
TFG: Análisis Emocional para la Inclusión Digital



**Gema Eugercios Suárez
Paloma Gutiérrez Merino
Elena Kaloyanova Popova**

**Departamento de Ingeniería del Software e Inteligencia Artificial
Facultad de Informática
Universidad Complutense de Madrid**

Mayo 2018

Documento maquetado con TEXIS v.1.0.

Este documento está preparado para ser imprimido a doble cara.

TFG: Análisis Emocional para la Inclusión Digital

Informe técnico del departamento
Ingeniería del Software e Inteligencia Artificial
IT/2009/3

Versión 1.0

**Departamento de Ingeniería del Software e Inteligencia
Artificial**
Facultad de Informática
Universidad Complutense de Madrid

Mayo 2018

Copyright © Marco Antonio y Pedro Pablo Gómez Martín

ISBN 978-84-692-7109-4

Índice

1. Introducción	1
1.1. Definición	1
1.2. Dificultades	1
1.3. Estructura	1
2. Estado del arte	3
2.1. Emociones	3
2.2. Computación Afectiva	4
2.2.1. Métodos para el marcado de texto emocional	6
2.3. Diccionarios Afectivos	7
2.4. Servicios Web	10
2.4.1. Características de los Servicios Web	10
2.4.2. Tipos de Servicios Web	11
2.4.3. Arquitectura de los Servicios Web	11
2.4.4. Ventajas e inconvenientes	12
2.5. Scrum	12
2.5.1. Roles	13
2.5.2. Artefactos	13
2.5.3. Eventos	13
2.6. Integración Continua	15
3. Herramientas	17
3.1. Herramientas básicas	17
3.2. Diccionario	18
3.3. Django	19
3.4. Trello	19
3.5. Doctest y Jenkins	20
3.6. SpaCy y PyStemmer	21
4. Análisis del Contenido Afectivo de un Texto	23
4.1. Base para los servicios web	23

4.2. Análisis afectivo de una palabra	25
4.2.1. Servicio web porcentajes de una palabra	25
4.2.2. Servicio web emoción consensuada de una palabra . . .	26
4.2.3. Servicio web emoción mayoritaria de una palabra . . .	26
4.3. Análisis afectivo de una frase	27
4.3.1. Porcentajes de una frase	27
4.3.2. Emoción mayoritaria de una frase	28
4.4. Análisis afectivo de un texto	28
5. Diseño de la aplicación web	29
5.1. Prototipos iniciales	29
5.1.1. Pantalla principal	29
5.1.2. Pantalla de análisis	29
5.2. Prototipo común	30
5.2.1. Pantalla principal	30
5.2.2. Pantalla de análisis	30
5.3. Interfaz final	30
6. Desarrollo de la aplicación web	35
7. API de la aplicación web	37
8. Resultados	39
8.1. Resultados de los casos de prueba	39
8.2. Resultados evaluación con usuarios finales	39
A. Así se hizo...	41
A.1. Sprint 1	41
A.2. Sprint 2	41
A.3. Sprint 3	41
A.4. Sprint 4	42
A.5. Sprint 5	43
A.6. Sprint 6	44
A.7. Sprint 7	44
B. Casos de prueba	49
B.1. Cuentos	49
B.2. Noticias	49
B.3. Entradas de Blog	49
Bibliografía	51

Índice de figuras

2.1. Flujo Scrum	14
3.1. Fragmento de la adaptación del diccionario ANEW traducido.	18
3.2. Sprint inicial que ilustra la estructura.	19
3.3. Final del sprint inicial.	20
4.1. Vista general de la lista de palabras.	24
4.2. Vista detalle de la palabra «Alegre».	25
4.3. Respuesta al buscar los porcentajes de la palabra «alegre» . .	25
4.4. Respuesta al buscar los porcentajes de una palabra que no está en el diccionario.	26
4.5. Respuesta al encontrar la emoción consensuada.	26
4.6. Respuesta si la palabra no tiene emoción consensuada.	27
4.7. Respuesta al encontrar la emoción mayoritaria.	27
4.8. Respuesta si hay dos emociones mayoritarias.	28
5.1. Pantalla inicial del prototipo de Elena	30
5.2. Pantalla inicial del prototipo de Gema	30
5.3. Pantalla inicial del prototipo de Paloma	31
5.4. Pantalla de análisis del prototipo de Elena	31
5.5. Pantalla de análisis del prototipo de Gema	31
5.6. Pantalla de análisis del prototipo de Paloma	32
5.7. Pantalla inicial por defecto del prototipo final	32
5.8. Menú de opciones de la interfaz	32
5.9. Pantalla inicial alternativa del prototipo final	33
5.10. Pantalla de análisis con los porcentajes abajo	33
5.11. Pantalla de análisis con los porcentajes derecha	33
5.12. Pantalla de análisis con gráfico	34
5.13. Pantalla de análisis mostrando las emociones que aparecen . .	34
5.14. Pantalla de análisis mostrando sólo la mayoritaria	34
A.1. Planificación del sprint 1 (14/11/2017)	42

A.2. Final del sprint 1 (28/11/2017)	42
A.3. Planificación del sprint 2 (28/11/2017)	43
A.4. Final del sprint 2 (19/12/2017)	43
A.5. Planificación del sprint 3 (19/12/2017)	44
A.6. Final del sprint 3 (09/01/2017)	44
A.7. Planificación del sprint 4 (09/01/2017)	45
A.8. Final del sprint 4 (21/02/2017)	45
A.9. Planificación del sprint 5 (21/02/2017)	45
A.10.Final del sprint 5 (07/03/2017)	46
A.11.Planificación del sprint 6 (07/03/2017)	46
A.12.Final del sprint 6 (21/03/2017)	46
A.13.Planificación del sprint 7 (21/03/2017)	47

Índice de Tablas

- | | |
|--|----|
| 3.1. Fragmento de la adaptación del diccionario ANEW traducido | 18 |
|--|----|

Capítulo 1

Introducción

RESUMEN: En este capítulo se va a realizar una introducción al trabajo que vamos a presentar. En primer lugar, en la sección 1.1 se explicará en qué consiste el trabajo. En la sección 1.2 se expondrán las principales dificultades a las que nos hemos enfrentado y las decisiones que se han tomado para afrontarlas. Por último, en la sección 1.3 se presentará la estructura del trabajo.

1.1. Definición

El trabajo se centra en la detección automática del contenido emocional de un texto. Esto nos va a ayudar a evitar ambigüedades emocionales a la hora de interpretar un texto facilitando su lectura a personas con algún déficit en la percepción de emociones, como aquellas con Trastornos del Espectro Autista. Se implementarán servicios web que permitan analizar emocionalmente un texto, identificando las emociones básicas que contiene y su intensidad, y hacer más explícito el significado emocional para hacer más fácil su comprensión a las personas con algún tipo de discapacidad emocional.

1.2. Dificultades

A lo largo de la realización del trabajo nos hemos ido encontrando una serie de problemas que se explican a continuación.

1.3. Estructura

El capítulo 2 presenta los aspectos más importantes de la computación afectiva e introduce a los servicios web, la metodología Scrum y la integración continua; metodología y tecnología que se va a utilizar. El capítulo 3 describe

las herramientas utilizadas a lo largo del trabajo. El capítulo 4 explica cada uno de los servicios web desarrollados durante el trabajo. El capítulo 5 se centra en el diseño y el desarrollo de la aplicación web que hará uso de los servicios web. El capítulo 6 contiene la API de la aplicación desarrollada. El capítulo 7 presenta los resultados obtenidos al finalizar el proyecto.

Capítulo 2

Estado del arte

RESUMEN: En este capítulo se van a tratar los aspectos más importantes tanto de la computación afectiva como de las diferentes tecnologías y metodologías que se van a utilizar. En primer lugar, en la sección 2.1, se define la computación afectiva y sus posibles aplicaciones. También se explican los distintos diccionarios afectivos ya existentes que permiten la marcación emocional de textos. En la sección 2.2 se introduce la tecnología que se va a utilizar para implementar el trabajo, los servicios web. En la sección 2.3 se fijan los conceptos relacionados con la metodología Scrum, la cual hemos seguido durante todo el trabajo. En la sección 2.4 se explican las bases de la integración continua aplicada al desarrollo de software y cómo se va a aplicar en este trabajo.

2.1. Emociones

Las emociones son reacciones afectivas que surgen súbitamente ante un estímulo, duran un corto espacio de tiempo y comprenden una serie de repercusiones psicocorporales (Francisco, 2008). Podríamos definir las emociones como reacciones automáticas que nuestro cuerpo experimenta ante un determinado estímulo. Sin embargo, todas ellas derivan en sentimientos más prolongados en el tiempo. Son experimentadas de manera particular por cada individuo debido a experiencias pasadas, su carácter, aprendizaje... No obstante, podemos afirmar que son estados afectivos y por ello indican estados internos personales, motivaciones, deseos, necesidades e incluso objetivos.

Existen muchas emociones, las cuales podemos representar gracias a gestos y movimientos; al igual que las emociones, podemos centrarnos en conocer la categorización de estas. Entre estas categorías se definen 5 categorías emocionales básicas (Cowie, 2000):

- **Miedo:** Angustia por un riesgo o daño real o imaginario (RAE, 2018).

- **Sorpresa:** Pillar desprevenido, conmover, suspender o maravillar con algo imprevisto, raro o incomprensible (RAE, 2018).
- **Enfado:** Impresión desgradable y molesta que hacen en el ánimo algunas cosas (RAE, 2018).
- **Alegría:** Sentimiento grato y vivo que suele manifestarse con signos exteriores (RAE, 2018).
- **Tristeza:** Doloroso, enojoso, difícil de soportar (RAE, 2018).

Las categorías básicas no son la única forma de clasificar las emociones. Otro método muy importante es la clasificación mediante dimensiones emocionales, aspectos esenciales de una emoción. Generalmente se utilizan tres para formar un espacio tridimensional que permitirá identificar la emoción según los valores que tome para cada dimensión. Las dimensiones que se tienen en cuenta son las siguientes (Lang et al., 1999):

- **Evaluación:** Representa como de positiva o negativa es la emoción.
- **Activación:** Representa en qué lugar de la escala activa (exitación) / pasiva (calma) se encuentra la emoción.
- **Control:** Representa el control que tiene la emoción sobre una persona.

2.2. Computación Afectiva

La computación afectiva (Picard, 1997) es el estudio y el desarrollo de sistemas y dispositivos capaces de percibir, medir e interpretar las emociones humanas.

Esta rama de la computación permite un avance notable en la inteligencia artificial, hasta tal punto que los ordenadores lleguen a adaptarse a los humanos, sus necesidades y estados de ánimo. Los seres humanos están rodeados de emociones, en cualquier ámbito de su vida, tanto de las suyas propias como las de las personas con las que se comunican. Tanta importancia tienen para nosotros que influyen no sólo en nuestra comunicación, sino también en nuestro aprendizaje y toma de decisiones. Por ello, resulta artificial y en ocasiones incluso frustrante intentar comunicarse con una máquina que no es capaz de expresar sentimientos. La computación afectiva pretende mejorar la interacción hombre-máquina haciéndola más natural y asequible en este sentido.

El funcionamiento de este tipo de sistemas se basa en identificar el estado emocional del sujeto a través de diferentes fuentes (voz, expresiones, señales fisiológicas, palabras...) y procesar la información para clasificarla y

aprender de ella. Clasificar la información de entrada puede resultar complicado ya que se suelen recibir varias señales diferentes a la vez, lo que hace necesario utilizar técnicas de priorización para determinar cuáles son las que más aportan a la hora de analizar y gestionar la información. Una vez identificada la emoción predominante, el sistema responde adecuándose a ella. La salida dependerá del tipo de sistema y las herramientas de las que dispone este para expresar su respuesta (colores, sonidos, emoticonos...) En casos más complejos aplicados a robótica o modelado, la salida producida es una simulación de la respuesta que produciría un ser humano ante los estímulos recibidos imitando su expresión corporal, voz o gestos faciales. Por ejemplo los robots Geminoid¹.

La computación afectiva tiene multitud de aplicaciones, ya que como se ha mencionado antes, las emociones están presentes en todos los ámbitos de la vida de una persona. Por lo tanto, basándonos en el artículo de Sandra Baldasarri 2016, puede aplicarse a áreas muy diferentes entre sí como:

- **Seguridad:** Poder analizar las emociones como el estrés, el aburrimiento o la distracción pueden ser muy interesantes de analizar para tareas repetitivas como conducir, trabajar en una fábrica o ser controlador.
- **Marketing:** Actualmente una de sus aplicaciones más explotadas y área que más se preocupa por su desarrollo. Poder evaluar la reacción emocional de alguien ante un anuncio o producto es una estrategia comercial que ya está siendo utilizada cada vez por más empresas.
- **Salud:** Actualmente, se emplea principalmente para la detección del estrés y así minimizar sus efectos y aprender a controlarlo. Es posible inferir el nivel de estrés de una persona midiendo sus señales fisiológicas (ritmo cardíaco, respiración...) y si este nivel es demasiado alto se reaccionará en consecuencia según el tipo de sistema. Se puede aplicar de forma similar a las fobias.
- **Entretenimiento:** La industria de los videojuegos ha crecido mucho en los últimos años e introducir este tipo de tecnología permite a las compañías crear juegos más adaptables y cercanos al jugador, lo que atrae a más público y mejora la experiencia de juego (Ng et al., 2012).
- **Robótica:** El mayor problema de los robots diseñados para interactuar con humanos es la carencia de emociones. Algunos de ellos llegan a producir una sensación de incomodidad. Dotar a este tipo de robots de cierta “humanidad” no sólo haría más cómodo el tratar con ellos sino que podrían realizar tareas como el acompañamiento de personas mayores (Riek et al., 2010).

¹<http://www.geminoid.jp/en/index.html>

- **Accesibilidad:** Existen numerosas aplicaciones terapéuticas para ayudar a personas con problemas emocionales. El paradigma actual obliga a cualquier usuario a adaptarse a las máquinas sin tener en cuenta las dificultades particulares que pueda tener. Es complicado crear una máquina capaz de adaptarse a todas las circunstancias. El uso de computación afectiva permite facilitar la inversión del paradigma para que las máquinas utilicen las emociones del usuario para adaptarse a él.

En este trabajo vamos a centrarnos en el último área, la accesibilidad. En particular, en facilitar a personas que padecen Trastornos del Espectro Autista (TEA) el entendimiento de las emociones de un texto, es decir, su carga emocional. Trataremos de analizar un texto de entrada para identificar las emociones predominantes y en qué medida se presentan, evitando las posibles ambigüedades que pueda haber. Las emociones con las que trataremos son las básicas: alegría, tristeza, miedo, sorpresa y enfado.

2.2.1. Métodos para el marcado de texto emocional

Los métodos existentes para el marcado de texto emocional podrían clasificarse en cinco categorías básicas: keyword spotting, afinidad léxica, procesamiento estadístico del lenguaje natural, métodos basados en el conocimiento del mundo real y métodos manuales (Francisco, 2008).

- **Keyword spotting:** El marcado del texto se basa en la detección de palabras clave. Es decir, consiste en detectar la aparición de palabras emocionales como *happy*, *sad*... Las principales desventajas de este método son: que causa problemas cuando aparece una negación en la frase y la dependencia del método en aspectos superficiales cuando en la práctica existen muchas frases cuyo contenido emocional no se encuentra en los adjetivos.
- **Afinidad léxica:** Este método no solo detecta palabras obviamente emocionales como los adjetivos, sino que asigna al resto de palabras una afinidad con las distintas emociones. Este método tiene dos problemas fundamentales: emplea tan solo las palabras sin tener en cuenta el contexto en el que se encuentran por lo que puede fallar con la aparición de las negaciones y además este tipo de métodos suelen entrenarse con un corpus lo que dificulta el desarrollo de un modelo reutilizable e independiente del dominio.
- **Procesamiento estadístico:** Consiste en alimentar a un algoritmo de aprendizaje con un varios textos marcados emocionalmente. Este método ha sido empleado en el proyecto de Webmind (Goertzel et al., 2000) entre otros.

- **Métodos manuales:** Estos métodos implican el modelado de distintos estados emocionales en términos de modelos afectivos basados en teorías psicológicas sobre las necesidades, los deseos y las metas de los seres humanos. El modelo DAYDREAMER (Mueller y Dyer, 1985), una computadora que puede generar emociones, emplea este método.
- **Métodos basados en el conocimiento del mundo real:** Este método no solo mira los aspectos superficiales del texto, sino que va más allá, evalúa la calidad afectiva de la semántica subyacente que contiene el texto. Un ejemplo es el marcador llevado a cabo por Liu, Lieberman y Selker (Liu et al., 2002)

Las técnicas basadas en el procesamiento estadístico del lenguaje natural solo funcionan con un texto de entrada lo suficientemente extenso, los métodos manuales precisan un amplio análisis y entendimiento de los textos y hacen muy difícil la generalización y los métodos basados en el conocimiento del “mundo real” necesitan un amplio conocimiento denominado “mundo real” (Francisco y Gervás, 2006).

En este trabajo, emplearemos el método keyword spotting, ya que emplearemos un diccionario con palabras emocionales que nos servirán para analizar todo el texto.

2.3. Diccionarios Afectivos

Un diccionario afectivo es un diccionario en el que las palabras se encuentran marcadas con etiquetas afectivas. Generalmente las etiquetas que se usan son emociones básicas o las categorías emocionales.

En las siguientes secciones presentamos los diccionarios afectivos más representativos.

2.3.0.1. Lasswell Value Dictionary

Fue a comienzos de los años 60 cuando diversos investigadores comenzaron a construir diccionarios en los que las palabras estaban marcadas con etiquetas afectivas. Uno de estos investigadores era Lasswell. El diccionario Lasswell (Peterson et al., 1965) clasifica sus palabras de manera binaria y en ocho categorías básicas (riqueza, poder, respeto, rectitud, habilidad, iluminación, afecto y bienestar) que a su vez pueden estar divididas en subcategorías (ganancias, pérdidas, participantes, fines y arenas). Además de estas, el esquema de este diccionario distingue entre objetivos sustantivos y los elementos y atributos del proceso de distribución de valores, es decir, evaluación personal y asignación social. La clasificación social del contenido mantiene estas distinciones. El diccionario cuenta con 378 palabras de riqueza, 1266 palabras de poder, 245 palabras de respeto, 310 palabras de

rectitud, 257 palabras de habilidad, 835 palabras de iluminación, 196 palabras de afecto y 487 palabras de bienestar; en total 3974 palabras. Por ejemplo la palabra ***ascender*** se encuentra en poder-ganancias ya que indica un aumento de poder. Las palabras ***adoración*** o ***perdón*** se encuentran como rectitud-ganancias, ya que la rectitud se refiere a los valores morales y cuando adoras o se perdona se ganan valores morales.

2.3.0.2. Diccionario de Hatzivassiloglou y McKeown

Este diccionario fue creado en 1997, y trata de encontrar etiquetas de polaridad positiva o negativa de manera automatizada a través del análisis de un corpus. Hatzuvassiloglou y McKeown 1997 tomaron una serie de adjetivos que aparecían de manera frecuente y decidieron darles una orientación. Empleando un análisis estadístico fueron capaces de obtener un diccionario de adjetivos clasificados como positivos y negativos. A través de este análisis se predice la polaridad de los adjetivos, analizándolos en pares unidos por conjunciones como: *and*, *or*, *but*, *either-or*, *or p neither-nor*. Es decir, si una palabra con polaridad conocida aparece unida por la conjunción "and" (and en inglés) a una segunda palabra concluyo que la nueva palabra tendrá una polaridad similar. En cambio si vienen unidas por la conjunción "pero" (but en inglés) la polaridad de la nueva palabra será opuesta. Por ejemplo, si aparece ***hermosa y divertida*** y sabemos que la palabra ***hermosa*** es positiva la palabra ***divertida*** pasará a ser positiva también; sin embargo si aparece ***hermosa pero malvada*** la palabra ***malvada*** tomará la orientación de negativa.

2.3.0.3. Diccionario de Turney y Littman

Turney y Littman (Turney y L.Littman, 2003) encontraron una manera más eficiente para decidir la orientación semántica de la palabra (positiva o negativa) junto con el grado (leve a fuerte). Dado un conjunto de palabras, que previamente ellos conocían como positivas o negativas, probaron como de frecuentemente podía aparecer una palabra en un contexto con un conjunto de palabras positivas o negativas. Empleando la información de estudios cercanos y estadísticas clasificaron como positivas todas aquellas palabras que aparecían de forma más significativa con un conjunto de palabras positivas y como negativas aquellas palabras que aparecían con un conjunto de palabras negativas. El método se prueba experimentalmente con 3596 palabras de las cuales 1614 son positivas y 1982 negativas.

2.3.0.4. Clairvoyance Affect Lexicon

Este diccionario fue desarrollado a mano (Huettner y Subasic, 2000). Las entradas de este diccionario se dividen en cinco campos que caracterizan la

palabra.

- Lema.
- Categoría gramatical.
- Clase afectiva: Entre un pequeño conjunto de categorías emocionales.
- Centralidad: Valor entre 0.0 y 1.0 que trata de medir la relación de la palabra con la clase afectiva que se le ha asignado.
- Intensidad: Mide la intensidad emocional de la palabra.

Cada palabra aparece en el diccionario una vez por cada clase afectiva a la que pertenece. La centralidad varía en las diferentes apariciones mientras que la intensidad se mantiene igual.

Por ejemplo la palabra *jubiloso* tiene una centralidad de 0.7 para felicidad y una centralidad de 0.3 para excitación, pero en ambos casos tiene la misma intensidad 0.6.

2.3.0.5. Diccionario de Grefenstette

Grefenstette et al. 2006 crearon un diccionario afectivo a partir de patrones sintácticos. Tomaron 21 verbos: *appear, appears, appeared, appearing, feel, feels, feeling, felt, are, be, is, was, were, look, looked, looks, looking, seem, seems, seemed, seeming* y les añadieron uno de estos 5 adverbios: *almost, extremely, so, too, very*. A partir de todas las combinaciones posibles obtuvieron 105 patrones. Buscaron los patrones en toda la web para encontrar las palabras que comúnmente aparecen a continuación de ellos y, una vez encontradas filtraban las que tuviesen un componente afectivo clasificándolas en positivas o negativas.

2.3.0.6. ANEW

La Affective Norms for English Words (Bradley y Lang, 1999) contiene un gran número de palabras inglesas medidas en términos de las dimensiones emocionales comentadas anteriormente (evaluación, activación y control). Para obtener este diccionario, se pidió a una serie de personas que marcasen un conjunto de palabras fuera de contexto con los valores para cada una de las tres dimensiones de una emoción en una escala de 9 puntos. Para ello contaron con la ayuda del sistema de marcación SAM (Lang, 1980).

Se trata de un modelo espacial de varias palabras con carga emocional que contribuye a facilitar la clasificación de textos arbitrarios y el estudio de las emociones.

2.3.0.7. Traducción de ANEW al castellano

Se trata de la traducción del diccionario ANEW al castellano (Redondo et al., 2007). En él participaron 720 estudiantes de Psicología de entre 18 y 25 años, 560 mujeres y 160 hombres. Estos participantes evaluaron las 1034 palabras que contiene ANEW. Cada palabra debía ser marcada con las dimensiones emocionales: evaluación, activación y control, en una escala de 9 puntos.

Cada entrada en el diccionario contiene: un número que identifica a la palabra, de manera que esta numeración coincide con el número que dicha palabra tiene en el ANEW; la palabra inglesa (E-word), la palabra original en la base de datos ANEW; la palabra española (S-word); y las evaluaciones afectivas, los valores medios y la desviación estándar para cada dimensión emocional.

En nuestro trabajo emplearemos un diccionario similar a la Traducción de ANEW al castellano ya que tendremos todas las palabras evaluadas con una media de la carga afectiva de todas las evaluaciones realizadas.

2.4. Servicios Web

En la actualidad, la definición más general de Servicio Web es la que lo define como un conjunto de aplicaciones o tecnologías capaces de inter-operar en la Web. Estas aplicaciones o tecnologías intercambian datos con el fin de ofrecer unos servicios. Los proveedores ofrecen sus servicios como procedimientos remotos y los usuarios solicitan un servicio llamando a estos procedimientos a través de la Web ².

2.4.1. Características de los Servicios Web

Un servicio web, como su nombre indica, debe poder ser accesible a través de la web. Para ello debe utilizar protocolos de transporte estándares como HTTP, y codificar los mensajes en un lenguaje estándar que pueda conocer cualquier cliente que quiera utilizar el servicio.

Un servicio web, además, debe contener una descripción de sí mismo. De esta forma, una aplicación podrá saber cuál es la función de un determinado servicio, un servicio web debe poder ser localizado. Deberemos tener algún mecanismo que nos permita encontrar un servicio web que realice una determinada función. De esta forma tendremos la posibilidad de que una aplicación localice el servicio que necesite de forma automática, sin tener que conocerlo previamente el usuario. ³

²<https://www.w3c.es/Divulgacion/GuiasBreves/ServiciosWeb>

³<http://www.jtech.ua.es/j2ee/publico/servc-web-2012-13/sesion01-apuntes.html>

2.4.2. Tipos de Servicios Web

A nivel técnico, los servicios pueden implementarse de varias formas. En este sentido, podemos distinguir dos tipos de servicios web (Dto. de Ciencia de la Computación e Inteligencia Artificial):

- SOAP: utilizan mensajes XML para intercomunicarse que siguen el estándar SOAP (Simple Object Access Protocol), un lenguaje XML que define la arquitectura y formato de los mensajes. Dichos sistemas normalmente contienen una descripción legible por la máquina de la descripción de las operaciones ofrecidas por el servicio, escrita en WSDL (lenguaje basado en XML para definir las interfaces sintácticamente).
- Servicios web RESTful: utilizan estándares conocidos como HTTP, SML, URI, MIME, y tienen una infraestructura “ligera” que permite que los servicios se construyan utilizando herramientas de forma mínima.

2.4.3. Arquitectura de los Servicios Web

Los servicios web se componen de varias capas entre las que destacan: servicios de transporte (constituidos por los protocolos de nivel más bajo, que codifican la información independientemente de su formato y que pueden ser comunes a otros servicios), servicios de mensajería, de descripción y de descubrimiento. En las siguientes subsecciones explicaremos más en detalle cada una de estas capas⁴ ⁵.

- Servicios de transporte: Capa que se encarga de transportar los mensajes entre el servidor y el cliente. Normalmente se utiliza el protocolo HTTP para este transporte.
- Servicios de mensajería: Capa encargada de la codificación de los mensajes en formato estándar para que pueda ser interpretado en cualquiera de los nodos de la red. Decide la tipología del mensaje que se envía, puede ser en *XML* o *JSON*.
- Servicios de descripción: Capa encargada de definir la interfaz pública de un determinado servicio. Esta definición se realiza mediante WSDL (Web Service Description Language), tipo de documento XML que describe lo que hace un servicio web, dónde se encuentra y la forma de ser invocado.
- Servicios de descubrimiento: Capa encargada del registro centralizado de servicios, permitiendo que estos sean anunciados y localizados. Para

⁴<https://diego.com.es/introduccion-a-los-web-services>

⁵<http://www.jtech.ua.es/j2ee/publico/servc-web-2012-13/sesion01-apuntes.html>

ello se utiliza el protocolo UDDI, el cual define la especificación para construir un directorio distribuido de servicios web, donde los datos se almacenan en XML. Este registro también almacena información sobre las organizaciones que los proporcionan, la categoría en la que se encuentran, y sus instrucciones de uso.

2.4.4. Ventajas e inconvenientes

Las principales ventajas de los servicios web son:

- Aportan interoperabilidad entre aplicaciones de software independientemente de sus propiedades o de las plataformas sobre las que se instalan.
- Fomentan los estándares y protocolos basados en texto, de tal manera que facilitan su comprensión y accesibilidad.
- Permiten la combinación e integración de distintos servicios y software independientemente de su ubicación.
- Al apoyarse en HTTP, los servicios web se pueden aprovechar de los de seguridad firewall sin necesidad de cambiar las reglas de filtrado.
- Disminuyen el tiempo de desarrollo de las aplicaciones: Pues gracias a la filosofía de orientación a objetos utilizada, el desarrollo se convierte más bien en una labor de composición.

Las dos principales desventajas de los servicios web son:

- Bajo rendimiento si se compara con otros modelos de computación distribuida.
- Al apoyarse en HTTP, pueden esquivar medidas de seguridad basadas en firewall cuyas reglas tratan de bloquear la comunicación entre programas.

2.5. Scrum

Scrum es una metodología ágil para gestionar el desarrollo de software. Fue definida por *Ikujiro Nonaka* e *Hirotaka Takeuchi* a principios de los 80 (ScrumManager, 2016).

Scrum divide el trabajo en diferentes unidades llamadas *sprints*, que tienen una duración preestablecida de entre dos y cuatro semanas obteniendo siempre al final una versión del software con nuevas prestaciones listas para ser usadas. En cada *sprint* se ajusta la funcionalidad y se añaden nue-

vas prestaciones priorizando aquellas que aporten más valor al producto⁶ (ScrumManager, 2016).

2.5.1. Roles

Esta metodología hace mucho énfasis en el «equipo de trabajo». Este equipo está formado por diferentes roles:

- **Product Owner:** Representa al cliente. Se encarga de definir los objetivos y de dar valor al producto.
- **Scrum Master:** Encargado de asegurar que se cumplen las buenas prácticas y valores descritos en el modelo Scrum.
- **Scrum Team:** Equipo encargado de desarrollar y entregar el producto. Es autogestionado y multidisciplinar.

2.5.2. Artefactos

Los artefactos definidos en Scrum son:

- **Product backlog:** Lista realizada por el Product Owner que contiene los requisitos del producto, ordenados por prioridad. A lo largo del desarrollo crece y evoluciona. En Scrum los requisitos se definen mediante historias de usuarios. Una historia de usuario es la descripción de una funcionalidad que debe incorporar un sistema de software, y cuya implementación aporta valor al cliente.
- **Sprint backlog:** Descomposición en tareas de las historias de usuario del product backlog seleccionadas para que el equipo las realice durante el sprint.

2.5.3. Eventos

Por último, Scrum también define una serie de eventos:

- **Sprint:** Cada una de las iteraciones del desarrollo.
- **Sprint Planning:** Se trata de la reunión de planificación del sprint. Este evento consta de dos partes. En la primera parte el Product Owner presenta lo que quiere que se haga en el sprint y resuelve dudas acerca de la historia de usuario y explica la importancia de dicha tarea para que el grupo de trabajo tenga clara la prioridad. Después,

⁶<https://www.softeng.es/es-es/empresa/metodologias-de-trabajo/metodologia-scrum/proceso-roles-de-scrum.html>

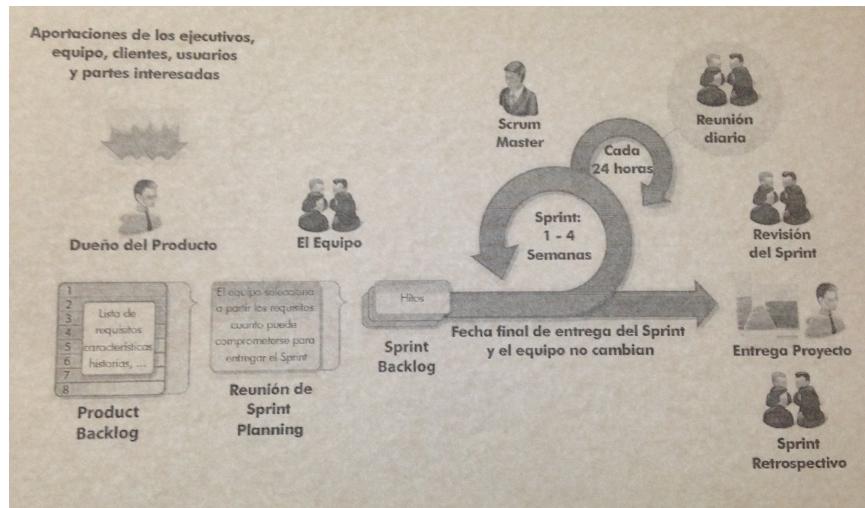


Figura 2.1: Flujo Scrum

en la misma reunión, el equipo estima el tiempo que llevará a desarrollar cada una de las diferentes historias de usuario propuestas en el Product Owner y deciden cuantas van a implementar en el sprint, para posteriormente crear el Sprint Backlog. En la segunda parte, el equipo desarrollador divide las historias de usuario en pequeñas tareas ordenadas por prioridad en el Sprint Backlog.

- **Daily Scrum:** Reunión diaria como máximo de quince minutos, de pie, donde cada componente del equipo informa sobre cómo va en sus tareas, lo que hizo el día anterior, lo que hará ese día y los problemas que ha encontrado o los que cree que se va a encontrar.
- **Sprint Review:** Reunión que se realiza al concluir el sprint centrándose en el producto. Se presenta el producto creado en el sprint al Product Owner y este lo analiza y da su feedback.
- **Sprint Retrospective:** Reunión donde se habla de cómo ha funcionado el equipo en el sprint y qué cosas se pueden mejorar para el siguiente a nivel de proceso o metodología.

En la Figura 2.1 podemos ver un diagrama del flujo Scrum. Sería el siguiente: el Product Owner crea el Product Backlog con los requisitos y características por orden de prioridades. A continuación, en un Sprint Planning, se presenta el Product Backlog y se decide qué actividades van a desarrollar en el sprint. En esta reunión se elabora el Sprint Backlog con todas las historias de usuario que van a realizar divididas en tareas. A continuación, se comienza el sprint con el tiempo establecido, cada componente del grupo se

asigna una tarea y en cuanto acabe seguirá con la siguiente actividad que no esté hecha. Es muy importante el orden de las tareas, ya que el cliente ha ordenado en el Product Backlog las actividades por prioridad y el equipo desarrollará estas siguiendo dicho parámetro. Cada día, el equipo se reunirá en una Daily Scrum. Al final del sprint se realiza la Sprint Review donde se entrega el incremento del producto al cliente. Por último, todo el equipo se vuelve a reunir para hacer el Sprint Retrospective.

2.6. Integración Continua

En general, las metodologías ágiles y en particular la metodología Scrum enfoca su objetivo en tener versiones del producto al finalizar cada etapa; es por esto por lo que la integración continua es vital en proyectos que empleen estas metodologías. Empleado la integración continua, los desarrolladores se aseguran de encontrar los errores a tiempo para que el producto esté listo para la entrega al cliente o para ponerlo en producción si este así lo desea.

La integración continua (Fowler, 2006), se basa en que los desarrolladores combinen todos los cambios que realicen en el código en un repositorio común de forma periódica (cada pocas horas, al menos una vez al día), de tal forma que una vez subidos estos cambios, se ejecutan una serie de pruebas automáticas sobre estos con el fin de validarlos y detectar errores de integración tan pronto como sea posible.

Las principales ventajas de la integración continua:

- **Detección de errores:** Cada vez que el código cambia se compila y se somete a pruebas de forma inmediata para garantizar que no hay errores. Este proceso aumenta la calidad del software y minimiza los riesgos del proceso ya que se tiene control sobre las versiones en todo momento.
- **Visibilidad del proceso:** Todos los pasos que se realizan en el desarrollo son visibles a todo el equipo, que tiene una estrategia común muy bien definida.
- **Mejora del equipo:** Los desarrolladores no solo tienen una visión muy clara y estructurada del proceso sino que también aprenden a realizar todo tipo de pruebas, lo que les hace mejorar a nivel profesional.

Lo primero para poder utilizar integración continua es tener definido un pipeline, es decir, un conjunto de fases por las que tiene que pasar el software y que están automatizadas. Se establecen criterios para que el código pase de una fase a otra y estrategias para gestionar errores que puedan surgir en las diferentes fases (control de versiones). Es importante tener bien definidas las pruebas que se van a realizar sobre cada fase y que estas puedan garantizar la

máxima corrección posible sin tardar mucho, ya que se necesita un feedback rápido para poder seguir avanzando en el proceso. Cada fase es un grupo de pruebas y cada subida de código es un pipeline distinto que avanza de forma independiente por las fases. Por lo tanto se sabe en todo momento en qué punto se encuentra una versión específica. Esto permite tener una visión general de todo el proceso facilitando notablemente la detección de errores en fases y pipelines concretos.

Para el correcto funcionamiento de esta práctica tiene que haber pequeñas integraciones de forma frecuente, una vez al día por ejemplo. Cuantos menos cambios haya más fácil es la integración en el producto general y solucionar los posibles errores que esta pueda generar. Cabe destacar que aunque una parte de código funcione de forma independiente no implica que vaya a funcionar al integrarlo en un programa más grande, por ello cuanto más frecuentes sean las integraciones mejor.

Capítulo 3

Herramientas

RESUMEN: En este capítulo se profundizará en las herramientas que utilizaremos a lo largo del trabajo. En la sección 3.1 se muestran las herramientas básicas que vamos a usar. En la sección 3.2 se presenta el diccionario que vamos a utilizar para el marcado emocional. En la sección 3.3 se introduce el framework que vamos a utilizar para el desarrollo de los servicios web, Django. En la sección 3.4 se explica cómo vamos a utilizar Trello para seguir la metodología Scrum. En la sección 3.5 se expone la forma de realizar las pruebas utilizando Jenkins y Doctest. En la sección 3.6 se presentan SpaCy y PyStemmer, las herramientas que se utilizarán para poder procesar las palabras que forman una frase para poder realizar el análisis emocional sobre ella.

3.1. Herramientas básicas

- **Repositorio:** Se utilizará un repositorio común de *GitHub* en el que se subirán todos los cambios realizados en el código. Al ser un equipo de desarrollo pequeño y estar utilizando la metodología Scrum en principio todos los miembros del equipo estarán trabajando en la misma rama. Esto puede llegar a bloquear el proceso mientras una versión acabe de pasar las pruebas.
- **Pruebas manuales:** Algún miembro del equipo realizará las tareas de “tester”. Después de las pruebas automáticas se realizará otra serie de pruebas planificadas de antemano para probar mejor la funcionalidad del código. El “tester” no será siempre la misma persona.
- **Pruebas automáticas:** Usaremos un software especial con el fin de controlar la ejecución de pruebas y la comparación entre los resultados obtenidos y los resultados esperados. El uso de estas pruebas, nos

permite incluir pruebas muy repetitivas y necesarias, dado que habrá pruebas que realizarlas de manera manual nos podrá ser muy costoso.

3.2. Diccionario

Este diccionario, basado en el diccionario ANEW traducido al castellano, es una adaptación en el que se han elegido las 5 emociones básicas (tristeza, miedo, alegría, enfado y sorpresa) y la neutralidad para no expresar ninguna emoción. Para crearlo participaron 10 personas, y cada una de ellas marcó cada una de las 1034 palabras con una emoción básica o la neutralidad. Cada palabra puede representar varias emociones por lo que se ha creado un número de confianza c , entre 0 y 1, que determina con qué certeza la palabra corresponde a esa emoción. Este número indica el tanto por uno de evaluadores que asignaron dicha emoción a la palabra. La suma de estos números debe sumar 1.

En la Tabla 3.1 podemos ver los valores obtenidos para las palabras “abandonado”, “aborto” y “abeja”. En este ejemplo se puede ver que tristeza fue la emoción asignada a la palabra “abandonado” por todos los evaluadores, mientras que con “abejas” y “aborto” no hubo acuerdo. En el caso de “aborto”, un 67% de los anotadores le asignaron tristeza, mientras que un 17% consideró que la palabra no tenía emoción asociada y un 17% le asignaron miedo.

S-Word	Tisteza	Miedo	Alegría	Enfado	Sorpresa	Neutral
abandonado	1,00	0,00	0,00	0,00	0,00	0,00
abejas	0,00	0,50	0,17	0,00	0,00	0,33
aborto	0,67	0,17	0,00	0,00	0,00	0,17

Tabla 3.1: Fragmento de la adaptación del diccionario ANEW traducido

3.3. Django

Toda la implementación del trabajo se hará utilizando Django, un framework para aplicaciones web gratuito y open source escrito en Python. El framework de Django nos proporciona un servidor web, en el que se almacena la base de datos. Esta base de datos, contiene las palabras con sus respectivas probabilidades para cada emoción y la neutralidad modeladas mediante su lexema y los grados de certeza para cada emoción. La base de datos permitirá hacer las consultas necesarias. Para realizar las diferentes consultas sobre las palabras disponibles existen una serie de clases que implementan los diferentes métodos de un servicio web REST típico: **GET, POST, DELETE**.

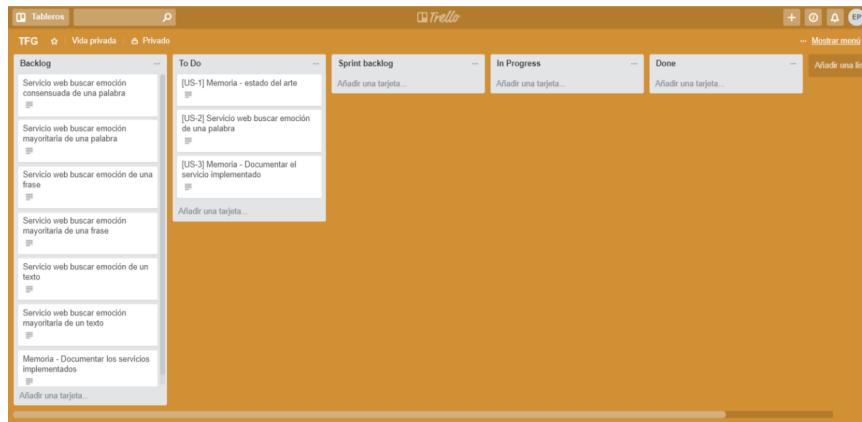


Figura 3.1: Sprint inicial que ilustra la estructura.

Cada una de las diferentes clases nos aportarán una manera diferente de acceder a la información, como pueden ser: acceso a todo el diccionario de palabras, a una palabra concreta o a un campo de una palabra concreta. Los resultados serán devueltos en formato JSON.

3.4. Trello

Trello es una aplicación web que permite organizar proyectos y actividades. Para representar las tareas y las historias de usuario se usan tarjetas virtuales. En la Figura 3.2 podemos ver el estado inicial del proyecto. Se observa el **Product Backlog**, del que product owner saca la cantidad de las historias de usuario que quiere que se realicen durante el sprint y estas pasan a la lista **To Do**. En la Figura 3.3 se puede ver un ejemplo más avanzado en el que se puede ver como las historias de usuario han sido divididas en tareas para formar el **Sprint Backlog**, del que van saliendo en orden para estar **En Progreso** y, una vez acabadas, **Done**.

3.5. Doctest y Jenkins

Como ya se comentó en el capítulo 2.4 utilizaremos Jenkins como parte de la integración continua del proyecto. Esto nos permitirá asegurarnos de que la unificación es correcta y realizar las pruebas automáticas. Esto último se llevará a cabo mediante una orden shell que Jenkins ejecutará cada vez que se detecte un cambio en el repositorio. La orden únicamente se encarga de ejecutar el script de pruebas que contendrá las llamadas a los diferentes programas de pruebas que se desarrollen.

Los programas de pruebas utilizarán Doctest para hacer las pruebas.

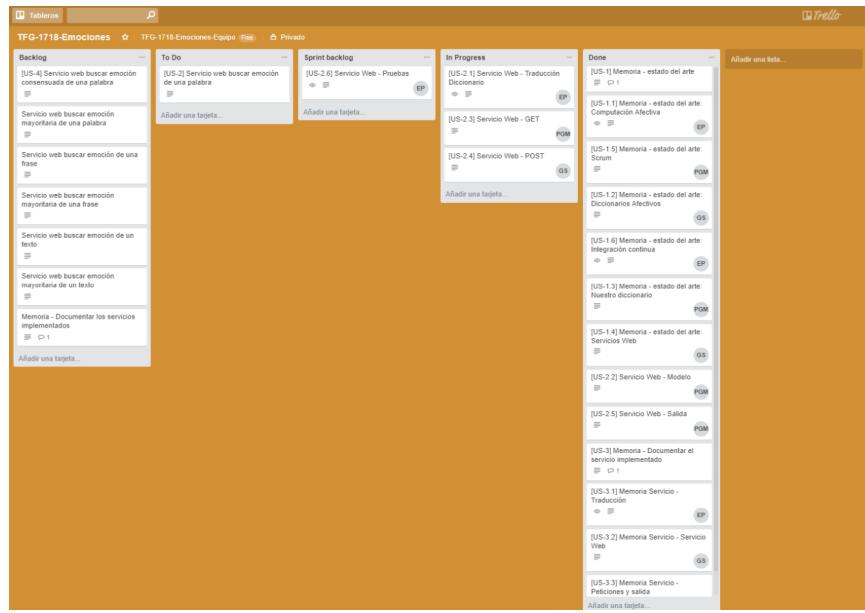


Figura 3.2: Final del sprint inicial.

Doctest es un módulo incluido en la librería estándar de Python. Su funcionamiento se basa en definir la función que se quiera probar y, dentro de un comentario al inicio de esta, poner una serie de llamadas y el resultado que se espera obtener de ellas. Tiene una función “testmod” que realiza las pruebas y devuelve el número de fallos y el resultado de todas las pruebas. Si el número de fallos es mayor que cero provocamos una excepción que Jenkins detectará para notificar a todo el equipo que hay algún fallo. Los resultados de las pruebas se muestran por consola al acabar y Jenkins los guardará para ayudar a encontrar el problema.

3.6. SpaCy y PyStemmer

El objetivo final es llegar a interpretar la emoción de textos enteros, no sólo palabras. Para ello se necesita una herramienta que nos facilite trabajar con frases, etiquetando cada una de las palabras que las forman. **SpaCy** es una librería open source escrita en Python y dedicada al Procesamiento de Lenguajes Naturales. Soporta el español y nos permite etiquetar las palabras para poder buscar sólo aquellas que puedan tener carga emocional. SpaCy recibirá el texto plano, en este caso una serie de frases, y devolverá un objeto de tipo “Doc”, propio de la librería, que contendrá la frase con una serie de anotaciones sobre cada una de las palabras que la forman (lema, etiqueta, dependencias sintácticas, forma...).

Una vez que hemos filtrado la frase para quedarnos con las palabras que nos interesan para el análisis emocional tenemos que obtener el lema de cada una de ellas. Para ello utilizaremos la librería de Python **PyStemmer** que consiste en una adaptación de Snowball para Python. Snowball es un pequeño lenguaje de procesamiento que permite crear algoritmos de lematización. PyStemmer soporta el español y nos ofrece mejores resultados que Spacy a la hora de obtener los lemas de las palabras, por eso vamos a combinar ambas herramientas para procesar las palabras.

Capítulo 4

Análisis del Contenido Afectivo de un Texto

RESUMEN: En este capítulo se explicarán los servicios web que se han desarrollado con el fin de analizar el contenido afectivo de un texto. Primero, en la sección 4.1, se explica la base sobre la que trabajarán todos los servicios. En la sección 4.2, se presentarán los servicios orientados a identificar las emociones predominantes en una palabra concreta. Después, en la sección 4.3, se explica como el análisis de la palabra se aplica a determinar la emoción de una frase entera para finalmente, en la sección 4.4, aplicar todo al análisis de un texto arbitrario completo.

4.1. Base para los servicios web

Como ya se comentó en la sección 3.1, para el desarrollo de este trabajo se utilizará un diccionario afectivo en español. Este diccionario, inicialmente en formato Excel, ha sido convertido a un fichero CSV para que resulte más sencillo obtener los datos.

Los datos estarán almacenados en el servidor Django.

- **Modelo:** Los campos que tiene el modelo para cada palabra, son: la propia palabra, el lema de esta y el porcentaje de certeza de cada una de las emociones básicas y la neutralidad. Cada una de las palabras que tenemos recogidas se encapsulan en este modelo para ser almacenadas en el servidor.
- **Consultas:** Se realizarán mediante declaración de clases “vista” según las funcionalidades que requiera cada uno de los servicios web que vamos a desarrollar. Inicialmente disponemos de una vista general (Figura 4.1) que muestra la lista entera de palabras almacenadas y una vista detalle (Figura 4.2) para cada una de ellas.

```

GET /emociones/
HTTP 200 OK
Allow: GET, POST, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

[
  {
    "id": 1,
    "palabra": "abandonado",
    "lexema": "abandon",
    "porcentajes": "[100, 0, 0, 0, 0, 0]"
  },
  {
    "id": 2,
    "palabra": "abejas",
    "lexema": "abej",
    "porcentajes": "[0, 50, 17, 0, 0, 33]"
  },
  {
    "id": 3,
    "palabra": "aborto",
    "lexema": "abort",
    "porcentajes": "[67, 17, 0, 0, 0, 17]"
  },
  {
    "id": 4,
    "palabra": "abrasador",
    "lexema": "abras",
    "porcentajes": "[0, 50, 0, 17, 0, 33]"
  }
]

```

Figura 4.1: Vista general de la lista de palabras.

Una vez que se tiene el servidor y todo el código necesario para el funcionamiento de los servicios web se procede a publicar/subir las palabras que recoge nuestro diccionario. Para ello se ha desarrollado un programa en *Python* cuya función es leer el fichero CSV, interpretar cada una de sus líneas y almacenar/subir la información que obtiene al servidor. Una vez que todas las palabras estén almacenadas/subidas ya se pueden realizar las consultas necesarias.

4.2. Análisis afectivo de una palabra

Hemos implementado tres servicios web diferentes para mostrar la carga afectiva de una palabra, dependiendo de la información buscada en cada momento.

4.2.1. Servicio web porcentajes de una palabra

Dada una palabra, este servicio nos devuelve la información respectiva a los porcentajes de cada emoción que posee. Para ello se realiza una petición **GET** al servidor, y se devuelve el resultado en JSON al usuario.

```
GET /alegre/
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

{
    "id": 42,
    "palabra": "alegre",
    "lexema": "alegr",
    "porcentajes": "[0, 0, 100, 0, 0, 0]"
}
```

Figura 4.2: Vista detalle de la palabra “Alegre”.

```
GET /porcentajes/alegre/
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

"Tristeza:0% || Miedo:0% || Alegría:100% || Enfado:0% || Sorpresa:0% || Neutral:0%"
```

Figura 4.3: Respuesta al buscar los porcentajes de la palabra “alegre”.

Como entrada tendríamos la palabra de la cual queremos conocer sus emociones; existirán dos posibilidades, bien que la palabra esté en nuestro diccionario, por lo que la salida sería el porcentaje de cada emoción (Figura 4.3), o bien que la palabra no exista en nuestro diccionario, que devolverá un en la respuesta un mensaje 404 de “NOT FOUND” (Figura 4.4).

4.2.2. Servicio web emoción consensuada de una palabra

Dada una palabra, este servicio nos devuelve a través de una petición **GET** al servidor, si la palabra tiene o no emoción consensuada y en caso de tenerla cuál. Hablamos de emoción consensuada cuando el grado de certeza que se tiene de una de sus emociones es 1, es decir, el porcentaje para dicha emoción es 100.

Como entrada tendríamos la palabra de la cual queremos conocer su emoción consensuada, mientras que la salida será la emoción consensuada en caso de tenerla (Figura 4.5) o un mensaje de que no la tiene (Figura 4.6).

4.2.3. Servicio web emoción mayoritaria de una palabra

De nuevo, dada una palabra el servicio nos devuelve, recibiendo una petición **GET** en el servidor, la emoción mayoritaria de una palabra. Hablamos de emoción mayoritaria cuando el porcentaje de certeza de una emoción es mayor al del resto de emociones.

Como entrada tendríamos la palabra de la cual queremos conocer su emoción mayoritaria, mientras que la salida será la emoción mayoritaria (Fi-

```
GET /emocion/te/percentages/
HTTP 404 Not Found
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

{
    "detail": "Not found."
}
```

Figura 4.4: Respuesta al buscar los porcentajes de una palabra que no está en el diccionario.

```
GET /consensuada/alegre/
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

"Consensuada: Alegría"
```

Figura 4.5: Respuesta al encontrar la emoción consensuada.

gura 4.7), o mayoritarias (en caso de que haya dos con el mismo porcentaje) (Figura 4.8).

4.3. Análisis afectivo de una frase

El análisis emocional de una frase se sustenta en los servicios web desarrollados para hallar la carga afectiva de las palabras. Procesaremos las palabras que forman la frase utilizando las herramientas comentadas en el apartado 3.5 para obtener una lista de palabras candidatas acompañadas de sus lemas. Para cada una de las palabras de la lista obtenemos su información emocional mediante los servicios web explicados en el apartado anterior y, combinando los resultados obtenidos, podemos mostrar la carga afectiva de la frase.

```
GET /consensuada/afianzar/
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

"No hay emoción consensuada"
```

Figura 4.6: Respuesta si la palabra no tiene emoción consensuada.

```
GET /mayoritaria/alegre/
HTTP 200 OK
Allow: GET, HEAD, OPTIONS
Content-Type: application/json
Vary: Accept

"Mayoritaria: Alegría con un 100%"
```

Figura 4.7: Respuesta al encontrar la emoción mayoritaria.

4.3.1. Porcentajes de una frase

Los porcentajes de cada emoción para una frase se obtienen haciendo una media con los porcentajes de las palabras que la forman. Por ejemplo, si la frase es: *Estoy alegre y feliz.*, las palabras emocionales son “alegre” y “feliz” y se haría una media entre las dos para todos los porcentajes. En este caso ambas son 100 % alegres y 0 % el resto de emociones, así que el resultado sería que la frase es 100 % alegre. Si la frase, en cambio, es: *Estoy alegre y triste*, como “alegre” es 100 % alegre y “triste” es 100 % triste, la frase será 50 % de cada una de estas emociones.

La media que realizamos es ponderada, dando más peso a los verbos ya que estos son el núcleo de la frase.

4.3.2. Emoción mayoritaria de una frase

Para hallar la emoción mayoritaria de una frase se cuentan las apariciones de una emoción entre las palabras emocionales como mayoritaria y se hace una media con los porcentajes de todas estas apariciones, dando más importancia a los porcentajes de los verbos. Una vez que se tienen los porcentajes definitivos para cada emoción se comparan para hallar el mayor y la emoción a la que corresponde será la mayoritaria de la frase.

```
GET /mayoritaria/afianzar/  
  
HTTP 200 OK  
Allow: GET, HEAD, OPTIONS  
Content-Type: application/json  
Vary: Accept  
  
"Mayoritarias: Alegría y Neutral con un 50%"
```

Figura 4.8: Respuesta si hay dos emociones mayoritarias.

4.4. Análisis afectivo de un texto

Analizar un texto consiste, básicamente, en partirlo en frases y obtener la carga afectiva de cada una de ella. Las operaciones realizadas durante el análisis son, por lo tanto, las mismas que para analizar una frase: medias ponderadas para hallar los porcentajes y la emoción mayoritaria a partir de la información devuelta por cada frase. Analizar un texto consiste, básicamente, en partirlo en frases y obtener la carga afectiva de cada una de ella. Las operaciones ralizadas durante el análisis son, por lo tanto, las mismas que para analizar una frase: medias ponderadas para hallar los porcentajes y la emoción mayoritaria a partir de la información devuelta por cada frase.

Capítulo 5

Diseño de la aplicación web

RESUMEN: En este capítulo se explicará el proceso seguido para diseñar la interfaz que tendrá la aplicación web que desarrollaremos. Primero, en la sección 5.1, se presentan los prototipos que diseñamos cada una de las tres. En la sección 5.2, se muestra el prototipo que obtuvimos tras poner los tres en común. Finalmente, en la sección 5.3, se presenta la interfaz que finalmente tendrá la aplicación.

5.1. Prototipos iniciales

Cada una de nosotras diseñó un prototipo distinto para la aplicación web. Los prototipos resultantes son bastante similares entre sí. A continuación se muestran las principales secciones de la aplicación, la pantalla principal y la pantalla que muestra el resultado del análisis, y sus diferencias en los tres prototipos.

5.1.1. Pantalla principal

La pantalla principal en todos los casos es bastante simple, un cuadro de texto donde introducir el texto a interpretar y un botón para ejecutar el interprete (Figura 5.3). Puede incluir los emoticonos que representen las distintas emociones con las que trabajamos debajo del cuadro (Figura 5.2) o a la derecha de este (Figura 5.1).

5.1.2. Pantalla de análisis

Los resultados se muestran de forma similar en los tres prototipos. Todos ellos utilizan emoticonos

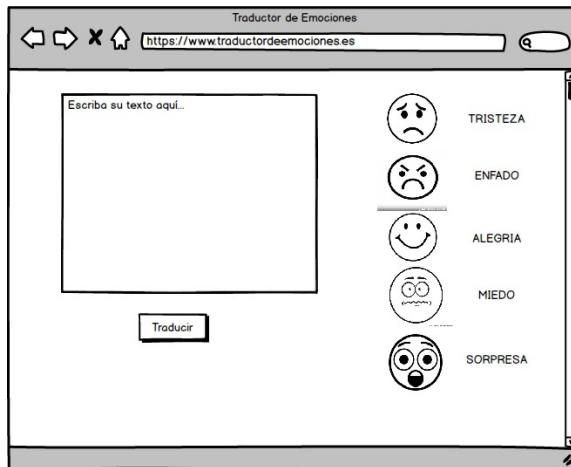


Figura 5.1: Pantalla inicial del prototipo de Elena

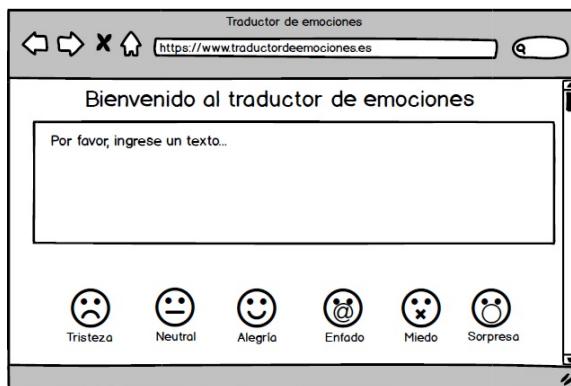


Figura 5.2: Pantalla inicial del prototipo de Gema

5.2. Prototipo común

Tras comparar los tres prototipos se llegó a uno común.

5.2.1. Pantalla principal

5.2.2. Pantalla de análisis

5.3. Interfaz final

Tras reunirnos con la asociación y teniendo en cuenta la información que nos proporcionó acordamos un prototipo final sobre el que realizar la interfaz de la aplicación.

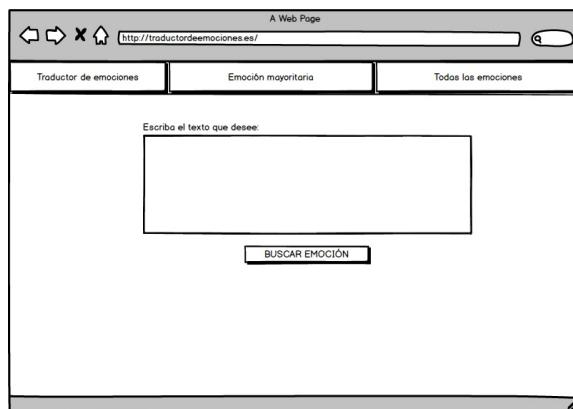


Figura 5.3: Pantalla inicial del prototipo de Paloma

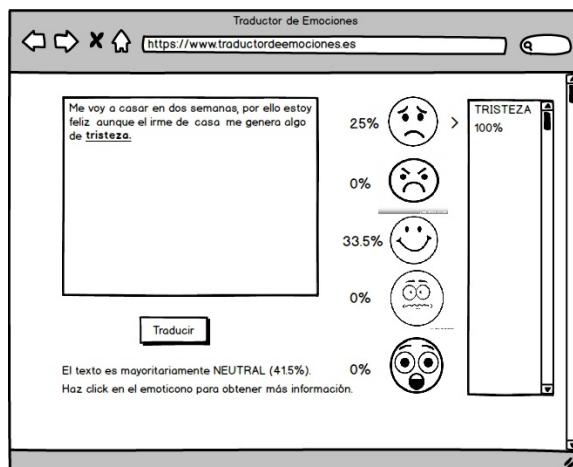


Figura 5.4: Pantalla de análisis del prototipo de Elena

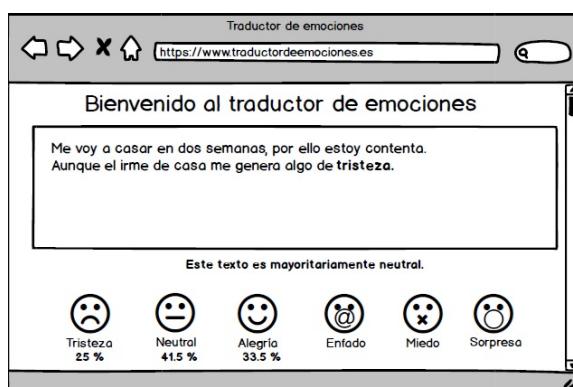


Figura 5.5: Pantalla de análisis del prototipo de Gema

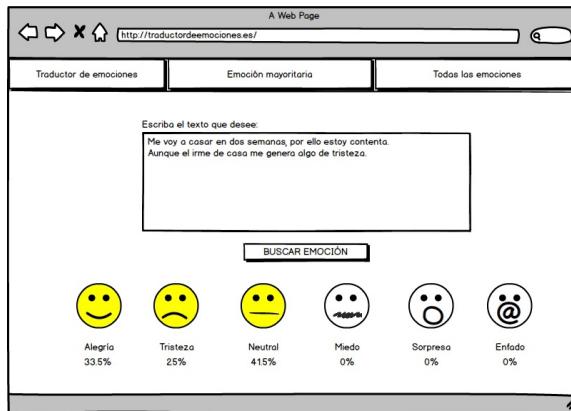


Figura 5.6: Pantalla de análisis del prototipo de Paloma



Figura 5.7: Pantalla inicial por defecto del prototipo final

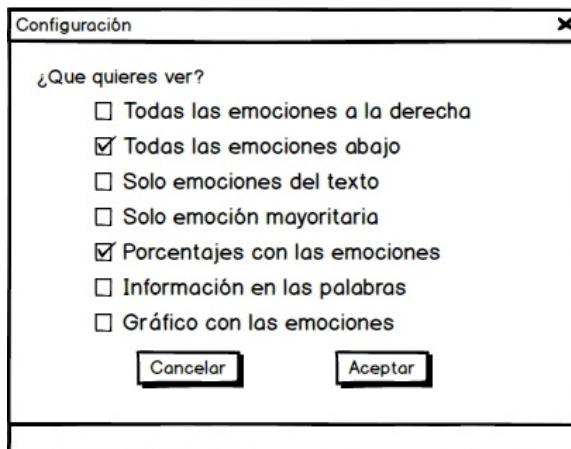


Figura 5.8: Menú de opciones de la interfaz

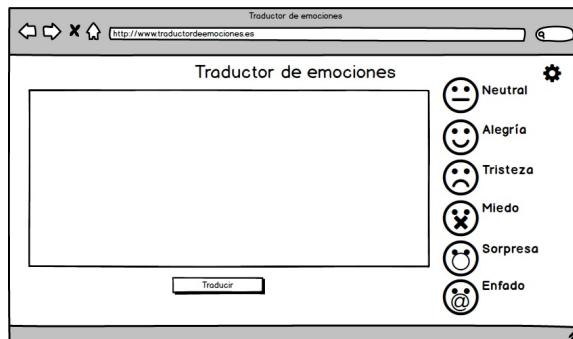


Figura 5.9: Pantalla inicial alternativa del prototipo final

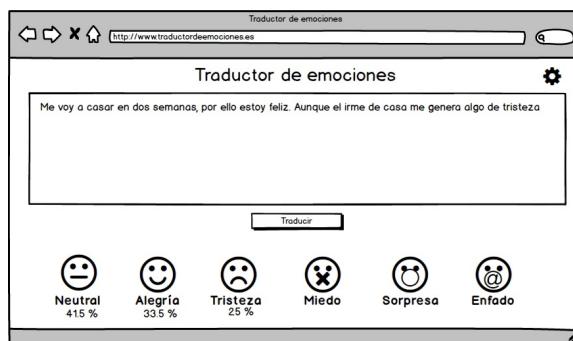


Figura 5.10: Pantalla de análisis con los porcentajes abajo

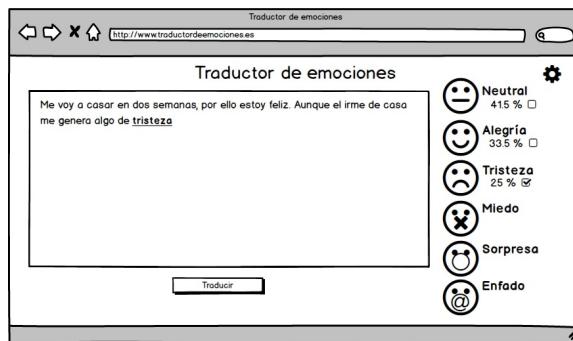


Figura 5.11: Pantalla de análisis con los porcentajes derecha

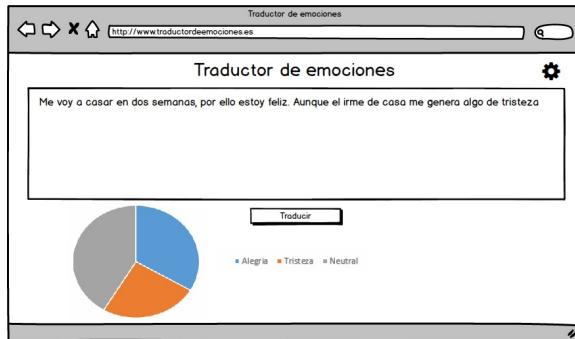


Figura 5.12: Pantalla de análisis con gráfico

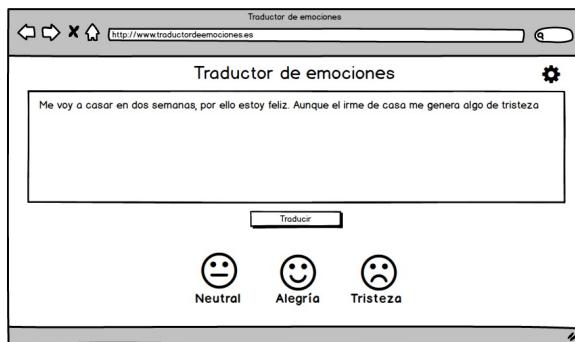


Figura 5.13: Pantalla de análisis mostrando las emociones que aparecen

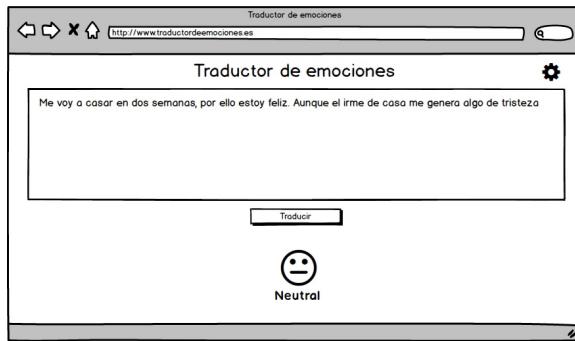


Figura 5.14: Pantalla de análisis mostrando sólo la mayoritaria

Capítulo 6

Desarrollo de la aplicación web

RESUMEN: En este capítulo se explicará el proceso de desarrollo de la aplicación web.

Capítulo 7

API de la aplicación web

RESUMEN: En este capítulo se presentará una API de la aplicación desarrollada.

Capítulo 8

Resultados

RESUMEN: En este capítulo se presentarán los resultados obtenidos tras acabar el proyecto. En la sección 8.1 se presentan los resultados obtenidos de los casos de prueba mientras que en la sección 8.2 se muestran los resultados de la evaluación con usuarios.

8.1. Resultados de los casos de prueba

Los casos de prueba que hemos utilizado se encuentran en el Apéndice B. Los resultados obtenidos son los siguientes:

8.2. Resultados evaluación con usuarios finales

Tuvimos la oportunidad de realizar una evaluación con usuarios finales de la aplicación y gracias a ella obtuvimos los siguientes resultados:

Apéndice A

Así se hizo...

RESUMEN: Como ya se comentó en el capítulo 2, durante el proyecto se ha seguido la metodología Scrum. En este apéndice se presentan todos los sprints que han tenido lugar a lo largo del desarrollo.

A.1. Sprint 1

El primer sprint (Figura A.1) consistió mayoritariamente en investigar el estado del arte y las herramientas que vamos a utilizar y añadir la información obtenida a la memoria (Sección 2). También se comenzó a realizar el primer servicio web, obtener los porcentajes de cada emoción para una palabra, y documentarlo (Sección 4.2.1).

Al acabar el sprint (Figura A.2) quedaba por acabar el servicio y ejecutar las pruebas pertinentes sobre él.

A.2. Sprint 2

El segundo sprint (Figura A.3) consistió en terminar el servicio web que se comenzó en el sprint 1 y desarrollar uno nuevo que nos permita obtener la emoción consensuada de una palabra (Sección 4.2.2).

Al acabar el sprint (Figura A.4) quedaba realizar las pruebas sobre los servicios desarrollados, tanto manuales como automáticas, y resolver un problema que teníamos al subir las palabras al servidor, no reconocía las tildes correctamente.

A.3. Sprint 3

El tercer sprint (Figura A.5) consistió en realizar las pruebas sobre los servicios web que ya teníamos dejándolos así acabados, desarrollar uno nuevo

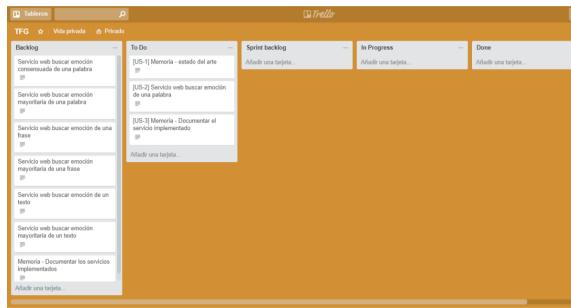


Figura A.1: Planificación del sprint 1 (14/11/2017)

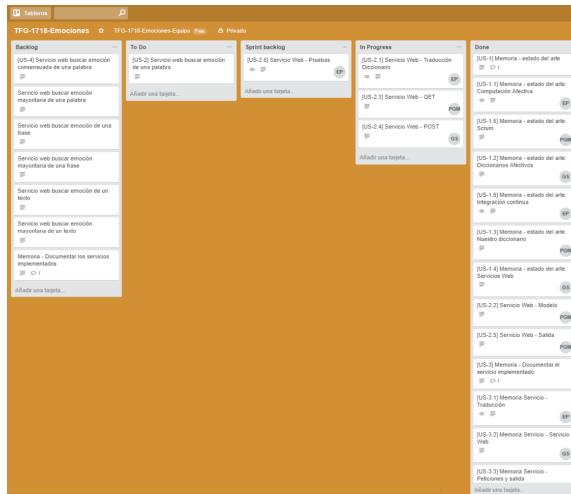


Figura A.2: Final del sprint 1 (28/11/2017)

que nos proporcione la emoción mayoritaria de una palabra (Sección 4.2.3) y documentar todos los servicios. Antes de añadir los servicios nuevos a la memoria tuvimos que corregirla tras una primera revisión.

Al acabar el sprint (Figura A.6) quedaba por acabar las pruebas sobre los tres servicios utilizando Jenkins para acabar así el pipeline de integración continua.

A.4. Sprint 4

El cuarto sprint (Figura A.7) consistió en acabar las pruebas sobre todos los servicios realizados hasta el momento y desarrollar los servicios relacionados con el análisis emocional de una frase, obtener porcentajes y emoción mayoritaria (Sección 4.3).

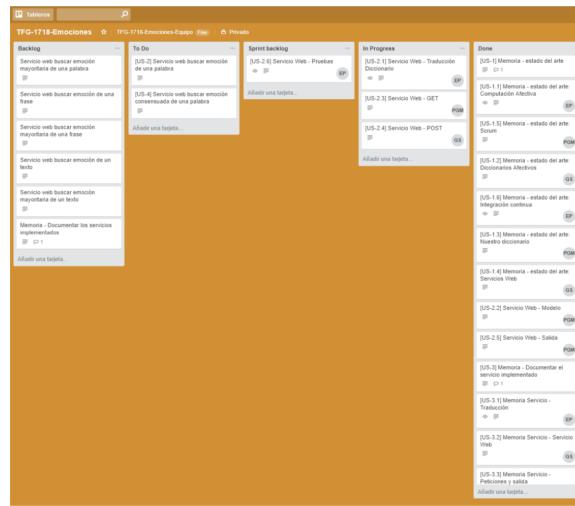


Figura A.3: Planificación del sprint 2 (28/11/2017)

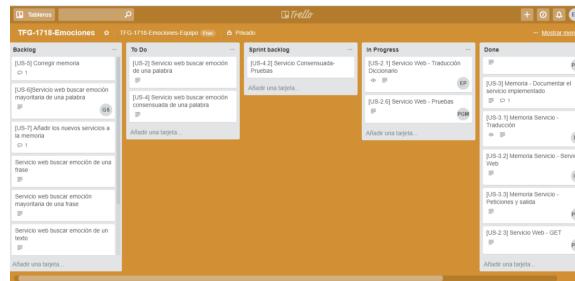


Figura A.4: Final del sprint 2 (19/12/2017)

Al acabar el sprint (Figura A.8) quedaba por desarrollar y probar el servicio que nos proporcionase la emoción mayoritaria de una frase.

A.5. Sprint 5

El quinto sprint (Figura A.9) consistió en terminar el servicio sobre la emoción mayoritaria de un texto, corregir la memoria tras la segunda revisión y diseñar tres prototipos, uno cada una, de la interfaz de la aplicación web a desarrollar (Sección 5.1).

Al acabar el sprint (Figura A.10) quedaba corregir la memoria.

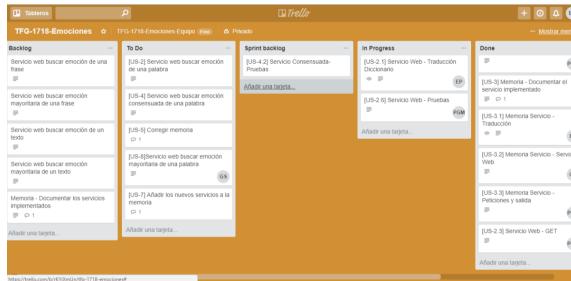


Figura A.5: Planificación del sprint 3 (19/12/2017)

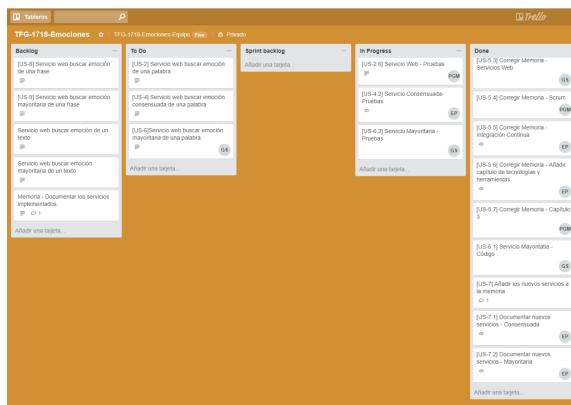


Figura A.6: Final del sprint 3 (09/01/2017)

A.6. Sprint 6

El sexto sprint (Figura A.11) consistió en terminar de corregir la memoria, revisar la codificación de los servicios y hacer ciertas modificaciones (como añadir más peso a los verbos) para depurarlos, buscar una serie de textos para usarlos como casos de prueba y generar el prototipo final para presentarselo a la asociación Asperger Madrid (Sección 5.2).

Al acabar el sprint (Figura A.12) quedaba por analizar los resultados obtenidos al interpretar los casos de prueba.

A.7. Sprint 7

El séptimo sprint (Figura A.13) consistió en realizar el análisis de los casos de prueba y añadirlos a la memoria (Apéndice B), desarrollar la aplicación web y documentar todo (Sección 6) después de haber corregido la memoria tras la tercera revisión.

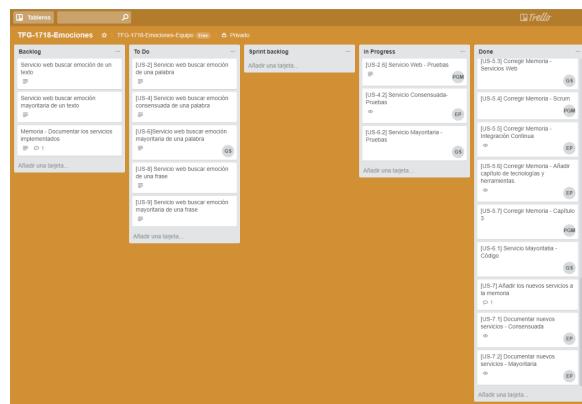


Figura A.7: Planificación del sprint 4 (09/01/2017)

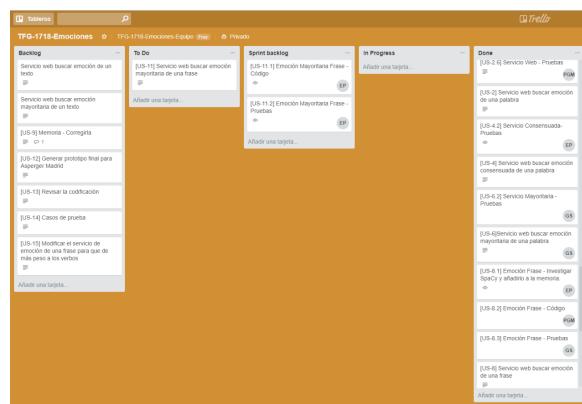


Figura A.8: Final del sprint 4 (21/02/2017)

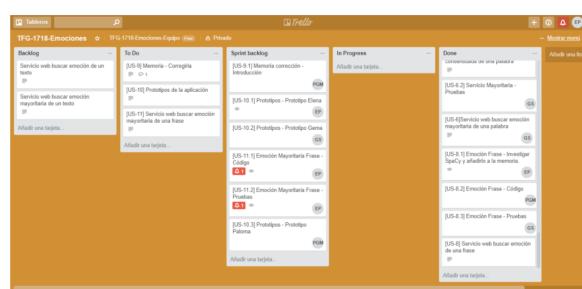


Figura A.9: Planificación del sprint 5 (21/02/2017)

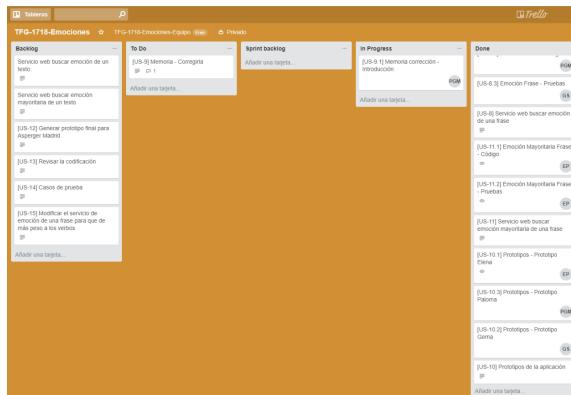


Figura A.10: Final del sprint 5 (07/03/2017)

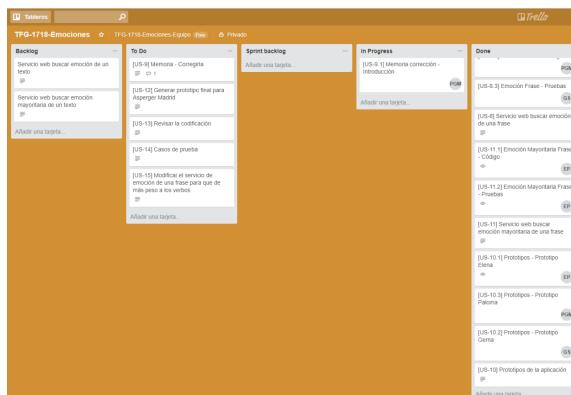


Figura A.11: Planificación del sprint 6 (07/03/2017)

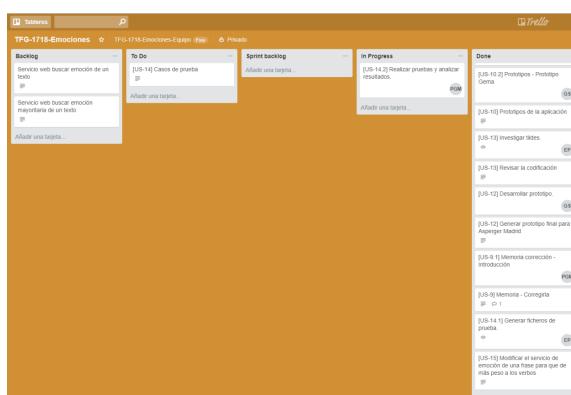


Figura A.12: Final del sprint 6 (21/03/2017)

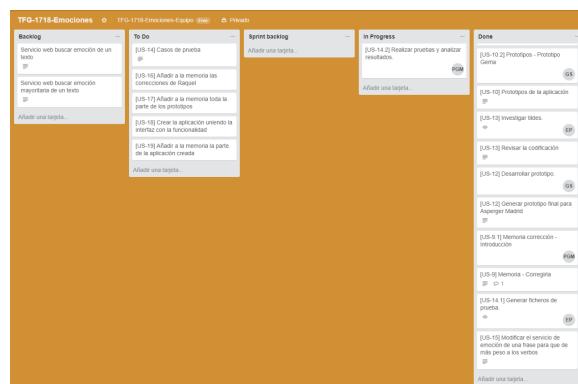


Figura A.13: Planificación del sprint 7 (21/03/2017)

Apéndice B

Casos de prueba

RESUMEN: En este apéndice aparecen los textos que hemos utilizado como casos de prueba para realizar las distintas pruebas sobre el proyecto. Hemos utilizado cuentos, noticias y entradas de blogs.

B.1. Cuentos

Hemos utilizado 12 cuentos cortos.

B.2. Noticias

Hemos utilizado 3 noticias de distintos ámbitos.

B.3. Entradas de Blog

Hemos utilizado 1 entrada como ejemplo.

Bibliografía

*Y así, del mucho leer y del poco dormir,
se le secó el celebro de manera que vino
a perder el juicio.*

Miguel de Cervantes Saavedra

RAE real academia española. 2018. Disponible en <http://www.rae.es> (último acceso, Marzo, 2018).

BALDASARRI, S. Computación afectiva: tecnología y emociones para mejorar la experiencia de usuario. *Revista Institucional de la Facultad de Informática Universidad Nacional de La Plata (Argentina)*, 2016.

BRADLEY, M. M. y LANG, P. J. Affective norms for english words (ANEW): Instruction manual and affective ratings. 1999. Disponible en <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.306.3881&rep=rep1&type=pdf> (último acceso, Marzo, 2018).

COWIE, R. Describing the emotional states expressed in speech. *Speech and Emotion*, 2000.

FOWLER, M. Continuous integration. 2006. Disponible en <https://www.martinfowler.com/articles/continuousIntegration.html> (último acceso, Marzo, 2018).

FRANCISCO, V. *Identificación Automática del Contenido Afectivo de un Texto y su Papel en la Presentación de Información*. Phd, Universidad Complutense de Madrid, 2008.

FRANCISCO, V. y GERVÁS, P. Automated mark up of affective information in english texts. En *Text, Speech and Dialogue (TSD 2006)*, página 375–382. Springer Verlag, Springer Verlag, Brno, Czech Republic, 2006.

GOERTZEL, B., SILVERMAN, K., HARTLEY, C., BUGAJ, S. y ROSS, M. *The Baby Webmind Project*. 2000.

- GREFENSTETTE, G., QU, Y., EVANS, D. A. y SHANAHAN, J. G. Validating the coverage of lexical resources for affect analysis and automatically classifying new words along semantic axes. 2006.
- HATZIVASSILOGLOU, V. y MCKEOWN, K. R. *Predicting the semantic orientation of adjectives*. 1997.
- HUETTNER, A. y SUBASIC, P. Fuzzy typing for document management. 2000.
- DTO. DE CIENCIA DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL, U. D. A. Servicios web. ???? Disponible en <http://www.jtech.ua.es/j2ee/publico/servc-web-2012-13/sesion01-apuntes.html> (último acceso, Marzo, 2018).
- LANG, P. Behavioral treatment and bio-behavioral assessment: Computer applications. 1980.
- LANG, P., BRADLEY, M. y CHUTHBERT, B. International affective picture system (IAPS): Technical manual and affective ratings. 1999.
- LIU, H., LIEBERMAN, H. y SELKER, T. Adaptive linking between text and photos using common sense reasoning. *Proceedings of the 2n International Conference on Adaptive Hypermedia and Adaptive Web Based Systems*, 2002.
- MUELLER, E. T. y DYER, M. G. Towards a computational theory of human daydreaming. *Porceedings of the Seventh Annual Conference of the Cognitive Science Society*, 1985.
- NG, Y., KHONG, C. y THWAITES, H. A review of affective design towards video games. *Procedia-Social and Behavioral Sciences*, vol. 51, 2012.
- PETERSON, R., BREWER, T. y LASSWELL, H. *The Lasswell Value Dictionary*. 1965.
- PICARD, R. W. *Affective Computing*. 1997.
- REDONDO, J., FRAGA, I., PADRÓN, I. y COMESAÑA, M. The spanish adaptation of ANEW (Affective Norms for English Words. *Behavior Research Methods*, 2007. Disponible en <https://link.springer.com/article/10.3758/BF03193031> (último acceso, Marzo, 2018).
- RIEK, L. D., PAUL, P. C. y ROBINSON, P. When my robot similes at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, 2010.

SCRUMMANAGER. *Scrum Manager v. 2.6.* 2016. Disponible en http://scrummanager.net/files/scrum_manager.pdf (último acceso, Marzo, 2018).

TURNEY, P. D. y L. LITTMAN, M. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information System (TOIS)*, 2003.

*-¿Qué te parece desto, Sancho? – Dijo Don Quijote –
Bien podrán los encantadores quitarme la ventura,
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

*-Buena está – dijo Sancho –; fírmela vuestra merced.
–No es menester firmarla – dijo Don Quijote–,
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

