToMake ontology. WhatToMake ontology is separated into three components; Food, FoodOn, and Ingredient. The FoodKG recommendation system provides the user the ability to search by parameters such as ingredients, cook time, course type, and meal type. It also provides an ingredient substitution based on the user's allergies.[18]

- FoodOn focuses on packaged food. It is a farm to fork ontology is an ontology built to name animals or plants parts, fungai, and so forth which bear a food role for livestock and humans, as well as obtained food products and the exercise to produce them. FoodOn was initially motivated by the need to create a standard for the reporting of foodborne diseases. FoodOn focuses more on food safety and security, and it covers all the aspects from food gathering, production, processing, packaging to delivery. [19]

- The BBC Food Ontology is a simple vocabulary for describing recipes, ingredients, menus, and diets. It covers aspects from food preparation to categorization and it originated from a specific BBC use case. [20]

- The Bionutrition Ontology creates and sustains a repository of biomedical ontologies and vocabularies. They have also built a web service to enable the use of ontologies and terminologies in expert research. It consists of 100 classes, the major ones being Diet Plans, Protocol Development, and Nutrition Procedure, further subclassified into representing the core concepts of nutrients specific to diet plans. [21]

- ISO-FOOD ontology links and harmonizes different knowledge repositories concerning isotopic data contained in Food Science. It focuses on the metadata and source data that is required to describe isotopic measurements and is associated with existing ontologies, such as Food, Nutrient, and Units of Measurements Ontologies. [22]

- FoodWiki is a mobile system for safe food consumption which uses the Ontology-Driven approach and focuses on compounds that are added to the processed food to add color, texture, thickness, flavor, and so forth to reduce such ingredients in manufactured foods as they have side effects causing health issues such s heart diseases, cancer, diabetes. [23]

We defined that our developed Ontology will be able to map data coming from multiple sources like websites, blogs. This means that our ontology needs to be specific enough to accommodate different conceptual representations on the source site. Secondly, Our ontology needs to be flexible towards different recipe-related applications, thus we have to include the capability for such extensions to a reasonable extent. Thirdly, Attribute and condition-based search for recipes shall be possible, meaning that our ontology needs to be generic to allow searches based on Ingredient, Actions, Nutrition, e.t.c.

None of the existing ontologies fulfilled our scope. We can make use of the existing ones by adding new nodes in the graph, then we will end up with too many nodes, blowing up the graph. Since re-usability was not an option. We took inspiration from the existing ontologies in making our modeling choices.

3.3 Ontology development

Ontologies are a formal and explicit description of i) concepts in a particular domain (classes or entities), ii) properties of each class describing its attributes and their features (relationships) and, iii) any limitation or restrictions on the properties (rules). The mapping of an ontology's instances on their respective instances of the classes forms a knowledge base. The following section shows, which articles we researched to learn to develop our ontology?

- Ontology Development 101 is a good starting point for developing your first Ontology, it uses Protégé-2000 for developing a wine and food example: which basically guides and helps a beginner in developing an ontology. The paper uses a knowledge-representation system based on a description-logics approach, stating hierarchy, designing classes and relations, different approaches and methodologies using Protégé. [8]

- An Ontology Design Pattern for Cooking Recipes shares an experimental class activity used for ontology development. The model for this ontology was based on acquiring knowledge of the domain from cookbooks and highlight the conceptualization process, breaking it into different parts of ontology development: defining the scope by identifying the concepts and their attributes, further classification of the groups of concepts, defining the properties, instance creation, and creating the end product. [24]

- FOODS is also an Ontology-Driven System in the Food domain and it takes you through the process of defining and developing a food and menu planning system. The systems comprise a food ontology and an application designed on top of the ontology with an interface for novice and expert users. The expert system assists in finding the appropriate meal based on different attributes. [25]
- Food Ontology uses the word embedding technique for the automated ontology population from an existing ontology. They proposed a semi-supervised framework and applied this on the food domain, trained from Wikipedia corpus, and compared it with the foodOn ontology. [26]
- A Process for Building a Domain Ontology: an Experience in Developing a Government Budgetary Ontology presents a process of describing data semantics, characteristics and providing schema for the development of an ontology in the Budget domain. [27]
- The Ontology in the Products and services domain to help consumers and enterprises to search for suitable suppliers. It publishes data in ways that maximize reuse for e-commerce on the semantic web by analyzing the product description and covering the representation needs of business scenarios. [28]

3.4 Application Areas

In this section, we focus on the research areas with more focus towards the applications developed in the food domain to which our proposed Recipe Ontology can play a role as a proposed solution.

- Smart Chef: Evolving Recipe, EvoChef: Show me What to Cook! Artificial Evolution of Culinary Arts, and AutoChef: Automated Generation of Cooking Recipes extend each other and follow the same model of evolution in culinary arts by autonomously evolving novel and human-readable recipes through ingredient substitution, actions replacement, and recombining the instructions. [29] [30]

- FOODS mentions an expert system based on top of the food ontology with a user interface for the novice and expert users to search for recipes on the basis of ingredients, nutrition facts, region, menu. [25]
- Recommending Food: Reasoning on Recipes and Ingredients focuses on recognizing the implementation of recommendation techniques in the domain of food area, it presents an IT solution in health management which assists users in planning a healthy and fascinating meal, and aim to shorten the work required of users to modify their diet. [31]
- Computational Creativity in the Culinary Arts present and explains a system that creates an artificial chef programmed to produce novel recipes but only limited to the salad domain with minimal human assistance. [32]
- A Neural Network System for Transformation of Regional Cuisine Style proposes an innovative system that can change a chosen recipe from any region into another selected regional style using ingredient substitution through an extended word2vec model. [33]

Table 3.1: An overview of different food systems

System Name	**Type of Food**	**Technology**	**Features and Coverage**
FoodWiki	*Packaged Food*	*OWL, RDF*	Mobile App, Limited only to food products included in the ministry's database.
AGROVOC	*All kind of Food*	*OWL, RDF, and SKOS/-XL*	Comprehensive coverage for international food but complex. Only experts can use it.
FoodOn	*Packaged Food*	*OWL, RDF*	Very extensive, doesn't cover the cooking process. For organization use only.
Food-KG	*All kind of Food*	*OWL, RDF, SPARQL*	Web-based. Limited parameters, doesn't cover all recipe attributes, missing actions and ingredients classification for evolution

Due to the limitations mentioned in the table 3.1 we decided to build our own ontology by taking the inspiration from the existing systems and making it fit our use case. By having a well-defined Ontology, we will be able to demonstrate how it can be applied to existing recipe data by aggregating and mapping the data sets from different sources to present: (i) A case for this ontology to act as a knowledge base for intelligent systems i.e. Recipe Evolution (ii) provide answers to questions which cannot be queried through existing systems.

Chapter 4

Approach

In this chapter, we present our Approach by explaining the Ontology, its components and its development process.

4.1 Ontology

Now a days Ontologies are widely used. Ontologies come from the branch of philosophy which deals with the concepts of "existence" and "reality". It determines the formation of a system of categories that exncompase classification if all entities. In Information science, an Ontology gives a way of constituting the relationships in a domain area and how they are affiliated to a group of defined concepts and categories that represent the entities in the subject domain. In simple terms, an explicit and formal specification of concepts in a domain. The Ontology needs to have:

- Conceptualization - abstract model (domain, relevant concept, relations)
- Explicit - All entities and attributes are defined.
- Shared - Available for re-usability.
- Formal - Machine understandable.

Sharing common understanding of the structure of information, Enabling reuse, Splitting up the domain and the operational knowledge and Analyzing them separately are most common uses of the Ontology. The major components of ontologies include:

- Individuals - Instances or objects of a particular thing.

- Classes - Sets, group of entities, collections of thing.
- Attributes - Parameters, or properties associated with an entity.
- Relations - defines how entities and attributes are related to one another.
- Restrictions and Rules - Constraints and logical rules applied on class and objects.
- Events - Attributes and relation change.

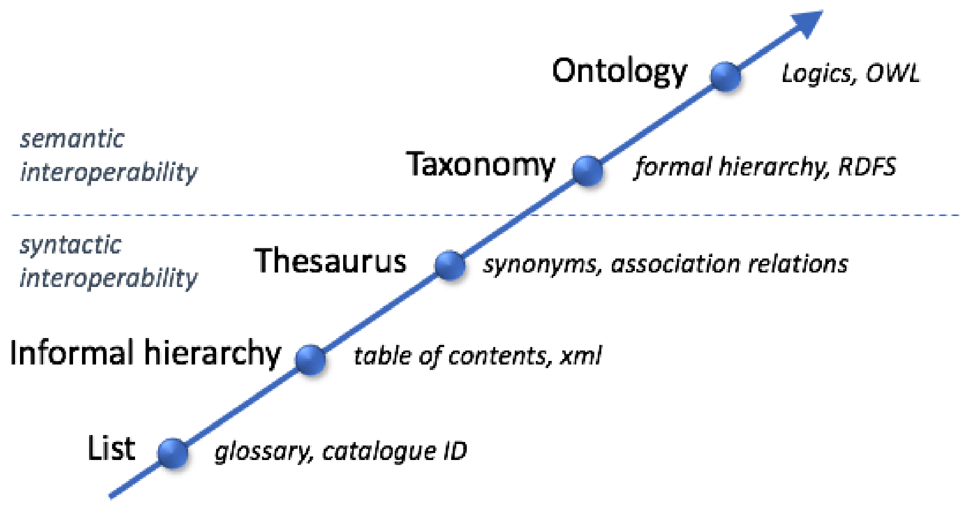


Figure 4.1: An Ontology Spectrum

4.1.1 Ontology description Languages

Ontology languages are formal languages used to create an ontology. They allow the encapsulation of knowledge about a particular domains and frequently include logical rules that help in the processing of that knowledge. Here is the classification of Ontology Languages:

- Traditional syntax ontology languages - KIF (Knowledge Interchange Format), OCML (Operational Conceptual Modelling Language), F-Logic (Frame logic).

- Markup ontology languages - Web Ontology Language (OWL), Resource Description Framework (RDF) , RDF Schema (RDFS)
- Natural languages (controlled) - English.

We won't be discussing the frame-based languages and first-order Traditional syntax ontology languages, rather we will be focusing on the most recent Markup Ontology languages that have found their practical usage in the field of semantic web and became a standard.

Resource Description Framework

RDF (Resource Description Framework) is a data model originally used for metadata for web resources, then generalized. It Encodes structured information into universal, machine readable exchange format in expressions known as triples, of the form subject–predicate–object. A set of RDF statements represents a labeled, and directed multi-graph.[34]

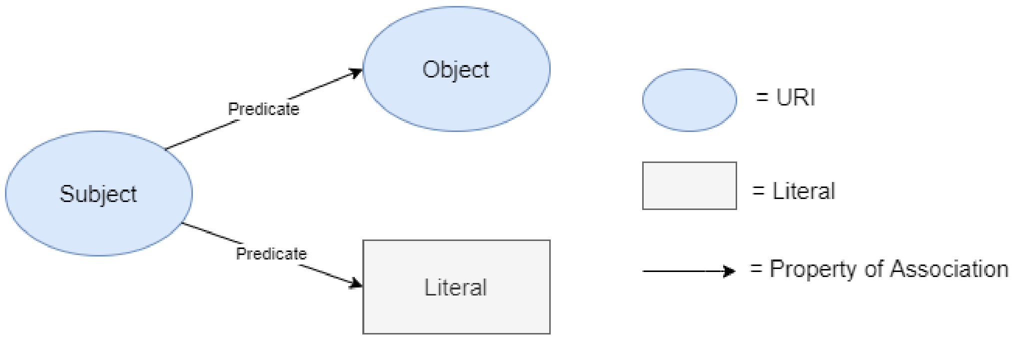


Figure 4.2: RDF triple

The RDF triple representation in figure 4.2 is an example of the RDF data model. The main Parts of RDF graph are:

- Subject - URI (Used to reference resources unambiguously)
- Object - URI, Literal (describe data values), Blank nodes (Facilitate existential quantification for an individual with certain properties without naming it)
- Predicate - URI for Property/relation definition between resources.

Several different serialization formats are widely used for RDF including: Turtle, N-Triple, JSON-LD, RDF/XML. Here is an example of RDF triple respresentation in Turtle Syntax.

$$\begin{gathered}\langle http://dbpdeia.org/resource/Leipzig\rangle \\ \langle https://www.w3.org/2000/01/rdfschema\#label\rangle \backslash Leipzig"@\end{gathered} \tag{4.1}$$

Resource Description Framework Schema

RDFS (Resource Description Framework Schema) contains a set of classes with certain properties built as an extension using the RDF representation data model, providing basic elements for the description of ontologies.[35]

Along with all the vocabulary of RDF, RDFS constructs are classes, associated properties, and utility properties which allow the flexibility for class hierarchy, range/domain and restrictions definition. Main concepts of RDF Schema are:

- Class/ Subclass
- Property/ Sub-Property
- Domain/Range
- Restriction

Figure: 4.3 is an example of RDF Schema, which shows how we can use these constructs to implement a Acid and base representation System.

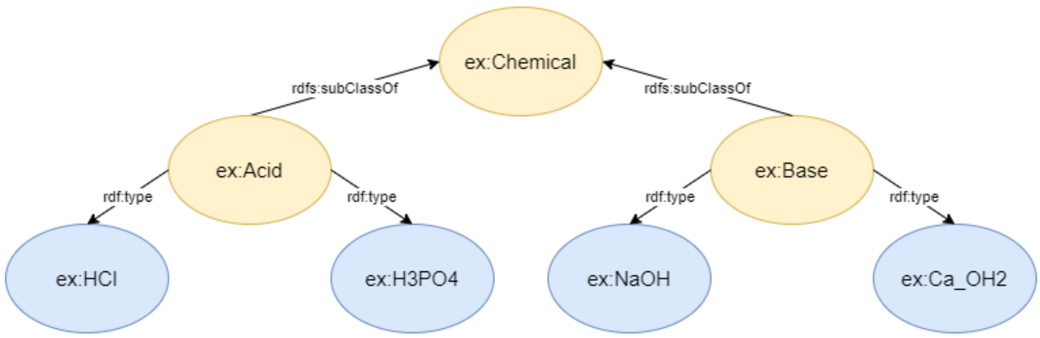


Figure 4.3: RDF-Schema representation of Chemical system

Web Ontology Language

RDF Schema extends RDF with special vocabulary for terminological knowledge, and can be used as a lightweight language for defining a vocabulary. However, RDF Schema has some limitations regarding the possibilities for formulating ontologies. RDF Schema miss expressivity, it can not define the negation of an expression, define cardinality, define set of classes, or define metadata of the schema. Therefore we will be looking at OWL now. [36]

Web Ontology Language (OWL) is a family of knowledge representation languages used for describing ontologies. The OWL languages are characterized by formal semantics built upon W3C standards for RDF. Since OWL is built on top of RDF/S, it supports the same syntax with stronger vocabulary and greater machine interpretability.

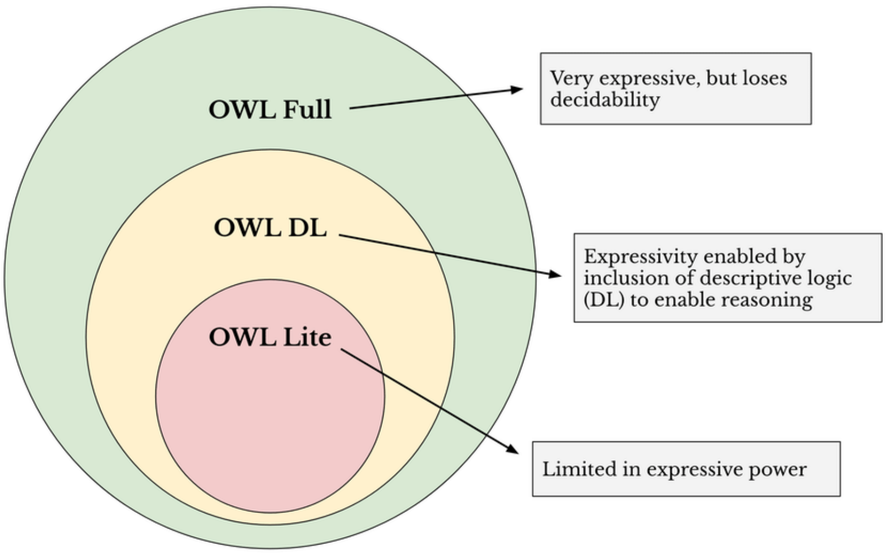


Figure 4.4: OWL Sub-languages

OWL specification include three variants, Figure: 4.4 based on their level of expressiveness. Each of the sub-languages is a syntactic extension of its predecessor.

- OWL Lite ontology is a subset of OWL DL ontology.
- OWL DL ontology is a subset of OWL Full ontology.
- OWL Lite conclusion is a subset OWL DL conclusion.
- OWL DL conclusion is a subset of OWL Full conclusion.

4.2 Ontology development Process

There is no one accurate or perfect technique or methodology for ontology development. The Ontology creation Process, shown in Figure: 4.5 is an repetitive process that iterates continuously between steps to refine the ontology. There are many distinct approaches for designing and developing an ontology. In practice, the application area of the underlying ontology decides about the modelling approach. The Ontologies have mainly three different types:

- Domain Ontology - Domain specific definition of terms.
- Upper Ontology - Shared relations and objects across wide range of domain ontologies.
- Hybrid Ontology - A Gellish Ontology.

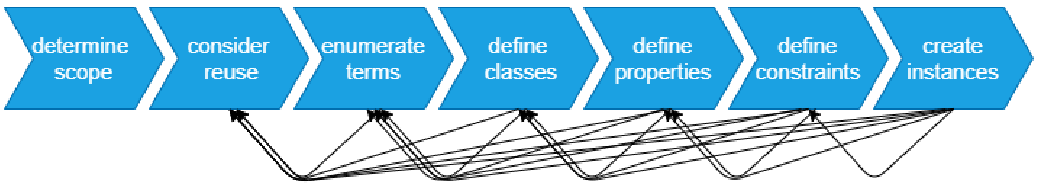


Figure 4.5: Ontology development process

Determine Scope

Determine Scope process includes the literal scope and determining the subject area to be covered by the Ontology i.e. Our ontology is covering the "Food Recipe Domain". It also deals with questions like: What will be the use of the ontology? What kind of knowledge will be represented by it? What application will be using that knowledge base? This step also includes the formulation of Competence Questions which help define the Scope better.We will look at the competence questions for our ontology in the section: 5.2.

Consider reuse

Ontology reuse can be defined as the process in which available ontologies are used fully or in parts to develop new ontologies. Depending on the domain overlapping and use case similarities, we can choose between ontology merging and integration. Ontology reuse is recommended because it play a key role in cost and efforts for the developed ontology. The decision whether to reuse or newly build an ontology is made by weighing the cost and efforts. We have already discussed in the section: 3.2 why we are developing from scratch and not reusing the existing ontologies.

Enumerate terms

In this step, we mention all the terms related from our domain one by one. We discussed how we identified the terms of our domain by analyzing the existing vocabularies, analyzing data from websites in the sections: 3.1, 5.1 and 5.2.

Define classes

Classes are concepts in the designated domain. They represent a collection of objects with similar properties i.e. "Instruction", "Ingredient" are classes in our Recipe ontology. In this step we also model the class hierarchies using one of the following approaches: top-down, bottom-up or middle-out approach. We have discussed our class definition step in the next chapter's section: 6.1.

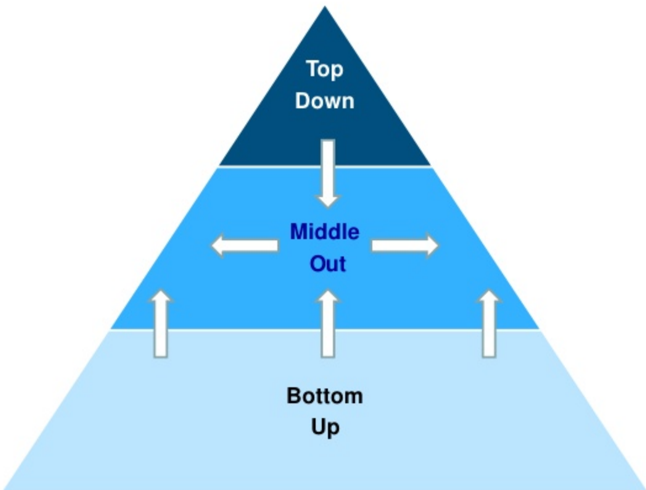


Figure 4.6: Class hierarchy approaches

Define properties

Properties in a class definition describe attributes of the instances i.e. "name", "cookTime" are properties related to the "Recipe" class in our ontology.

DatatypeProperty and ObjectProperty, describe what kind of values a triple with the associated property hold. Datatype properties relate individuals to literals(strings, numbers) whereas object properties relate individuals to other individuals. We have discussed the properties associated with our classes in the next chapter's section: 6.2.

Create instance

Instances are "type" of a class defined in our ontology i.e. "BreakFast" is a type of instance for the class "MealType". Figure: 4.7 shows Simba to be an Instance of type AfricanLion and its super class Animal as well. We have discussed the Instances type of our classes in the next chapter's section: 6.1.

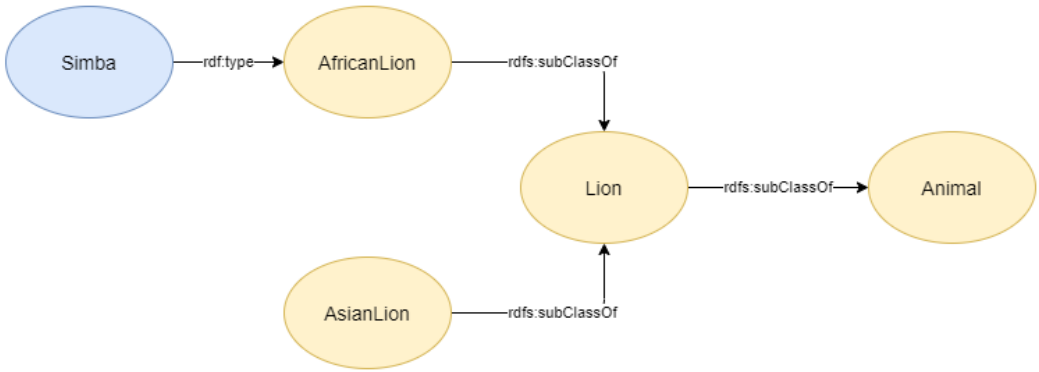


Figure 4.7: Example: Instance of Lion

4.2.1 Ontology Editors

There are a number of tools available for ontology development and use, we have used the following tools in our thesis.

- Protégé Ontology editor, developed by Standford Medical Informatics for development.
- WebVOWL: Web-based Visualization of Ontologies, developed by University of Stuttgart for visualization.

Chapter 5

Methodology

5.1 Data Sources for knowledge base

The cookbooks on paper in the past have don't take after the ones we have and are using in the modern era. Those cookbooks consisted of a lists of home remedies and chef's personal recipes and favorite dishes, or special Mom recipes passed down through generations. However, the digital powers that be are trying to make this facet of analog life obsolete and will ultimately replace the hard copy cookbook completely.

A large number of individuals who search for recipes use either a dedicated food website or a favorite recipe blog to look for recipes. There are countless recipes all over the Internet. Because documents on the web come from a variety of sources without any set standards to which they conform, differences will occur among the data set.Since no absolute standard exists for how recipes are presented, we need to set some criteria for data quality purposes to ensure that all the recipes would have a similar composition in terms of the data.

We identified the schema for our data set by analyzing the key factors in a recipe and consulting the linked open vocabulary available at www.schema.org. Our ontology will be able to map all available recipes, so for the next criterion we selected top 50 recipe websites and used them for our data extraction and storing it in a structured form.

5.1.1 Data Extraction

Information retrieval from web pages can be achieved through web scraping (web harvesting or web data extraction, a process in which data is automatically retrieved from the Web using software tools to extract information. The result of web scraping is a transformation of unstructured data into a more structured and machine-understandable format which can be stored in a database and further analyzed.

Various web scrapping solutions exist, for this project we use Visual Web Ripper to extract data from targeted websites. It is a powerful web scrapping tool that is used to scrape data from web according to our requirements. We provided the targeted websites and blogs into Visual web ripper and saved the attributes as elements in content tab based on the schema we had defined before. The extracted data was exported into 3 formats; XML, Excel worksheet and SQL database file. Table 5.1 has the major recipe attributes and their descriptions that were required to be extracted.

5.1.2 Data Transformation and Consolidation

Once the data was successfully extracted and exported, we converted the data into a SQL database file using SQLite Studio after data formatting. Once the changes were done, we uploaded the data to a SQL server to perform data cleansing and normalization by using MYSQL scripts and MYSQL workbench. The following steps were performed on the data:

- Inserting new column "ID" and correcting the naming convention of the columns.
- Separating duplicates and unique records from the recipe table into separate table.
- Take recipe website attributes and consolidate into the recipe table.
- Consolidate the Ingredients, Instructions, and Nutrition table.
- Ensure Primary and foreign relationship between all the tables.

5.2 Competency Questions and Use Cases

Making competency questions while analyzing the data sources improves the definition scope of the problem and helps with modelling the ontology. Competency questions are natural language queries, performed on the knowledge base for data retrieval.They help in further specification of use cases and potential extensions.

In Recipe domain, the following can be possible competency questions:

- Sugar-Free desserts.

- How to make low-fat food?
- Low-cholesterol fried chicken recipes?
- Top Indian recipes?.
- Breakfast dishes with eggs, potatoes and bread?
- Simple 10 minute recipes?
- 5 minute snack recipe?
- A side dish with beans, onions, nuts and bacon.

The competency questions indicate that we need to focus on retrieval of cooking instructions, enable search by different attributes, i.e. Ingredients, Nutrition, Meal Type, Meal Category, Cuisine, Difficulty level, Time consumption.

- **Purpose:** The purpose of FoodRecipe Ontology is to provide a reference model for the food recipe by describing its primary and secondary attributes.
- **Scope:** The ontology will focus on the food recipe process.
- **Implementation Language:** The ontology has to be implemented in a description logic language which allows encoding of knowledge about a specific domain.
- **Intended Users:**
  - Novice Users
  - Product Developers
  - Food Domain Experts
- **Intended Uses**
  - To describe recipes from different classifications.
  - To help identify the short comings and optimize the recipe representation process.
    * Ingredient categorisation and substitution.
    * Automated recipe generation through mutation, or recipe cross over.

- To improve the user interaction
    - In recipe attribute-based search.
    - In condition-appropriate search.
- To allow extension and integration with existing systems
    - Integration with Nutrition databases

- **Ontology Requirements**
  - Non-functional requirements (not applicable)
  - Functional requirements: Groups of competency questions
    - Group1: Recipe classification related competency questions
      - What features classify a recipe into a category?
      - How many different classifications are there?
      - Do the classifications have sub-categories?
      - Are there any classifications based on time of the day?
      - Does region, area, ethnicity have an impact on cooking style?
      - Can a same dish have two different classifications?
      - What are recipe courses?
      - ...
    - Group 2: Primary attributes related competency questions
      - What are the primary attributes of a recipe?
      - Can ingredients be further divided into different types?
      - Can ingredients be further classified into categories?
      - Do ingredients need to be in a certain quantity?
      - Does the order of the instruction matter?
      - Can instructions be parsed into actions?
      - Can actions be divided into preparation, cooking and presentation phase?
      - Are instructions time dependant?
      - Can the time be divided into prep time, cook time and chill time?
      - ...
    - Group 3: Secondary attribute related competency questions
      - What are the secondary attributes of a recipe?
      - Is there a utensil dependency?

· Can a recipe have different difficulty levels?
· Does the quantity of ingredients matter?
· Can we calculate the nutritional value?
· Does the quantity have an impact on nutrition value?
· Is there a source for the recipe? author? rating?
· . . .

∗ Group 4: User Interaction related competency questions
  · User search on the base of primary attributes only?
  · Can user search by secondary attributes as well?
  · Can user search by time, difficulty?
  · Can user search by nutrients?
  · Can user search by cuisine, meal time, or course?
  · ...

∗ Group 5: Evolution related competency questions
  · Can the recipe representation be optimized?
  · Can the ingredients be substituted?
  · How to identify the best ingredient replacement?
  · Can two or more recipes be merged?
  · Can a recipe be created from scratch?
  · ...

∗ Group 6: Extension related competency questions
  · Can we use existing food systems to optimize recipe search?
  · Can nutrition value be calculated for every recipe using ingredients quantity?
  · ...

Now Looking at the data at hand and the competency questions, we re-evaluate our modeling choices and make sure that the Ontology we produce is enable to handle all possible queries while allowing extension towards possible future inclusions. The decision process regarding what to model now versus later is called scoping. We Identified the key elements, and relationships to model based on analyzing the existing ontologies, their short comings, available linked open vocabularies, data sources, data set and our use cases. The details of which will come in the next section.

Recipe Property	Description
url	a recipe url that tells the address of the webpage
name	a recipe's name that identifies it
image url	a url that links us to the image attached
author	a person or organization, who produces the content of the recipe
cookTime	Cook Time tells the duration it takes to cook a recipe
prepTime	Prep Time is the time to do all the initial preparations
totalTime	Total Time is the time it took to perform all steps
marinateTime	Marinate time, is used in the recipes where meat takes time to get marinated
recipeCuisine	Recipe Cuisine is the cuisine of the recipe, for example, Italian pasta
recipeYield	Recipe Yield denotes how many people does a recipe serve
yield	Yield to indicate the quantity it will produce
cookingMethod	Cooking Method tells the procedure we undertake to cook the recipe i.e frying, steaming, etc.
commentCount	indicates the number of comments the recipe has received so far
dateCreated	Date Created is the date on which recipe was created
datePublished	Date Published is the date when the recipe was published and introduced to the public
dateModified	Date Modified is the recent date when the recipe was edited
keywords	Keywords are the tags that are used to search
aggregateRating	Aggregate Rating represents the overall ratings of a recipe
contentRating	Content rating, the official rating the recipe has gotten
nutrition	Nutrition indicates the nutrient content a recipe e.g. calories, fat content, e.t.c
estimatedCost	Estimated cost shows the cost of raw materials or supplies
recipeIngredients	a list of all the ingredients used to make the recipe
recipeIngredeint	a single ingredient from the ingredient list
recipeInstruction	all the steps that make up the "how-to" part of the recipe
step	step is a single direction from the recipe instructions list

Table 5.1: Major recipe attributes

Chapter 6

Implementation

In this chapter, we explain the development process of our Recipe Ontology.

6.1 Recipe Ontology Design

6.1.1 Recipe

Recipe is the core of our ontology, so it is the first choice in our model. By definition, recipe means "something which is likely to lead to a particular outcome" but in context of cooking, a recipe is considered as "a set of instructions for preparing a particular dish, including a list of the ingredients required".

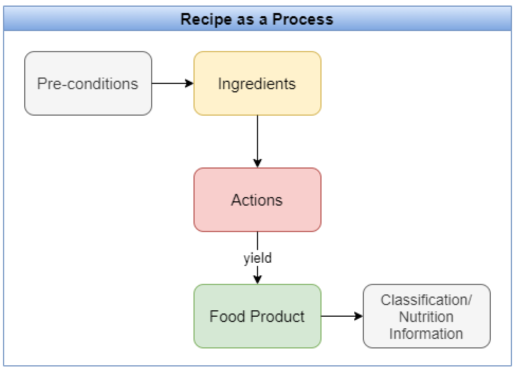


Figure 6.1: Recipe as a process

Recipe can be viewed as a process, which takes ingredients as input, performs a series of instructions on it, and produce a food product as an output. Alternatively, it could be viewed as an entity which has properties such as ingredients, actions and time required associated with it. We have modelled recipe as both a process and an entity by including all its associated attributes and properties, defining its core functionality in the ontology by establishing its relationship with other classes.

The "Recipe" class modelling is based after brain storming on the previous work and gathered data. Recipe is document with a title, published by an author and extracted from a Source URL. It has information regarding the preparation, cooking and presentation phase of a food product, each of these phases have a set of instructions, which contain actions to be performed on the ingredients. The Recipe includes nutrition information which is calculated on the base of ingredient quantity and yield. It also has some more information such as time required, utensils needed, and classification of recipe based on; meal type, cuisine, recipe type.

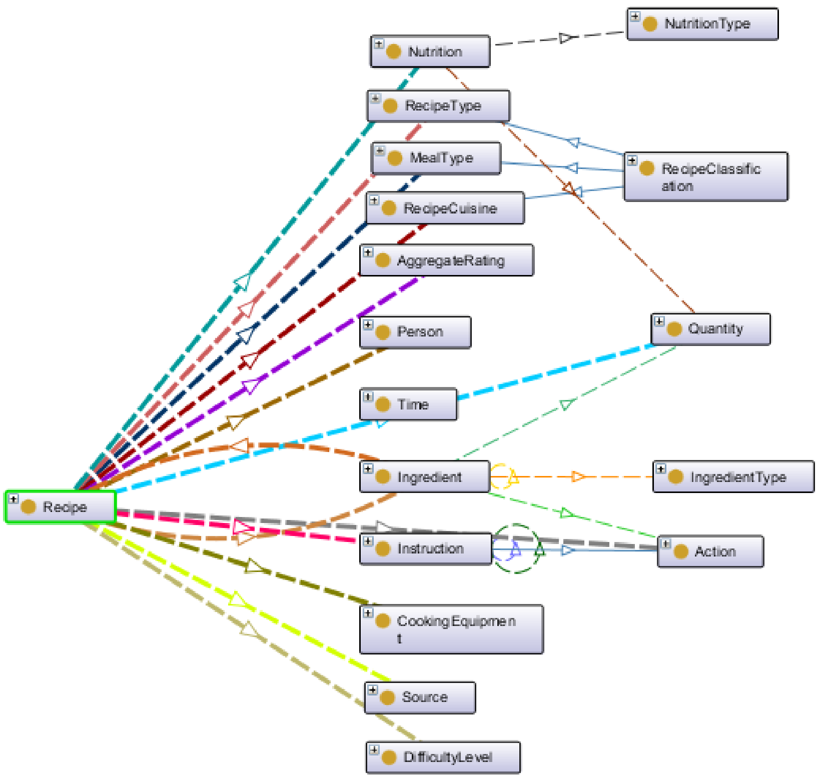


Figure 6.2: Recipe Class model

6.1.2 Ingredients

Ingredient is a "component part or element of something",and as part of cooking recipe "any of the foods or substances that are combined to make a particular dish". Ingredients play a vital role in our recipe for their nutritional, functional and sensory characteristics. They also add flavour, colour or texture to the food.

In our recipe concept model, "Ingredient" class has the following relationships:

- an ingredient can have a substitute of type "Ingredient".
- an ingredient has a "Quantity" associated with it.
- an ingredient is used in a "Recipe".
- an ingredient has an "Action" performed on it.
- an ingredient belongs to a type "Ingredient-Type".

Sub Classes of Ingredient

Ingredient is the super class, every ingredient in our recipe model will belong to a sub-class which were derived on the bases of categories. This sub classification of taken from the cook's thesaurus [37] will help us to narrow down the search for a suitable substitute ingredient, along with playing an important role in the recipe evolution and mutation step.

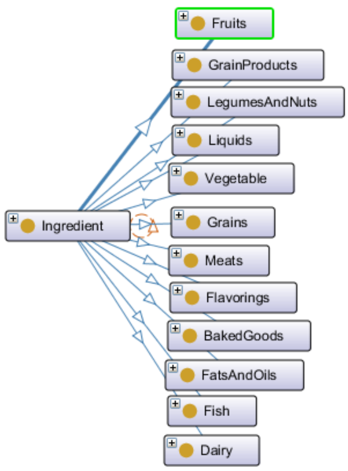


Figure 6.3: Ingredients: Sub-classes

Vegetables The Vegetables category includes many of the edible plants or parts of a plant, used as food such as roots, cabbages, beans, leaves and so forth. Fruits also contain seeds and come from flower of a plant, we have included those fruits which aren't too sweet into our vegetable categorization such as Tomatoes, eggplants are fruits but referred as vegetables. Vegetables are further divided into sub-classes as well, which help in better classifying the ingredients of type vegetables.

Figure 6.4: Vegetables: Sub-classes

BakedGoods BakedGoods category include food ingredients made from a dough or batter that is baked i.e different types of breads. The baked products are further divided into sub-classes to better classify an ingredient.

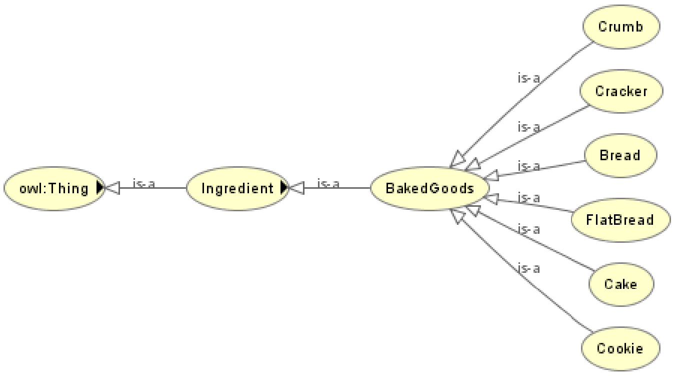


Figure 6.5: BakedGoods: Sub-classes

Fish Fishes and sea food are famous as meals and consumed in all around the globe. Fishes have a distinct flavor and provide protein and other nutrients to humans, thus included in many recipe plans as their Main ingredient. The Fish category is further divided into sub-classes to better classify an ingredient.

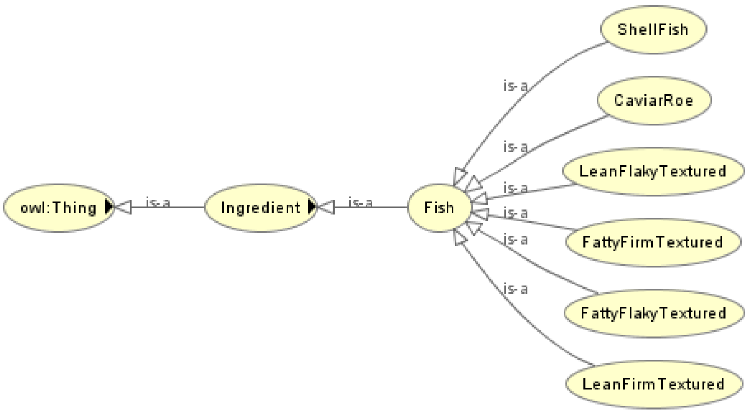


Figure 6.6: Fish: Sub-classes

Dairy Dairy products are a type of food produced from or contains milk from dairy animals, most commonly cattle, sheep, and goats. The dairy products are further divided into sub-classes to better classify an ingredient.

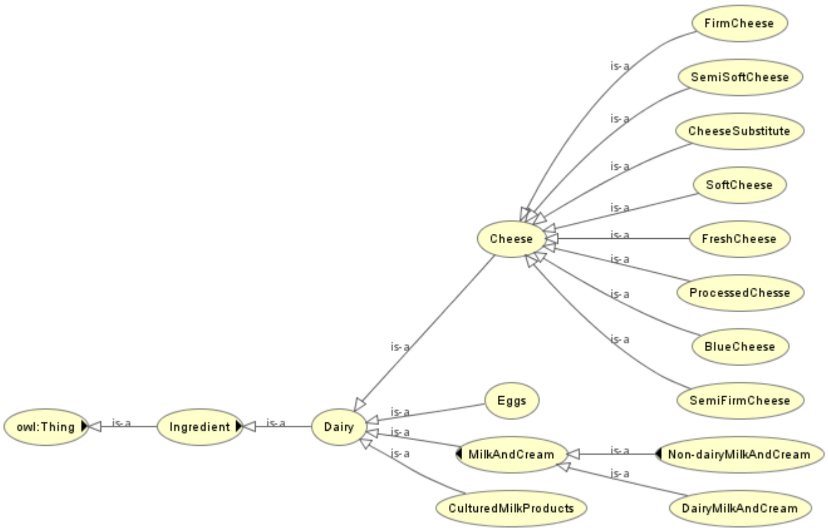


Figure 6.7: Dairy: Sub-classes

GrainProducts Any type of food prepared using the grains of rice, wheat, barley, oats, cornmeal, or some other grain is called a grain product. Wheat flour is consumed as part of everyday meal in many regions of the world. The GrainProducts category is further divided into sub-classes to better classify an ingredient.

Figure 6.8: GrainProducts: Sub-classes

Flavouring In order to add a particular taste, scent, color or aruma to the food an additional additive is added to to food or drink called Flavouring. Flavors are used all around us in both cooked and packaged foods. They are extracted from flowers, leaves, fruits, spices, vegetables, and even some trees. The Flavouring category is further divided into sub-classes to better classify an ingredient.

Figure 6.9: Flavouring: Sub-classes

Fruits Fruits also contain the seeds like vegetables and grow on plants. Many fruits are generally sweet and consumed by animals and humans. Most of the fruits don't need to be cooked and can be eaten directly, but they are also used as part of many recipes or individually to prepare juices, milk shakes, sweeteners, preserves and many more. The Fruits category is further divided into sub-classes to better classify an ingredient.

Figure 6.10: Fruits: Sub-classes

FatsAndOils Fats and Oils are used in preparing a lot of food dishes because they provide energy, transport fat soluble vitamins in the blood and insulate organs. They are the most abundant lipids, A triglyceride in its solid state is called a fat and oil when it is in its liquid state. The fatsAndOils category is further divided into sub-classes to better classify an ingredient.

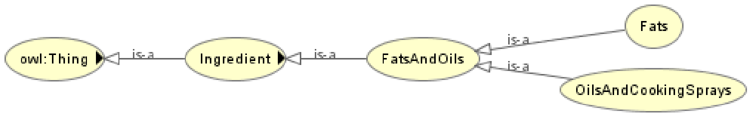


Figure 6.11: FatsAndOils: Sub-classes

Grains Grains are edible seeds, they do taste bland but they contain the highest amount of nutrients in them, have the lowest fat, and very cheap. The Food Guide Pyramid has grains at its base. The Grains category is further divided into sub-classes to better classify an ingredient.

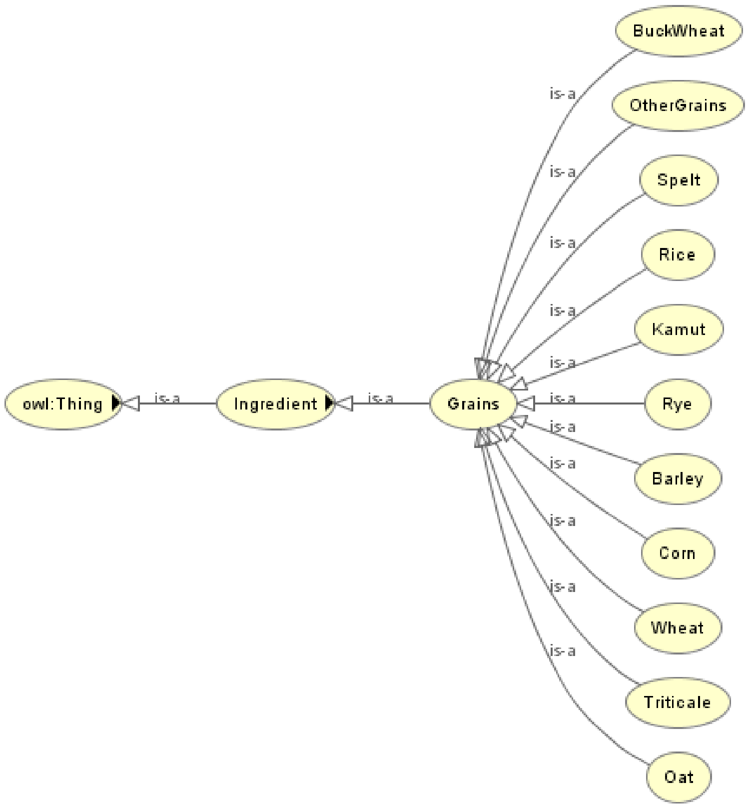


Figure 6.12: Grains: Sub-classes

Liquids Liquids are used very commonly in our daily meals, whether it is water used to boil or milk for a milk-shake. We use different liquids as ingredients to prepare solid food or new liquids drinks. The Liquids category is further divided into sub-classes to better classify an ingredient.

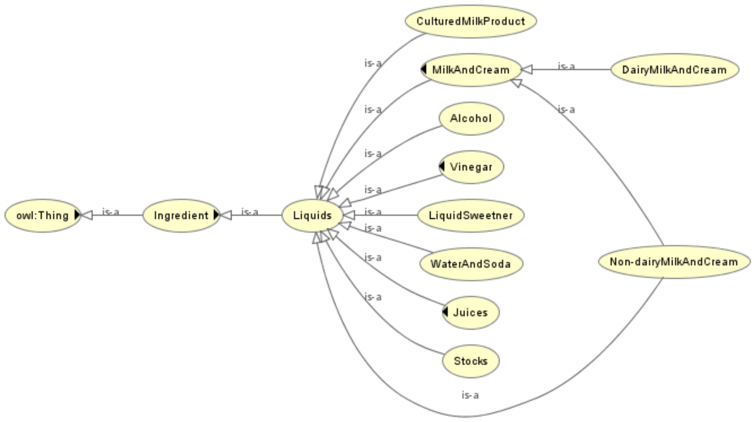


Figure 6.13: Liquids: Sub-classes

Meats Meat is the consumable flesh of animals used by humans as food. Gathering food and hunting animals for meat has been a practice since prehistoric times. According to nutritionists, Meat contains high protein and fat concentrations. A huge number of recipes focus on cooking meat based diet. The Meat category is further divided into sub-classes to better classify an ingredient.

Figure 6.14: Meat: Sub-classes

LegumesAndNuts A legume comes from the plant species, any plant that has a seed pods when ripe opens up along its sides is called a legume. Lentils, peas, peanuts, Beans, and soybeans are some of the most common legumes used in meals. A composite of the seed and the fruit which doesn't split to release the seed inside is called a nut. The LegumesAndNuts category is further divided into sub-classes to better classify an ingredient.

Figure 6.15: LegumesAndNuts: Sub-classes

Ingredient Type

The ingredients in recipe are assigned a role in the recipe, and that role is specified by Ingredient type. Ingredients can have one of the following types:

- Main Ingredient
- Side Ingredient
- Spice

A typical recipe will have at least one main ingredient, a number of side ingredients, and spices. This definition of role will help us identify and replace

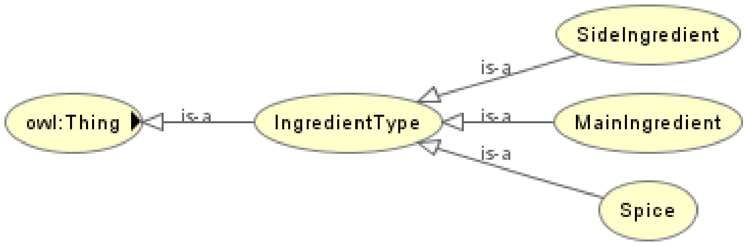


Figure 6.16: IngredienType: Sub-classes

the ingredients better in Recipe Mutation process. Figure **??** shows how this role assignment can help us create a tree of our recipe.

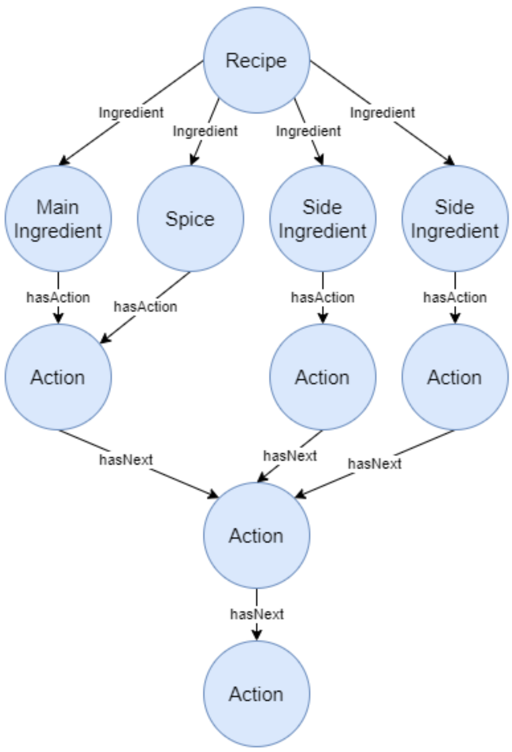


Figure 6.17: Ingredient - Action Recipe Tree

6.1.3 Instruction

Instructions are also a key attribute in recipe, without them a recipe is incomplete. Instructions are a set of directions with a list of ingredients for making or preparing something, especially food thus making it an important class in our recipe model. Instructions consist of individual actions which are performed on ingredients in either of the preparation, cooking or presentation phase.

Action

Action is a sub-class of Instructions.The Actions will play an important role in the mutation of recipe, as we can construct a recipe graph based on these actions and respective ingredients. Actions is further classified into following:

- Heating
  - DryHeat
    - Frying
    - Roasting
  - MoistHeat
  - CombinitionHeat
- Preparation
- OtherActions

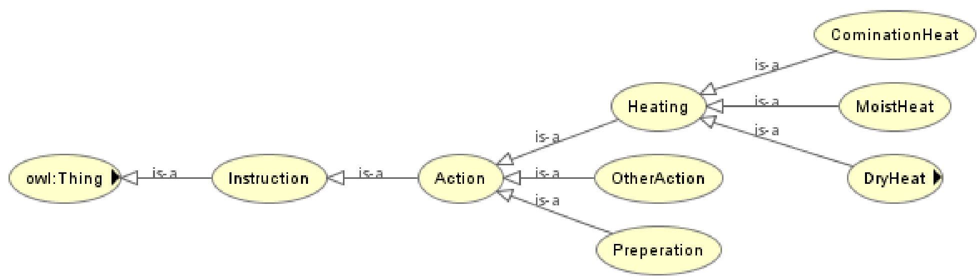


Figure 6.18: Instruction and its sub-classes

6.1.4 RecipeClassification

Recipes in our model are classified for better searching of recipe on the basis of time of day, cuisine style and recipe course. RecipeClassisification is another important class in our model which is further divided into the following sub-classes. Figure: 6.19 shows the RecipeClassification and its sub-classes.

- RecipeCourse

- MealType
- RecipeCuisine

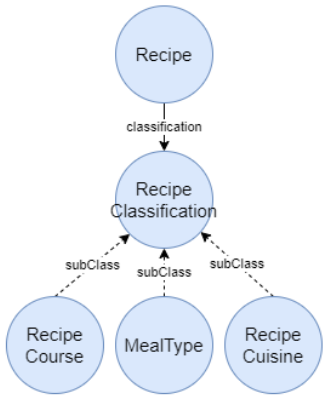


Figure 6.19: RecipeClassification and its sub-classes

RecipeCourse

In recipe, a course is a selection of food items that are served in combination during a meal. A full course meal consist of multiple dishes served over time, it generally begins with an appetizer like a soup which is served alone, followed by a selection of dishes from the main course and a side course i.e. salad along with some drinks. The recipe course in our model is divided into sub-classes for better classification.

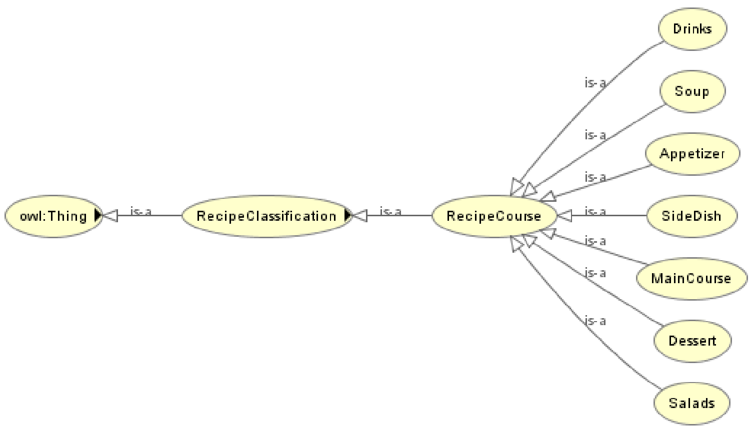


Figure 6.20: RecipeCourse: sub-classes

MealType

The Food or dish consumed with respect to a certain time of the day is called a meal. For example, the food eaten in the morning is called breakfast, afternoon is lunch, and night is dinner. The MealType classification is based on the time of the day. We have further divided it into sub-classes for better classification.

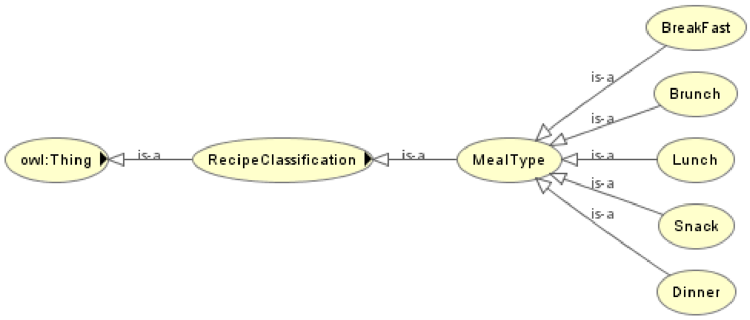


Figure 6.21: MealType and its sub-classes

RecipeCuisine

RecipeCuisine defines a certain method or style of cooking, mainly as characteristic of a particular geographical area, region, or a country. For example: A dish called biryani comes under Indian cuisines of recipe category. We have further divided it into sub-classes for better classification.

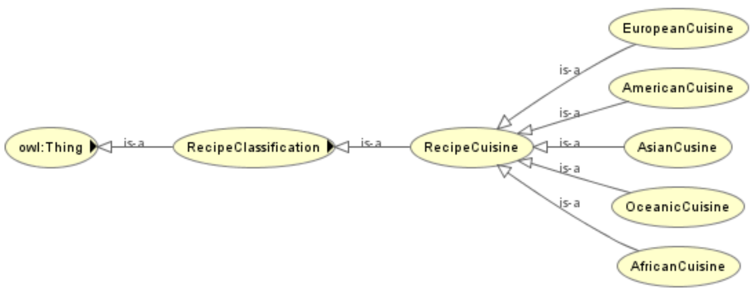


Figure 6.22: RecipeCuisine and its sub-classes

6.1.5 Nutrition and Nutrients

We get most of our nutrition through food, where the consumed food is digested and transformed into body tissues and energy for our body that's why it is recommended to eat regularly and have a balanced diet so our body has its nutrition requirements met and we look and feel our best. A recipe contains nutrient information based on the ingredients it contain, thus a choice in our recipe model to allow users the recipe search by nutrient value.

The Department of Health has classified the nutrients into six major types namely , vitamins, proteins, carbohydrates, minerals, water and fats (lipids). Each nutrient type has its own role to play in maintaining a good health and all the nutrients are required to work together to contribute to a healthy body.

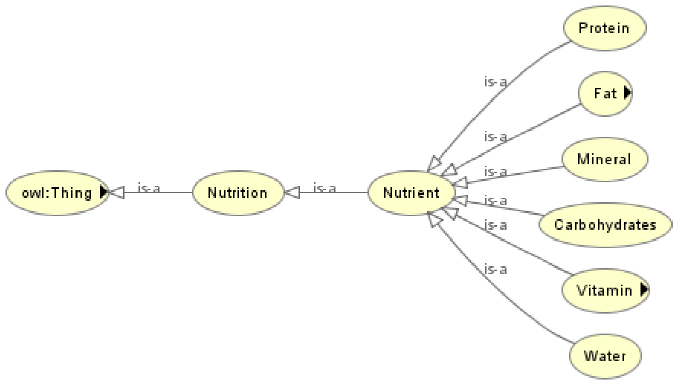
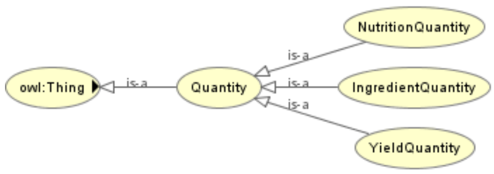


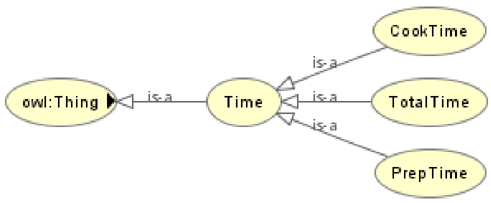
Figure 6.23: Nutrition and Nutrients

6.1.6 Other Classes

The Quantity class has relationship with Recipe, Ingredient and Nutrition Class. The quantity class defines the amount and unit used in recipe where as the Time class defines the TotalTime to cook a recipe along with CookTime and PrepTime.



(a) Quantity: Sub classes



(b) Time: Sub classes

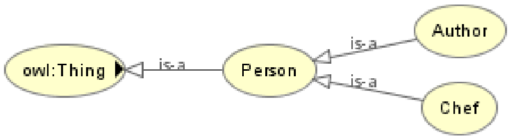
Figure 6.24: Quantity and Time

A recipe has other information associated with as well; the Source URL of the recipe, the Author or Publisher or Chef, DifficultyLevel, AggregateRating and the CookingUtensil. We took the classification of Cooking utensils from the cook's Theseus.

(a) Source

(b) AggregateRating

(c) DifficultyLevel



(d) Person

Figure 6.25: Source, AggregateRating, DifficultyLevel and Person

This information is helpful in grouping recipe based on difficulty. Searching Recipes based on Author and Publisher. And Sorting the queried results on the Aggregate Rating. We have also included this information in our recipeModel.

Figure 6.26: CookingUtensil

6.2 Recipe Ontology Properties

OWL is very similar to other OOP syntax languages as it allows class and property definition, however it is different as it focuses on semantics definition: relationships between classes, associations between classes and properties through interference, logical rules and constraints. For this reason ontologies are more property oriented rather than object oriented. In OWL we have recognized two main types of properties.

- Object Property (class owl:ObjectProperty): the object property relates two individuals.
- Datatype property (class owl:DatatypeProperty): the data property relates an individual to a data value.

owl:AnnotationProperty and owl:OntologyProperty classes are two other types of properties that are used for documentation of ontology instead of reasoning. All of these are subclass of rdf:Property.

The main component of an OWL ontology is a set of axioms. Axioms tell if a statement is true or false in a subject area. Axioms and the logical rules makes the overall speculation that the ontology is designed to show in its application domain. The property axioms are as follow:

- Old RDFS
  - rdfs:subPropertyOf
  - rdfs:domain
  - rdfs:range
- Property relation
  - owl:equivalentProperty
  - owl:inverseOf
- Cardinality constraint (global)
  - owl:FunctionalProperty
  - owl:InverseFunctionalProperty
- Logical charactersitics
  - owl:SymmetricProperty

- owl:TransitiveProperty
- owl:AsymmetricProperty
- owl:ReflexiveProperty

- Property Chain
  - owl:propertyChainAxiom

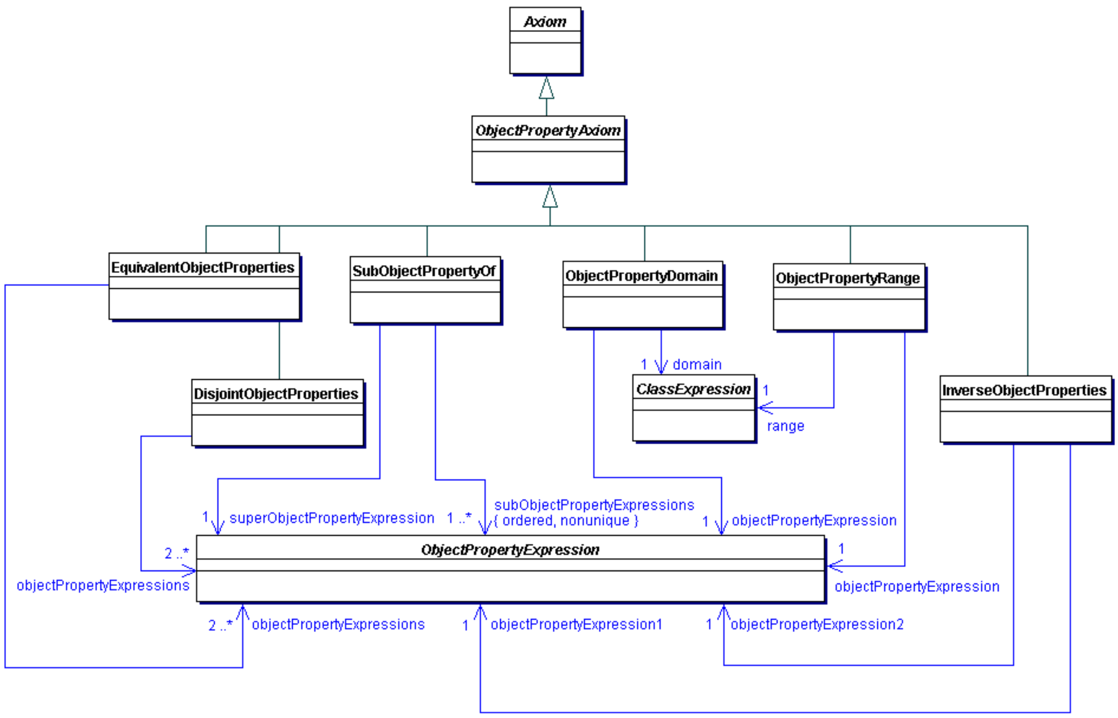


Figure 6.27: ObjectProperty Axiom

6.2.1 Object Property

Objects in an ontology can be described by relating them to attributes or independent things. Each attribute can be a class or an individual. The relation between an object and an attributes expresses a fact that is specific to the object. Object Properties specify how an object is related to another object. They connect two individuals (a subject and object) using a predicate. Figure: 6.28 shows the object properties defined in our ontology in Protege and table gives the details of the range and domain of each object property individually.

owl:Object Property	**rdfs:domain**	**rdfs:range**
aggregateRating	Recipe	AggregateRating
author	Recipe	Person
hasDifficultyLevel	Recipe	DifficultyLevel
hasIngredientAction	Ingredient	Action
hasIngredientQuantity	Ingredient	Quantity
hasIngredientType	Ingredient	IngredientType
hasMealType	Recipe	MealType
hasNutritionContent	Recipe	Nutrition
hasNutritionQuantity	Nutrition	Quantity
hasRecipeAction	Recipe	Action
hasRecipeCuisine	Recipe	RecipeCuisine
hasRecipeInstructions	Recipe	Instruction
hasNext	Instruction	Instruction
hasPrevious	Instruction	Instruction
hasRecipeType	Recipe	RecipeCourse
hasSubstitute	Ingredient	Ingredient
hasTotalTime	Recipe	TotalTime
hasCookTime	TotalTime	CookTime
hasPrepTime	TotalTime	PrepTime
recipeIngredient	Recipe	Ingredient
recipeYield	Recipe	Quantity
tool	Recipe	CookingUtensil
usedInRecipe	Ingredient	Recipe
wasDerivedFrom	Recipe	Source

Table 6.1: Recipe Ontology: List of Object Properties

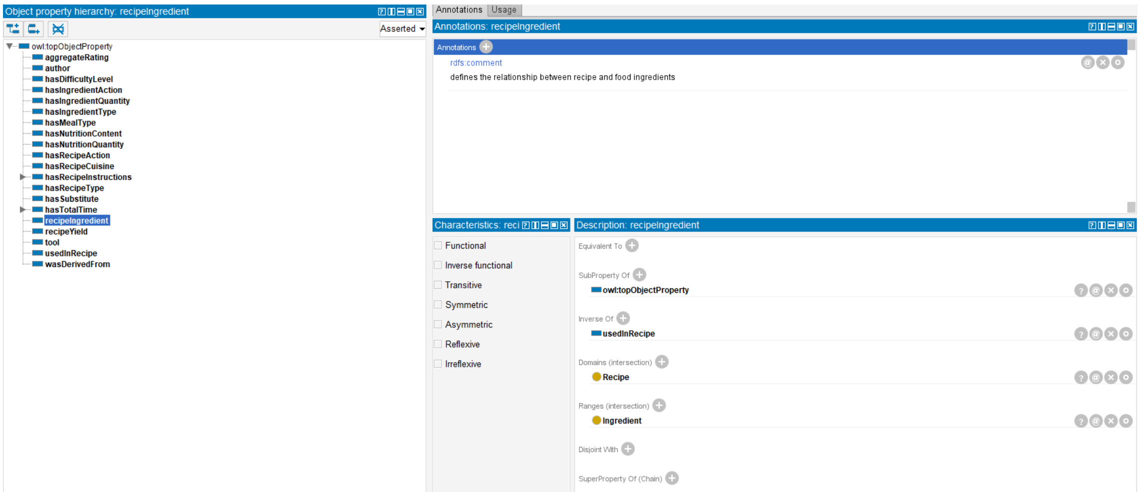


Figure 6.28: Protege: Object Properties

6.2.2 Data Property

Data properties have defined datatypes including string, integer, date, date-time, or boolean. The data properties have their domain as an individual where as their range is a literal connected via a predicate. The figure 6.29 shows the representation of data properties in Protege and table shows all the data properties used in our ontology.

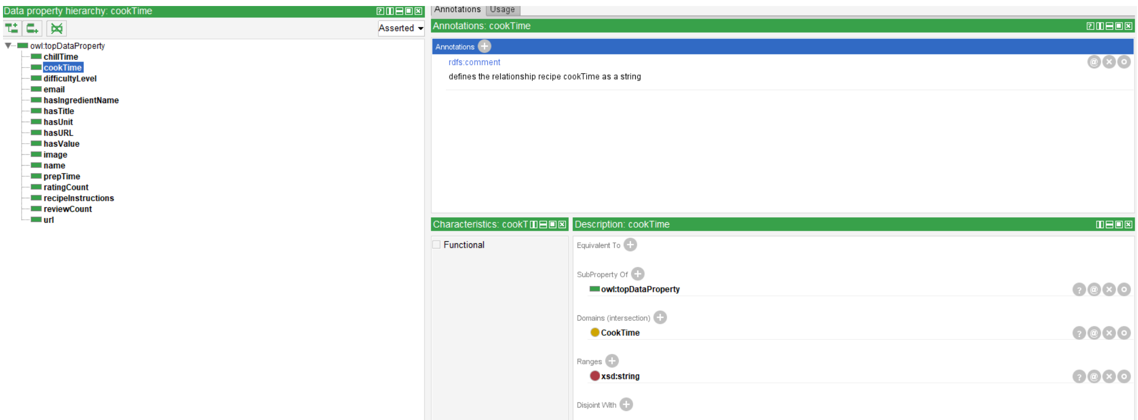


Figure 6.29: Protege: Data Properties

owl:dataProperties	**rdfs:domain**	**Rdfs:range**
chillTime	CookTime	xsd:string
cookTime	CookTime	xsd:string
difficultyLevel	DifficultyLevel	xsd:string
email	Person	xsd:string
hasIngredientName	IngredientType	xsd:string
hasTitle	Recipe	xsd:string
hasUnit	Quantity	xsd:string
hasURL	Person	xsd:string
hasValue	Quantity	xsd:string
image	Recipe	xsd:string
name	Person	xsd:string
prepTime	PrepTime	xsd:string
ratingCount	AggregateRating	xsd:integer
recipeInstructions	Instruction	xsd:string
reviewCount	AggregateRating	xsd:integer
url	Source	xsd:anyURI

Table 6.2: Recipe Ontology: List of Data Properties

Chapter 7

Evaluation

In this chapter, we evaluate the developed ontology through different scenarios and test cases.

7.1 Ontology Validation

We developed our ontology using the Protégé Editor tool. As discussed earlier, the ontology development process is a recurring process, where each stage is connected to every other stage. Once we were satisfied that our ontology is capable of dealing with the problem at hand, we mapped a few hand picked recipes from different websites belonging to different categories to evaluate its performance. Our ontology is capable of mapping recipes from all categories and sources. The final ontology metrics is shown in figure: 7.1

Metrics

Axiom	2048
Logical axiom count	1686
Declaration axioms count	276
Class count	231
Object property count	25
Data property count	19
Individual count	12
Annotation Property count	4

Figure 7.1: Ontology Metrics

Although, Protégé Editor tool is a strong tool for ontology development but it lacks visualization and without visualization verifying the connections between different entities and their relationships can be tough. We used different plugins (OWLViz, OntoGraph and VOWL) to visualize the components of our ontology. Apart from using the plugins we also used other ontology development tools to see the structure of our ontology and understand it better. Figure 7.2 shows the structure of our Recipe Ontology using the "WebVOWL: Web-based Visualization of Ontologies" tool.

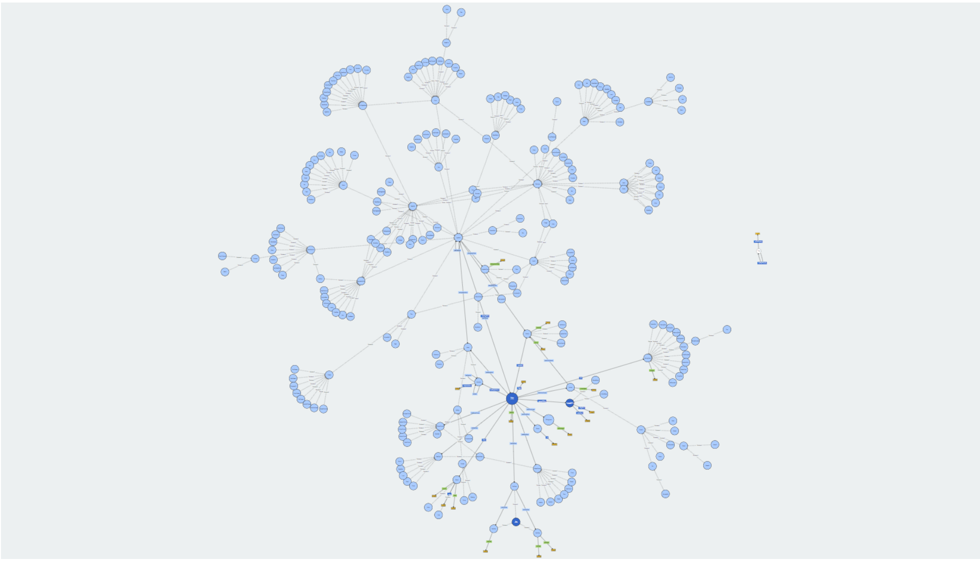


Figure 7.2: Ontology Representation - WebVowl

7.2 Recipe Ontology - Knowledge Graph

The next step include modeling the recipe data we extracted recipe data from different websites in Section: 5.1.1. Ontology rely on RDF/XML schema where data can be defined and linked using RDF and OWL so that there is more effective discovery, automation, integration and reuse across different applications. In the Semantic Web, RDF/OWL and linked open data is considered as a standard format for data integration, publication and visualization.

The data we extracted is in tabular form, and the tabular data does not conserve the domain structure and semantics, thus making it difficult to understand, combine and envision the data. There are a number of tools available that aim to facilitate the process of converting the tabular data in semantically structured interlinked data. We used the Open Refine tool to semantically construct the gather information stored in tabular format obtained from web scraping mentioned in the section: 5.1.1. The recipe data set needed to be cleaned up in ordered for the information to be accepted into the ontology. The clean up tasks included normalizing removing foreign characters, trimming extra spaces. The Figure 7.3 shows the columns of a small sample taken from the original data set and imported into the Open refine tool.

id	name	author	image	url	prepTime	cookTime	totalTime	yield	recipeYield
90054502	Home made Pita Bread Recipe	jkstern	http://s2.grouprecipes.com/images/recipes/200/8efc2f1f1423acb87367c5a29b790afa.jpg	http://www.grouprecipes.com/1000/home-made-pita-bread.html	NULL	NULL	8 minutes	NULL	8
90081408	Schnitz Pie Recipe	bondc	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/10000/schnitz-pie.html	NULL	NULL	120 minutes	NULL	8
90120134	melted cheese Recipe	cookingmasta12	http://s2.grouprecipes.com/images/recipes/200/6015526521.jpg	http://www.grouprecipes.com/100000/melted-cheese.html	NULL	NULL	15 minutes	NULL	1
90122576	N/A Recipe	grumpymomma	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100001/na.html	NULL	NULL	-	NULL	-
90104974	Pan Bagnet Recipe	wonderland_tina	http://s1.grouprecipes.com/images/recipes/200/3627356686.jpg	http://www.grouprecipes.com/100003/pan-bagnet.html	NULL	NULL	-	NULL	5
90039376	Nonnas Lemon Ricotta Biscuits Recipe	mrsc543	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100004/nonnas-lemon-ricotta-biscuits.html	NULL	NULL	20 minutes	NULL	12
90086934	Butterflied Lamb In Pomegranate Juice Recipe	improvcook	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100006/butterflied-lamb-in-pomegranate-juice.html	NULL	NULL	60 minutes	NULL	10
90007072	Crisp Lacy Cheese chips or baskets Recipe	jjals	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100007/crisp-lacy-cheese-chips-or-baskets.html	NULL	NULL	15 minutes	NULL	68
90075055	Pacific Salmon Recipe	22566	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100009/pacific-salmon.html	NULL	NULL	20 minutes	NULL	2
90066111	Orecchiette With Peppers And Sausages Recipe	darmorrow	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/10001/orecchiette-with-peppers-and-sausages.html	NULL	NULL	20 minutes	NULL	4
90055206	Loaf Dessert Recipe	22566	http://s1.grouprecipes.com/images/recipes/200/recipe.png	http://www.grouprecipes.com/100010/loaf-dessert.html	NULL	NULL	60 minutes	NULL	6
90029675	Shredded Buffalo Chicken Sandwiches Recipe	keni	http://s2.grouprecipes.com/images/recipes/200/7715319669.jpg	http://www.grouprecipes.com/100011/shredded-buffalo-chicken-sandwiches.html	NULL	NULL	360 minutes	NULL	8

Figure 7.3: Data Set sample in Open Refine

Figure 7.4 shows the set up for the mapping of the tabular data in our data set onto their respective predicates from our ontology.

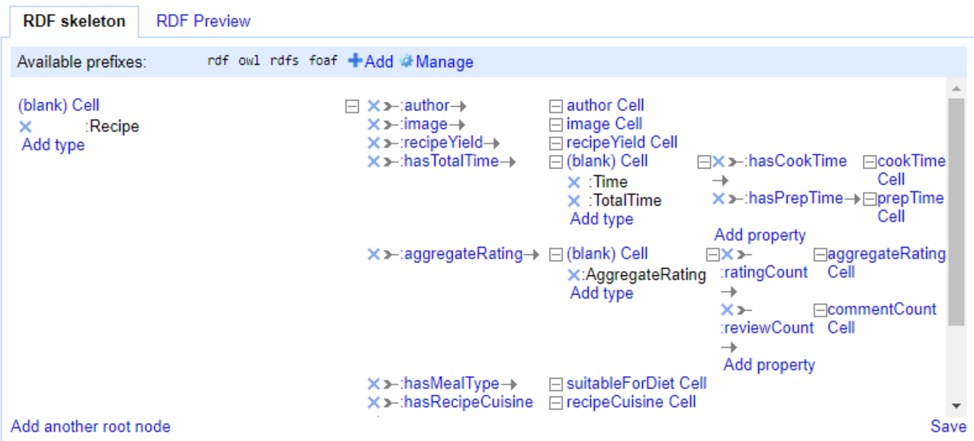


Figure 7.4: OpenRefine: RDF Skeleton

7.3 Recipe Ontology - Usability Test

In order to test the usability of the ontology and validate the functioning, we hand picked a few recipes from different sources, with different classification and wrote their RDF using the Turtle editor in order to mapped their attributes on to our ontology in the Protégé Editor tool as instances of the classes. Figure: 7.5 shows the split view of turtle editor where we can write the RDF in turtle format and it is visualized at the same time.

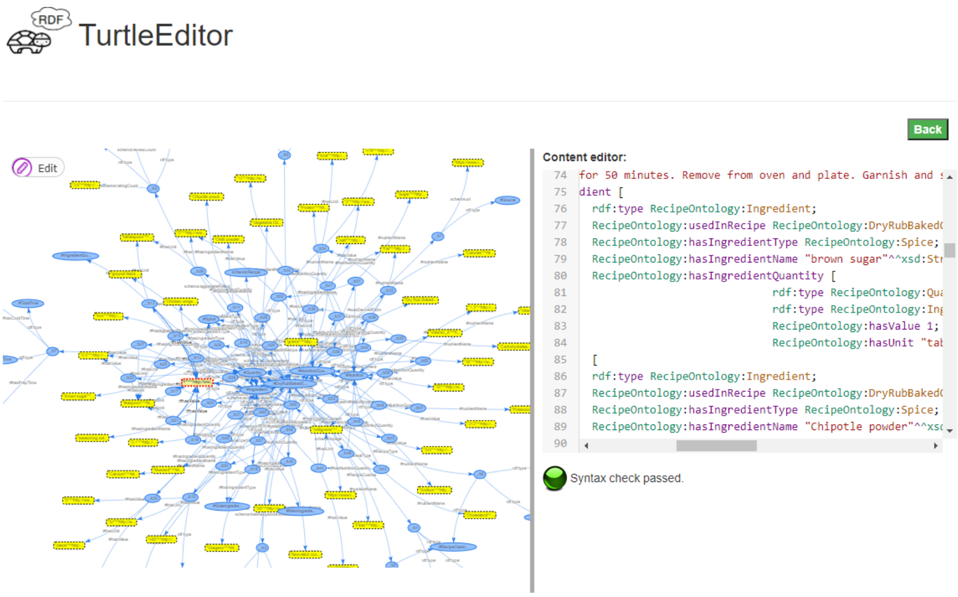


Figure 7.5: Turtle Editor - Split View

7.3.1 Recipe Ontology - Attribute based search

After reviewing the domain concepts and looking back at the competency questions mentioned in the section: 5.2, we queried our ontology using the SPARQL Query Plugin for Protégé Editor on different attributes of recipe *(Difficulty level, Recipe Cuisine, Meal Type, Recipe Course, Ingredients, Main Ingredient, Side Ingredient, Spices, Tool, Time, Rating)*

- **Recipes with "Moderate" difficulty level.**

```
SELECT ?subject ?name ?author
WHERE {
    ?subject rdf:type schema:Recipe.
     ?subject RecipeOntology:hasTitle ?name.
     ?subject  RecipeOntology:wasDerivedFrom ?source.
     ?source schema:url ?url.
     ?subject schema:author ?person.
     ?person schema:name ?author.
```

```
        ?subject RecipeOntology:hasDifficultyLevel
        RecipeOntology:Moderate.
    }
```

name	author	url
"Fried Rice"	"Gimme some Oven"	"https://www.gimmesomeoven.com/fried-rice-recipe
"Japanee Cheesecake Aka. Cotton Cheesecake"	"Baketo the root"	"https://www.yummly.com/recipe/Japanese-Cheese
"Spanakopita Twice-Baked Potatoes"	"SARAH WHARTON"	"https://www.simplyrecipes.com/twice-baked-potato
"Philly Cheesesteak"	"Nick Evans"	"https://www.simplyrecipes.com/philly-cheesesteak-

Figure 7.6: Moderate difficulty Recipes - Result

- **"Easy" Recipes for "Dinner".**

```
SELECT ?name ?url
WHERE {
     ?subject rdf:type schema:Recipe.
    ?subject RecipeOntology:hasTitle ?name.
     ?subject  RecipeOntology:wasDerivedFrom ?source.
     ?source schema:url ?url.
      ?subject RecipeOntology:hasMealType ?mt.
     ?mt rdf:type RecipeOntology:Dinner.
     ?subject RecipeOntology:hasDifficultyLevel
     RecipeOntology:Easy.
 }
```

name	url
"Easy Grilled Salmon"	"https://www.simplyrecipes.com/recipes/easy_grilled_salmon/?print"
"All American Beef Slider"	"https://www.agoudalife.com/wprm_print/9409"

Figure 7.7: Easy Dinner Recipes - Result

- **"American Cuisine" style "Appetizer" Recipes.**

```
SELECT ?name ?url
WHERE {
     ?subject rdf:type schema:Recipe.
    ?subject RecipeOntology:hasTitle ?name.
     ?subject  RecipeOntology:wasDerivedFrom ?source.
```

```
        ?source schema:url ?url.
        ?subject RecipeOntology:RecipeCuisine ?rc.
        ?rc rdf:type RecipeOntology:AmericanCuisine.
        ?subject RecipeOntology:RecipeType ?rt.
        ?rt rdf:type    RecipeOntology:Appetizer.
    }
```

name	url
"Baked Oatmeal with Mixed Berries"	"https://www.simplyrecipes.com/baked-oatmeal-with-mixed-berries-recipe-5185886?print"
"Dry Rub Baked Chicken Wings"	"https://www.lordbyronskitchen.com/wprm_print/19181"

Figure 7.8: American cuisine style appetizer - Result

- **Recipes with "Beef" in them.**

```
SELECT ?name ?url
WHERE {
        ?subject rdf:type schema:Recipe.
        ?subject RecipeOntology:hasTitle ?name.
        ?subject  RecipeOntology:wasDerivedFrom ?source.
        ?source schema:url ?url.
        ?subject schema:recipeIngredient ?ing.
        ?ing RecipeOntology:hasIngredientName ?ingname.
        FILTER regex(?ingname, "beef", "i")
    }
```

name	url
"All American Beef Slider"	"https://www.agoudalife.com/wprm_print/9409"

Figure 7.9: Beef Recipes - Result

- **Recipes with "Chicken" as "Main Ingredient".**

```
SELECT ?name ?url
WHERE {
        ?subject rdf:type schema:Recipe.
        ?subject RecipeOntology:hasTitle ?name.
```

```
    ?subject  RecipeOntology:wasDerivedFrom ?source.
    ?source schema:url ?url.
    ?subject schema:recipeIngredient ?ing.
    ?ing RecipeOntology:hasIngredientName ?ingname.
    ?ing RecipeOntology:hasIngredientType ?type.
    ?type rdf:type RecipeOntology:MainIngredient.
    FILTER regex(?ingname, "chicken", "i")
}
```

name	url
"Dry Rub Baked Chicken Wings"	"https://www.lordbyronskitchen.com/wprm_print/19181"

Figure 7.10: Chicken as Main Ingredient Recipe - Result

- **Top three highest "Rated" Recipes with their "Prep Time".**

```
SELECT ?name ?preptime ?rating
WHERE {
    ?subject rdf:type schema:Recipe.
    ?subject RecipeOntology:hasTitle ?name.
    ?subject schema:aggregateRating ?ar.
    ?ar schema:ratingCount ?rating.
    ?subject RecipeOntology:hasTotalTime ?time.
    ?time RecipeOntology:hasPrepTime ?prep.
    ?prep RecipeOntology:prepTime ?preptime.
}
ORDER BY DESC (?rating) LIMIT 3
```

name	preptime	rating
"Pinwheel Cookies"	"15 Minutes"	"5"^^<http://www.w3.org/2001/XMLSchema
"Fried Rice"	"5 Minutes"	"4.9"^^<http://www.w3.org/2001/XMLScher
"Japanee Cheesecake Aka. Cotton Cheesecake"	"30Minutes"	"4.5"^^<http://www.w3.org/2001/XMLScher

Figure 7.11: Top three recipe with their prep time - Result

7.3.2 Recipe Ontology - Condition appropriate search

In order to test the ontology usability for condition based searches, we executed some SPARQL queries on the base of Nutrition and nutrients.

- **Recipes with lowest "Calories"**

```
SELECT ?name ?value ?unit
WHERE {
     ?subject rdf:type schema:Recipe.
    ?subject RecipeOntology:hasTitle ?name.
     ?subject RecipeOntology:hasNutritionContent ?nc.
     ?nc RecipeOntology:nutrientName "Calories".
     ?nc RecipeOntology:hasNutritionQuantity ?nq.
     ?nq RecipeOntology:hasValue ?value.
     ?nq RecipeOntology:hasUnit ?unit.
}
ORDER BY ASC (?value) LIMIT 3
```

name	value	unit
"Japanee Cheesecake Aka. Cotton Cheesecake"	"380"^^<http://www.w3.org/2001	"kcal"
"Dry Rub Baked Chicken Wings"	"479"^^<http://www.w3.org/2001	"kcal"
"Easy Grilled Salmon"	"507"^^<http://www.w3.org/2001	"kcal"

Figure 7.12: Recipes with lowest Calories - Result

7.4 Recipe Ontology - Evolution

In order to show the potential of the Recipe ontology, we defined two test cases in a controlled environment and achieved recipe evolution. In the first case, we used the Ingredient substitution method and in the second case, we used recipe mutation to create new recipes.

7.4.1 Case 1: Ingredient Substitution

In our ontology, we have defined Ingredient class hierarchy. Each Ingredient belongs to a classification of Ingredient. Lets take potato, potato is a vegetable, and in vegetables potato belongs to Tuber class. Similarly, Pork is a Meat, it belongs to Curved Meat class and further sub classified into Bacon class.

In the table: 7.2 we have a recipe to cook chickpeas with chicken eggs in indian cusine style, we have performed ingredient substitution on it to generate a new recipe.

Table 7.1: Sample Recipe

Indian chickpeas with poached chicken eggs
Ingredients: *Yellow Pepper*, *Garlic*, *Rapeseed Oil*, *Red Chilli*, *Spring Onion*, Cumin seeds, Coriander, Tomatoes, Chickpeas, Reduced Salt powder, *Chicken Eggs* **Instructions** 1. Heat the oil, add the pepper, garlic, spring onions, chilli, and fry it for 5 minutes over a medium to high heat. 2. Add the tomatoes, coriander, spices, and the chickpeas to the cooking pan and sauté it for 1 to 2 more minutes, then stir it in the bouillon powder. 3. Boil water, crack in the eggs and poach for 2 minutes. Stir the spring onions with the chickpeas, scatter the reserved coriander, and top with the eggs.

Using the class hierarchy of the Ingredients categorization we can achieve evolution through Ingredient substitution. In the recipe shown in table: 7.2 we have multiple ingredients which can be substituted with ingredients belonging to same class.

- Rapeseed Oil *"is a"* Ingredient *"is a"* FatsAndOils *"is a"* OilsAndCookingSprays.

  Substitute: Canola Oil, Lear Oil.

- Yellow Pepper *"is a"* Ingredient *"is a"* Vegetable *"is a"* SweetPepper.

  Substitute: Bell Pepper, Red Pepper, Green Pepper, Cubanelle.

- Red Chilli *"is a"* Ingredient *"is a"* Vegetable *"is a"* FruitVegetable *"is a"* FreshChile

  Substitute: Hot pepper, Chilli pepper

- Spring Onion *"is a"* Ingredient *"is a"* Vegetable *"is a"* Onion.

  Substitute: Sweet onion, Fresh onion.

- Chicken Eggs *"is a"* Ingredient *"is a"* Dairy *"is a"* Eggs

  Substitute: Duck eggs.

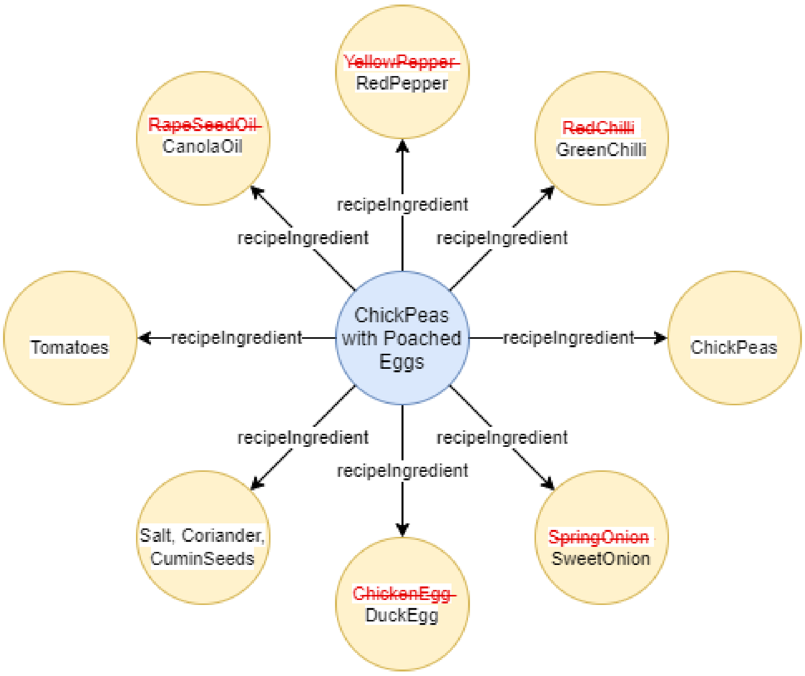


Figure 7.13: Indian chickpeas with Eggs- Ingredients substituion

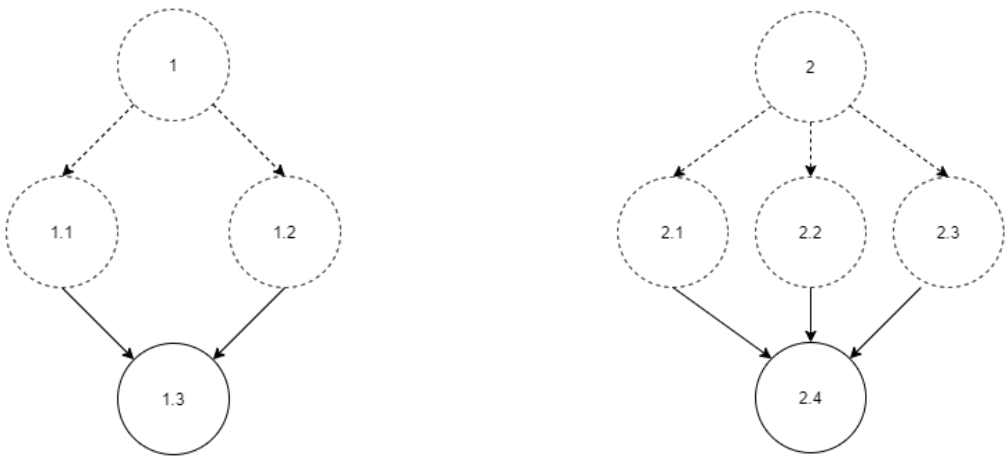
Figure: 7.13 shows our recipe ingredients in a graph being replaced with their equivalent ingredients and In the table: 7.3 we can see the resultant recipe after the ingredients are substituted with other ingredients belonging to same class.

Table 7.2: Recipe after Ingredient substitution

Indian chickpeas with poached duck eggs
Ingredients:
Canola Oil, *Garlic*, *Red Pepper*, *Green Chilli*, *sweet Onion*, Cumin seeds, Coriander, Tomatoes, Chickpeas, Reduced Salt bouillon powder, *Duck Eggs*
Instructions
1. Heat the oil, add the garlic, pepper, chilli and the sweet onions, and fry it for 5 minutes over a medium to high heat.
2. Add the tomatoes, coriander, spices, and the chickpeas to the cooking pan and sauté it for 1 to 2 more minutes, then stir it in the bouillon powder.
3. Boil water, crack in the eggs and poach for 2 minutes. Stir the spring onions with the chickpeas, scatter the reserved coriander, and top with the eggs.

7.4.2 Case 2: Recipe Mutation

In our ontology, we have defined IngredientType classes *(MainIngredient, SideIngredient, Spice)*. Each Ingredient belongs to a type based on the role it is playing in the recipe. For example: Potato are Main ingredient in one recipe: Potato fries, where as they are used a side ingredient in another recipe: Beef steak with mashed potatoes. Similarly, other ingredients can also play different roles in a recipe depending on the type of recipe. Using the Ingredient type classification we can mutate two recipes to create a new recipe by swapping the ingredients and their respective instructions based on their role.



(a) Mashed Sweet Potatoes (b) Potato Fried Rice

Figure 7.14: Sample Recipe for Mutation

Figure 7.14 shows the graph representation of two recipes which we have chosen for mutation represented as a graph. The nodes show the instructions and the edge show the next instruction in line.

In the table: 7.3 we have two recipes, where potato is used as a main ingredient.We have performed mutation by swapping the ingredients and their respective instruction to create a child recipe. We have taken the main ingredient, side ingredients and their respective cooking instructions from parent recipe 01 and included or replaced them with their respective instructions from parent recipe 02. We performed this experiment in a controlled environment, where the recipes were pre-selected and normalized as per the requirements in order to show the potential of our developed ontology. The inspiration and idea of this experiment was taken from a similar experiment performed in EvoChef: Show me what to cook.[30]

Table 7.3: Recipes for Mutation

Parent recipe # 01	*Parent recipe # 02*
Mashed Sweet Potatoes	**Potato Fried Rice**
Ingredients:	**Ingredients:**
Sweet Potatoes, Milk, Butter, Maple Syrup	Avocado Oil, Potatoes, Bell pepper, Garlic, Cooked Brown rice, Coconut aminos, liquid aminos, rice vinegar
Instructions:	Instructions:
1. Take a large pot of salted water and boil the potatoes for 20 to 30 minutes until tender then remove and mash them.	1. In a small bowl, whisk together the coconut aminos, liquid aminos, and rice vinegar. Add the avocado oil to a large skillet over medium to high heat.
2. Set the electric mixer on low, and blend the potatoes. with 1/2 a cup milk. Use more milk if required to achieve desired texture. Add maple syrup and butter as per taste, keep blending until it becomes smooth.	2. Once the oil is warm, add in the potatoes, chopped bell pepper, and scallions. Cook, stir often, until cooked.
3. Serve warm.	3. Add in the brown rice and sauce. Keep the heat on high and cook it til the rice is glossy, warmed through, and has absorbed all the flavors in the sauce.
	4. Serve warm.

We have taken the sweet Potatoes from the Parent recipe 01, and we have included it into the Parent recipe 02 as a result we have a recipe with sweet mashed potatoes with rice instead of fried potatoes with rice. The respective instruction 1.1 associated with the sweet potatoes is also included in the child recipe. We can also do the vice versa in some cases but for that we need to work on better training our model in selecting the correct recipes, ingredient replacement and instructions swapping for better and more accurate mutation to create novel recipes.

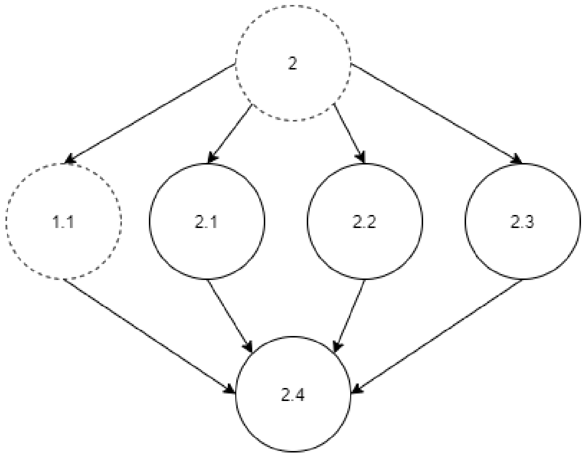


Figure 7.15: Child recipe after mutation

The figure: 7.15 shows the graphical representation of the new child recipe created through mutation shown in Table: 7.4.

Table 7.4: Child Recipes through Mutation

Child recipe # 01

Mashed Sweet Potatoes Rice
Ingredients:
Sweet Potatoes, Milk, Butter, Maple Syrup, Avacado Oil, Bell pepper, Garlic, Coconut aminos, liquid aminos, rice vinegar, Cooked brown rice
Instructions:
1. Take a large pot of salted water and boil the potatoes for 20 to 30 minutes until tender then remove and mash them.
2. In a small bowl, whisk together the coconut aminos, liquid aminos, and rice vinegar. Add the avocado oil to a large skillet over medium to high heat.
3. Once the oil is warm, add in the potatoes, chopped bell pepper, and scallions. Cook, stir often, until cooked.
4. Serve warm

Chapter 8

Conclusions and Future Work

8.1 Conclusion

Our research introduced a new methodology for the representation and modelling of the Food recipe process. The novelty of our research lays in our use of ontology base approach for modeling the Cooking process and its use for improving the recipe recommendation and automated recipe generation systems. The goal of our research was to develop a common vocabulary of concepts and their relationships to model the Recipe domain while making it machine understandable.

The proposed ontology consists of two parts, It standardizes the representation method for recipes, and makes the recipes machine understandable. Thus improving the existing recipe recommendation systems. Additionally, modelling of these concepts allows the possibility to create and generate new recipes computationally.

Our research efforts focused mainly on the design and modeling of food recipe. We covered the core aspects of a cooking process so it can be reused later on, as well as added depth to our classes for better expression and future extensions. A number of information resources and data sets were used to gather information in order to define the domain and model the ontology as per our need. After data gathering the schema structure was created, we focused on creating classes, properties, and relationship amongst the entities. After the modelling, the data was converted to RDF and mapped on the ontology for answering the competency questions. We passed our ontology through different test cases to analyze its performance and capabilities. Our test cases included, mapping of recipes from different sources to create a knowledge graph. Querying the knowledge graph on different attributes of the recipe and a controlled experiment to showcase the evolution of recipes.

We did our best efforts and we are confident that our research proved fruit full and covered the points mentioned in our problem statement. This research created an ontology that describes the domain and range of the recipe by (1) enabling the mapping of all kinds of recipes available over the internet to acts as a knowledge base for recipes, (2) re-usability of the knowledge base to evolve the recipes automatically and (3) selecting condition-appropriate meals.

Table 8.1 shows a comparison between the Food ontologies available in the market with our FoodRecipe ontology.

Table 8.1: Comparison of different 'food ontologies'

System Name	**Type of Food**	**Technology**	**Features and Coverage**
FoodWiki	*Packaged Food*	*OWL, RDF*	Mobile App, Limited only to food products included in the ministry's database.
AGROVOC	*All kind of Food*	*OWL, RDF, and SKOS/-XL*	Comprehensive coverage for international food but complex. Only experts can use it.
FoodOn	*Packaged Food*	*OWL, RDF*	Very extensive, doesn't cover the cooking process. For organization use only.
Food-KG	*All kind of Food*	*OWL, RDF, SPARQL*	Web-based. Limited parameters, doesn't cover all recipe attributes, missing actions and ingredients classification for evolution
FoodRecipe	*All kind of Food*	*OWL, RDF, Turtle, and SPARQL*	Full recipe attributes coverage, attribute-based and condition appropriate search capability, extension, and evolution possible.

8.1.1 Challenges and Limitations: Data set

When gathering and analyzing recipe data, there are issues to be addressed. Not all recipes have the same representation in a uniformed manner, For example time representation is different in recipes from different websites. This requires that the data after collection needs to be normalized which can require lots of effort. Sometimes the ingredient list isn't complete and the use of spices is mentioned in the instructions.

Another issue is that every recipes doesn't contain all the required information. Some blogs are very descriptive but some have just the minimal information required to cook food. For example, Some recipes have nutrition information missing. Issues like these are common when working with large data sets. Creating and using controlled vocabulary is one way of addressing this issue but it will require creating a standard after normalizing the data set first through NLP, or manual efforts.

8.1.2 Limitation of Study

The large part ontology design is inspired after researching the existing systems, available recipe websites and talks with domain experts. Although the design is simple and fulfil all the requirements but as we have mentioned before that an ontology is as good as its mapping on the use case. One ontology can be very good for one use case but bad for another use cases. Further discussion with domain experts, and its use can show certain limitation which we are unaware of at the time being.

Another limitation is related to ontology testing over large data sets. Keeping the limitation mentioned in section: 8.1.1 we normalized a very small subset of the data from our collection and tested our ontology on that. Further sub classification can be added to different classes bases on new feedback and new use cases to increase the potential for extension.

8.2 Future work

As we have identified some limitation in the section: 8.1, our future work include focusing on those limitations: data normalization and using our ontology as a knowledge base. Afterwards we can test the performance of our ontology on the complete knowledge base and generate better results.

Secondly, we are taking forward this ontology and building an automated novel recipe generation and recommendation system. The actual use of this ontology will yield better understanding of our needs in evolution, thus giving us more room for growth and finer tuning to answer more questions in this domain.

Another hope is to make our ontology more user friendly, the current ontology requires knowledge of OWL, XML and semantic web to understand it and answer the question. After extending the ontology, we will design a web application or a mobile app that will be easy to use and more interactive, they system will communicate with the knowledge base through APIs.

Finally, we are working on a research article to publish in the upcoming food conferences and present our ontology there for better reach and recognition. Our ontology will be available for discussion and feed backs from domain experts will show us different perspectives and areas for improvement. Our further research include the food conferences where we can participate, better data normalization techniques and new application area for extension.

Although an abundance of ontologies in different domains are available on the internet, but ontologies dedicated to food are few in numbers and for the purpose of this research they are even more reduced if we consider the subject area as Recipes. So we think that our foodRecipe ontology will be a great addition to the existing food related ontologies available.

Bibliography

[1] L. Byron, “Dry rub chicken wings- recipe.” Retrieved from: www.lordbyronskitchen.com, Accessed: July 2021.

[2] E. Simperl, R. Krummenacher, and L. Nixon, “A coordination model for triplespace computing,” pp. 1–18, 06 2007.

[3] G. Trends, “Interest over time ”recipe” vs ”food near me”,” 2021. Retrieved from: https://trends.google.com, Accessed: July 2021.

[4] GoogleTrends, “Think with google,” 2021. Retrieved from: https://www.thinkwithgoogle.com, Accessed: July 2021.

[5] Hunter, “Hunter food study special report,” 2021. Retrieved from: https://www.hunterpr.com/, Accessed: July 2021.

[6] O. recipe, “Oxford learners dictionary - recipe.” Retrieved from: https://www.oxfordlearnersdictionaries.com/, Accessed: July 2021.

[7] K. Hub, “Ontology definition.” Retrieved from: https://www.ontotext.com/knowledgehub/fundamentals/what-are-ontologies/, Accessed: July 2021.

[8] N. Noy and D. Mcguinness, “Ontology development 101: A guide to creating your first ontology,” *Knowledge Systems Laboratory*, vol. 32, 01 2001.

[9] W3Org, “Swt definition.” Retrieved from: https://www.w3.org/standards/semanticweb/, Accessed: July 2021.

[10] SWT, “Swt challenges.” Retrieved from: https://semwebtec.wordpress.com/, Accessed: July 2021.

[11] LOV, “Linked open vocabularies.” Retrieved from: https://lov.linkeddata.es/dataset/lov/about, Accessed: July 2021.

[12] USDA, “Food and nutrition — usda.” Retrieved from: https://www.usda.gov/topics/food-and-nutrition, Accessed: July 2021.

[13] euroFIR, “Food database — eurofir.” Retrieved from: https://www.eurofir.org/food-information/, Accessed: July 2021.

[14] Agrovoc, “Linked data — agrovoc.” Retrieved from: http://www.fao.org/agrovoc/linked-data, Accessed: July 2021.

[15] A. McCann, J. Pennington, E. C. Smith, J. Holden, D. Soergel, and R. C. Wiley, “Fda’s factored food vocabulary for food product description.,” *Journal of the American Dietetic Association*, vol. 88 3, pp. 336–41, 1988.

[16] J. Marín, A. Biswas, F. Ofli, N. Hynes, A. Salvador, Y. Aytar, I. Weber, and A. Torralba, “Recipe1m+: A dataset for learning cross-modal embeddings for cooking recipes and food images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, pp. 1–1, 07 2019.

[17] Yummly, “Data set — yummly.” Retrieved from: http://lherranz.org/local/datasets/yummly28k/Yummly28K.rar, Accessed: July 2021.

[18] S. Haussmann, O. Seneviratne, Y. Chen, Y. Ne’eman, J. Codella, C.-H. Chen, D. Mcguinness, and M. Zaki, *FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation*, pp. 146–162. 10 2019.

[19] D. Dooley, E. Griffiths, G. Gosal, P. L. Buttigieg, R. Hoehndorf, M. Lange, L. Schriml, F. Brinkman, and W. Hsiao, “Foodon: a harmonized food ontology to increase global food traceability, quality control and data integration,” *npj Science of Food*, vol. 2, 12 2018.

[20] BBC, “Bbc - ontologies - food ontology,” 2014. Retrieved from: https://www.bbc.co.uk/ontologies, Accessed: July 2021.

[21] M. Musen, N. Noy, N. Shah, P. Whetzel, C. Chute, M. Storey, and B. Smith, “The national center for biomedical ontology,” *Journal of the American Medical Informatics Association : JAMIA*, vol. 19 2, pp. 190–5, 2012.

[22] T. Eftimov, G. Ispirova, D. Potočnik, N. Ogrinc, and B. Seljak, “Iso-food ontology: A formal representation of the knowledge within the domain of isotopes for food science,” *Food Chemistry*, vol. 277, 10 2018.

[23] D. Celik Ertuğrul, "Foodwiki: Ontology-driven mobile safe food consumption system," *TheScientificWorldJournal*, vol. 2015, p. 475410, 07 2015.

[24] M. Sam, A. A. Krisnadhi, C. Wang, J. Gallagher, and P. Hitzler, "An ontology design pattern for cooking recipes - classroom created," in *WOP*, 2014.

[25] C. Snae and M. Brückner, "Foods: A food-oriented ontology-driven system," pp. 168 – 176, 03 2008.

[26] J. Youn, T. Naravane, and I. Tagkopoulos, "Using word embeddings to learn a better food ontology," *Frontiers in Artificial Intelligence*, vol. 3, 11 2020.

[27] G. Brusa, M. Caliusco, and O. Chiotti, "A process for building a domain ontology: an experience in developing a government budgetary ontology," vol. 72, 01 2006.

[28] M. Hepp, "Goodrelations: An ontology for describing products and services offers on the web," in *EKAW*, 2008.

[29] H. Jabeen, J. Weinz, and J. Lehmann, "Autochef: Automated generation of cooking recipes," pp. 1–7, 07 2020.

[30] H. Jabeen, N. Tahara, and J. Lehmann, *EvoChef: Show Me What to Cook! Artificial Evolution of Culinary Arts*, pp. 156–172. 03 2019.

[31] J. Freyne and S. Berkovsky, "Recommending food: Reasoning on recipes and ingredients," vol. 6075, pp. 381–386, 06 2010.

[32] E. Cromwell, "Computational creativity in the culinary arts (abstract only)," pp. 715–715, 02 2015.

[33] M. Kazama, M. Sugimoto, C. Hosokawa, K. Matsushima, L. Varshney, and Y. Ishikawa, "A neural network system for transformation of regional cuisine style," *Frontiers in ICT*, vol. 5, 07 2018.

[34] W. S. Web, "W3c semantic web— resource description framework (rdf)." Retrieved from: https://www.w3.org/2001/sw/wiki/RDF, Accessed: August 2021.

[35] W. S. Web, "W3c semantic web— rdf schema (rdfs)." Retrieved from: https://www.w3.org/2001/sw/wiki/RDFS, Accessed: August 2021.

[36] W. S. Web, "W3c semantic web— web ontology language (owl)." Retrieved from: https://www.w3.org/2001/sw/wiki/OWL, Accessed: August 2021.

[37] L. Alden, "The cook's thesaurus," 2006. Accessed: 2021-06-01.