



Práctica 3: Sistemas de Pregunta-Respuesta (SPR)

Ingeniería Lingüística

Luna Jiménez Fernández
Alejandro Muñoz Navarro

MUIA - UPM
Curso: 2020/2021

Índice general

1. Introducción	1
2. Diseño del sistema	2
2.1. Selección de documentos	2
2.1.1. Receta 1	2
2.1.2. Receta 2	4
2.2. Extracción de la tabla de dimensiones	5
2.2.1. Análisis de documentos	5
2.2.2. Tabla de dimensiones	7
2.3. Preguntas tipo identificadas	9
2.4. Parafraseo de preguntas tipo	14
2.5. Diseño de reglas	18
3. Implementación y prueba del sistema	24
3.1. Formato de las recetas	24
3.2. Funcionamiento del sistema	26
3.2.1. Carga y procesamiento de los documentos	27
3.2.2. Pre-procesamiento de la pregunta	28
3.2.3. Interpretación y respuesta a la pregunta	29
3.3. Implementación y uso del sistema	30
3.3.1. Implementación y dependencias	30
3.3.2. Manual de usuario del sistema	30
3.4. Pruebas de funcionamiento del sistema	30
4. Conclusiones	34
A. Contenidos del fichero entregable	35

Índice de figuras

2.1. Análisis de las dimensiones de la receta	6
3.1. Prueba del sistema - Inicialización	31
3.2. Prueba del sistema - Pregunta y respuesta (1)	31
3.3. Prueba del sistema - Preguntas y respuestas (2)	32
3.4. Prueba del sistema - Interpretación de variabilidad (1)	32
3.5. Prueba del sistema - Interpretación de variabilidad (2)	33
3.6. Prueba del sistema - Pregunta no reconocida	33
3.7. Prueba del sistema - Fin de la ejecución	33

Capítulo 1

Introducción

El objetivo de esta práctica es el desarrollo de un **sistema de pregunta-respuesta** (*Query-Answering System* en inglés), un tipo de **sistema de extracción de información** encargado de responder a preguntas del usuario respecto a un **espacio de información** previamente definido.

Para el desarrollo de esta práctica se ha optado por definir un sistema de pregunta-respuesta con dominio de **recetas de cocina**, trabajando con **dos** recetas.

La estructura de la memoria es la siguiente:

- En el **Capítulo 2** se detallará el **diseño del sistema** realizado, explicando cada uno de los pasos propuestos en el enunciado:
 1. Selección y análisis de los documentos
 2. Creación de la tabla de dimensiones
 3. Creación y parafraseo de las preguntas tipo para cada dimensión
 4. Diseño de reglas para la identificación de las preguntas por parte del sistema
- En el **Capítulo 3** se expondrá la **implementación y prueba del sistema** que se ha llevado a cabo, centrándose en:
 - Las **estructuras de datos** utilizadas por la implementación.
 - La **implementación del sistema** llevada a cabo.
 - El **manual de usuario del sistema**, para indicar su uso adecuado.
 - Algunas **pruebas del correcto funcionamiento del sistema**, junto a observaciones.
- En el **Capítulo 4** se finalizará con unas **conclusiones** obtenidas tras el desarrollo y testeo del sistema, referentes al proceso y a su eficiencia.

Además, se incluye un apéndice “**Contenidos del fichero entregable**” donde se desglosan las carpetas y ficheros disponibles en el entregable junto a esta memoria.

Capítulo 2

Diseño del sistema

En este capítulo se detallará todo el proceso de **diseño** del sistema de pregunta-respuesta previo a la implementación.

Para esto se comenzará con la **selección de los documentos** para formar el repositorio, seguido del **análisis** de los documentos para la extracción de la **tabla de dimensiones**. Tras esto, se plantearán **preguntas tipo** para cada dimensión propuesta (junto a sus **parafraseos** más esperados) y finalmente se diseñarán las **reglas** para identificar las preguntas y las respuestas por el sistema.

2.1. Selección de documentos

Como se mencionó en la introducción, se ha optado por definir un espacio de información centrado en la **cocina**, concretamente en las **recetas de cocina**. Se ha optado por este tipo de documentos por considerarse documentos **simples** (suelen ser relativamente cortos y concisos) y **muy estructurados** (lo que facilita notablemente la tarea de análisis y extracción de la información).

Concretamente, se ha optado por usar **dos** recetas de Alberto Chicote, disponibles en el enlace <http://www.albertochicote.com/recetas/>. Para cada receta se mostrará el enlace específico a su localización junto a la receta, adaptada al formato de este documento.

Las recetas seleccionadas son las siguientes:

2.1.1. Receta 1

Terrina de Foie Grass con sobaos pasiegos y confitura de zanahorias (disponible haciendo click aquí).

Ingredientes:

- 4 sobaos pasiegos
- aceite de oliva virgen extra
- sal maldon

Para la terrina de Foie-gras:

- 800 g de hígado de pato desvenado
- 3 cl de brandy de Jerez
- 2 g de pimienta negra recién molida
- 13 g de sal

Para la confitura de zanahorias:

- 250 g de zanahorias
- 500 g de naranjas
- 75 g de azúcar
- sal

PROCEDIMIENTO:

Terrina de foie-gras:

1. Colocar el hígado en una bandeja de acero y aliñar con la sal, la pimienta negra y el brandy. Dejarlos 4 horas a temperatura ambiente tapados con film transparente.
2. Colocar los hígados en un molde de papel de aluminio, procurando que las caras exteriores queden hacia el molde.
3. Precalentar el horno a 200 0C y meter el molde dentro, con un termómetro pinchado en el hígado.
4. Apagar el horno inmediatamente y dejar los foie-gras dentro hasta que el termómetro nos indique que la temperatura en el corazón del molde es de 52 0C.
5. Tapar y enfriar rápidamente. Conservar en la nevera.

Confitura de zanahorias:

1. Poner las zanahorias y las naranjas peladas con el azúcar en un cazo con agua fría.
2. Llevar a ebullición y mantener a fuego suave hasta que las zanahorias estén bien tiernas.
3. Escurrir y triturar, añadiendo poco a poco el líquido de la cocción hasta que la confitura esté fina y brillante.
4. Poner a punto de sal y enfriar rápido.

PRESENTACIÓN:

Cortar los sobaos en lonchas de unos 0,7 centímetros de grosor y tostar en el horno justo antes de usarlos. Colocar un poco de confitura de zanahorias sobre los sobaos tostados y encima un trozo de foie-gras. Poner un poco más de confitura de zanahorias. Aliñar con unas gotas de aceite de oliva virgen extra y unas escamas de sal.

2.1.2. Receta 2

Chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas (disponible haciendo click aquí).

Ingredientes:

- 200 gramos de harina tamizada
- 1 cucharadita de moka de sal fina
- 10 Vueltas de molinillo de pimienta negra
- 175 gramos de mantequilla
- 50 gramos de aceite de oliva virgen extra
- 200 gramos de chocolate negro
- 440 gramos de azúcar moreno
- 4 huevos

Para la salsa acida de cacao:

- $\frac{1}{2}$ dcl de Vinagre de Jerez
- 50 gramos de cacao en polvo

PROCEDIMIENTO:

1. Calentamos el horno a 170°C
2. Calentamos el vinagre de jerez un poco, añadimos el cacao en polvo y mezclamos muy bien. Reservamos.
3. Mezclamos la sal y el harina y tamizamos juntas con un colador fino Reservamos.
4. Fundimos el chocolate con la mantequilla y el aceite al baño maria, hasta que la mezcla quede homogénea. Añadimos la mitad del azúcar y mezclamos bien. Retiramos del calor.
5. Ponemos los huevos enteros con el azúcar restante y batimos, sin montarlos. Separamos esta mezcla en dos partes mas o menos iguales.
6. Mezclamos una de las mitades de la mezcla de huevos con el chocolate, mezclando con una espátula hasta que la mezcla quede lisa.
7. Montamos, ahora si, la otra mitad hasta que la mezcla quede espumosa y blanquecina.
8. Mezclar los dos batidos con cuidado para que no se baje mucho. Poco a poco, con una espátula y de abajo arriba.
9. Añadimos la pimienta y mezclamos al fin la mezcla de harina y sal con mucho cuidado.

10. Enmondamos poniendo mantequilla y harina si fuese necesario, cocemos unos 20 minutos. El corazón ha de quedar cremoso como un brownie.

Una vez cocido, lo dejamos atemperar, lo desmoldamos y lo refrigeramos.

PRESENTACIÓN:

Para servirlo, cortamos una ración en frío, unos tres trozos pequeños y lo calentamos unos 10 segundos en un microondas. Servimos con las flores escarchadas y la salsa alrededor.

2.2. Extracción de la tabla de dimensiones

En esta sección se describirá la **extracción** de información realizada para la creación de la tabla de dimensiones. Para esto se comentará el **análisis** aplicado a los documentos para extraer la información relevante, y posteriormente se presentará la **tabla de dimensiones** obtenida de ambos documentos.

2.2.1. Análisis de documentos

De cara al análisis de documentos, las recetas utilizadas tienen una estructura **altamente definida** (todas las recetas siguen una estructura con divisiones claras entre cada punto) y **similar entre ellas** (al provenir todas de la misma página web). Estos puntos facilitan notablemente la extracción de información (al encontrarse esta ya directamente categorizada en el propio documento).

En la Figura 2.1 se puede observar una de las recetas con las **dimensiones identificadas**. De forma **explícita** se han identificado las siguientes categorías:

- **Receta:** El título de la receta.
- **Ingredientes:** Los ingredientes contenidos en la receta. Estos pueden pertenecer a la receta directamente o a alguna de las partes de la receta.
- **Pasos:** Los pasos seguidos para cocinar la receta. Estos pueden pertenecer a la receta directamente o a alguna de las partes de la receta.
- **Presentación:** Las indicaciones para presentar el plato una vez cocinado.
- **Fecha:** Fecha en la que se publicó la receta.
- **Cantidades:** Cantidad que se necesita de cada ingrediente para la receta.
- **Partes de la receta:** Partes que forman la receta. Se entiende como **parte** de la receta cada uno de los elementos del plato en la que se divide la receta. Por ejemplo, en la Figura 2.1 se puede ver como la *Terrina de foie-gras con sobaos pasiegos y confitura de zanahorias* se divide en dos partes, la *terrina de foie-gras* y la *confitura de zanahorias*.
- **Duraciones:** Duraciones indicadas durante la preparación de las recetas, principalmente tiempos de cocinado.

- **Temperaturas:** Temperaturas indicadas durante la preparación de las recetas, principalmente a la hora de cocinar ingredientes.

RECETA

Terrina de Foie Grass con sobaos pasiegos y configura de zanahorias

★★★★★

26 Diciembre 2014

Comparta esta receta: [Twitter](#) [Facebook](#) [Google+](#) [Pinterest](#) [Email](#)

FECHA

PROCEDIMIENTO: PASOS

PARTES DE LA RECETA

Terrina de foie-gras

- Colocar el hígado en una bandeja de acero y aliñar con la sal, la pimienta negra y el brandy. Dejarlos 4 horas a temperatura ambiente tapados con film transparente.
- Colocar los hígados en un molde de papel de aluminio, procurando que las caras exteriores queden hacia el molde.
- Precalentar el horno a 200 OC y meter el molde dentro, con un termómetro pinchado en el hígado.
- Apagar el horno inmediatamente y dejar los foie-gras dentro hasta que el termómetro nos indique que la temperatura en el corazón del molde es de 52 OC.
- Tapar y enfriar rápidamente. Conservar en la nevera.

Confitura de zanahorias

- Poner las zanahorias y las naranjas peladas con el azúcar en un cazo con agua fría.
- Llevar a ebullición y mantener a fuego suave hasta que las zanahorias estén bien tiernas.
- Ecurrir y triturar, añadiendo poco a poco el líquido de la cocción hasta que la confitura esté fina y brillante.
- Poner a punto de sal y enfriar rápido.

INGREDIENTES

Ingredientes:

- 4 sobaos pasiegos
- aceite de oliva virgen extra
- sal maldon

Para la terrina de Foie-gras

- 800 g de hígado de pato desvenado
- 3 cl de brandy de Jerez
- 2 g de pimienta negra recién molida
- 13 g de sal

Para la confitura de zanahorias

- 250 g de zanahorias
- 500 g de naranjas
- 75 g de azúcar
- sal

TEMPERATURAS

PRESENTACIÓN

PRESENTACIÓN:

Cortar los sobaos en lonchas de unos 0,7 centímetros de grosor y tostar en el horno justo antes de usarlos. Colocar un poco de confitura de zanahorias sobre los sobaos tostados y encima un trozo de foie-gras. Poner un poco más de confitura de zanahorias. Aliñar con unas gotas de aceite de oliva virgen extra y unas escamas de sal.

Figura 2.1: Análisis de las dimensiones de la receta

De forma **implícita** se ha identificado también una dimensión:

- **Autor:** Autor de la receta. En este caso, *Alberto Chicote*.

Es importante destacar que, si bien puede haber otras dimensiones que son importantes para recetas (como puede ser el **tiempo de preparación**), al no estar incluidas en las recetas elegidas no se tendrán en cuenta. Se entiende que, de añadir una receta al sistema que incluyese dimensiones nuevas, estas se incluirían.

Además, es necesario mencionar que es probable que existan errores a la hora de la **clasificación de los datos**, al no tener experiencia en el campo de la cocina (lo que puede inducir a errores en el análisis). Estos errores se solventarían con la ayuda de un experto en el campo.

2.2.2. Tabla de dimensiones

La tabla de dimensiones obtenida a partir del análisis de los documentos (como se puede ver en la Tabla 2.1) es la siguiente:

Dimensión	Valores	Observaciones
Receta	'Terrina de foie-gras con sobaos pasiegos y confitura de zanahorias', 'Chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas'	Nombre del plato a cocinar
Partes de las recetas	'Terrina de foie-gras', 'Confitura de zanahorias', 'Salsa acida de cacao'	Partes contenidas en una receta. No todas las recetas tienen partes.
Ingredientes	'sobaos pasiegos', 'aceite de oliva virgen extra', 'sal maldon', 'harina tamizada', 'moka de sal fina', 'pimienta negra', 'mantequilla', 'chocolate negro', 'azúcar moreno', 'huevos', 'hígado de pato desvenado', 'brandy de Jerez', 'pimienta negra recién molida', 'sal', 'zanahorias', 'naranjas', 'azúcar', 'vinagre de Jerez', 'cacao en polvo'	Ingredientes de las recetas. Los ingredientes pertenecen a una receta y, opcionalmente, a la parte de una receta.
Cantidades	'4', '200 gramos', '1 cucharadita', '10 vueltas de molinillo', '175 gramos', '50 gramos', '200 gramos', '440 gramos', '800 g', '3 cl', '2 g', '13 g', '250 g', '500 g', '75 g', '1/2 dcl', '50 gramos'	Cantidades necesarias de cada ingrediente para una receta. La cantidad esta siempre asociada a un ingrediente en una receta o parte.
Temperaturas	'170°C', '200°C', '52°C', 'ambiente'	Temperaturas a las que se deben cocinar los ingredientes. Las temperaturas se asocian a los ingredientes.
Tiempos	'4 horas', '10 segundos', '20 minutos'	Tiempos necesarios para cocinar los ingredientes. Los tiempos se asocian a los ingredientes

Table 2.1 continued from previous page

Dimensión	Valores	Observaciones
Procedimiento	<p>'Calentamos el horno a 170 °C',</p> <p>'Calentamos el vinagre de jerez un poco, añadimos el cacao en polvo y mezclamos muy bien. Reservamos.',</p> <p>'Mezclamos la sal y el harina y tamizamos juntas con un colador fino. Reservamos.',</p> <p>'Fundimos el chocolate con la mantequilla y el aceite al baño maria, hasta que la mezcla quede homogénea. Añadimos la mitad del azúcar y mezclamos bien. Retiramos del calor.',</p> <p>'Ponemos los huevos enteros con el azúcar restante y batimos, sin montarlos.</p> <p>Separamos esta mezcla en dos partes mas o menos iguales.',</p> <p>'Mezclamos una de las mitades de la mezcla de huevos con el chocolate, mezclando con una espátula hasta que la mezcla quede lisa.',</p> <p>'Montamos, ahora si, la otra mitad hasta que la mezcla quede espumosa y blanquecina.',</p> <p>'Mezclar los dos batidos con cuidado para que no se baje mucho. Poco a poco, con una espátula y de abajo arriba.',</p> <p>'Añadimos la pimienta y mezclamos al fin la mezcla de harina y sal con mucho cuidado.',</p> <p>'Enmondamos poniendo mantequilla y harina si fuese necesario, cocemos unos 20 minutos.</p> <p>El corazón ha de quedar cremoso como un brownie.',</p> <p>'Una vez cocido, lo dejamos atemperar, lo desmoldamos y lo refrigeramos.',</p> <p>'Colocar el hígado en una bandeja de acero y aliñar con la sal, la pimienta negra y el brandy. Dejarlos 4 horas a temperatura ambiente tapados con film transparente.',</p> <p>'Colocar los hígados en un molde de papel de aluminio, procurando que las caras exteriores queden hacia el molde.',</p> <p>'Precalentar el horno a 200 °C y meter el molde dentro, con un termómetro pinchado en el hígado.',</p> <p>'Apagar el horno inmediatamente y dejar los foie-gras dentro hasta que el termómetro nos indique que la temperatura en el corazón del molde es de 52 °C.',</p> <p>'Tapar y enfriar rápidamente. Conservar en la nevera.',</p> <p>'Poner las zanahorias y las naranjas peladas con el azúcar en un cazo con agua fría.',</p> <p>'Llevar a ebullición y mantener a fuego suave hasta que las zanahorias estén bien tiernas.',</p> <p>'Ecurrir y triturar, añadiendo poco a poco el líquido de la cocción hasta que la confitura esté fina y brillante.',</p> <p>'Poner a punto de sal y enfriar rápido.'</p>	<p>Instrucciones a seguir para cocinar una receta</p>

Table 2.1 continued from previous page

Dimensión	Valores	Observaciones
Presentación	<p>'Para servirlo, cortamos una ración en frio, unos tres trozos pequeños y lo calentamos unos 10 segundos en un microondas. Servimos con las flores escarchadas y la salsa alrededor.'</p> <p>'Cortar los sobaos en lonchas de unos 0,7 centímetros de grosor y tostar en el horno justo antes de usarlos. Colocar un poco de confitura de zanahorias sobre los sobaos tostados y encima un trozo de foie-gras. Poner un poco más de confitura de zanahorias. Aliñar con unas gotas de aceite de oliva virgen extra y unas escamas de sal.'</p>	Indicaciones para la presentación del plato tras ser cocinado.
Autor	'Alberto Chicote'	Autor de la receta
Fecha	'26/12/2014'	<p>Fecha en la que se publica la receta.</p> <p>La fecha se espera en formato dd/mm/yyy</p>

Cuadro 2.1: Tabla de dimensiones del sistema

2.3. Preguntas tipo identificadas

En esta sección se detallarán las **preguntas tipo** obtenidas a partir de la tabla 2.1 en la sección anterior.

En general, en vez de detallar la pregunta tipo para cada valor de cada dimensión (al haber una gran cantidad de elementos), se ha optado por utilizar **variables** para representar **cualquier valor de la dimensión**. Por tanto, se presentarán las preguntas tipo **genéricas** que se realizarán sobre cada dimensión, junto a un ejemplo de cada pregunta y su respuesta.

Además, es necesario saber que algunas recetas contienen partes dentro de si mismas (como si de subrecetas se tratasen). Por esto, debido a la flexibilidad de nuestras preguntas, siendo validas tanto para las recetas como para sus partes, se ha optado por no duplicar las preguntas y usar directamente *< nombre_del_plato >* para denotar ambos conceptos. Esto significa que, cuando sea relevante, la pregunta devolverá una respuesta apropiada dependiendo de si se ha pedido información sobre la receta (el plato entero) o una parte de la receta (solo la información que concierne a dicha parte).

Las preguntas tipo identificadas para cada dimensión de la tabla son las siguientes:

Preguntas sobre las recetas

1. ¿Qué recetas hay?

Esta pregunta da como resultado el nombre de todas las recetas.

Como resultado obtendríamos los siguientes:

- *Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias*
- *Chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas*

2. ¿Qué recetas publicó *< nombre_del_autor >*?

Esta pregunta da como resultado el nombre de todas las recetas que publicó un autor dado. Un ejemplo con nuestras recetas podría ser: ¿Qué recetas publicó Alberto Chicote?

Como resultado obtendríamos las siguientes recetas:

- *Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias*
- *Chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas*

Preguntas sobre las partes de las recetas

1. ¿Qué partes tiene la receta *< nombre_de_la_receta >*?

Esta pregunta da como resultado el nombre de todas las partes de una receta dada. Un ejemplo con nuestras recetas podría ser: ¿Qué partes tiene la receta Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias?

Como resultado obtendríamos los siguientes:

- *Terrina de foie-gras*
- *Confitura de zanahorias*

Preguntas sobre los ingredientes

1. ¿Cuáles son los ingredientes de *< nombre_del_plato >*?

Esta pregunta da como resultado los ingredientes necesarios para preparar el plato dado. Un ejemplo con nuestras recetas podría ser: ¿Cuáles son los ingredientes de la terrina de Foie-gras?

Como resultado obtendríamos los siguientes ingredientes:

- *hígado de pato desvenado*
- *brandy de Jerez*
- *pimienta negra recién molida*
- *sal*

2. **¿Qué platos llevan** *< nombre_del_ingrediente >*?

Esta pregunta da como resultado los platos cuya preparación requiere del ingrediente dado. Un ejemplo con nuestras recetas podría ser: ¿Qué platos llevan naranjas?

Como resultado obtendríamos los platos siguientes:

- *Confitura de zanahorias*

3. **¿Qué platos puedo preparar con** [*< nombre_del_ingrediente >*, *< nombre_del_ingrediente >*, *< nombre_del_ingrediente >*, ...]?

Esta pregunta da como resultado los platos cuya preparación se podría realizar con todos o unos pocos ingredientes de la lista dada. Un ejemplo con nuestras recetas podría ser: ¿Qué plato puedo preparar con brandy de Jerez, pimienta negra recién molida, sal, zanahorias, naranjas, azúcar, vinagre de Jerez y cacao en polvo?

Como resultado obtendríamos los platos siguientes:

- *Confitura de zanahorias*
- *Salsa ácida de cacao*

Preguntas sobre las cantidades

1. **¿Qué cantidad de** *< nombre_del_ingrediente >* **se necesita para** *< nombre_del_plato >*?

Esta pregunta da como resultado la cantidad necesaria de un ingrediente dado para cocinar un plato dado. Un ejemplo con nuestras recetas podría ser: ¿Qué cantidad de sal se necesita para la terrina de Foie-gras?

Como resultado obtendríamos: *13 g*

Preguntas sobre las temperaturas

1. **¿A qué temperatura debo calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?**

Esta pregunta da como resultado la temperatura a la que deberíamos calentar el horno para este plato específico.

Como resultado obtendríamos: *170 °C*

2. **¿A qué temperatura debo precalentar el horno para la terrina de foie-gras?**

Esta pregunta da como resultado la temperatura a la que deberíamos precalentar el horno para este plato específico.

Como resultado obtendríamos: *200 °C*

3. **¿A qué temperatura debe estar el molde de la terrina de foie-gras?**

Esta pregunta da como resultado la temperatura a la que debería estar el molde para este plato específico.

Como resultado obtendríamos: *52 °C*

4. **¿A qué temperatura debo dejar la terrina de foie-gras?**

Esta pregunta da como resultado la temperatura a la que deberíamos dejar este plato específico.

Como resultado obtendríamos: *ambiente*

Preguntas sobre los tiempos

1. **¿Cuánto tiempo debo dejar la terrina de foie-gras a temperatura ambiente?**

Esta pregunta da como resultado el tiempo al que deberíamos dejar este plato específico en temperatura ambiente.

Como resultado obtendríamos: *4 horas*

2. **¿Cuánto tiempo debe estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?**

Esta pregunta da como resultado el tiempo al que deberíamos dejar este plato específico en el microondas.

Como resultado obtendríamos: *10 segundos*

3. **¿Cuánto tiempo debo cocer el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?**

Esta pregunta da como resultado el tiempo al que deberíamos dejar cocer este plato específico.

Como resultado obtendríamos: *20 minutos*

Preguntas sobre el procedimiento

1. **¿Cuál es el procedimiento para cocinar *< nombre_del_plato >*?**

Esta pregunta da como resultado el procedimiento para cocinar un plato dado. Un ejemplo con nuestras recetas podría ser: ¿Cuál es el procedimiento para cocinar la confitura de zanahorias?

Como resultado obtendríamos los pasos siguientes:

- *Poner las zanahorias y las naranjas peladas con el azúcar en un cazo con agua fría.*
- *Llevar a ebullición y mantener a fuego suave hasta que las zanahorias estén bien tiernas.*
- *Escurrir y triturar, añadiendo poco a poco el líquido de la cocción hasta que la confitura esté fina y brillante.*
- *Poner a punto de sal y enfriar rápido.*

Preguntas sobre la presentación

1. ¿Cómo se presenta *< nombre_del_plato >*?

Esta pregunta da como resultado la presentación del plato dado. Un ejemplo con nuestras recetas podría ser: ¿Cómo se presenta el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?

Como resultado obtendríamos: *Cortar los sobaos en lonchas de unos 0,7 centímetros de grosor y tostar en el horno justo antes de usarlos. Colocar un poco de confitura de zanahorias sobre los sobaos tostados y encima un trozo de foie-gras. Poner un poco más de confitura de zanahorias. Aliñar con unas gotas de aceite de oliva virgen extra y unas escamas de sal.*

Preguntas sobre los autores

1. ¿Quién es el autor de la receta *< nombre_del_plato >*?

Esta pregunta da como resultado el nombre del autor que publicó la receta dada. Un ejemplo con nuestras recetas podría ser: ¿Quién es el autor de la receta Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias?

Como resultado obtendríamos: *Alberto Chicote*

2. ¿Qué autores hay?

Esta pregunta da como resultado el nombre de todos los autores.

Como resultado obtendríamos: *Alberto Chicote*

Preguntas sobre las fechas

1. ¿Cuándo se publicó la receta *< nombre_del_plato >*?

Esta pregunta da como resultado la fecha en la que se publicó la receta dada. Un ejemplo con nuestras recetas podría ser: ¿Cuándo se publicó la receta Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias?

Como resultado obtendríamos: *26/12/2014*

2. ¿Cuándo se publicaron las recetas de *< nombre_del_autor >*?

Esta pregunta da como resultado las fechas en la que se publicaron las recetas de un autor dado. Un ejemplo con nuestras recetas podría ser: ¿Cuándo se publicaron las recetas de Alberto Chicote?

Como resultado obtendríamos:

- *Terrina de Foie Gras con sobaos pasiegos y confitura de zanahorias - 26/12/2014*
- *Chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas - 26/12/2014*

2.4. Parafraseo de preguntas tipo

A la hora de enfrentarse el sistema a las preguntas del usuario, es probable que se encuentre con una gran **variabilidad lingüística** a la hora de interpretar las preguntas, ya que es posible expresar una misma idea de muchas formas.

Por esto, se ha optado por ofrecer **posibles parafraseos** de todas las preguntas tipos planteadas, para ver formas en las que el usuario podría pedir la información. Hay que destacar que esto es solo un subconjunto de posibles preguntas, sin tener en cuenta variaciones morfológicas. Una descripción exhaustiva de todos los posibles parafraseos sería imposible o prácticamente inabordable por la enorme cantidad de combinaciones, sinónimos, expresiones...

Preguntas sobre las recetas

1. ¿Qué recetas hay?

- ¿Qué recetas hay disponibles?
- ¿Qué recetas puedo cocinar?
- ¿Qué recetas tiene el sistema?
- ¿Qué recetas contiene el sistema?

2. ¿Qué recetas publicó *< nombre_del_autor >*?

- ¿Qué recetas escribió *< nombre_del_autor >*?
- ¿Qué recetas tiene *< nombre_del_autor >*?
- ¿Qué recetas hay de *< nombre_del_autor >*?
- ¿Cuáles son las recetas que escribió *< nombre_del_autor >*?
- ¿Cuáles son las recetas que tiene *< nombre_del_autor >*?
- ¿Cuáles son las recetas que publicó *< nombre_del_autor >*?

Preguntas sobre las partes de las recetas

1. ¿Qué partes tiene la receta *< nombre_de_la_receta >*?

- ¿Qué partes contiene la receta *< nombre_de_la_receta >*?
- ¿Qué partes forman parte de la receta *< nombre_de_la_receta >*?
- ¿Cuáles son las partes que contiene la receta *< nombre_de_la_receta >*?
- ¿Cuáles son las partes contenidas en la receta *< nombre_de_la_receta >*?
- ¿Cuáles son las partes de la receta *< nombre_de_la_receta >*?

Preguntas sobre los ingredientes

1. ¿Cuáles son los ingredientes de *< nombre_del_plato >*?

- ¿Qué lleva el plato *< nombre_del_plato >*?
- ¿Qué tiene el plato *< nombre_del_plato >*?
- ¿Qué ingredientes necesita el plato *< nombre_del_plato >*?

- ¿Qué ingredientes forman parte del plato $\langle \text{nombre_del_plato} \rangle$?
 - ¿Qué ingredientes lleva el plato $\langle \text{nombre_del_plato} \rangle$?
 - ¿Con qué ingredientes puedo hacer el plato $\langle \text{nombre_del_plato} \rangle$?
 - ¿Con qué ingredientes puedo cocinar el plato $\langle \text{nombre_del_plato} \rangle$?
2. **¿Qué platos llevan $\langle \text{nombre_del_ingrediente} \rangle$?**
- ¿Qué platos necesitan $\langle \text{nombre_del_ingrediente} \rangle$?
 - ¿Qué platos tienen $\langle \text{nombre_del_ingrediente} \rangle$?
 - ¿Cuáles son los platos que necesitan $\langle \text{nombre_del_ingrediente} \rangle$?
 - ¿Cuáles son los platos que llevan $\langle \text{nombre_del_ingrediente} \rangle$?
 - ¿Cuáles son los platos que tienen $\langle \text{nombre_del_ingrediente} \rangle$?
3. **¿Qué platos puedo preparar con $[\langle \text{nombre_del_ingrediente} \rangle, \langle \text{nombre_del_ingrediente} \rangle, \langle \text{nombre_del_ingrediente} \rangle, \dots]$?**
- ¿Qué puedo cocinar con $[\langle \text{nombre_del_ingrediente} \rangle, \langle \text{nombre_del_ingrediente} \rangle, \langle \text{nombre_del_ingrediente} \rangle, \dots]$?

Preguntas sobre las cantidades

1. **¿Qué cantidad de $\langle \text{nombre_del_ingrediente} \rangle$ se necesita para $\langle \text{nombre_del_plato} \rangle$?**
- ¿Qué cantidad de $\langle \text{nombre_del_ingrediente} \rangle$ es necesaria para el plato $\langle \text{nombre_del_plato} \rangle$?
 - ¿Cuál es la cantidad de $\langle \text{nombre_del_ingrediente} \rangle$ necesaria para el plato $\langle \text{nombre_del_plato} \rangle$?
 - ¿Cuántos $\langle \text{nombre_del_ingrediente} \rangle$ necesita el plato $\langle \text{nombre_del_plato} \rangle$?

Preguntas sobre las temperaturas

1. **¿A qué temperatura debo calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?**
- ¿A qué temperatura tengo que calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿A qué temperatura necesito calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿Cuál es la temperatura a la que necesito calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿Cuál es la temperatura a la que debo calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿Cuál es la temperatura a la que tengo que calentar el horno para el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
2. **¿A qué temperatura debo precalentar el horno para la terrina de foie-gras?**
- ¿A qué temperatura tengo que precalentar el horno para la terrina de foie-gras?
 - ¿A qué temperatura necesito precalentar el horno para la terrina de foie-gras?

- ¿Cuál es la temperatura a la que necesito precalentar el horno para la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que debo precalentar el horno para la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que tengo que precalentar el horno para la terrina de foie-gras?
3. **¿A qué temperatura debe estar el molde de la terrina de foie-gras?**
- ¿A qué temperatura tiene que estar el molde de la terrina de foie-gras?
 - ¿A qué temperatura necesita estar el molde de la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que necesita estar el molde de la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que debe estar el molde de la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que tiene que estar el molde de la terrina de foie-gras?
4. **¿A qué temperatura debo dejar la terrina de foie-gras?**
- ¿A qué temperatura tengo que dejar la terrina de foie-gras?
 - ¿A qué temperatura necesito dejar la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que necesito dejar la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que debo dejar la terrina de foie-gras?
 - ¿Cuál es la temperatura a la que tengo que dejar la terrina de foie-gras?

Preguntas sobre los tiempos

1. **¿Cuánto tiempo debo dejar la terrina de foie-gras a temperatura ambiente?**
- ¿Cuánto tiempo tengo que dejar la terrina de foie-gras a temperatura ambiente?
 - ¿Cuánto tiempo necesito dejar la terrina de foie-gras a temperatura ambiente?
2. **¿Cuánto tiempo debe estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?**
- ¿Cuánto tiempo tiene que estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?
 - ¿Cuánto tiempo necesita estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?
 - ¿Cuál es el tiempo necesario que debe estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?
 - ¿Cuál es el tiempo necesario que tiene que estar el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas en el microondas?
3. **¿Cuánto tiempo debo cocer el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?**
- ¿Cuánto tiempo tengo que cocer el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿Cuánto tiempo necesito cocer el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
 - ¿Cuál es el tiempo necesario para cocer el chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?

Preguntas sobre el procedimiento

1. **¿Cuál es el procedimiento para cocinar** *< nombre_del_plato >?*
 - ¿Qué procedimiento debo seguir para cocinar *< nombre_del_plato >?*
 - ¿Qué procedimiento tengo que seguir para cocinar *< nombre_del_plato >?*
 - ¿Qué procedimiento necesito seguir para cocinar *< nombre_del_plato >?*
 - ¿Cómo cocino *< nombre_del_plato >?*
 - ¿Cómo se cocina *< nombre_del_plato >?*
 - ¿Cómo se hace *< nombre_del_plato >?*
 - ¿Cuál es el procedimiento que debo seguir para cocinar *< nombre_del_plato >?*
 - ¿Cuál es el procedimiento que tengo que seguir para cocinar *< nombre_del_plato >?*
 - ¿Cuál es el procedimiento que necesito seguir para cocinar *< nombre_del_plato >?*

Preguntas sobre la presentación

1. **¿Cómo se presenta** *< nombre_del_plato >?*
 - ¿Cómo debería presentarse *< nombre_del_plato >?*
 - ¿Cómo puedo presentar *< nombre_del_plato >?*
 - ¿De qué forma hay que presentar *< nombre_del_plato >?*
 - ¿Cómo se emplata *< nombre_del_plato >?*

Preguntas sobre los autores

1. **¿Quién es el autor de la receta** *< nombre_del_plato >?*
 - ¿Quién escribió la receta *< nombre_del_plato >?*
 - ¿Quién publicó la receta *< nombre_del_plato >?*
 - ¿Cuál es el autor de la receta *< nombre_del_plato >?*
2. **¿Qué autores hay?**
 - ¿Qué autores hay disponibles en la base de datos?
 - ¿Qué autores hay disponibles en el sistema?
 - ¿De qué autores hay disponibles recetas?

Preguntas sobre las fechas

1. **¿Cuándo se publicó la receta** *< nombre_del_plato >?*
 - ¿En que fecha se publicó la receta *< nombre_del_plato >?*
2. **¿Cuándo se publicaron las recetas de** *< nombre_del_autor >?*
 - ¿En que fecha se publicaron las recetas de *< nombre_del_autor >?*

2.5. Diseño de reglas

A partir de las preguntas tipo y sus parafraseos propuestos previamente, se diseñará un **conjunto de reglas** que usará el sistema pregunta-respuesta para identificar las preguntas del usuario. Concretamente, servirá para **interpretar qué información quiere el usuario** y, si es oportuno, **sobre qué** (que variable) quiere la información.

Hay que comentar que, debido a la gran cantidad de parafraseos existentes, las reglas se verán limitadas principalmente a las preguntas tipo, considerando algunos parafraseos cuando se vea oportuno.

Las reglas serán presentadas en forma de **implicación lógica**, donde si todos los elementos del implicante (la parte izquierda de la regla) son ciertos, el resultado será el consecuente (la parte derecha de la regla). En concreto, se usarán los siguientes conceptos:

- Un **elemento** puede ser una palabra o una variable (un ingrediente, un nombre de receta...)
- Un elemento es **cierto** si el elemento aparece en la pregunta del usuario (es decir, las reglas indican que elementos deben aparecer y se buscan en la pregunta para identificarlas).
- El orden de los elementos en la regla no es relevante.

Si bien esto puede llevar a que se interpreten preguntas formuladas de forma errónea, es preferible esto a cambio de la mayor flexibilidad para interpretar correctamente preguntas parafraseadas.

- \wedge indica una **conjunción lógica** (los elementos a ambos lados de \wedge deben ser ciertos para que \wedge sea cierto).
- \vee indica una **disjunción lógica** (si uno de los elementos a ambos lados de \vee es cierto, \vee también es cierto).

Reglas sobre las recetas

1. ¿Qué recetas hay?

Se puede ver que, en todos los parafraseos, siempre está en común las palabras **qué recetas**. El resto de la frase realmente no aporta información adicional al sistema, por lo que no es necesaria para su reconocimiento. Por tanto, la regla resultante es:

$$que \wedge recetas \implies [lista \text{ de recetas en el sistema}]$$

Como se observará posteriormente, esta regla puede ser **ambigua** (al haber reglas parecidas). Por tanto, el orden en el que se analicen las reglas por parte del sistema será importante, teniendo que seguir un orden de más específico a más genérico. Esto se discutirá posteriormente en la implementación.

2. ¿Qué recetas publicó $\langle nombre_del_autor \rangle$?

De forma muy parecida a la regla anterior, todas las preguntas incluyen la palabra **recetas** como un elemento clave. Otra palabra importante a incluir es el **qué/cuáles**, al ser importante para distinguir a la pregunta.

Finalmente, es indispensable en este caso indicar **el nombre de un autor en la pregunta**, siendo esta la principal diferencia respecto a la regla anterior. La regla resultante es:

$$(que \vee cuales) \wedge recetas \wedge \langle nombre_autor \rangle \\ \implies [lista\ de\ recetas\ de\ \langle nombre_autor \rangle\ en\ el\ sistema]$$

Reglas sobre las partes de las recetas

1. ¿Qué partes tiene la receta $\langle nombre_de_la_receta \rangle$?

La palabra más importante en esta pregunta y en todos sus parafraseos es **partes**, indicando de forma muy clara y evidente que se está preguntando por las partes de una receta. También es importante el pronombre (siendo los más usados **qué** y **cuales**) y, por supuesto, la **receta** sobre la que se busca la información. Por tanto, la regla resultante es:

$$(que \vee cuales) \wedge partes \wedge \langle nombre_receta \rangle \\ \implies [lista\ de\ recetas\ de\ \langle nombre_autor \rangle\ en\ el\ sistema]$$

Reglas sobre los ingredientes

1. ¿Cuáles son los ingredientes de $\langle nombre_del_plato \rangle$?

Si bien hay parafraseos que no la incluyen, la palabra **ingredientes** es la que mejor representa esta pregunta para el sistema. Junto al **plato** (receta o parte de receta) de la que se busca la información y un pronombre para distinguir (**qué** o **cuales**), el sistema tiene suficiente información para reconocer la pregunta. La regla resultante es:

$$(que \vee cuales) \wedge ingredientes \wedge \langle nombre_plato \rangle \\ \implies [lista\ de\ ingredientes\ contenidos\ en\ \langle nombre_plato \rangle]$$

2. ¿Qué platos llevan $\langle nombre_del_ingrediente \rangle$?

La palabra **platos** o **recetas**, junto a un **ingrediente** da suficiente información al sistema para indicarle que el usuario busca los platos que llevan dicho ingrediente. Los pronombres (**qué** y **cuales** son parte de todas las preguntas planteadas, y por tanto también se incluyen. La regla resultante es:

$$(que \vee cuales) \wedge (platos \vee recetas) \wedge \langle nombre_ingrediente \rangle \\ \implies [lista\ de\ platos\ que\ contienen\ \langle nombre_ingrediente \rangle]$$

3. **¿Qué platos puedo preparar con** [$\langle \text{nombre_del_ingrediente} \rangle, \langle \text{nombre_del_ingrediente} \rangle, \dots \rangle$]?

Esta pregunta es, en esencia muy parecida a la anterior. La principal diferencia es que, en vez de un único ingrediente, se pasa **una lista de dos o más ingredientes**. Además, las palabras **preparar** o **cocinar** definen la diferencia (la pregunta no es sobre platos que contienen el ingrediente, sino sobre platos que se podrían cocinar enteros con los ingredientes propuestos) por lo que se considerarán también. La regla resultante es:

$$(que \vee cuales) \wedge (platos \vee recetas) \wedge (cocinar \vee preparar) \wedge \langle \text{lista_ingredientes} \rangle (2 \text{ o mas}) \\ \implies [\text{lista de platos que contienen todos los ingredientes en } \langle \text{lista_ingredientes} \rangle]$$

Reglas sobre las cantidades

1. **¿Qué cantidad de** $\langle \text{nombre_del_ingrediente} \rangle$ **se necesita para** $\langle \text{nombre_del_plato} \rangle$?

En este caso, como era de esperar, la palabra **cantidad** junto a un **ingrediente** y una **receta** (y su pronombre asociado, en este caso **qué** o **cuál**) son los elementos esenciales para que el sistema sea capaz de reconocer la pregunta adecuadamente con toda la información necesaria para procesarlo.

Ahora bien, el parafraseo '*¿Cuánto* $\langle \text{nombre_del_ingrediente} \rangle$ *necesita* el plato $\langle \text{nombre_del_plato} \rangle$?' es una forma muy frecuente de realizar esta pregunta y no queda recogido en la regla anterior. Por tanto, sería interesante plantear una regla adicional para este caso.

Por tanto, las dos reglas propuestas son:

Regla 1:

$$(que \vee cual) \wedge cantidad \wedge \langle \text{nombre_ingrediente} \rangle \wedge \langle \text{nombre_plato} \rangle \\ \implies \text{cantidad de } \langle \text{nombre_ingrediente} \rangle \text{ en } \langle \text{nombre_plato} \rangle$$

Regla 2:

$$(cuanto \vee cuantos \vee cuanta \vee cuantas) \wedge \langle \text{nombre_ingrediente} \rangle \wedge \langle \text{nombre_plato} \rangle \\ \implies \text{cantidad de } \langle \text{nombre_ingrediente} \rangle \text{ en } \langle \text{nombre_plato} \rangle$$

Reglas sobre las temperaturas

Si bien existen cuatro preguntas tipo distintas sobre temperaturas, observándolas se puede ver que todas siguen un patrón claro:

- La construcción '*A qué temperatura*'.
- Un **verbo** (calentar, precalentar, estar, dejar...)
- Opcionalmente, un **utensilio** (horno, molde...)
- El **plato** (receta o parte de receta) al que se aplica la temperatura.

Por tanto, la regla general que cubre todas las preguntas tipo es la siguiente:

$$a \wedge que \wedge temperatura \wedge (calentar \vee precalentar \vee estar \vee dejar) \wedge \langle nombre_plato \rangle$$

$$\implies \text{temperatura asociada para la acción en } \langle nombre_plato \rangle$$

En la práctica, la gran mayoría de estas reglas son muy específicas de la receta y difíciles de interpretar por el sistema, por lo que estas reglas sufrirán una gran cantidad de modificaciones conforme se añadan recetas.

Reglas sobre los tiempos

De forma parecida a las temperaturas, las preguntas sobre tiempos siguen un patrón constante pese a su especificidad:

- La construcción '*Cuánto tiempo*'.
- Un **verbo** (dejar, estar, cocer...)
- El **plato** (receta o parte de receta) al que se aplica la temperatura.

Por tanto, la regla general que cubre todas las preguntas tipo es la siguiente:

$$cuanto \wedge tiempo \wedge (dejar \vee estar \vee cocer) \wedge \langle nombre_plato \rangle$$

$$\implies \text{tiempo asociada a la acción en } \langle nombre_plato \rangle$$

Igual que en el caso anterior, en la práctica la mayoría de estas preguntas tipo son muy específicas para cada receta y complicadas de interpretar por el sistema, por lo que será necesario modificarlas continuamente con la adición de nuevas recetas.

Reglas sobre el procedimiento

1. ¿Cuál es el procedimiento para cocinar $\langle nombre_del_plato \rangle$?

La palabra **procedimiento** junto al **nombre del plato** indica al sistema claramente qué información busca el usuario (cómo cocinar un plato concreto). Además, todas las preguntas planteadas incluyen un pronombre (**cuál** o **qué**) junto al verbo **cocinar**, por lo que también se incluyen en la regla (para discriminar mejor).

Ahora bien, existe otro parafraseo típico para esta pregunta ('¿Cómo cocino $\langle nombre_del_plato \rangle$?') que también debe tenerse en cuenta, por lo que se representará mediante una regla adicional.

Las reglas planteadas son las siguientes: **Regla 1:**

$$(cual \vee que) \wedge procedimiento \wedge cocinar \wedge \langle nombre_plato \rangle$$

$$\implies [\text{lista de pasos para cocinar } \langle nombre_plato \rangle]$$

Regla 2:

$$como \wedge (cocino \vee cocina \vee hace) \wedge \langle nombre_plato \rangle$$

$$\implies [\text{lista de pasos para cocinar } \langle nombre_plato \rangle]$$

Reglas sobre la presentación

1. ¿Cómo se presenta $\langle \text{nombre_del_plato} \rangle$?

La definición de una regla es más compleja en este caso, al no haber demasiada similitud entre los parafraseos propuestos. El elemento más común es el pronombre **cómo** junto a una variante del verbo **presentar** y un **plato**, por lo que se utilizarán estos elementos para discriminar la pregunta. La regla planteada es la siguiente:

$$\begin{aligned} & \text{como} \wedge (\text{presenta} \vee \text{presentar} \vee \text{emplata} \vee \text{emplatar}) \wedge \langle \text{nombre_plato} \rangle \\ \implies & \text{indicaciones para la presentación de } \langle \text{nombre_plato} \rangle \end{aligned}$$

Reglas sobre los autores

1. ¿Quién es el autor de la receta $\langle \text{nombre_del_plato} \rangle$?

Si bien hay muchas formas de expresar esta pregunta, hay un elemento común a todas ellas: el pronombre **quién**. Al ser la única pregunta propuesta con ese pronombre, esa información junto al **plato** es suficiente para identificar la petición del usuario. Ahora bien, se puede añadir información extra (como la palabra **autor**, **escribió**, **publicó**...) para discriminar mejor.

La regla diseñada es la siguiente:

$$\begin{aligned} & \text{quien} \wedge (\text{autor} \vee \text{escribió} \vee \text{publicó} \vee \text{escrito} \vee \text{publicado}) \wedge \langle \text{nombre_plato} \rangle \\ \implies & \text{autor de } \langle \text{nombre_plato} \rangle \end{aligned}$$

2. ¿Qué autores hay?

Todos los parafraseos de la pregunta incluyen, de una forma u otra, las palabras **qué autores**. Por tanto, se usarán para distinguir de forma adecuada la pregunta. La regla propuesta es la siguiente:

$$\begin{aligned} & \text{que} \wedge \text{autores} \\ \implies & \text{autores en el sistema} \end{aligned}$$

Reglas sobre las fechas

1. ¿Cuándo se publicó la receta $\langle \text{nombre_del_plato} \rangle$?

Para expresar este tipo de pregunta, se han identificado dos expresiones características: el pronombre **cuando** o la construcción **en que fecha**. Por tanto, cualquiera de esas expresiones junto al nombre de la receta servirá para identificar adecuadamente la pregunta. La regla propuesta es la siguiente:

$$\begin{aligned} & (\text{cuando} \vee (\text{en} \wedge \text{que} \wedge \text{fecha})) \wedge \langle \text{nombre_plato} \rangle \\ \implies & \text{fecha en la que se publicó la receta de } \langle \text{nombre_plato} \rangle \end{aligned}$$

2. **¿Cuándo se publicaron las recetas de $\langle nombre_del_autor \rangle$?**

Esta pregunta sigue exactamente la misma estructura que la pregunta anterior, con la única diferencia de que se pregunta por un autor en vez de por un plato. Por tanto, la regla será idéntica teniendo en cuenta únicamente esa diferencia. La regla propuesta es la siguiente:

$$(cuando \vee (en \wedge que \wedge fecha)) \wedge \langle nombre_autor \rangle$$

\Rightarrow [lista con las recetas publicadas por $\langle nombre_autor \rangle$ junto a sus fechas de publicación]

Capítulo 3

Implementación y prueba del sistema

En este capítulo se tratarán todos los temas relacionados con la **implementación** del sistema de pregunta-respuesta, junto a algunas **pruebas** del funcionamiento adecuado del sistema.

Concretamente, se empezará detallando el **formato** en el que se almacenan las recetas para el uso por parte del sistema. Tras esto, se explicará el **funcionamiento interno** del sistema, se comentará su **implementación** y finalmente se realizarán algunas **pruebas** del funcionamiento adecuado del sistema.

3.1. Formato de las recetas

Para el funcionamiento del sistema, es necesario que la información extraída en el capítulo anterior (la tabla de dimensiones, las reglas definidas...) se encuentre almacenada en el sistema de alguna forma. Si bien es posible incorporar directamente la información durante la implementación, se ha optado por almacenar las recetas en formato *JSON* y cargarlas al arranque del sistema por los siguientes motivos:

- **Simple, rápido y fiable:** Frente a un parseo de los textos para extraer la información necesaria (un proceso muy lento, complejo de implementar y con posibles errores), utilizar archivos *JSON* permite a los diseñadores estructurar la información de forma fácil y precisa y al sistema procesarla de forma rápida.
- **Estandarizable:** Es posible definir un estándar para las recetas, de forma que cualquier persona (no necesariamente los diseñadores) pueda aportar recetas al sistema.
- **Modular:** En vez de añadir directamente la información al programa, optar por almacenar todas las recetas en una carpeta en un formato estandarizado permite, en el futuro, añadir fácilmente recetas adicionales al sistema.

Un ejemplo simple del formato usado para almacenar las recetas en *JSON* es el siguiente:

```

{
  "nombre": "receta de ejemplo",
  "autor": "luna y alejandro",
  "fecha": "06/01/2021",
  "partes":
  [
    "parte 1",
    "parte 2"
  ],
  "ingredientes": {
    "general": [
      {
        "nombre": "ingrediente 1",
        "cantidad": "5 gramos"
      }
    ],
    "parte 1": [
      {
        "nombre": "ingrediente de la parte 1",
        "cantidad": "800 gramos"
      }
    ],
    "parte 2": [
      {
        "nombre": "ingrediente de la parte 2",
        "cantidad": "400 gramos"
      }
    ]
  },
  "procedimiento": [
    {
      "parte": "general",
      "pasos":
      [
        "Paso 1",
        "Paso 2"
      ]
    },
    {
      "parte": "parte 2",
      "pasos":
      [
        "Paso 1 de la parte 2",
        "Paso 2 de la parte 2"
      ]
    }
  ],

```

```

{
  "parte": "parte 1",
  "pasos":
  [
    "Paso 1 de la parte 1",
    "Paso 2 de la parte 1"
  ]
},
"presentacion": "Indicaciones para la presentación"
}

```

En general, todas las secciones se corresponden directamente con la dimensión que representan, con algunos detalles a tener en cuenta:

- Todos los nombres (de la receta, de las partes, de los autores...) deben estar en **minúscula** y **no contener acentos**.
- Los ingredientes y los procedimientos están asociados siempre a una parte. Para los ingredientes y procedimientos que no pertenezcan a una parte directamente, su parte se indica como **general**.
- La lista de **procedimiento** indica también el orden en el que se deben realizar (el primer bloque en la lista se realiza primero, luego el segundo y así sucesivamente).
- Se ha optado por no incluir información sobre **temperaturas** o **tiempos** para simplificar el trabajo y por considerarse que necesitaría preguntas demasiado específicas.

Las recetas usadas por el sistema se encuentran disponibles en la carpeta *Recetas*. Se recomienda su lectura para comprender mejor el formato.

3.2. Funcionamiento del sistema

En esta sección se describirá el **funcionamiento interno** del sistema de pregunta-respuesta. Este sistema funcionará en bucle continuamente hasta que se cierre o se de una señal de salida (que será posteriormente indicada en el manual de usuario).

Concretamente, el proceso que sigue el sistema es el siguiente:

- Carga y procesamiento de los documentos.
- Mientras el usuario realice preguntas al sistema:
 - Pre-procesamiento de la pregunta del usuario.
 - Interpretación y respuesta a la pregunta del usuario.

Estos puntos serán descritos a continuación.

3.2.1. Carga y procesamiento de los documentos

Para poder trabajar con las recetas planteadas en formato *JSON*, es necesario **cargarlas** a la memoria del sistema. Ahora bien, el formato en el que se almacenan las recetas puede resultar complejo a la hora de trabajar con él, por lo que además se **procesará** la información de dichos documentos para facilitar el trabajo al sistema.

Los documentos se cargarán en el sistema uno a uno, incorporando su información a una estructura de **diccionario** con la siguiente forma:

```
[<nombre de la receta>] =>
  ["author"] => <autor de la receta>
  ["date"] => <fecha de la receta>
  ["parts"] => [<lista con las partes de la receta>]
  ["ingredients"] => <conjunto con TODOS los ingredientes de todas las partes>
  ["part_ingredients"] =>
    [<nombre de parte>] => <conjunto con los ingredientes que pertenecen a esa
                           parte>
    ...
  ["quantities"] =>
    [<nombre de parte>] =>
      [<ingrediente>] => <cantidad del ingrediente>
      ...
  ["procedure_order"] => <lista con el orden de las partes en el procedimiento>
  ["procedure"] =>
    [<procedimiento de la parte 1>, <procedimiento de la parte 2>...]
  ["presentation"] => <indicaciones de presentación>
```

Donde:

- [*clave*] indica una **clave** en el diccionario.
- < *variable* > indica una **variable** en el diccionario.

Además del diccionario, se crean **cuatro conjuntos** (siendo un conjunto una lista sin repeticiones) de valores identificados en las recetas leídas, siendo estos:

- Autores.
- Nombres de recetas.
- Partes de recetas.
- Ingredientes.

Estos conjuntos representan todos los valores de esas cuatro dimensiones capaces de ser reconocidos por el sistema, y serán usados posteriormente durante el preprocesamiento de la pregunta del usuario para identificar las posibles variables que haya en la pregunta.

3.2.2. Pre-procesamiento de la pregunta

Antes de intentar interpretar la pregunta para darle una respuesta, es necesario realizar un breve **preprocesamiento** de cara a reducir la variabilidad. En este caso el preprocesamiento es simple, teniendo dos pasos:

- Se eliminan las **mayúsculas** y los **acentos**. Este cambio se hace para reducir la variabilidad, al considerar que tanto las letras mayúsculas como los acentos introducen únicamente complejidad innecesaria sin aumentar la capacidad del sistema.

Si bien se podría pensar que los símbolos no alfanuméricos (como las interrogaciones, los guiones...) también son superfluos y se podrían eliminar, hay algunos casos en los que pueden ser importantes (como el guión en foie-gras) y por tanto se ha optado por conservarlos.

- Se **extraen las variables** de la pregunta del usuario. Como se ha comentado antes, se extraen variables de las siguientes cuatro dimensiones:
 - Autores.
 - Nombres de recetas.
 - Partes de recetas.
 - Ingredientes.

Esta extracción se hace comprobando si los elementos de los conjuntos descritos en la sección anterior aparecen en la pregunta del usuario.

Si bien este proceso puede parecer simple, hay que tener en cuenta un detalle importante: **los nombres de recetas pueden contener partes e ingredientes en su interior**. Por ejemplo, si se hace una pregunta sobre la *terrina de foie-gras con sobaos pasiegos y confitura de zanahorias* el sistema podría identificar erróneamente la parte de *terrina de foie-gras* y los ingredientes *sobaos pasiegos* y *zanahorias*. Por tanto, el proceso que se sigue es el siguiente:

1. Se identifican los **autores** y las **recetas** contenidas en las preguntas.
2. Se identifican las **partes de recetas** contenidas en la pregunta, y se descartan aquellas que están contenidas en alguna receta identificada y que solo aparecen una vez.

Si solo aparecen una vez, se entiende que es la misma vez que aparece en el título de la receta. Ahora bien, si aparecen dos o más veces se entiende que aparece por separado, y por tanto hay que identificarla.

3. Se identifican los **ingredientes** contenidos en la pregunta, y se descartan aquellos que están contenidos en alguna receta o parte identificada y que solo aparecen una vez.

De nuevo, si aparecen una única vez se entiende que forman parte de la receta o de la parte. Ahora bien, si aparecen dos veces (por ejemplo, "*¿Cuántas zanahorias se necesitan para cocinar la confitura de zanahorias?*") se entiende que es necesario identificar al ingrediente como una variable.

3.2.3. Interpretación y respuesta a la pregunta

Una vez preprocesada la pregunta del usuario, se **interpretará y se ofrecerá una respuesta**. Este proceso es simple, constando de dos pasos:

1. Se **identifica la pregunta** como una de las preguntas tipo identificadas, usando las reglas planteadas. Esto se hace comprobando si las palabras indicadas en la regla se encuentran en la pregunta del usuario.
2. Una vez identificada la pregunta, se **responde** buscando la información relevante en el diccionario.

Sobre este proceso se puede realizar una serie de comentarios:

- El sistema **responde a la primera pregunta que identifica**. En caso de que una pregunta del usuario encajase en varias reglas, se tomará la primera de ellas.
- Las reglas son procesadas (probadas) en **el orden en el que se han propuesto** en la memoria con una excepción: la primera regla ("*¿Qué recetas hay?*") es la última regla en ser procesada.

Esto se debe a que es una regla más genérica que otras reglas planteadas, por lo que acabaría englobándolas. Por tanto, se ha intentado seguir un orden de reglas más específicas a menos específicas.

- Las reglas sobre **temperaturas y tiempos** están implementadas **de forma directa**.

Al no contener la información el fichero *JSON*, se ha introducido esta información de forma directa en el sistema para tener la posibilidad de responderlas.

- En caso de identificar varias variables del mismo tipo, **se toma siempre la primera variable**.

Por ejemplo, si el sistema reconoce una pregunta del tipo "*¿Qué recetas publicó autor?*" y reconoce varios autores, considerará únicamente el primer autor.

- El sistema es incapaz de reconocer variables que no estén exactamente en su base de conocimiento.

Por tanto, el sistema **no será capaz de reconocer una pregunta** que, por ejemplo, necesite una variable ingrediente, si no se ha proporcionado un ingrediente conocido el sistema. El sistema también será incapaz de reconocer variables mal escritas, por lo que el usuario debe utilizar exactamente las formas descritas en la tabla de dimensiones.

- En caso de que el sistema no sea capaz de reconocer ninguna pregunta, **se mostrará un mensaje de error** y se permitirá al usuario introducir una nueva pregunta.

3.3. Implementación y uso del sistema

En esta sección se describirán los detalles sobre la **implementación** del sistema (centrándose en las dependencias necesarias para su correcto funcionamiento) y se ofrecerá el **manual de usuario del sistema**, para su correcto uso.

3.3.1. Implementación y dependencias

El sistema de pregunta-respuesta implementado se encuentra disponible en el fichero *spr.py*. Este programa ha sido implementado en **Python 3.7.6**, y utiliza las siguientes librerías externas:

- *unidecode* v1.1.2.

El programa ha sido probado únicamente con las versiones descritas. Si bien es posible que funcione con otras versiones de Python o de las librerías, no se puede garantizar su compatibilidad.

La implementación como tal se encuentra documentada en detalle en el fichero indicado previamente, por lo que se recomienda encarecidamente su lectura para observar los detalles técnicos.

3.3.2. Manual de usuario del sistema

El programa se utiliza mediante el siguiente comando:

```
python spr.py
```

Al ejecutar el comando se lanzará el sistema de pregunta-respuesta y se cargarán las recetas en formato JSON que se encuentran en la carpeta *Recetas*. La ruta a la carpeta conteniendo las recetas se encuentra en la sección *"Variables Globales"* del código fuente, y puede ser modificada si se almacenan en otro sitio las recetas.

Una vez las recetas están cargadas, el sistema pedirá una pregunta al usuario. El usuario plantea la pregunta por teclado vía la consola de comandos, y el sistema ofrecerá la respuesta bajo la pregunta. Este proceso se repetirá en bucle hasta que el usuario indique el fin de la sesión escribiendo **STOP** en vez de una pregunta.

3.4. Pruebas de funcionamiento del sistema

Para comprobar el funcionamiento adecuado del sistema, se realizará una **serie de pruebas** con el fin de probar que el sistema cumple con las expectativas. Estas pruebas se documentarán a través de **capturas de pantalla** del sistema en ejecución.

Inicialización del programa

Una muestra de la **inicialización del sistema** se puede ver en la Figura 3.1. El sistema muestra un título y el tiempo necesario para cargar los ficheros en memoria. Tras esto, se muestra un *prompt* al usuario pidiéndole una pregunta.

```

SISTEMA PREGUNTA-RESPUESTA - RECETARIO

Cargando la informacion en el sistema
Informacion cargada (0.03953s)

Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.1: Prueba del sistema - Inicialización

Proceso de pregunta-respuesta

Una muestra de cómo el sistema **responde a una pregunta del usuario** se puede ver en la Figura 3.2. Como se puede observar, el sistema es capaz de interpretar correctamente la pregunta y la receta sobre la que se busca la información, y se imprime la respuesta adecuada. Tras la respuesta, se vuelve a mostrar el *prompt*.

```

Introduce tu pregunta (o "STOP" para salir del programa): ¿Cuál es el procedimiento para cocinar la terrina de foie-gras
con sobaos pasiegos y confitura de zanahorias?

Mostrando los pasos para cocinar terrina de foie-gras con sobaos pasiegos y confitura de zanahorias

TERRINA DE FOIE-GRAS
1- Colocar el hígado en una bandeja de acero y aliñar con la sal, la pimienta negra y el brandy. Dejarlos 4 horas a temp
eratura ambiente tapados con film transparente.
2- Colocar los hígados en un molde de papel de aluminio, procurando que las caras exteriores queden hacia el molde.
3- Precalentar el horno a 200 grados y meter el molde dentro, con un termómetro pinchado en el hígado.
4- Apagar el horno inmediatamente y dejar los foie-gras dentro hasta que el termómetro nos indique que la temperatura en
el corazón del molde es de 52 grados.
5- Tapar y enfriar rápidamente. Conservar en la nevera.

CONFITURA DE ZANAHORIAS
1- Poner las zanahorias y las naranjas peladas con el azúcar en un cazo con agua fría.
2- Llevar a ebullición y mantener a fuego suave hasta que las zanahorias estén bien tiernas.
3- Escurrir y triturar, añadiendo poco a poco el líquido de la cocción hasta que la confitura esté fina y brillante.
4- Poner a punto de sal y enfriar rápido.

Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.2: Prueba del sistema - Pregunta y respuesta (1)

Se puede observar además que, antes de la respuesta, se indica qué se va a responder (en este caso, *pasos para cocinar terrina de foie-gras con sobaos pasiegos y confitura de zanahorias*). Esto se ha añadido para **facilitar el trabajo al usuario**: si, por cualquier razón, el sistema respondiese algo inesperado a su pregunta, así será capaz de interpretar donde ha fallado el sistema y refinar la búsqueda.

Se pueden ver otros ejemplos de preguntas y respuestas correctas en la Figura 3.3.

```

Introduce tu pregunta (o "STOP" para salir del programa): ¿Qué autores hay disponibles en el sistema?
Mostrando los autores contenidos en el sistema
* alberto chicote
Introduce tu pregunta (o "STOP" para salir del programa): ¿Cuántas zanahorias necesita la confitura de zanahorias?
Mostrando la cantidad de zanahorias necesaria en confitura de zanahorias
250 gramos
Introduce tu pregunta (o "STOP" para salir del programa): ¿Quién es el autor del chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas?
Mostrando el autor de chocolate, aceite de oliva virgen extra, cacao ácido y flores escarchadas
alberto chicote
Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.3: Prueba del sistema - Preguntas y respuestas (2)

Interpretación de variabilidad en las preguntas

El sistema tiene cierta capacidad de **afrentar la variabilidad** de las preguntas del usuario por cómo se han estructurado las reglas. Esto se puede observar en la Figura 3.4, donde se pregunta de formas diversas la **cantidad de cacao en polvo** necesaria en parte de una receta. Como se puede ver, las preguntas tienen estructuras distintas, pero el sistema ha sido capaz de identificarlas y responderlas adecuadamente en todo momento.

```

Introduce tu pregunta (o "STOP" para salir del programa): ¿Cuanto cacao en polvo lleva la salsa acida de cacao?
Mostrando la cantidad de cacao en polvo necesaria en salsa acida de cacao
50 gramos
Introduce tu pregunta (o "STOP" para salir del programa): ¿Qué cantidad de cacao en polvo lleva la salsa acida de cacao?
Mostrando la cantidad de cacao en polvo necesaria en salsa acida de cacao
50 gramos
Introduce tu pregunta (o "STOP" para salir del programa): ¿Con cuanto cacao en polvo puedo hacer la salsa acida de cacao?
Mostrando la cantidad de cacao en polvo necesaria en salsa acida de cacao
50 gramos
Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.4: Prueba del sistema - Interpretación de variabilidad (1)

Se puede observar otro ejemplo de esta capacidad en la Figura 3.5.

```

Introduce tu pregunta (o "STOP" para salir del programa): ¿Quién es el autor de la terrina de foie-gras con sobaos pasiegos y confitura de zanahorias?

Mostrando el autor de terrina de foie-gras con sobaos pasiegos y confitura de zanahorias

alberto chicote

Introduce tu pregunta (o "STOP" para salir del programa): ¿Quién publicó la receta de la terrina de foie-gras con sobaos pasiegos y confitura de zanahorias?

Mostrando el autor de terrina de foie-gras con sobaos pasiegos y confitura de zanahorias

alberto chicote

Introduce tu pregunta (o "STOP" para salir del programa): ¿Quién ha escrito la guía para la terrina de foie-gras con sobaos pasiegos y confitura de zanahorias?

Mostrando el autor de terrina de foie-gras con sobaos pasiegos y confitura de zanahorias

alberto chicote

Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.5: Prueba del sistema - Interpretación de variabilidad (2)

Pregunta no reconocida

Hay ocasiones en los que el sistema **no es capaz de reconocer una pregunta**. En estos casos, el sistema mostrará una respuesta como la que se puede ver en la Figura 3.6. Como se puede ver, la respuesta es una ayuda para el usuario, recordándole el error más frecuente.

```

Introduce tu pregunta (o "STOP" para salir del programa): ¿Es capaz de entender esta pregunta?

No se ha podido reconocer la pregunta
Asegurate de que los nombres de autores, recetas, partes o ingredientes esten escritos adecuadamente

Introduce tu pregunta (o "STOP" para salir del programa):

```

Figura 3.6: Prueba del sistema - Pregunta no reconocida

Fin de la ejecución

El usuario es capaz de **finalizar la ejecución** introduciendo la palabra **STOP** (respetando las mayúsculas) en vez de una pregunta. Tras esto, el programa simplemente se cerrará y volverá al nivel superior de la consola de comandos. Un ejemplo de esto se puede ver en la Figura 3.7.

```

Introduce tu pregunta (o "STOP" para salir del programa): STOP

D:\Users\Luna\Desktop\Universidad y documentos\Master\Asignaturas\Ingenieria Linguistica\Practicas\Practica 3\jimenez-muñoz>

```

Figura 3.7: Prueba del sistema - Fin de la ejecución

Capítulo 4

Conclusiones

En este trabajo se ha descrito el desarrollo de un **sistema de pregunta-respuesta** centrado en recetas de cocina, detallando todos los aspectos del **diseño** del sistema: la **selección** de documentos, la **extracción de la tabla de dimensiones**, el planteamiento de **preguntas tipo**, el **parafraseo** de estas preguntas y el **diseño de reglas** para cubrir estas preguntas.

Tras esto, se ha comentado la **implementación** del sistema, mencionando tanto el **formato** en el que se almacenan las recetas como el **funcionamiento y uso** del sistema. Finalmente, se han mostrado algunas **pruebas** del funcionamiento adecuado del sistema.

Algunas conclusiones que se pueden extraer de este trabajo son:

- El **diseño del sistema** realizado, en todos sus aspectos, es adecuado. Como se puede observar en las pruebas realizadas, el sistema funciona correctamente y es capaz de reconocer y responder satisfactoriamente a las preguntas de los usuarios.
- Las **preguntas tipo planteadas** han sido adecuadas, cubriendo la gran mayoría de información que un usuario podría preguntar de forma natural al sistema.
- El **diseño de reglas realizado** ha sido correcto, siendo capaz de identificar preguntas con cierta variabilidad debido a la forma de reconocimiento utilizada. Ahora bien, el sistema **está limitado a las reglas planteadas** y no es capaz de identificar preguntas que no entren en los patrones de dichas reglas.

Algunas posibles líneas futuras para este trabajo serían:

- Incorporar en la estructura de las recetas información sobre **temperaturas y tiempo**. Actualmente, la información se ha introducido manualmente al sistema, pero sería interesante estudiar el funcionamiento del sistema si se estandarizase dicha información en el formato de las recetas.
- Añadir **un número de recetas mayor**. Usar únicamente dos recetas implica sesgos y carencias de información que pueden aparecer en otras recetas. Un sistema diseñado con un mayor número de recetas (de fuentes variadas) seguramente incluiría un mayor número de dimensiones y preguntas tipo a considerar.

Apéndice A

Contenidos del fichero entregable

En el fichero comprimido entregado, se encuentran disponibles los siguientes elementos:

- Memoria del trabajo en formato **PDF**, con nombre *Ingenieria Linguistica - Practica 3.pdf*.
- Código fuente del **Sistema de Pregunta-Respuesta** en **Python** en el fichero *spr.py*.
- Todos las recetas seleccionadas, ya procesadas, en formato **JSON** en el directorio "*Recetas*".