

Doble Grado en Ingeniería Informática y Administración y
Dirección de Empresas
2019-2020

Trabajo Fin de Grado

Desarrollo y uso de un software de
generación de noticias deportivas con
sentimiento en la mejora de procesos en
el ámbito periodístico

Carlos del Rey García

Tutores

Irene Albarrán Lozano

José María Álvarez Rodríguez

Directora externa

Covadonga Gijón

Septiembre 2020, Colmenarejo



Esta obra se encuentra sujeta a la licencia Creative Commons **Reconocimiento - No Comercial - Sin Obra Derivada**

RESUMEN

La generación de texto dejó de ser del dominio exclusivo de los humanos hace años. Hoy en día, existen sistemas de generación de lenguaje natural que escriben resúmenes de documentos; código para construir aplicaciones y textos de todo tipo, incluyendo noticias. Además, cada vez un número mayor de organizaciones son conscientes de los avances que se producen en este campo y adoptan tecnología relacionada para reducir el tiempo que sus trabajadores emplean, recortar sus gastos, etc.

Aquí se presenta un trabajo multidisciplinar, en el entorno de una Cátedra en la que colaboran la Universidad Carlos III de Madrid y la Corporación de Radio y Televisión Española, S. A. En él se exponen el desarrollo de una herramienta de generación de noticias deportivas capaz de redactar el texto en función de la afición a la que vaya dirigida la noticia, y la guía para la integración de la mencionada herramienta en la corporación siguiendo el Ciclo de Mejora de los Procesos de negocio.

Palabras clave: NLG, generación de lenguaje natural, periodismo computacional, noticias, RTVE, fútbol, mejora de procesos, CMP.

ABSTRACT

Text generation ceased to be exclusive human domain years ago. Currently, there exists Natural Language Generation (NLG) system that synthesize abstracts, generate code for building applications, and write other texts, including news. Additionally, more and more organizations are becoming aware of the progress that is being made in the field and are including NLG technology into their structures to reduce expenses and help employees save time.

This document displays a multidisciplinary thesis within the fellowship involving Universidad Carlos III de Madrid and Corporación de Radio y Televisión Española S. A. The presented work depicts the development of an automatic sports news generator that can tailor the text depending on the sports fans the text is intended to. Furthermore, a guide is provided on how to incorporate such tool in RTVE following CMP, a continuous business process improvement methodology.

Keywords: NLG, natural language generation, computational journalism, news, RTVE, football, process improvement, CMP.

En estas líneas quiero aprovechar para agradecer.

Agradecer a mi padre y a mi madre por su paciencia, su guía, y su apoyo económico y personal. Sin ellos no habría llegado hasta aquí.

Agradecer mi hermano por ser, pese a las diferencias, la persona cercana a la que poder acudir sin problema.

Agradecer al resto de mi familia por estar también ahí, porque cuento con casi un ejercito.

Agradecer a mis tutores todo su apoyo, paciencia e implicación en mi proyecto. Han sido, sin dudarlo, una parte esencial.

Agradecer a mi compañera de Cátedra por hacerme llegar la oferta de la que ha salido este trabajo, por confiar en mí para ello.

Agradecer al resto de compañeros de The Reuse Company por facilitarme la vida durante el transcurso de la beca.

Agradecer a RTVE la oportunidad brindada.

Agradecer a todos los que ahora considero amigos. Hayáis aparecido en la etapa de mi vida en la que lo hayáis hecho y sin importar la razón por la que hayamos conectado.

Gracias por acompañarme hasta aquí y espero seguir con vosotros por mucho más.

And last but not even least. I want to thank all of you that are scatered around the world. For so many things. You literary left a mark on me.

Agradeceros a todos por haberme permitido crecer y convertirme en quien soy.

ÍNDICE GENERAL

I	Introducción	1
1.	MOTIVACIÓN Y OBJETIVOS	3
2.	ESTRUCTURA DEL DOCUMENTO	5
II	Desarrollo de la aplicación	6
3.	ESTADO DEL ARTE	7
3.1.	Contextualización	7
3.1.1.	Subproblemas de NLG	8
3.2.	Soluciones actuales en el sector de los medios de comunicación	12
3.2.1.	Tecnología más reciente	16
3.3.	Solución contemporánea similar	16
3.3.1.	Personalized Automated Soccer text System	16
3.3.2.	PASS en múltiples idiomas	18
4.	ANÁLISIS DEL SISTEMA	19
4.1.	Contexto de la herramienta	19
4.2.	Definición del problema	19
4.3.	Descripción de la solución	20
4.4.	Especificación de requisitos	21
4.4.1.	Requisitos funcionales	22
4.4.2.	Requisitos no funcionales	29
4.5.	Casos de uso	32
4.5.1.	Diagrama de casos de uso	32
4.5.2.	Definición de casos de uso	33
5.	DISEÑO DEL SISTEMA.	37
5.1.	Infraestructura del desarrollo	37
5.2.	Arquitectura del sistema.	39
5.2.1.	Arquitectura de alto nivel	39

5.2.2. Vista de desarrollo	40
5.2.3. Vista de procesos	48
5.2.4. Base de conocimiento	48
5.2.5. Proceso de generación de texto	49
6. IMPLEMENTACIÓN	52
6.1. Web y servicios	52
6.2. Generación de texto	56
6.3. Pruebas	56
6.3.1. Evaluación de la generación de textos	59
6.4. Problemas encontrados	59
7. PLANIFICACIÓN Y PRESUPUESTO	61
7.1. Planificación	61
7.2. Presupuesto del proyecto realizado	62
7.2.1. Presupuesto del proyecto a futuro	63
8. MARCO REGULADOR Y ENTORNO SOCIOECONÓMICO	65
8.1. Marco regulador	65
8.2. Entorno socioeconómico	65
III Mejora de procesos	67
9. PUNTO DE PARTIDA	68
9.1. Descripción de la empresa	68
9.2. Análisis del negocio y su entorno	70
9.2.1. Business Model Canvas	70
9.2.2. Análisis PESTEL	79
9.3. Descripción del proceso	81
9.4. Descripción del problema	83
10. MARCO TEÓRICO	86
10.1. Reingeniería de procesos	86
10.1.1. Principios clave	86
10.1.2. Rasgos característicos	87
10.1.3. Herramientas utilizadas	88

10.2. Mejora continua de los procesos	89
10.2.1. Requisitos	89
10.2.2. Ciclo de Deming	89
10.2.3. Seis Sigma.	90
10.2.4. Ciclo de mejora de los procesos de negocios.	92
10.3. Selección del método.	94
11. SOLUCIÓN	95
11.1. Propuesta de aplicación del CMP	95
11.1.1. Sensibilización	95
11.1.2. Identificación	96
11.1.3. Planificación	96
11.1.4. Análisis	97
11.1.5. Diseño	98
11.1.6. Implantación	99
11.1.7. Control.	100
11.1.8. Capacitación	100
11.2. Plan de Marketing hipotético	101
IV Conclusiones	102
Referencias	106

ÍNDICE DE FIGURAS

3.1	Desglose de NLP	7
3.2	Ejemplo de Perspective API	13
3.3	Proceso de Juicer	14
3.4	Publicidad de Heliograf	14
3.5	Ejemplo de Wordsmith	15
3.6	NewsWhip	15
4.1	Diagrama de casos de uso	32
5.1	Arquitectura de alto nivel	39
5.2	Diagrama de componentes	41
5.3	Ejemplos de clústeres dentro de la base de conocimiento.	49
5.4	Ejemplo de patrón	49
5.5	Diagrama de actividades	50
5.6	Patrón sin dependencias	50
5.7	Diagrama de actividades	51
6.1	Captura de la web 1	52
6.2	Captura de la web 2	53
6.3	Captura de la web 3	53
6.4	Captura de la web 4	54
6.5	Captura de la web 5	55
6.6	Captura de la web 5	55
7.1	Diagrama de Gantt	62
9.1	División del Business Model Canvas	71
9.2	Business Model Canvas del área web y de redes sociales de RTVE.	78
9.3	Evolución de la tasa de desempleo en España	79
9.4	Pirámide de población en España	80

9.5	Proceso de redacción	83
10.1	Esquema del ciclo PDCA	90
10.2	Esquema DMAIC para Seis Sigma	91
10.3	Representación del CMP	92
11.1	Mapa del proceso As-Is	98
11.2	Mapa del proceso To-Be	99

ÍNDICE DE TABLAS

4.1	Plantilla de requisito.	21
4.2	Requisito RF-01.	22
4.3	Requisito RF-02.	22
4.4	Requisito RF-03.	23
4.5	Requisito RF-04.	23
4.6	Requisito RF-05.	23
4.7	Requisito RF-06.	24
4.8	Requisito RF-07.	24
4.9	Requisito RF-08.	24
4.10	Requisito RF-09.	25
4.11	Requisito RF-10.	25
4.12	Requisito RF-11.	26
4.13	Requisito RF-12.	26
4.14	Requisito RF-13.	27
4.15	Requisito RF-14.	27
4.16	Requisito RF-15.	28
4.17	Requisito RF-16.	28
4.18	Requisito RNF-01.	29
4.19	Requisito RNF-02.	29
4.20	Requisito RNF-03.	30
4.21	Requisito RNF-04.	30
4.22	Requisito RNF-05.	30
4.23	Requisito RNF-06.	31
4.24	Requisito RNF-07.	31
4.25	Requisito RNF-08.	31
4.26	Plantilla de caso de uso.	33
4.27	Caso de uso CU-01.	35
4.28	Caso de uso CU-02.	36

5.1	Plantilla de componente	42
5.2	Componente Interfaz	43
5.3	Componente NewsController	44
5.4	Componente VerbaliserController	45
5.5	Componente News	46
5.6	Componente Verbaliser	47
5.7	Componente Cake	47
6.1	Plantilla de prueba.	56
6.2	Prueba P-01.	57
6.3	Prueba P-02.	57
6.4	Prueba P-03.	58
7.1	Costes de personal	62
7.2	Costes de material	63
7.3	Costes totales	63
7.4	Costes de personal del supuesto	64
7.5	Costes de material del supuesto	64
7.6	Costes totales del supuesto	64

GLOSARIO

Aquí se explicarán términos utilizados a lo largo del trabajo que puedan dar lugar a confusión o que necesiten clarificación.

Términos

- **Árbol.** En lo referido a estructuras de datos, este tipo de representación se caracteriza por contar con un nodo raíz del que penden el resto de nodos enlazados.
- **Base de conocimiento.** Tipo especial de base de datos en la que se gestiona el conocimiento para una aplicación.
- **Corpus.** Recopilación de escritos.
- **Influencer.** Persona que destaca en una red social y otro canal de comunicación.
- **Leasing.** También conocido como arrendamiento financiero. Consiste en un contrato de alquiler que incorpora una opción de compra a favor del arrendatario.
- **Redes de neuronas recurrentes.** Son aquellas en las que se incluyen bucles de realimentación. Es decir, permiten que una neurona transmita información a otra en distintas iteraciones.
- **Stakeholder.** Hace referencia a los grupos de interés en una empresa o transacción.
- **Startup.** Empresa emergente, normalmente sobre una base tecnológica innovadora y con gran capacidad de crecimiento.
- **Treebank.** También conocido como ‘Corpus parseado’. Es un corpus en el que a cada frase se le ha incluido su estructura sintáctica.

Acrónimos

- **NLP.** *Natural Language Processing.*
- **NLU.** *Natural Language Understanding.*
- **NLG.** *Natural Language Generation.*
- **RTVE.** RadioTelevisión Española.
- **API.** *Application Programming Interface.*
- **EUFA.** *Union des Associations Européennes de Football,*

- **FIFA.** *Fédération Internationale de Football Association.*
- **REST.** *Representational State Transfer.*
- **CMP.** Ciclo de Mejora de los Procesos de negocios.

Parte I

Introducción

La realidad en la que vivimos es innegablemente digital. Cada vez más y de formas más diversas. La necesidad o comodidad de poder realizar tareas en remoto, contactar con personas cuándo y cómo se necesite, no tener que almacenar a mano y de manera física cualquier tipo de registro (contable, de pacientes, de programación, etc.) son parte del presente que una vez fue considerado como futuro. Estas y otras tareas como responder de manera automática a clientes que intentan contactar con una empresa o persona cuando estos no están disponibles, que un vehículo (cargado con mercancía o transportando personas) se desplace hasta cualquier determinado punto sin necesidad de conducción o supervisión humana son, entre otros, aspectos que son actualmente cubiertos por la tecnología.

Dentro de la digitalización en la que lleva inmersa el mundo los últimos años, la inteligencia artificial está jugando un papel muy importante. Cada vez más se buscan maneras de utilizar esta inteligencia para automatizar procesos que serían tediosos y largos si los tuviera que realizar una persona como buscar patrones en un conjunto de datos que una persona podría incluso no encontrar, o conseguir que componentes software puedan entender lenguajes propios de personas y comunicarse con humanos.

En medio de esta situación, todas las empresas que quieran prosperar o mantenerse deberán adaptarse, y los medios de comunicación no son una excepción. Desde vigilar qué noticias tienen el foco de atención en cada momento; incluyendo resumir textos o identificar de qué tema tratan; hasta detectar noticias falsas o generar textos; y crear contenidos listos para publicar e, incluso, publicarlos directamente. Estos son algunos ejemplos que ilustran lo mucho que puede hacer la tecnología y la inteligencia artificial por los medios de comunicación.

Una de las empresas de medios de comunicación más importante de España es la Corporación de Radio y Televisión Española, S.A., más conocida como RTVE (en adelante será referida como tal). Es la sociedad mercantil dedicada a dar servicio público de radio y televisión en toda España [1]. Cuenta con siete canales de televisión activos, seis cadenas de radio, su web, el Instituto RTVE, un coro y una orquesta. Con semejante despliegue de medios, se mantiene como uno de los grandes dentro de su sector en España, lo que no le exime de un margen de mejora, actualización y crecimiento.

Ante la necesidad de adaptación tecnológica, por parte de las empresas, mencionada anteriormente y como parte de una colaboración en forma de cátedra entre la Universidad Carlos III de Madrid y RTVE, en este trabajo se plantea una solución: el desarrollo de un sistema de generación automática de noticias con sentimiento para eventos deportivos. Más adelante se aportarán detalles explicando dicho sistema y también se mostrará cómo RTVE podría incluirlo dentro de su estructura actual para mejorar su proceso de generación de noticias y contenido.

1. MOTIVACIÓN Y OBJETIVOS

A lo largo de los últimos años el sector audiovisual y de los medios de comunicación ha sufrido, y sigue sufriendo, una continua renovación gracias al avance continuo de la tecnología; en particular, la informática y la inteligencia artificial. Esto ha permitido que las empresas se adapten cada vez más a sus usuarios y que estos demanden cada vez más un trato más individual. Desde contenido específico por localización geográfica, pasando por contenido promovido por el interés en otro distinto, hasta contenido basado en información demográfica. Toda esta información juega un papel relevante e innegable a la hora de generar y ofrecer contenido al público.

Dentro del panorama español en este sector, RTVE es una de las potencias más importantes y es un ente público que sirve a toda la comunidad española (incluso a otros colectivos de distintas comunidades hispanohablantes). Como tal debería adecuarse a los avances tecnológicos que sucedan y brindar a todo su público con una experiencia atractiva y digna. Además, sus trabajadores deberían poder contar con tecnología capaz de ayudarles en su rutina laboral, no con aquella que les entorpezca (tecnología obsoleta) y mucho menos que les supla. Dada la categoría pública de la organización, es deseable que el aprovechamiento de los recursos sea eficiente y que la calidad de los productos o servicios que se consigan sea la mejor posible.

Al tener RTVE una labor nacional, toda la extensión del territorio español debería estar cubierta por sus servicios, pero esto puede tornarse algo complicado por la falta de personal o presupuesto, o por que el personal actual ya esté ocupado con dar parte de los eventos que más prioridad tengan en el momento. Sectores como el deporte y, más concretamente, el fútbol tienen en España un interés y una cobertura más que notable. No obstante, la cantidad de partidos de las diferentes ligas; de información durante la pretemporada y la temporada; el mercado de fichajes de invierno y verano; las competiciones nacionales como la Copa del Rey, la Supercopa de España; e internacionales como la liga de Campeones de la UEFA (más conocida como “La Champions”) y la liga de Europa de la UEFA (más conocida como “La EUFA”), el Campeonato Europeo de la UEFA (más conocido como “La Eurocopa”), la Supercopa de la UEFA, la Copa Mundial de la FIFA, la Copa Mundial de Clubes de la FIFA, la Copa Intercontinental de Clubes Campeones; los partidos amistosos y benéficos son una cantidad muy alta de información a procesar. En respuesta a esta demanda constante de información y el volumen de esta que se maneja en los últimos años, se ha desarrollado tecnología que también permite tener datos sobre todos los eventos mencionados anteriormente de forma rápida, sino instantánea; y, muchas veces, pública.

A día de hoy, los periódicos y portales oficiales de los medios de comunicación siguen siendo un medio útil, porque permiten a los que publican y a las empresas utilizar el formato que ellos elijan en las publicaciones y también pueden conseguir monetizar

directamente todo el contenido que hagan público. Sin embargo, si realmente se quiere tener un gran impacto y llegar a una gran cantidad de público, no se pueden olvidar las redes sociales. Ya sea por la conveniencia del formato o porque permiten a sus usuarios compartir con sus contactos, la creación de contenido en las redes sociales es un factor a tener en cuenta dentro del sector y de las actividades a las que se dedica RTVE (que cuenta con perfiles en Instagram, Facebook y Twitter).

Utilizando las técnicas expuestas más adelante, en el apartado 3 de este documento, junto con algunas de las herramientas tecnológicas disponibles actualmente, se podría construir una herramienta que supusiera una mejora en el día a día de los periodistas de RTVE y para la empresa en general. Todo ello sin que para estos supusiera ningún problema o enfrentamiento, o sin que esta herramienta apareciera como una fuente de miedo e incertidumbre para los trabajadores.

Para aunar todo lo mencionado y como parte de la Cátedra UC3M-RTVE, en este documento se explicará el desarrollo de una herramienta de generación de noticias deportivas con un sentimiento añadido con el que pueda identificarse el lector y se propondrá a su vez una mejora del proceso productivo actual de generación de contenido en RTVE. Esta herramienta se centrará en noticias sobre partidos de fútbol de la Segunda División B española, concretamente en los equipos originarios de la región de Castilla la Mancha. Esto viene motivado por la escasez de cobertura de estos partidos que, siendo relevantes para parte de la población, serían costosos de cubrir enteramente por periodistas con otros deberes, pero podrían ser cubiertos de forma sencilla con ayuda de la herramienta que se propondrá más adelante. Además, el sentimiento añadido en la generación permitirá dirigir el texto según la audiencia, en este caso, según las aficiones involucradas en los partidos (después de una derrota el sentimiento será distinto que después de una victoria). De esta forma, se conseguirán noticias más personalizadas y con potencialmente mayor interés y ‘gancho’ para los lectores.

2. ESTRUCTURA DEL DOCUMENTO

La estructura de este documento, de ámbito mixto, constará de partes relacionadas con la Ingeniería Informática; con la Administración y Dirección de Empresas; y, también, albergará partes comunes debido a la naturaleza multidisciplinar del mismo.

Y será la siguiente:

- **Desarrollo de la aplicación.** En esta parte del documento se exponen los aspectos del trabajo relacionado con la Ingeniería Informática.
 - **Estado del arte.** Dentro de este apartado se podrá encontrar una contextualización tecnológica y académica de la herramienta, así como ejemplos de otras soluciones similares y utilizadas en el sector respectivamente.
 - **Análisis del sistema.** Las funcionalidades que deberá cubrir el sistemas serán identificadas en este apartado.
 - **Diseño del sistema.** Aquí se explicará cómo se llegará a cubrir las funcionalidades ya mencionadas.
 - **Implementación.** En este apartado se relatará cómo ha sido la puesta en marcha de la herramienta.
 - **Presupuesto y planificación.** Podrá observarse aquí una planificación del desarrollo de la herramienta junto con el presupuesto de su desarrollo.
 - **Marco regulador.** Como último apartado de la parte del desarrollo del sistema, aquí se describirá el entorno legal y económico que rodea a la herramienta.
- **Mejora de procesos.** En esta segunda parte del documento se centra en los aspectos del trabajo relacionados con la Administración y Dirección de Empresas.
 - **Punto de partida.** Aquí se profundizará en la empresa sobre la que se va a realizar la mejora de procesos.
 - **Marco teórico.** En este apartado se expondrá la teoría disponible sobre la mejora de procesos.
 - **Solución.** Finalmente, se concluirá la parte de administración y dirección de empresas con la descripción de una solución. También se aportará un breve plan de marketing hipotético.
- **Conclusiones.** Como cierre del documento, en esta parte se aportarán conclusiones, detalles en posibles trabajos futuros y limitaciones encontradas en la realización de este trabajo.

Parte II

Desarrollo de la aplicación

3. ESTADO DEL ARTE

Para revisar el estado del arte de la tecnología que se describirá durante el desarrollo de este documento, primero se procederá a contextualizar el problema tecnológica y académicamente dentro de la informática. Después, se presentarán herramientas de software utilizadas por empresas importantes en el sector de los medios de comunicación, así como los últimos avances en el campo. Por último, se describirá una solución similar a la propuesta para el problema planteado.

3.1. Contextualización

Cabe destacar de antemano que para el desarrollo de los siguientes párrafos se utilizará como guión parte el excelente trabajo realizado por Albert Gatt ¹ y Emiel Krahmel ² en [2].

Uno de los campos dentro de la informática en la que más se ha avanzado en los últimos años y que más despunta en la actualidad es la inteligencia artificial. Este se centra principalmente en la creación y desarrollo de algoritmos, programas y sistemas que muestren o sigan comportamientos inteligentes. Hoy en día, la inteligencia artificial se aplica en diferentes sectores y con diferentes finalidades, pero una de las grandes ramas dentro de este campo que está muy relacionada con la lingüística computacional [3] es el procesamiento de lenguaje natural (*NLP: natural language processing* en inglés). Como su propio nombre indica, este campo tiene como foco el lenguaje que utilizan los humanos para comunicarse y cómo conseguir que algoritmos, sistemas y aplicaciones puedan “entender”, procesar y utilizar dicho lenguaje.

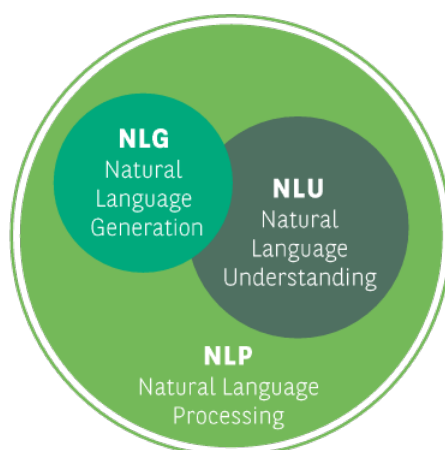


Fig. 3.1. Situación de NLU y NLG respecto a NLP [4]

¹<https://staff.um.edu.mt/albert.gatt/>

²<https://research.tilburguniversity.edu/en/persons/emiel-krahmer>

Dentro del procesamiento de lenguaje natural (obsérvese figura 3.1), se podrían dividir gran parte de los esfuerzos actuales en dos ramas: la comprensión de lenguaje natural (*NLU: natural language understanding* en inglés) y la generación de lenguaje natural (*NLG: natural language generation* en inglés) [5]. La primera orbita en torno a cómo los sistemas y aplicaciones actuales pueden obtener e interpretar el significado del lenguaje natural, mientras que la segunda tiene como fin la construcción de estructuras en lenguaje natural a partir de una información que haya sido obtenida o almacenada. Como una traducción de una definición más formal de la segunda rama, dado que este trabajo versará sobre ella: la generación de lenguaje natural se conoce como “el subcampo de la inteligencia artificial y la lingüística computacional encargado de la construcción de sistemas de ordenadores que puedan producir textos comprensibles en Inglés o cualquier otro idioma humano a partir de algún tipo de representación no lingüística relacionada de la información” [6, p. 1] (citado en [2, p. 6]). Téngase en cuenta que esta definición no se ajusta a toda la generación de lenguaje natural como podrá comprobarse a continuación, a pesar de ser una definición presente en uno de los estudios sobre métodos de generación del lenguaje natural más reconocidos [6].

Dependiendo de qué formato tenga la información que se transforma en lenguaje, se puede considerar generación de texto a texto (*text-to-text generation* en inglés) si se pretende generar un texto nuevo a partir de uno o fragmentos de uno existente, o generación de datos a texto (*data-to-text generation* en inglés) si la generación se hace a partir de datos sin aparente estructura lingüística [2]. Algunas aplicaciones del primer tipo de generación son: traducción automática [7]; fusión y resumen de oraciones relacionadas o textos [8]; simplificación de textos para hacerlos más accesibles a distintos tipos de público [9]; corrección automática de escritura, gramática y textos [10]; generación automática de revisiones realizadas por gente del medio [11]; generación de párrafos a partir de frases [12]; y generación automática de preguntas [13]. Como ejemplos del segundo tipo de generación se pueden observar: noticias de fútbol (como [14] y [15]) ; periódicos virtuales sobre datos extraídos por sensores [16] y noticias sobre polémicas actuales [17]; textos sobre preocupaciones medioambientales [18]; informes financieros y meteorológicos [19]; resúmenes de la información de un paciente en una clínica [20]; información interactiva sobre artefactos culturales [21]; y escritos que pretenden persuadir [22] o motivar un cambio de comportamiento [23]. Todos los ejemplos mencionados anteriormente y algunos más aparecen en [2, pp. 4-5].

3.1.1. Subproblemas de NLG

A la hora de generar textos en lenguaje natural, hay una división en seis subproblemas muy popular, pudiéndose encontrar alguno de ellos en prácticamente cualquier sistema de generación de lenguaje natural. Se describen a continuación los seis subproblemas de esta división, que se pueden profundizar en [6] (mencionado en [2, p. 9]).

Subproblema 1: Determinación de contenido

Del inglés *content determination*. En este subproblema se intenta dar respuesta a qué información se debe incluir en el texto que se quiere generar [2]. Normalmente hay más información de la que se necesita o quiere en la estructura previa al texto que se va a generar. Ya sea porque no se quiere ser redundante, porque hay hechos no relevantes, porque se tenga que ajustar la profundidad del texto y el detalle al conocimiento del lector, o porque la intención del texto sea una u otra, este suele ser uno de los primeros pasos a llevar a cabo a la hora de construir un sistema o aplicación de generación del lenguaje natural. Es muy habitual que la información se termine refinando como entidades abstractas en las que almacenar de manera más concisa y eficaz la información que se quiere utilizar posteriormente.

A pesar de lo básico o fundamental que pueda parecer para la generación de lenguaje natural, los distintos intentos de solucionar este subproblema suelen estar ligados al dominio en el que se localiza la aplicación. No obstante, recientemente se han probado muchas técnicas en las que predomina un enfoque basado en datos [24] y también se ha probado la utilización de clusters para redactar biografías [25].

Subproblema 2: Estructuración del texto

Del inglés *text structuring*. Posteriormente, se determina cuál es el orden que debe llevar la información que se va a incluir en el texto que se quiere generar [2]. Según el tipo de tema sobre el que verse el texto a generar y también el tipo de público que vaya a “consumir” ese documento, podría darse el caso de que se quisiera primero dar una introducción o contexto previo a la información relevante. O, también, saltarse cualquier tipo de información (o incluso palabra) que no sea necesaria y plasmar directamente los hechos. Por ejemplo, a la hora de generar un resumen de un partido de fútbol, tendría sentido empezar con contexto del partido e información de los dos equipos, para después dar paso a hechos concretos que hubieran sucedido durante el encuentro.

Estas y más restricciones se solventan, también, dependiendo del dominio de la aplicación. No obstante, de igual manera que en el anterior subproblema (y en el terreno de la informática en general), en este ámbito se han incluido técnicas de aprendizaje automático en los últimos años, incluso resolviendo el problema junto a la determinación de contenido [25].

Subproblema 3: Agregado de oraciones

Del inglés *sentence aggregation*. Esta es el área en la que se elige qué información debe ir en cada oración individual [2]. Distintas piezas independientes de información podrían tener cabida en la misma oración o podría darse el caso en el que tuvieran que aparecer en oraciones separadas. De este modo se pueden eliminar posibles redundancias

en el texto a generar y, también, se pueden combinar estructuras lingüísticas para dar lugar a formas más naturales, más humanas, de expresar la misma información. Además, se pueden distinguir agregado a nivel semántico (por significado) y a nivel de sintagma (por porción de información dentro de la frase).

De nuevo se observa que, en el trabajo temprano, la aproximación al problema era muy dependiente del dominio, con reglas prácticamente ‘artesanales’ para cada uno de ellos [26], y que los enfoques más recientes pivotan hacia técnicas basadas en datos [27].

Subproblema 4: Lexicalización

Del inglés *lexicalization*. A continuación, se escogen cuáles son las palabras y expresiones que mejor se ajustan a la información que se quiere expresar [2]. Y, con ello, entra en escena un inconveniente muy común que es que, para expresar un mismo suceso, cada idioma tiene un número (mayor que uno) de formas de describirlo sin variar sustancialmente su significado. Dicho inconveniente se puede acentuar dependiendo del grado de variabilidad que se quiera en los textos a producir. Cuanto menor repetición se persiga en el texto final (y con ello más humano se quiera que parezca ese texto), más atención deberá ponerse al mencionado inconveniente. También hay que tener en consideración que los mismos humanos muestran variabilidad entre ellos y a veces en un mismo individuo a la hora de expresar un mensaje. Esto, sumado a las connotaciones que pueda acarrear cierta expresión en un contexto determinado, hacen que este subproblema sea uno de los más complicados de solucionar de forma automática.

Subproblema 5: Generación de expresiones referenciales

Del inglés *referring expressions generation*. Por último, se determinan qué palabras y expresiones son las que mejor identifican los objetos de un dominio determinado [2]. Aunque pueda parecer similar al subproblema anterior, este se centra en determinar cuando es necesario distinguir una entidad de otras similares o cómo referirse a un sujeto cuando este ha sido mencionado previamente, entre otras. Uno de los puntos a decidir cuando se soluciona este subproblema es en qué forma se referencia al sujeto que se quiere mencionar (usando una descripción de este, un pronombre o un nombre propio). En el primero de estos casos se debe también lidiar con qué cantidad de información es la ajustada o necesaria para que la descripción no sea demasiado detallada, pero que contenga la suficiente información para que tampoco que se vuelva difícil distinguir al sujeto al que se quiere referir.

Como en otros subproblemas que se han comentado anteriormente, las formas de resolver la generación de expresiones referenciales también están ligadas al dominio de la aplicación o sistema que se esté desarrollando.

Subproblema 6: Realización lingüística

Del inglés *linguistic realisation*, en este subproblema se intenta dar respuesta a cómo deben combinarse las palabras y expresiones seleccionadas previamente, dando lugar a oraciones correctamente formadas y con sentido lingüístico [2]. Como partes de este apartado se pueden distinguir la ordenación de sintagmas dentro de cada oración; la selección de palabras y las formas correctas que respeten las concordancias entre palabras (en castellano: género, número y persona); la inclusión de preposiciones o auxiliares para dar lugar a estructuras lingüísticas completas y correctas; etc.

Entre la variedad de soluciones propuestas a este subproblema, se describirán a continuación tres de ellas: plantillas elaboradas por personas, sistemas basados en gramáticas elaboradas por personas y soluciones estadísticas.

Plantillas elaboradas por personas. Este tipo de soluciones suelen utilizarse cuando el dominio sobre el que trabaja la aplicación o sistema que se va a desarrollar es relativamente pequeño y la realización relativamente sencilla. Su utilización permite una calidad de los textos muy controlada y evitar la generación de estructuras que no son correctas desde un punto de vista gramatical [2]. Algunas variantes contemporáneas de este tipo de soluciones incluyen, además, información sintáctica en las plantillas como en [14].

Como viene indicado en su nombre, estas plantillas son generadas por personas y eso requiere un tiempo preciado dedicado únicamente a su definición. Por suerte, algunas soluciones más recientes han empezado a adoptar el aprendizaje automático para la generación de las plantillas [28].

Los sistemas basados en este tipo de propuesta (la utilización de plantillas), aunque a veces son tachados de no pertenecer estrictamente al campo de la generación de lenguaje natural por la forma en la que generan el texto, suelen producir textos que se consideran de alta calidad [29] generados de manera rápida [30] (como se comenta en [31]).

Sistemas basados en gramáticas elaboradas por personas. La alternativa a las plantillas que permite obtener una solución más independiente del dominio y de propósito más general son las gramáticas. Mediante su uso, varias o todas las decisiones que se toman durante el proceso de generación de lenguaje natural se basan en la gramática del lenguaje en el que se está generando. Este tipo de soluciones tienden a ser muy exigentes en cuanto al detalle de la entrada que reciben. Estas exigencias han promovido el desarrollo de motores o módulos que proporcionan funciones en los campos de la sintaxis y la morfología, pero que necesitan de desarrolladores para la toma de decisiones (por ejemplo, [32]).

Uno de los principales inconvenientes de este tipo de soluciones es la selección entre opciones relacionadas en las que la sensibilidad a la entrada y al contexto son difíciles de

determinar [6].

Soluciones estadísticas. Al contrario que las dos opciones anteriores, que suelen requerir de una labor considerable por parte de las personas, las propuestas más recientes se centran en la utilización de información estadística para resolver este problema. Para ello, se pueden observar dos principales enfoques [2]. El primero parte de alternativas generadas en forma de bosque por una sencilla gramática generada “a mano” entre las que un clasificador selecciona la mejor opción. La otra principal alternativa utiliza información estadística a la hora de tomar decisiones durante composición del texto, no únicamente al final del proceso para elegir el mejor candidato generado. Ambas alternativas comparten que la información estadística se utiliza para filtrar entre las posibles salidas (a un nivel u a otro), mientras que utilizan un generador base desarrollado por personas.

Otras alternativas, aún más recientes, han adoptado un enfoque completo basado en datos, en las cuales las reglas gramaticales son adquiridas en bancos de árboles (del inglés *treebanks*) en los que se almacenan frases con información sintáctica en forma de árboles [33].

3.2. Soluciones actuales en el sector de los medios de comunicación

Una vez descrito el contexto tecnológico y académico en el que se sitúa este trabajo, se va a dar paso a comentar las soluciones tecnológicas viables que se han encontrado en el sector de los medios de comunicación en la actualidad.

El periodismo computacional es una realidad a día de hoy, y muchos de los diarios, periódicos y agencias de información más relevantes del mundo ya son conscientes de ello. Nombres tan sonados como Forbes, el Washington Post o la agencia de noticias EFE son solo algunos de los muchos que ya utilizan y se apoyan en la tecnología para realizar varias de sus labores y funciones.

Una gran mayoría de los medios con considerable poder mediático y económico han optado por un desarrollo interno de esta solución y, por lo tanto, no se procederá a analizar su tecnología ni tipo de solución en mucha profundidad ya que no son accesibles al público y aportaría información a los competidores. Otros han optado por obtener este tipo de soluciones de empresas especializadas, de las que tampoco se tiene demasiada información sobre el desarrollo y el funcionamiento del software por su poca accesibilidad y transparencia motivadas por el poder que supone la información en su sector. No obstante, se considera relevante mencionar varios grandes gigantes del sector en los que se ha probado que este tipo de software puede ser de mucha utilidad. A continuación se presenta una lista con algunos ejemplos [34]:

1. **The New York Times.** En el conocido periódico neoyorquino llevan utilizando desde 2015 a *Editor*. Esta herramienta toma la forma de un editor de texto en el que el

periodista colabora con inteligencia artificial en forma de redes de neuronas recurrentes. Estas redes de neuronas desarrolladas a partir de código de Google [35], y alojadas como micro servicios, reciben como entrada las palabras que el periodista escribe y cuando detectan palabras a las que atribuir una etiqueta (fechas, nombres, lugares, etc.), estas aparecen de tal modo en el editor. Y no solo eso, también permite al periodista destacar partes que este considere interesantes y atribuirles etiquetas.

También utiliza software desarrollado por Jigsaw (parte de la empresa paterna de Google, Alphabet) llamado *Perspective API* [36] que permite identificar mediante aprendizaje automático qué tipo de impacto tendrá un comentario sobre un tema (cabe destacar que el diario español El País también usa este software). Un ejemplo que usan para ilustrarlo es ofrecer a los lectores comentarios sobre distintos temas ordenados por toxicidad (véase figura 3.2).

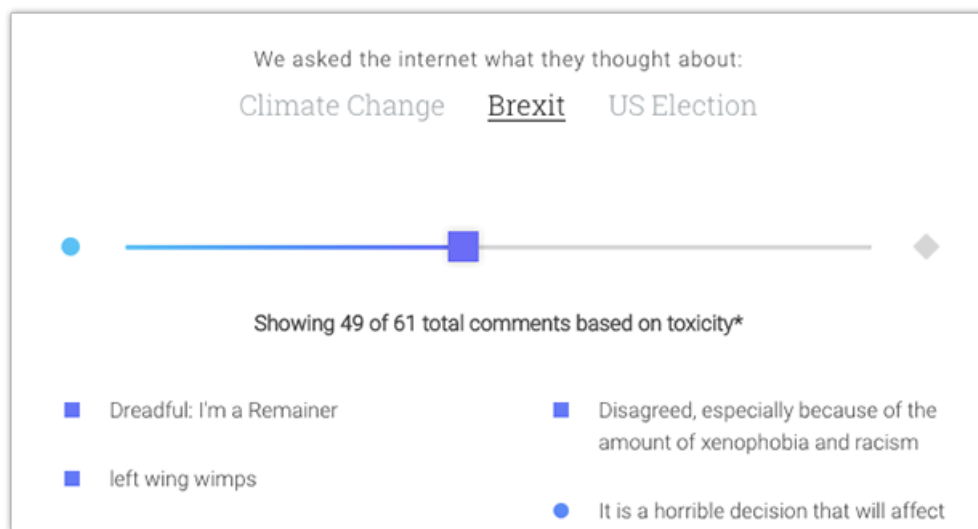


Fig. 3.2. Ejemplo información generada por Perspective API [37].

2. **BBC.** El gigante mediático británico también utiliza inteligencia artificial en un software llamado *Juicer* [38]. Esta API permite a los editores (y también a gente no relacionada con el periódico, previa autorización) buscar artículos por personas, lugares, organizaciones o cosas. Al estar conectado a un gran número de fuentes de artículos puede, según estas hayan publicado uno nuevo, procesarlo y asociarle las etiquetas por las que posteriormente puedan ser buscados dichos artículos (véase figura 3.3).
3. **The Washington Post.** El periódico más antiguo de la capital estadounidense utiliza también un software propio llamado *Heliograf* [39] (véase figura 3.4). El uso de esta herramienta permite a los redactores tener textos listos con información relevante para la noticia antes que nadie y cuyo contenido se especializa por geolocalización (lectores en distintos lugares recibirán, posiblemente, noticias ligeramente distintas), y también les permite modificar dichos textos a placer según lo que se



Fig. 3.3. Pasos que realiza el software Juicer [38].

haya generado o las necesidades que necesite cumplir el redactor. El primer gran debut de esta tecnología fue en el año 2016 para cubrir las olimpiadas de Río de Janeiro, Brasil.

Este diario cuenta con más herramientas como *Heliograf* disponibles en la plataforma *Arc Publishing*, que fue desarrollada y diseñada por personal del periódico para ajustarse a sus necesidades.

The advertisement for Heliograf features a blue background. On the left, the text reads: 'Heliograf' in large white font, followed by 'An Intelligent, Automated Storytelling Agent'. Below this, three bullet points describe its capabilities: 'Automatic story generation from real-time data sources', 'Delivery of channel-specific stories', and 'Personalized stories for readers'. On the right side, there is a colorful, retro-style robot with a red head and blue body. In the bottom right corner, the logo for 'The Washington Post' is visible.

Fig. 3.4. Publicidad de Heliograf [40].

4. **Yahoo!** Esta empresa de servicios por Internet más que conocida también se quiso sumar a la tendencia de la inteligencia artificial. En su sección de información deportiva, Yahoo! Sports, la compañía utiliza un software de generación y personalización de contenido para redactar noticias de distintos deportes (incluso de equipos inventados por los usuarios) para los usuarios [41]. Además, es capaz de sacar provecho de esta personalización de contenidos también a la hora de mostrar anuncios, ya que estos se ajustan más a los lectores.

Este software es propiedad de la empresa Automated Insights y está fundamentado en su herramienta *Wordsmith* (véase figura 3.5).

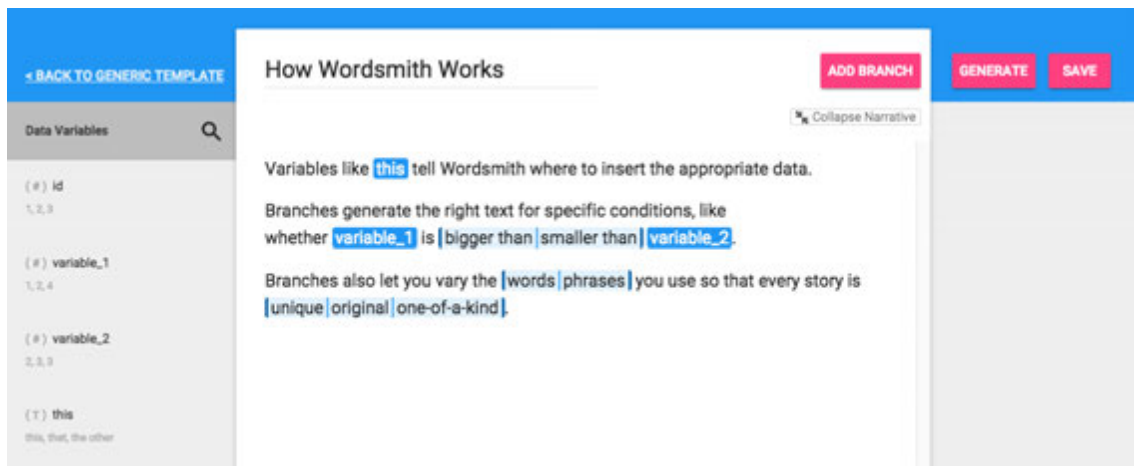


Fig. 3.5. Ejemplo información generada por Perspective API [42].

5. **Associated Press.** La cooperativa estadounidense de medios de comunicación utiliza también un software de otra compañía especializada, NewsWhip. Con ello consigue tener constancia de cómo lo están haciendo sus competidores a través de las redes sociales; cómo responde la audiencia en torno a palabras clave; y cómo están afectando *influencers* al rendimiento de ciertas marcas (véase figura 3.6).

Además de obtener tal cantidad de información, utilizan también una herramienta basada en *Wordsmith* (de Automated Insights) para redactar contenido de forma automática.

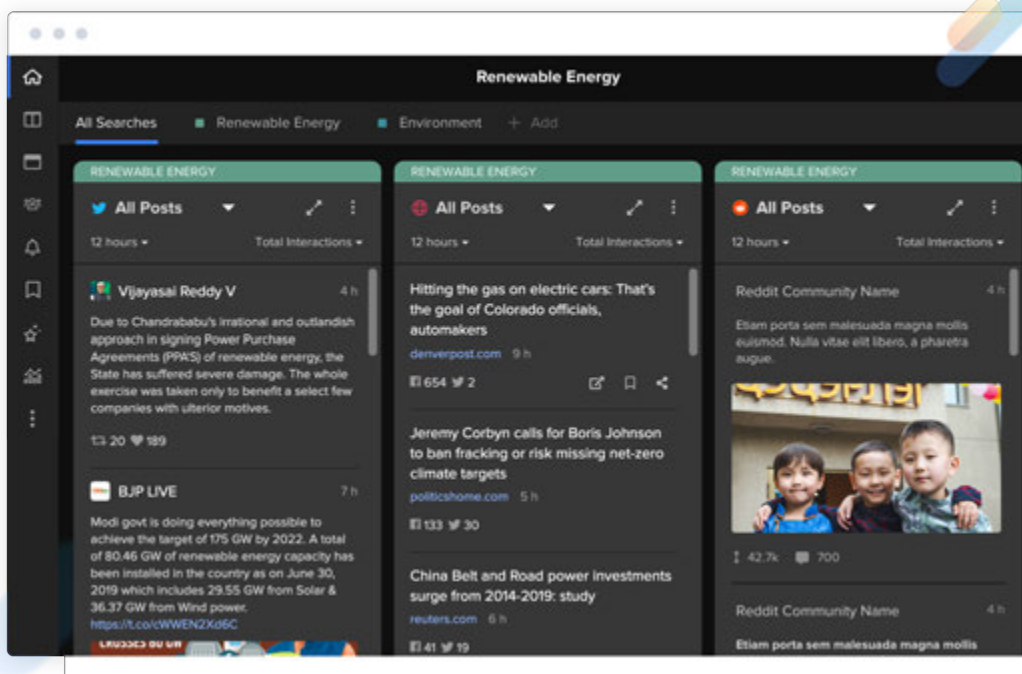


Fig. 3.6. Ejemplo de información generada por una herramienta de NewsWhip [43].

3.2.1. Tecnología más reciente

En este apartado se comentarán dos de los avances más recientes y más sonados recientes en el campo de NLG. Esto muestra la permanente innovación en este campo y la atención que ha atraído en los últimos años.

Turing-NLG

Desarrollado por Microsoft, Turing-NLG [44] es una herramienta de generación de lenguaje natural que sentó un antes y un después en este campo. Su tecnología se basa en un modelo de lenguaje que cuenta con diecisiete mil millones de parámetros. Esto le ha permitido adelantarse a todos sus competidores escribiendo, incluso, el blog sobre su propio éxito. La herramienta se utiliza para aplicaciones de pregunta y respuesta (pudiendo dar respuesta, incluso, a preguntas en las que no se ha aportado contexto), y generación de resúmenes como lo haría una persona y no cogiendo fragmentos del texto como hacen otros competidores.

GPT-3

En la misma línea que el ejemplo de Microsoft, *GPT-3* [45] es otro modelo de lenguaje desarrollado por OpenAI, una compañía de investigación sobre inteligencia artificial sin ánimo de lucro. Este modelo del lenguaje, de posterior publicación que el de su competidor Microsoft, fue entrenado con la asombrosa cifra de ciento setenta y cinco mil millones de parámetros, estableciéndose como la más avanzada en este campo. Su uso permite crear herramientas que generan aplicaciones a partir de una descripción de una funcionalidad; construyen elementos web también a partir de una descripción y aplicaciones que mantienen conversaciones con los usuarios.

3.3. Solución contemporánea similar

Teniendo en cuenta el contexto académico del problema y el contexto tecnológico, y las tendencias en el sector, se describirá a continuación una solución muy similar de generación de texto para eventos deportivos.

3.3.1. Personalized Automated Soccer text System

También conocido como *PASS* [31], este sistema desarrollado por Chris van der Lee^{III}, Emiel Kraemer^{IV} y Sander Wubben^V es un sistema centrado en redactar noticias para par-

^{III}<https://research.tilburguniversity.edu/en/persons/chris-van-der-lee>

^{IV}<https://research.tilburguniversity.edu/en/persons/emiel-kraemer>

^V<https://research.tilburguniversity.edu/en/persons/sander-wubben>

tidos de fútbol de las ligas holandesas a partir de datos sobre los partidos (generación de texto a datos, mencionada en el apartado 3.1 de este documento). Ante sistemas contemporáneos cerrados e inaccesibles al público, y sistemas antiguos obsoletos, los autores proponen esta solución basada en plantillas hechas a mano. Con estas plantillas pretenden emular textos redactados por periodistas y generar dichos textos dirigidos a las diferentes audiencias de los diferentes equipos implicados en un partido. Esto pudieron realizarlo ya que la información sobre la que se tiene que redactar era relativamente constante y fácil de conseguir. Además, dentro de la generación de textos sobre deportes, su sistema sirve para la generación de resúmenes de partidos, no de comentarios de los mismos (una narración más parecida a lo que se comentaría al ver un partido en tiempo real).

Respecto al software en sí, utiliza unas plantillas que los autores desarrollaron a partir de frases en *MeMo FC* [46], un corpus de noticias sobre partidos recogidas de las páginas web de los diferentes equipos. De ahí se extraen las frases dividiéndolas en diferentes bases de datos según si se refieren a una victoria, a un empate o a una derrota y agrupándolas por el tipo de evento que describían (un gol, una falta, etc.). Una vez separadas las frases, se comprobaba si había información en Goal.com sobre ese evento (para posteriormente rellenar plantillas de esa frase con información de esa web) y se comparaba también si la información hubiera estado presente en *GoalGetter* (un sistema antiguo de generación de noticias de deporte que obtenía información del teletexto y en el que se basaron los autores de *PASS*). Aquellas frases que cumplían todos los requisitos eran transformadas y guardadas como plantillas, pero sin ningún tipo de información semántica en ellas. Esto quiere decir que, si había un ligero cambio en una frase (el ejemplo que aportan los autores en [31] es “<anotador> anotó un gol” y “<anotador> anotó su segundo gol”), esto requería guardar ambas frases como plantillas distintas de categorías distintas.

El sistema cuenta con un diseño modular, porque los autores querían proporcionar a terceros la facilidad de poder tomar el sistema que desarrollaron (o los módulos que necesitaran) y poder modificarlos a placer. Durante todo el proceso de la generación de texto hay un módulo central que toma todas las decisiones. El primer módulo que forma parte de ese proceso se encarga de la identificación de eventos que ocurren en cierto partido y los ordena cronológicamente para que el siguiente, el módulo de búsqueda, obtenga todas las plantillas que contienen esos eventos.

A continuación, el módulo encargado del set de reglas clasifica las plantillas obtenidas por el módulo anterior, según si la categoría a la que pertenecen es válida o no en esa generación en particular, y el módulo de selección de plantilla escoge, entre aquellas válidas, la que se ajusta más al caso concreto (una plantilla menos genérica será más atractiva ya que es por la que un redactor humano se decantaría). Una vez las plantillas candidatas han sido seleccionadas, el siguiente módulo se encarga de rellenar los espacios que haya con la información del partido en cuestión. Después hay un módulo encargado de combinar las plantillas en un orden que tenga sentido (por orden: el título, introducción, transcurso del juego y por último un pequeño resumen).

Por último, se encuentran una serie de módulos encargados de mantener una mínima calidad dentro de los textos generados. Por un lado, un módulo se encarga de que la información que aparezca no sea redundante a lo largo del texto. Otro, tiene como función encontrar nombres repetidos y, si es necesario, reemplazar alguna de las apariciones de ese nombre por un pronombre o una expresión equivalente. Para que no se escojan las mismas plantillas siempre, hay un último módulo encargado de llevar la cuenta de qué plantillas se han usado y cuántas veces.

3.3.2. PASS en múltiples idiomas

Al año de haber publicado *PASS* [31], se desarrollo una versión ampliada del sistema, pero esta vez ampliando la generación también al inglés [47].

En esta segunda versión, los autores se plantean el añadir la generación de los mismos textos, pero también en inglés (*PASS* solo generaba textos en neerlandés). La generación indistinta de un mismo contenido en varios idiomas es algo que en su momento no era una práctica tan común. Una de las razones de que no tuviera el foco de atención es que la generación en varios idiomas no consiste solo en cambiar unas palabras por otras; también puede implicar reglas gramaticales distintas, entre otras diferencias. Sin embargo, en la generación basada en plantillas este problema es menor porque no suelen requerir un léxico muy amplio ya que el dominio del problema tiende a ser relativamente pequeño. Por otra parte, el dominio de los deportes suele requerir cambios en el léxico cada cierto tiempo ya que los equipos, los jugadores cambian, e incluso pueden añadirse elementos tecnológicos nuevos (como la introducción del *Video Assistant Referee* o VAR) para ayudar en el transcurso y la calidad de los partidos.

Para resolver estos problemas, los autores primero añadieron la opción de generar los textos en inglés, intentando variar lo mínimo el sistema ya construido. Dada la similitud entre el neerlandés y el inglés, los cambios fueron mínimos y pocos cambios gramaticales fueron necesarios. Respecto al problema de que el léxico necesario varíe de forma periódica, los autores cambiaron la base de conocimiento con la que contaban (hecha “a mano” por ellos) por una que contaba con datos de Wikidata (una base de datos colaborativa en varios idiomas en la que humanos y máquinas pueden modificar información [48]). De esta forma también pudieron añadir más información a la generación de los textos, ya que el tamaño de Wikidata era mayor que el de la anterior base de conocimiento.

4. ANÁLISIS DEL SISTEMA

4.1. Contexto de la herramienta

La herramienta sobre la que van a versar las siguientes páginas forma parte de un sistema mayor y, en este apartado, se le van a dedicar unas líneas para aportar contexto a la misma.

El desarrollo de este software está encuadrado en la Cátedra UC3M-RTVE, con la que la Universidad Carlos III de Madrid realiza proyectos piloto para RTVE. De esta forma, los alumnos cuentan con la posibilidad de realizar proyectos de aplicaciones reales en un entorno controlado.

Para este caso, el proyecto se trataba de construir un sistema que obtuviera información de un proveedor externo de información sobre partidos de fútbol (por el momento solo de equipos de Castilla-La Mancha que jueguen en 2ª división B), obtuviera información en Twitter sobre el ánimo en el que se encontrasen las dos aficiones afectadas por el encuentro y redactase una noticia sobre el partido con una emoción en la redacción acorde a la que se encontraba en la red social para cada una de las aficiones. De esta manera se podría personalizar el contenido dependiendo de la afición a la que se estuviera dirigiendo la noticia.

El proyecto se dividió en dos grandes bloques: uno centrado en la extracción de las emociones de Twitter y otro en la generación de la noticia una vez se hubieran procesado y almacenado dichas emociones y la información del partido. La parte sobre la que se centra este trabajo es la última de las mencionadas.

4.2. Definición del problema

Referido específicamente a la parte que atañe a este trabajo, el problema identificado consiste en la redacción automática de noticias con sentimiento para eventos deportivos. Estos eventos deportivos serán únicamente de equipos de Castilla-La Mancha que participan en el Campeonato Nacional de Liga de Segunda División B.

Para que se produzca dicha generación, se debe obtener información sobre el partido y sobre el sentimiento. Una vez se han obtenido los datos, el sistema deberá generar la noticia en el formato deseado, teniendo en cuenta las emociones de las aficiones y generando una versión de la noticia para cada una.

Además, los resultados de dicha generación deben ser accesibles a periodistas y redactores, los cuales también deben ser capaces de cambiar los parámetros de la generación si así lo desean.

4.3. Descripción de la solución

Para responder al problema planteado en el apartado anterior, se ha tenido en cuenta el estado del arte desarrollado en la sección 3. Al combinarlo con los recursos tecnológicos con los que se dispone en la Cátedra, se ha decidido que el sistema a desarrollar se basará en generación de texto a partir de plantillas. Esto requerirá de una base de conocimiento en la que se almacenarán las plantillas y la información con la que se rellenarán dichas plantillas. El sistema será capaz de generar texto a partir de estas, y mostrará su interfaz y el resultado a través de una web.

En este caso particular, se usará la tecnología de la Cátedra en forma de una herramienta que permite crear y manejar una base de conocimiento, generar y modificar plantillas y generar texto.

La información de los partidos se obtendrá a través de un proveedor de RTVE que facilitará datos para cada partido en la misma jornada en la que sucedan. La información sobre las emociones que experimentan las aficiones será proporcionada por la otra parte del sistema creado para la Cátedra.

Para la generación de texto, el trabajo se basará en el funcionamiento de una herramienta de la Cátedra, pero se explicarán los cambios que ha sido necesario realizar a la misma. Estos cambios también modificarán ligeramente el funcionamiento de la creación y modificación de plantillas, y se explicará más adelante.

Para que los profesionales puedan interactuar con el sistema, se creará una web en la que puedan acceder a los servicios. En este aspecto, se utilizará trabajo previo realizado en la Cátedra como base para la interfaz, pues se pretende mantener una continuidad de estilos en los trabajos que se realizan en la Cátedra con RTVE.

4.4. Especificación de requisitos

A continuación se procederá a listar los requisitos que se han identificado para este sistema. Antes de su especificación se ofrece una plantilla (véase tabla 4.1) y su correspondiente descripción.

ID	
Nombre	
Versión	
Prioridad	
Necesidad	
Estado	
Fuente	
Descripción	

Tabla 4.1. Plantilla de tabla de requisito.

Descripción de la plantilla de requisitos

La información presente en la plantilla de los requisitos se basa en los siguientes campos:

- **ID.** Sirve como identificador inequívoco del requisito. Utilizará un código siguiendo el formato: RF/RNF - XX. La primera parte identificará si es un requisito funcional (RF) o no funcional (RNF) y la segunda constará de dos dígitos para numerarlos.
- **Nombre.** Aquí se aportará una forma de identificar al requisito con un lenguaje más humano, sin necesidad de utilizar su identificador.
- **Versión.** Permite identificar en qué estadio está el requisito respecto al control de cambios y versiones.
- **Prioridad.** Establece la urgencia del requisito respecto al desarrollo del proyecto. Puede tomar los valores de *alta*, *media* o *baja*.
- **Necesidad.** Establece la importancia del requisito en el desarrollo del sistema. Puede tomar los valores de *esencial*, *deseado* o *opcional*.
- **Estado.** Refleja en qué fase del ciclo de vida está el requisito. Puede tomar los valores de *propuesto*, *aprobado*, *implementado* o *validado*.
- **Fuente.** Identifica el origen del requisito.
- **Descripción.** Detalla el contenido del requisito.

4.4.1. Requisitos funcionales

Los requisitos funcionales son aquellos que describen capacidades o funcionalidades que el sistema ofrecerá a los usuarios.

Generación de texto

La generación de texto estará definida en los requisitos recogidos en las tablas 4.2, 4.3, 4.4, 4.5, 4.6 y 4.7.

ID	RF-01
Nombre	Generación de titulares.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema generará dos titulares para cada noticia solicitada a partir de la especificación de dos equipos, una jornada y una emoción para cada equipo.

Tabla 4.2. RF-01. Generación de titulares.

ID	RF-02
Nombre	Generación de subtítulos.
Versión	2.0
Prioridad	Media
Necesidad	Deseado
Estado	Aprobado
Fuente	Tutor
Descripción	El sistema generará dos subtítulos para cada noticia solicitada a partir de la especificación de dos equipos, una jornada y una emoción para cada equipo.

Tabla 4.3. RF-02. Generación de subtítulos.

ID	RF-03
Nombre	Generación de los cuerpos de la noticia.
Versión	2.0
Prioridad	Media
Necesidad	Deseado
Estado	Aprobado
Fuente	Tutor
Descripción	El sistema generará dos cuerpos de la noticia para cada noticia solicitada a partir de la especificación de dos equipos, una jornada y una emoción para cada equipo.

Tabla 4.4. RF-03. Generación de los cuerpos de la noticia.

ID	RF-04
Nombre	Concordancia de género.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema generará textos manteniendo la concordancia de género entre los nombres utilizados y las palabras que se refieran a o que se relacionan con ellos.

Tabla 4.5. RF-04. Concordancia de género.

ID	RF-05
Nombre	Concordancia de número.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema generará textos manteniendo la concordancia de número entre los nombres utilizados y las palabras que se refieran a o que se relacionan con ellos.

Tabla 4.6. RF-05. Concordancia de número.

ID	RF-06
Nombre	Concordancia de persona.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema generará textos manteniendo la concordancia de persona entre los nombres utilizados y las palabras que se refieran a o que se relacionan con ellos.

Tabla 4.7. RF-06. Concordancia de persona.

Formato

Las tablas 4.8 y 4.9 reflejan los requisitos que especifican el formato de las generaciones.

ID	RF-07
Nombre	Formato de la noticia.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	RTVE
Descripción	El sistema mostrará la noticia con el formato descrito por RTVE: tema, titular, subtítular, fecha, imagen, cuerpo, metadatos.

Tabla 4.8. RF-07. Formato de la noticia.

ID	RF-08
Nombre	Formato de la web.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	RTVE
Descripción	El sistema mostrará la noticia del proveedor de información junto a la noticia generada.

Tabla 4.9. RF-08. Formato de la web.

Parámetros

Los requisitos relacionados con los parámetros de entrada se recogen en las tablas 4.10, 4.11, 4.12 y 4.13.

ID	RF-09
Nombre	Elección de equipos.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema permitirá elegir al usuario la jornada sobre la que quiere generar una noticia. Estos equipos deberán participar en el Campeonato Nacional de Liga de Segunda División B y ser de Castilla-La Mancha.

Tabla 4.10. RF-09. Elección de equipos.

ID	RF-10
Nombre	Elección de jornada.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema permitirá elegir al usuario la jornada sobre la que quiere generar una noticia. Estas jornadas deberán ser aquellas que hayan sucedido en el Campeonato Nacional de Liga de Segunda División B y deben estar disponibles por el proveedor de información. Por defecto se ofrecerá la jornada más reciente.

Tabla 4.11. RF-10. Elección de jornada.

ID	RF-11
Nombre	Elección de las emociones.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema permitirá elegir al usuario las emociones con las que quiere que se genere cada versión de la noticia (una para cada equipo). Las emociones disponibles serán las proporcionadas por el componente de detección de la Cátedra. Por defecto se asumirá la emoción 'blanca', en cuyo caso no se generará texto.

Tabla 4.12. RF-11. Elección de las emociones.

ID	RF-12
Nombre	Elección de la versión de la noticia.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema permitirá elegir al usuario visualizar la versión de la noticia de cada equipo participante en el partido (cada jornada tendrá un partido único por equipo).

Tabla 4.13. RF-12. Elección de la versión de la noticia.

Base de conocimiento

En las tablas 4.14, 4.15, 4.16 y 4.17 se recogen los requisitos funcionales relacionados con la base de conocimiento.

ID	RF-13
Nombre	Creación de la base de conocimiento.
Versión	1.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema permitirá crear al usuario una base de conocimiento en la que almacenar las plantillas y el léxico necesario para rellenarlas.

Tabla 4.14. RF-13. Creación de la base de conocimiento.

ID	RF-14
Nombre	Selección de la base de conocimiento.
Versión	1.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema permitirá al usuario seleccionar una base de conocimiento en la que almacenar las plantillas y el léxico necesario para rellenarlas.

Tabla 4.15. RF-14. Elección de la base de conocimiento.

ID	RF-15
Nombre	Modificación de la base de conocimiento.
Versión	1.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema permitirá modificar la base de conocimiento seleccionada por el usuario.

Tabla 4.16. RF-15. Modificación de la base de conocimiento.

ID	RF-16
Nombre	Eliminación de la base de conocimiento.
Versión	1.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema permitirá al usuario eliminar la base de conocimiento seleccionada por el usuario.

Tabla 4.17. RF-16. Modificación de la base de conocimiento.

4.4.2. Requisitos no funcionales

Los requisitos no funcionales son aquellos que describen restricciones impuestas sobre el sistema a la hora de diseñarlo o implementarlo y acotan las propiedades que debe tener.

Organizacionales

Las tablas 4.18 y 4.19 detallan los requisitos no funcionales organizacionales.

ID	RNF-01
Nombre	Marco de desarrollo.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema se desarrollará utilizando el framework ASP.NET, mediante el lenguaje de programación C#. Para el desarrollo de la interfaz web se utilizará el framework Angular.

Tabla 4.18. RNF-01. Marco de desarrollo.

ID	RNF-02
Nombre	Estructura del servicio web.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema de generación ofrecerá sus servicios a los usuarios a través de un servicio de tipo REST.

Tabla 4.19. RNF-02. Estructura del servicio web.

Externos

Las tablas 4.20, 4.21, 4.22, 4.23, 4.24 y 4.25 se plasman los requisitos no funcionales externos.

ID	RNF-03
Nombre	Gestor de la base de conocimiento.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema utilizará el programa llamado Knowledge Manager creado por The Reuse Company para gestionar la base de conocimiento.

Tabla 4.20. RNF-03. Gestor de conocimiento.

ID	RNF-04
Nombre	Herramienta de generación de textos.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema utilizará como base la herramienta del programa llamado Knowledge Manager creado por The Reuse Company para generar texto.

Tabla 4.21. RNF-04. Herramienta de generación de textos.

ID	RNF-05
Nombre	Base de datos.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema de gestión de bases de datos utilizado para almacenar la base de conocimientos será MySQL.

Tabla 4.22. RNF-05. Base de datos.

ID	RNF-06
Nombre	Fuente de las emociones.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema de generación obtendrá las emociones del módulo desarrollado en la Cátedra con dicho fin.

Tabla 4.23. RNF-06. Fuente de las emociones.

ID	RNF-07
Nombre	Fuente de información deportiva.
Versión	1.0
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Tutor
Descripción	El sistema de generación obtendrá la información deportiva del proveedor dicha información de RTVE.

Tabla 4.24. RNF-07. Fuente de información deportiva.

ID	RNF-08
Nombre	Integración con RTVE.
Versión	2.5
Prioridad	Alta
Necesidad	Esencial
Estado	Validado
Fuente	Becario
Descripción	El sistema se desplegará para que pueda ser usado por el personal de RTVE en su entorno de trabajo.

Tabla 4.25. RNF-08. Integración con RTVE.

4.5. Casos de uso

En base a la especificación de requisitos realizada en el apartado anterior, aquí se presentarán los casos de uso atribuidos al sistema. En ellos que se detallarán las interacciones que se han previsto de los distintos usuarios con el sistema.

4.5.1. Diagrama de casos de uso

En la figura 4.1 se muestra, de manera gráfica, las posibles interacciones de los distintos usuarios con el sistema. Se puede observar que solo se han identificados dos potenciales usos por parte de los distintos usuarios, pero actualmente el sistema solo sirve estos propósitos. No obstante, se contempla que, en un futuro, se amplíen estas posibilidades.

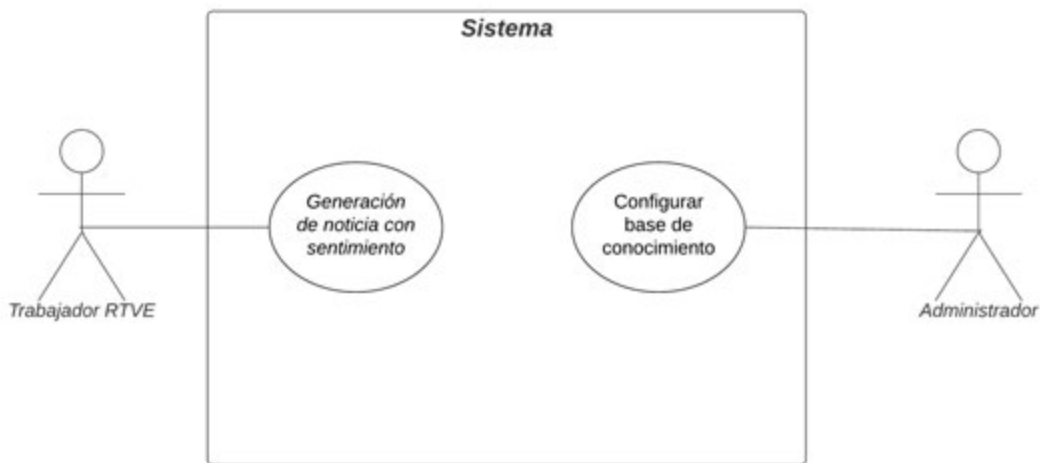


Fig. 4.1. Diagrama de casos de uso identificados para la herramienta.

El primer caso de uso, y el más obvio, que podemos observar es en el que el trabajador de RTVE accede al sistema para obtener una noticia. Para ello, el usuario primero deberá elegir el equipo para el que quiere generar la noticia. Posteriormente, el usuario podrá elegir la jornada y las emociones con las que quiere que la generación de texto se produzca. Por último, se le ofrecerá al usuario el resultado de la generación.

El otro caso de uso descrito es aquel en el que un administrador del sistema accede a la base de conocimiento para crearla, modificarla o bien eliminarla.

4.5.2. Definición de casos de uso

En el presente apartado se mostrarán las definiciones de los casos de uso ilustrados en la figura 4.1. Estos seguirán el formato de la plantilla que se describe en la tabla 4.26.

ID	
Nombre	
Actor	
Objetivo	
Precondiciones	
Poscondiciones	
Escenario básico	
Escenario alternativo	
Requisitos Funcionales	

Tabla 4.26. Plantilla de tabla de caso de uso.

Descripción de la plantilla de casos de uso

La información que se muestra en la plantilla de los casos de uso está formada de los siguientes campos:

- **ID.** Al igual que en la plantilla de requisitos ilustrada en la tabla 4.1, este identificador será la forma inequívoca de referirse al caso de uso en cuestión. Seguirá el formato CU-XX, siendo las “equis” dígitos para numerarlos.
- **Nombre.** Esta será una forma de identificar al caso de uso con una forma breve de lenguaje humano.
- **Actor.** Es el agente que interactúa con el sistema en el caso de uso descrito. En este trabajo se identifican el trabajador de RTVE y el administrador.
- **Objetivo.** Se enunciará la finalidad de la interacción del actor con el sistema.
- **Precondiciones.** Estas recogen cómo debe estar el sistema para que el caso de uso pueda darse.
- **Poscondiciones.** En la misma línea que el campo anterior, aquí se describe el estado del sistema una vez se haya producido la interacción completa.
- **Escenario básico.** Aquí se describe una actuación posible dentro del caso de uso contemplado.

- **Escenario alternativo.** En este campo se ofrece una actuación alternativa a la del campo anterior.
- **Requisitos funcionales.** Aquí serán enumerados los requisitos funcionales que se relacionen con el caso de uso.

En las tablas 4.27 y 4.28 se presentan las descripciones de los casos de uso identificados.

ID	CU-01
Nombre	Generación de noticia con sentimiento
Actor	Trabajador RTVE
Objetivo	Generar una noticia para una jornada, unos equipos y unas emociones seleccionados.
Precondiciones	El usuario ha accedido a la plataforma web y el sistema muestra por pantalla la interfaz de selección de parámetros.
Poscondiciones	El sistema muestra por pantalla la versión de la noticia acorde a los parámetros seleccionados.
Escenario básico	<ol style="list-style-type: none"> 1. El usuario selecciona uno de los equipos. 2. El sistema muestra las jornadas disponibles (mostrando la más reciente por defecto). También muestra las emociones disponibles (seleccionando la 'blanca' por defecto). Además, se muestran las versiones disponibles, local y visitante (estando por defecto seleccionada la versión local). 3. El sistema genera dos versiones de la noticia con los parámetros por defecto. 4. El sistema muestra por pantalla la versión seleccionada por defecto de la noticia generada.

Escenario alternativo	<ol style="list-style-type: none"> 1. El usuario selecciona uno de los equipos. 2. El sistema muestra las jornadas disponibles (mostrando la más reciente por defecto). También muestra las emociones disponibles (seleccionando la 'blanca' por defecto). Además, se muestran las versiones disponibles, local y visitante (estando por defecto seleccionada la versión local). 3. El usuario selecciona una jornada diferente. 4. El sistema muestra las emociones para esa jornada. 5. El sistema genera dos versiones de la noticia con los parámetros por defecto. 6. El usuario selecciona dos emociones para los distintos equipos. 7. El sistema muestra por pantalla la versión seleccionada por defecto de la noticia generada de acuerdo a los parámetros. 8. El usuario selecciona la versión del equipo visitante. 9. El sistema muestra por pantalla la versión seleccionada de la noticia generada.
Requisitos Funcionales	RF-01, RF-02, RF-03, RF-04, RF-05, RF-06, RF-07, RF-08, RF-09, RF-10, RF-11 y RF-12

Tabla 4.27. CU-01. Generación de noticia con sentimiento.

ID	CU-02
Nombre	Configurar base de conocimiento
Actor	Administrador
Objetivo	Acceder al gestor de la base de conocimiento para crear, modificar y eliminar.
Precondiciones	El usuario ha accedido herramienta de gestión de la base de conocimiento.
Poscondiciones	El sistema ha registrado los cambios realizados por el usuario en la correspondiente base de datos MySQL.

Escenario básico	<ol style="list-style-type: none"> 1. El usuario crea una base de conocimiento. 2. El usuario introduce información para rellenar plantillas. 3. El usuario introduce plantillas para la generación. 4. El usuario configura las plantillas para la generación 5. El usuario elimina la base de conocimiento.
Escenario alternativo	<ol style="list-style-type: none"> 1. El usuario crea una base de conocimiento. 2. El usuario introduce información para rellenar plantillas. 3. El usuario introduce plantillas para la generación. 4. El usuario configura las plantillas para la generación 5. El usuario modifica la información para rellenar plantillas. 6. El usuario modifica las plantillas para la generación. 7. El usuario configura de nuevo las plantillas para la generación.
Requisitos Funcionales	RF-13, RF-14, RF-15 y RF-16.

Tabla 4.28. CU-02. Configurar base de conocimiento.

5. DISEÑO DEL SISTEMA

En este capítulo se detallará el diseño propuesto para el caso descrito en el análisis anterior. Se aportará una descripción de la infraestructura utilizada, se desarrollará la arquitectura del sistema indicando diferentes vistas de la misma y se explicará el funcionamiento de la herramienta de la base de conocimiento, así como el entorno de la generación de textos.

5.1. Infraestructura del desarrollo

A continuación, se dará un breve resumen de las herramientas y *frameworks* utilizados durante el desarrollo del módulo del sistema sobre el que se basa este trabajo. El uso de todas ellas ha ido marcado por la infraestructura tecnológica de la Cátedra.

Git

Este ha sido el software de control de versiones que se ha utilizado para llevar un seguimiento del módulo del sistema mencionado. Este es un software libre que fue diseñado por Linus Torvalds [49].

Github

En combinación con *Git*, se ha dispuesto de la plataforma de desarrollo colaborativo GitHub [50] para alojar el proyecto. Al ser un desarrollo de un sistema con potencial y mediante un acuerdo con RTVE, se ha mantenido el código y la documentación en un repositorio privado.

C#

Ha sido el lenguaje de programación utilizado para la implementación de los servicios y la generación de textos. Es un lenguaje desarrollado por Microsoft orientado a objetos, que se basa en C y C++ y utiliza el modelo de objetos de .NET [51].

Angular

Angular es un *framework* desarrollado en TypeScript (un superconjunto de Javascript) y mantenido por Google, cuyo enfoque es el desarrollo de aplicaciones web [52]. Uno de sus objetivos principales es apoyar a los desarrolladores a la hora de desarrollar aplicaciones basadas en Modelo Vista Controlador (MVC) [53]. La versión 6.0 de este *framework*

ha sido la utilizada para desarrollar la web con la que interactúa el usuario y la que muestra el resultado de la generación de texto.

MySQL

Este sistema de gestión de bases de datos relacionales, de licencia pública general y también de licencia comercial por Oracle, es uno de los más populares a día de hoy [54]. Ha sido utilizado como pilar para la base de conocimiento, aunque se interactuaba con esta mediante la herramienta *Knowledge Manager* descrita más adelante. También se ha utilizado MySQL Workbench [55] para interactuar con la base de conocimiento cuando ha sido necesario.

Microsoft Visual Studio

Este ha sido el entorno de desarrollo integrado (*Integrated Development Environment*, IDE) utilizado para el desarrollo en C#. Está disponible para varios sistemas operativos (Windows y MacOS) y ofrece un gran soporte para implementación a lenguajes como C++, C#, Visual Basic, .NET, Java, Python, Ruby o PHP [56]. En este proyecto se ha utilizado el IDE en su versión Enterprise 2019 y el entorno ASP.NET MVC que ofrece para el desarrollo de servicios web REST.

Visual Studio Code

En combinación con el anterior, se ha utilizado el editor de código fuente Visual Studio Code [57] para el desarrollo de la plataforma web en Angular 6.

Knowledge Manager

Esta herramienta permite la gestión de bases de conocimiento y la generación de textos, entre otras, y es propiedad de The Reuse Company. En ella se puede introducir y gestionar un vocabulario, se pueden indicar patrones (plantillas) para la generación y el reconocimiento de texto, etc.

Para la generación de textos se han utilizado la API que ofrece el software detrás de esta herramienta. En ciertos casos (como se explicará más adelante en el documento), se ha tenido que modificar el código fuente de esta herramienta para variar o mejorar su comportamiento respecto a la generación original. El control de versiones del código de esta herramienta se lleva a cabo mediante Subversion [58].

Cabe destacar que el desarrollo de la solución planteada orbita en torno al uso de esta herramienta.

5.2. Arquitectura del sistema

Primero se mostrará una arquitectura de alto nivel del sistema completo para aportar contexto al trabajo realizado. Después se aportarán las vistas de desarrollo y de procesos, basándose en el modelos de vistas de arquitectura 4+1 [59]. No se aportará la vista física, ya que no procede describir la capa física del sistema. Tampoco se aportará la vista lógica del módulo, pues la mayor parte de este requeriría representar el software detrás de la herramienta *Knowledge Manager*. No obstante, se hará lo mejor posible para proporcionar una imagen clara del sistema descrito dadas las restricciones de información. Por último, los escenarios se ven reflejados en el apartado 4.5.2 de este documento.

5.2.1. Arquitectura de alto nivel

El trabajo realizado forma parte del sistema *Social Media News*, que pretende enriquecer las noticias deportivas a través del uso de la generación automática de texto. En esta generación, se tendrán en cuenta las emociones de las distintas aficiones afectadas por la noticia, además de las publicaciones más relevantes en redes sociales (en este caso, Twitter).

La arquitectura principal establecida en el proyecto se puede observar en la figura 5.1 y consta de los siguientes bloques:



Fig. 5.1. Representación a alto nivel de la arquitectura del sistema (*Social Media News*).

- **Adquisición de datos.** Este bloque se encarga de establecer la conexión con la plataforma de Twitter y de obtener información de ella.
- **Escalabilidad.** Aquí se dirigen los diferentes procesos del sistema.

- **Almacenamiento.** En este bloque se coordina el almacenamiento y la consulta de información.
- **Análisis.** Bloque central del sistema. En él se concentran el gestor de ontologías (encargado de la gestión de bases de conocimiento), el gestor de lenguaje natural (que se encarga de las tareas relacionadas con NLG), el gestor de confianza (en el que se asegura la relevancia de la información tratada en el sistema) y el gestor de emociones (parte en la que se determinan las emociones de las aficiones).
- **Servicio de generación de crónicas deportivas.** Servido del bloque de análisis descrito en el párrafo anterior, este bloque se encarga de producir las noticias enriquecidas.
- **Aplicaciones del cliente.** Como bloque más cercano al cliente, en él se muestran los resultados obtenidos y permite al usuario la interacción con el sistema.

5.2.2. Vista de desarrollo

En este apartado se explicará el sistema desde el punto de vista del desarrollo. Primero se ofrecerá una vista de los componentes en la figura 5.2 siguiendo el patrón Modelo Vista Controlador (MVC) [53]. Después se procederá a una descripción detallada de los componentes reflejados en el diagrama.

El patrón MVC se caracteriza en separar los componentes del sistema en tres categorías:

- **Modelo.** Esta categoría se caracteriza por contener aquellos componentes responsables del acceso a datos de la aplicación. Las actividades de acceso a cualquier tipo de base de datos están definidas aquí.
- **Vista.** Aquí se agrupan los componentes que proporcionan información al usuario y que le permiten la interacción con el sistema.
- **Controlador.** La última categoría recoge aquellos componentes que se encargan de la lógica del sistema e intermedian entre la Vista y el Modelo. Gestiona los eventos generados en la vista accediendo al modelo según sus reglas establecidas.

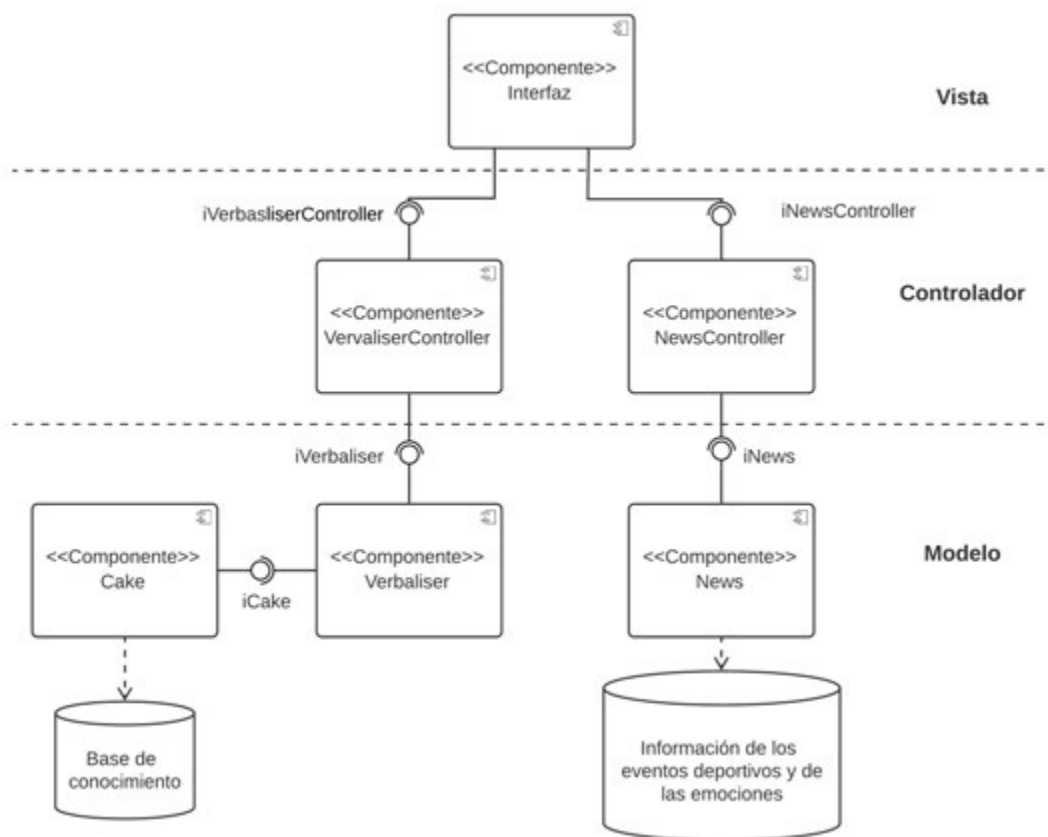


Fig. 5.2. Diagrama de componentes.

Antes de describir los componentes del diagrama, se explicará la plantilla ilustrada en la tabla 5.1 utilizada para su descripción.

Componente	
Propósito	
Dependencias	
Interfaz	
Contrato de operaciones	

Tabla 5.1. Plantilla de componente.

La composición de la tabla es:

- **Componente.** Aquí se especifica el nombre del componente que se describe.
- **Propósito.** Después se redacta su propósito.
- **Dependencias.** A continuación, se declaran sus dependencias.
- **Interfaz.** También se especifica la interfaz que ofrece.
- **Contrato de operaciones.** Por último, se detallan los métodos que contiene.

A continuación se aportarán las descripciones de los componentes reflejados en el diagrama.

Vista

En la tabla 5.2 se describe el único componente de la Vista.

Componente	Interfaz
Propósito	Ofrecer la interfaz de interacción gráfica al usuario.
Dependencias	iVerbaliserController, iNewsController
Interfaz	iInterfaz

<p>Contrato de operaciones</p>	<ul style="list-style-type: none"> ■ ObtenerJornadas(string equipo) <ul style="list-style-type: none"> ● Descripción. Obtiene y muestra las jornadas para el equipo seleccionado. ● Precondiciones. Se han seleccionado un equipo en la interfaz gráfica. ● Poscondiciones. Se muestran por pantalla las jornadas y partidos disputados por el equipo. ■ ObtenerEmociones(string equipo, int jornada) <ul style="list-style-type: none"> ● Descripción. Obtiene y muestra las emociones para el equipo y la jornada de entrada. ● Precondiciones. Se han seleccionado un equipo y una jornada en la interfaz gráfica. ● Poscondiciones. Se muestran por pantalla las emociones de los equipos local y visitante (un equipo y una jornada identifican un partido de manera unívoca). ■ ObtenerNoticia(string equipoLocal, string equipoVisitante, string emocionLocal, string emocionVisitante, int ganador) <ul style="list-style-type: none"> ● Descripción. Obtiene y muestra las emociones para el equipo y la jornada de entrada. ● Precondiciones. Se han seleccionado los parámetros en la interfaz gráfica y se ha calculado el ganador. ● Poscondiciones. Se obtiene un objeto con las versiones de la noticia seleccionada. ■ MostrarVersion(int equipo) <ul style="list-style-type: none"> ● Descripción. Muestra la versión seleccionada de la noticia. ● Precondiciones. Se ha seleccionado la versión que se desea a través de la interfaz gráfica. ● Poscondiciones. Se muestra por pantalla la versión seleccionada.
--------------------------------	---

Tabla 5.2. Descripción del componente Interfaz.

Controlador

Las tablas 5.3 y 5.4 muestran los componentes que son parte del Controlador.

Componente	NewsController
Propósito	Recibir y gestionar los eventos de la interfaz relacionados con obtener información de los partidos.
Dependencias	iNews
Interfaz	iNewsController
Contrato de operaciones	<ul style="list-style-type: none"> ■ getJornadas(string equipo) <ul style="list-style-type: none"> ● Descripción. Servicio web para obtener las jornadas. ● Precondiciones. El equipo seleccionado existe y no es nulo. ● Poscondiciones. Se devuelve un objeto con las jornadas. ■ getEmociones(string equipo, int jornada) <ul style="list-style-type: none"> ● Descripción. Servicio web para obtener las emociones para un partido en determinada jornada. ● Precondiciones. El equipo seleccionado existe y no es nulo. La jornada existe y no es nula. ● Poscondiciones. Se devuelve un objeto con las emociones para la información especificada. ■ getInformacionNoticia(string equipo, int jornada) <ul style="list-style-type: none"> ● Descripción. Servicio web para información que complementa a la noticia. ● Precondiciones. El equipo seleccionado existe y no es nulo. La jornada existe y no es nula. ● Poscondiciones. Se devuelve un objeto con la información del partido del proveedor.

Tabla 5.3. Descripción del componente NewsController.

Componente	VerbaliserController
Propósito	Recibir y gestionar los eventos de la interfaz relacionados la generación de textos.
Dependencias	iVerbaliser
Interfaz	iVerbaliserController

Contrato de operaciones	<ul style="list-style-type: none"> ▪ <code>getNoticia(string equipoLocal, string equipoVisitante, string emocionLocal, string emocionVisitante, int ganador)</code> <ul style="list-style-type: none"> • Descripción. Servicio web para obtener las versiones de la noticia. • Precondiciones. Los parámetros de los equipos y las emociones seleccionados existen y no son nulos. El parámetro ganador es -1, 0 o 1 (gana el local, empatan o gana el visitante). • Poscondiciones. Se devuelve un objeto con las versiones de la noticia.
-------------------------	---

Tabla 5.4. Descripción del componente VerbaliserController.

Modelo

Por último, los componentes del Modelo se detallan en las tablas 5.5, 5.6 y 5.7.

Componente	News
Propósito	Obtener información sobre los eventos deportivos y las emociones de la base de datos.
Dependencias	-
Interfaz	iNews

Contrato de operaciones	<ul style="list-style-type: none"> ■ <code>getInformaciónEvento(string equipo, string jornada)</code> <ul style="list-style-type: none"> ● Descripción. Método que accede a la base de datos para recuperar la información del encuentro para el equipo y la jornada seleccionados. ● Precondiciones. Los parámetros equipo y jornada seleccionados existen y no son nulos. ● Poscondiciones. Se devuelve un objeto con la información del evento especificado. ■ <code>getEventosEquipo(string equipo)</code> <ul style="list-style-type: none"> ● Descripción. Método que accede a la base de datos para recuperar las jornadas en las que ha participado un equipo. ● Precondiciones. El parámetro equipo existe y no es nulo. ● Poscondiciones. Se devuelve un objeto con las jornadas requeridas.
-------------------------	--

Tabla 5.5. Descripción del componente News.

Componente	Verbaliser
Propósito	Obtener los textos generados.
Dependencias	iCake
Interfaz	iVerbaliser

Contrato de operaciones	<ul style="list-style-type: none"> ■ <code>defineTemplate(string equipoLocal, string equipoVisitante, string emocionLocal, string emocionVisitante, int ganador)</code> <ul style="list-style-type: none"> ● Descripción. Método que da formato a la información en función de la plantilla que se vaya a utilizar. ● Precondiciones. Los parámetros de los equipos y las emociones existen y no son nulos. El ganador es -1, 0 o 1. ● Poscondiciones. Se devuelve la información para generar texto con la plantilla seleccionada. ■ <code>generateText(string info)</code> <ul style="list-style-type: none"> ● Descripción. Método que se comunica con el componente de generación de textos. ● Precondiciones. El parámetro <code>info</code> no es nulo. ● Poscondiciones. Se devuelve un texto generado de acuerdo a la información proporcionada.
-------------------------	--

Tabla 5.6. Descripción del componente Verbaliser.

Componente	Cake
Propósito	Generar el texto.
Dependencias	-
Interfaz	iCake
Contrato de operaciones	<ul style="list-style-type: none"> ■ <code>generate(string info)</code> <ul style="list-style-type: none"> ● Descripción. Método que genera el texto. ● Precondiciones. El parámetro <code>info</code> no debe ser nulo. ● Poscondiciones. Se devuelve la generación de texto requerida.

Tabla 5.7. Descripción del componente Cake.

5.2.3. Vista de procesos

Antes de mostrar el diagrama de actividades para explicar la generación de texto, se van a explicar conceptos preliminares del funcionamiento de la herramienta *Knowledge Manager* para facilitar la comprensión del posterior diagrama.

Nótese que durante todo este apartado se busca un equilibrio entre aportar suficiente información para que se entienda con cierta profundidad lo que aquí se explica y no aportar demasiada como para que otros pudieran “destripar” el funcionamiento de la herramienta *Knowledge Manager*.

5.2.4. Base de conocimiento

En la base de conocimiento se almacenarán las plantillas y la información que se contendrá en ellas. La información en esta base se estructura en torno a las siguientes figuras:

- **Términos.** Son las principal unidades de información y tienen la equivalencia a las palabras (o, a veces, expresiones) que utilizan los humanos para comunicarse. Como ejemplos de términos se aportan: José, bonito, hacer, etc. Estos términos se introducen en la base de conocimiento con parámetros como género, número y otra información.
- **Categorías gramaticales.** Las categorías gramaticales tienen la misma utilidad que en castellano y pretenden identificar la función que realizan los términos dentro de las frases y oraciones. Algunos ejemplos son: nombre, adjetivo, preposición, etc. Cada término solo puede ser asociado a una categoría.
- **Clústeres.** Esta es la forma de agrupar términos dentro de la base de conocimiento. La forma lógica de hacerlo es por significado (como puede observarse en la figura 5.3), pero pueden agruparse según desee el usuario. Cada palabra puede pertenecer a tantos clúster como se especifique. También pueden crearse clústeres dentro de otros (para especificar dentro de una familia de palabras, por ejemplo).
- **Patrones.** Los patrones son el equivalente a las plantillas. Se estructuran en *slots*, la unidad mínima que debe tener cada patrón. Los patrones pueden ser términos específicos; términos que pertenezcan a categorías gramaticales, *clústeres* o a ambos; otros patrones (en cuyo caso se le denomina subpatrón), o un texto definido (que no tiene porqué pertenecer a la base de conocimiento, pero será siempre el mismo). Hay otros valores de los *slots* como los comodines que no se explicarán en este trabajo por su complejidad.

En la figura 5.4 se puede observar un patrón compuesto por los siguientes *slots*: un término de la categoría gramatical determinante, otro del clúster equipo, seguido de dos términos del clúster alegría, siendo uno de la categoría gramatical verbo y



Fig. 5.3. Ejemplos de clústeres dentro de la base de conocimiento.



Fig. 5.4. Ejemplo de un patrón de la base de conocimiento.

otro de la categoría adverbio. Los últimos *slots* son el término ‘por’ y el término ‘goleada’. Este patrón podría instanciarse como la oración ‘El Villarrubia ganó alegremente por goleada’.

5.2.5. Proceso de generación de texto

En la figura 5.5 se puede apreciar el proceso de verbalización que se realiza en el sistema. En él se observan dos partes diferenciadas: la primera es la que se realiza a través de la propia interfaz de la herramienta *Knowledge Manager* y la segunda, pudiéndose realizar también a través de la herramienta, se realiza a través de la interfaz gráfica y los componentes descritos en la figura 5.2.

Para generar texto, el usuario primero debe crear los patrones de comprensión y de generación. Esto es porque en la herramienta *Knowledge Manager* se genera texto en un patrón siempre a través de otro. Esto se hace por si se quiere transmitir información al patrón desde el que se genera el texto. Los patrones de comprensión y generación son el mismo tipo de patrones, pero usados para fines distintos.

Se puede observar que, después del patrón, se pide construir el árbol de dependencias. Este paso consiste en que el usuario especifique al sistema qué dependencias se producen

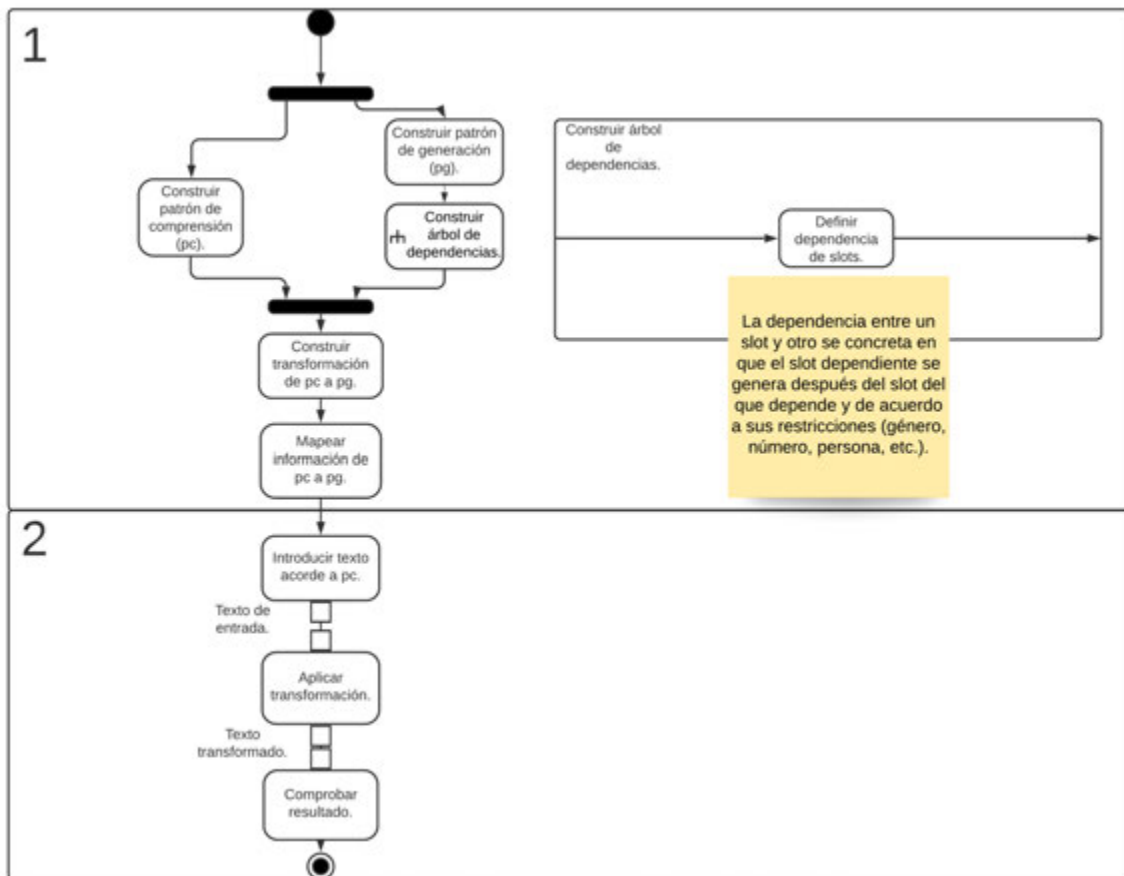


Fig. 5.5. Diagrama de actividades del proceso de generación de texto.

a la hora de la generación. En castellano existen las concordancias de género, número y persona entre algunas palabras y esto se ha modelado a través de estas dependencias. 'El' necesita ser masculino y singular porque 'Villarrubia' lo es. Por ello en la generación, el término dependiente se generará después. Esto puede comprenderse observando las figuras 5.6 y 5.7. El orden de generación es de izquierda a derecha.

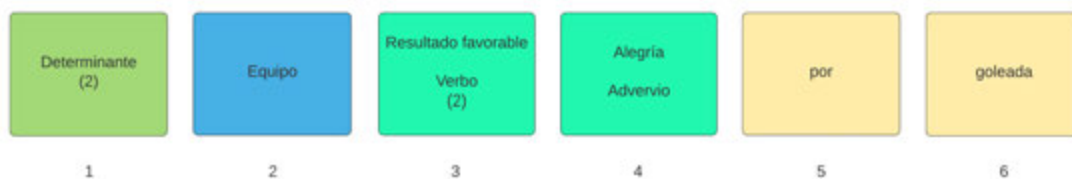


Fig. 5.6. Representación de un patrón no ordenado por dependencias.

Después debe construirse la transformación de el patrón de comprensión al de generación. Esto consiste, simplemente, en asignar qué parte de un patrón aporta información al otro.



Fig. 5.7. Representación de un patrón ordenado por dependencias.

Después solo debe introducirse el texto que corresponda con el patrón de comprensión y el sistema generará un texto acorde al patrón de generación.

6. IMPLEMENTACIÓN

En esta sección se explicarán los distintos módulos que se han implementado, las pruebas realizadas para comprobar su correcto funcionamiento y los problemas encontrados durante el desarrollo.

6.1. Web y servicios

Como ya se ha comentado en el apartado 5.1, la web se realiza utilizando Angular 6 y los servicios utilizando C#. El resultado de la web ha sido el que se observa en las figuras 6.1, 6.2, 6.3, 6.4, 6.5 y 6.6.

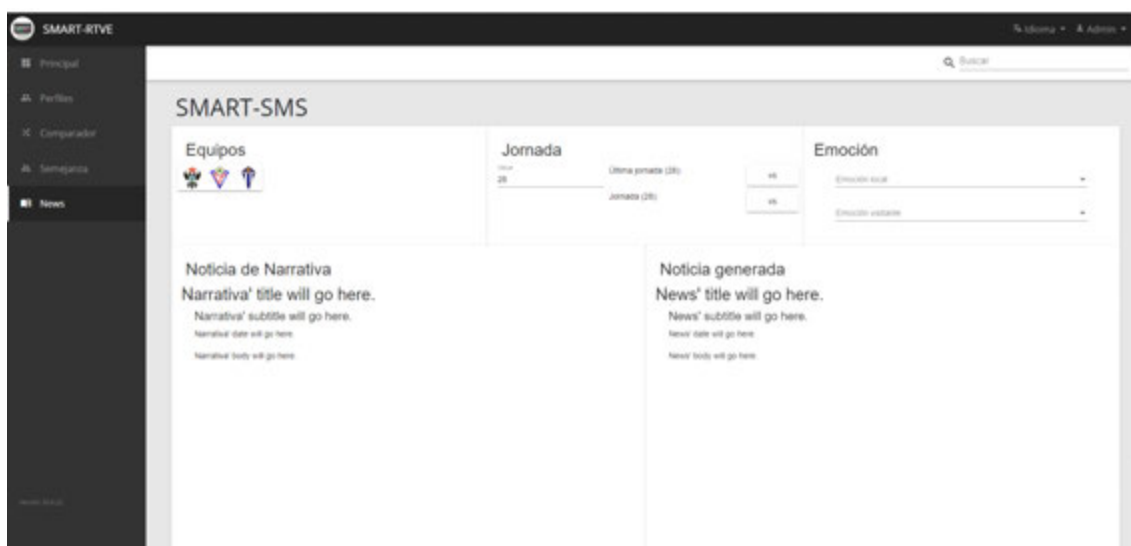


Fig. 6.1. Captura de la web sin realizar ninguna acción.

En la primera figura 6.1 se observa la web antes de que el usuario interactúe con ella. El menú de la izquierda es parte del entorno global y no fue realizado por el autor. Aquí el usuario puede elegir uno de los equipos de la izquierda.

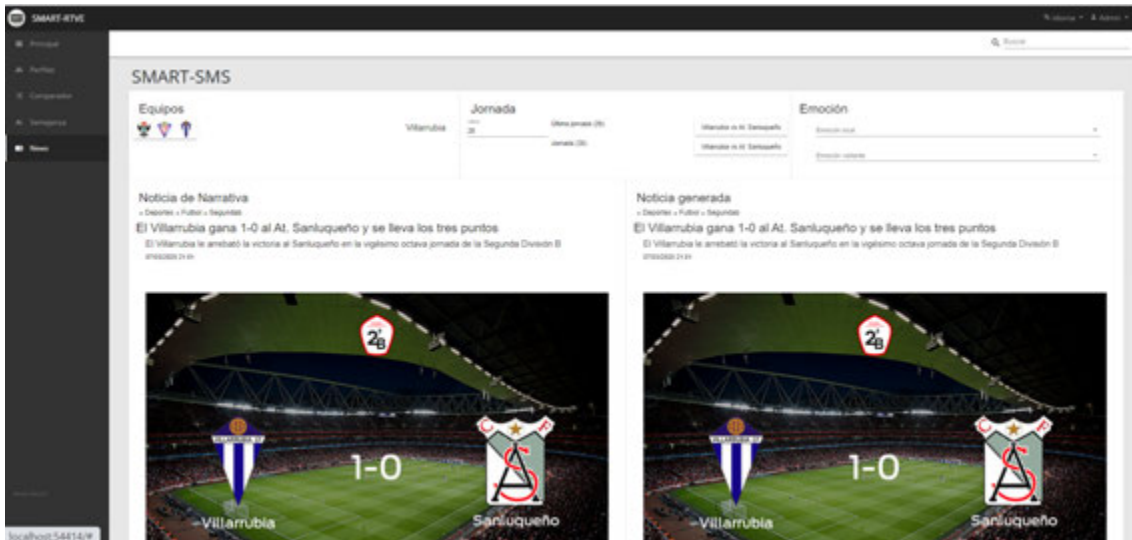


Fig. 6.2. Captura de la web después de haber seleccionado un equipo.

Una vez ha seleccionado un equipo, se cargan las jornadas disponibles para ese equipo (seleccionando la más reciente por defecto). Como se ha elegido la jornada por defecto, se cargan las emociones (seleccionándose la emoción 'blanca' por defecto). También se carga la noticia del proveedor de información en ambos lados, ya que no se ha elegido ninguna emoción.



Fig. 6.3. Captura de la web después de haber seleccionado un equipo (parte de abajo).

En la figura 6.3 se observa la parte de abajo de la página, en la que se observan los botones para elegir la versión de la noticia y aparece información de otros componentes del sistema. Estos últimos componentes no han sido realizados por el autor.

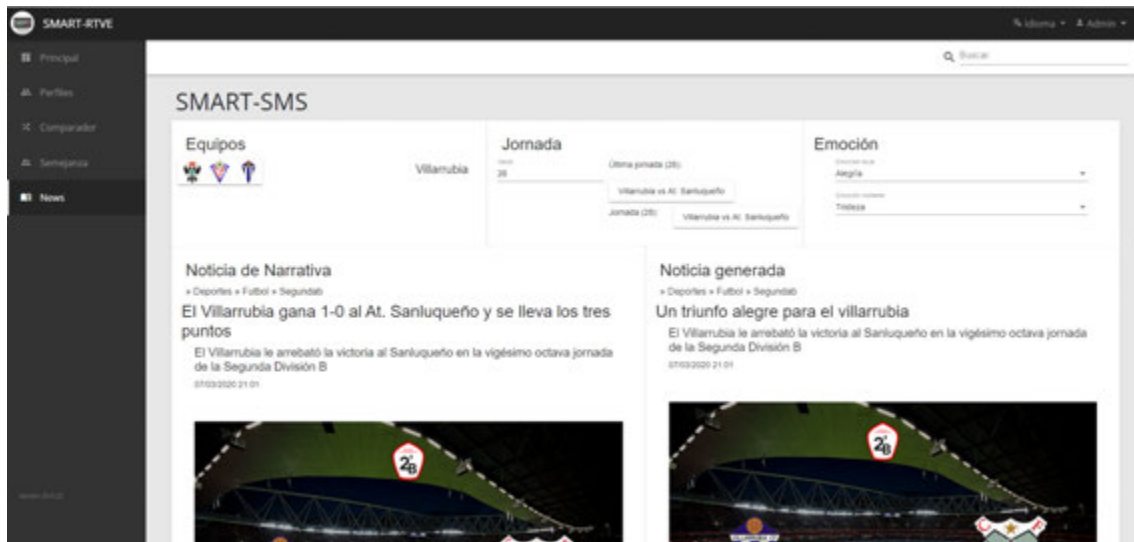


Fig. 6.4. Captura de la web después de haber seleccionado una versión.

Cuando se han seleccionado emociones distintas a la 'blanca' que se selecciona por defecto, se muestra a la derecha una versión de la noticia modificada (en este caso solo cambia en titular). Este caso puede observarse en la figura 6.4.

Por último, se muestra en las figuras 6.5 y 6.6 (en la página 55) el formato con el que se muestra la noticia en la web.

Noticia de Narrativa

» Deportes » Fútbol » Segundab

El Villarrubia gana 1-0 al At. Sanluqueño y se lleva los tres puntos

El Villarrubia le arrebató la victoria al Sanluqueño en la vigésimo octava jornada de la Segunda División B

07/03/2020 21:01

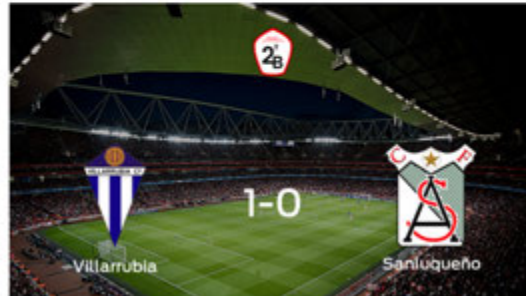


Fig. 6.5. Formato de la noticia en la web (primera parte).

El **Sanluqueño** no pudo vencer al **Villarrubia**, que ganó 1-0 durante el duelo disputado este sábado en el **Municipal**. Ambos equipos venían de perder sus últimos partidos. El **Villarrubia** llegó con ganas de volver al camino de la victoria tras sufrir una derrota por 5-0 en el partido anterior ante el **Don Benito**. Con respecto al equipo visitante, el **At. Sanluqueño** perdió por un resultado de 0-2 en el partido anterior frente al **Recreativo Granada** y llevaba una racha de cinco derrotas consecutivas. Con este resultado, el conjunto villarrubero es décimo, mientras que el **Sanluqueño** es decimoquinto tras la finalización del partido.

La primera parte del partido avanzó de cara para el equipo local, que se puso por delante gracias al gol de **Carlos Martínez**, cerrando así la primera mitad con un 1-0 en el luminoso.

Ninguno de los equipos estuvo afortunado de cara al gol en la segunda parte y por lo tanto el tiempo reglamentario acabó con el resultado de 1-0.

En el partido el árbitro amonestó con una tarjeta amarilla solamente al equipo local. Concretamente, se mostró tarjeta amarilla a **Algisi**.

Tras superar el partido, el **Villarrubia** se ubicó con 35 puntos en el décimo puesto de la tabla clasificatoria al finalizar el encuentro, mientras que el **At. Sanluqueño** se situó en decimoquinto puesto con 30 puntos.

[Villarrubia](#) [At. Sanluqueño](#) [Segunda División B](#) [España](#) [Carlos Martínez](#) [Algisi](#)

Fig. 6.6. Formato de la noticia en la web (segunda parte).

6.2. Generación de texto

Para que la generación de texto sea posible de acuerdo a los requisitos, primero se ha tenido que realizar un estudio del código existente de la herramienta *Knowledge Manager*. Al comprobar que la generación que se realizaba no tenía en cuenta las dependencias presentes en el castellano, se ha tenido que modificar su código fuente para que se produzcan en la generación. Antes de estas modificaciones, se podía generar texto gramaticalmente correcto en castellano, pero había que especificar exactamente todos los parámetros a la hora de definir los patrones. Dado que el propósito del sistema descrito es la generación de texto a partir de información conseguida de proveedores y redes sociales, se considera de gran utilidad los cambios realizados en el código base de la herramienta *Knowledge Manager*.

Como paso necesario para la generación de texto, se ha poblado la base de conocimiento con léxico relacionado con el entorno de generación (vocabulario básico, vocabulario de fútbol, vocabulario relacionado con las emociones, etc.). No obstante, no se considera relevante detallarlo en este trabajo ya que el vocabulario fue introducido a discreción del autor y no con ayuda de periodistas, lo que se considerará para futuras mejoras.

Cabe señalar que dada la complejidad del estudio de la herramienta *Knowledge Manager* y de la realización de los cambios, la generación de texto solo ha llegado hasta la generación de los titulares, pues se necesitaría más información sobre el desarrollo de la noticia y más detalles sobre la actual herramienta.

6.3. Pruebas

Las pruebas realizadas para comprobar que se han cumplido los requisitos especificados se detallan a continuación y siguen la plantilla de la tabla 6.1.

ID	
Requisitos implicados	
Descripción	
Acciones	
Resultado	

Tabla 6.1. Plantilla de tabla de prueba.

La tabla 6.1 consta de los siguientes campos:

- **ID.** Identificador único para cada prueba. Seguirá el formato P-XX, siendo cada X un dígito para su numeración.

- **Requisitos implicados.** Aquellos requisitos que se verifican en la prueba.
- **Descripción.** Breve detalle de lo que constará la prueba.
- **Acciones.** Secuencia realizada en la prueba.
- **Resultado.** Lo obtenido tras realizar la prueba.

Las pruebas realizadas se muestran en las tablas 6.2, 6.3 y 6.4.

ID	P-01
Requisitos implicados	RF-09, RF-10, RF-11 y RF-12
Descripción	Selección de parámetros en la web.
Acciones	<ol style="list-style-type: none"> 1. Cargar el entorno web. 2. Seleccionar los parámetros deseados (con la llamadas implícitas a los servicios pertinentes).
Resultado	El sistema muestra el texto de prueba correcto asociado a los parámetros introducidos

Tabla 6.2. P-01. Selección de parámetros en la web.

ID	P-02
Requisitos implicados	RF-01, RF-04, RF-05 y RF-06.
Descripción	Generación de titulares.
Acciones	<ol style="list-style-type: none"> 1. Cargar el entorno web. 2. Seleccionar los parámetros deseados (con la llamadas implícitas a los servicios pertinentes).
Resultado	El sistema muestra el titular generado acorde a los parámetros introducidos.

Tabla 6.3. P-02. Generación de titulares.

Los requisitos referidos a la base de conocimiento no se verifican con pruebas pues pertenecen a la herramienta ya operativa *Knowledge Manager*.

ID	P-03
Requisitos implicados	RF-07 y RF-08.
Descripción	Comprobación de formato.
Acciones	<ol style="list-style-type: none"> 1. Cargar el entorno web. 2. Seleccionar los parámetros de prueba de formato (con la llamadas implícitas a los servicios pertinentes).
Resultado	El sistema muestra la web con el formato requerido y ofrece una noticia con el formato de noticia deseado.

Tabla 6.4. P-03. Comprobación de formato.

6.3.1. Evaluación de la generación de textos

Para la evaluación de la generación de textos *per se* se puede apreciar un desacuerdo general en cuál es la mejor forma de hacerlo (hay una sección completa sobre este problema en [2]).

Primero, no hay un estilo o formato definido que deba tener la entrada en los sistemas de NLG [60] y [61]. Además, tampoco hay una convención establecida sobre qué salida es mejor [62].

No obstante, se han diferenciado diferentes técnicas para probar sistemas de NLG:

- **Métodos intrínsecos.** Se basan en características del propio texto.
 - **Evaluaciones subjetivas.** El sistema se prueba mostrando sus resultados a personas que puntúan los resultados. Pueden puntuar fluidez o facilidad de lectura [63] y [64]; precisión, adecuación, relevancia o si son correctos con relación a la entrada [65] o [66]; e incluso la efectividad del texto en ciertos propósitos.
 - **Parecido objetivo a redacciones humanas.** Más utilizado en sistemas que generan texto a través de un enfoque más basado en aprendizaje automático [67].
 - **Evaluaciones de género y efectividad de estilo.** En este tipo de pruebas se busca definir si el texto se adecúa a un determinado género [68].
- **Métodos extrínsecos.** Se basan en las respuestas que consigue el texto en los lectores [23].

Tampoco se tiene claro si se debería evaluar solo la salida, como una caja negra [23], o cada parte del proceso que realiza el sistema, como una caja de cristal [63]. No se aprecia tampoco relación entre los diferentes métodos, pues los resultados de las evaluaciones no muestran correlación ni causalidad de manera sustancial [69].

Teniendo en cuenta el panorama en el entorno académico respecto a la evaluación, la joven fase en la que se haya la herramienta respecto a la tarea de NLG y la falta de información durante la realización de este trabajo, se propone que se realicen pruebas de evaluaciones subjetivas más adelante. Se hubiera deseado contar con la colaboración de periodistas, pues ellos mejor que nadie conocen los aspectos para valorar una noticia, lo que aportaría valor a la herramienta.

6.4. Problemas encontrados

Durante la implementación de este proyecto se ha topado con problemas constantes de transparencia por parte de los periodistas. No se tenía información de los detalles de

la redacción y eso ha dificultado mucho la definición de patrones de calidad y del avance hacia la generación de la noticia completa. Este comportamiento se entiende dada la naturaleza del sector de los medios de comunicación en la que la información es poder.

En cuanto a la herramienta *Knowledge Manager*, al ser esta una herramienta propiedad de una empresa, no se ha podido obtener mucho soporte a través de Internet y se ha tenido que recurrir al conocimiento de los desarrolladores. Esto ha supuesto un problema de organización, pues se ha tenido que cuadrar con ellos las consultas. No obstante, prácticamente todas las consultas realizadas se solucionaron de manera muy satisfactoria.

Dentro de la generación de texto proporcionada por la herramienta *Knowledge Manager* también se han encontrado limitaciones a la hora de generar texto de la forma deseada en este caso en concreto. Esto provenía de que el uso que se le ha dado a la herramienta para este proyecto no se alineaba completamente con la funcionalidad original a la que estaba destinada. Sin embargo, esas asperezas podrían ser limadas aumentando la utilización de la herramienta en casos como el descrito en este trabajo.

7. PLANIFICACIÓN Y PRESUPUESTO

7.1. Planificación

En las siguientes líneas se detallará la planificación que fue determinada para el proyecto y hasta dónde se ha llegado en el desarrollo por el límite de la duración de la beca (el entorno tecnológico y temporal en el que la herramienta descrita ha sido diseñada e implementada).

Este proyecto ha constado de siete fases diferenciadas:

- **Definición del problema y de los objetivos.** En esta primera fase, los esfuerzos se centraron en identificar claramente el problema que se quería solucionar y concretar los objetivos que se pretendían conseguir.
- **Estudio del estado del arte.** Después, se dedicó un mes al estudio de la tecnología actual relacionada y relevante; alternativas usadas en el sector audiovisual y de los medios de comunicación, etc.
- **Análisis de sistema.** Esta fue la primera fase más cercana al desarrollo de la herramienta. En ella se centraron los esfuerzos en definir qué tenía que hacer la herramienta que se quería desarrollar. Además, como la realización del proyecto fue en el entorno de la Cátedra, en esta fase también se dedicaron notables esfuerzos en conocer y entender las herramientas y el marco en el que se iba a trabajar.
- **Diseño del sistema.** Una vez se conocieron el entorno de desarrollo y se tuvieron claras las funcionalidades que debe cumplir el sistema, se procedió a delinear cómo se quería que estuviera estructurada la herramienta.
- **Implementación.** Esta fase estuvo centrada en la creación del código en el que se basará la herramienta.
- **Pruebas.** Posteriormente, se comenzaron a crear pruebas para comprobar que la herramienta tenía el funcionamiento deseado.
- **Despliegue.** Por último, se procedió al despliegue de la herramienta en el entorno del cliente (RTVE en este caso).

En la figura 7.1 se puede apreciar un diagrama de Gantt en el que aparecen todas las fases junto a su duración temporal. La línea roja entre julio y agosto indica el final de la beca. Esto pretende ilustrar la estimación de lo que se hubiera necesitado para que el sistema que se propone en este trabajo estuviera completo.

Cabe señalar que las fases de implementación, pruebas y despliegue se solapan porque se considera que influyen las unas en las otras y, por tanto, los resultados obtenidos en una pueden provocar cambios en el resto.

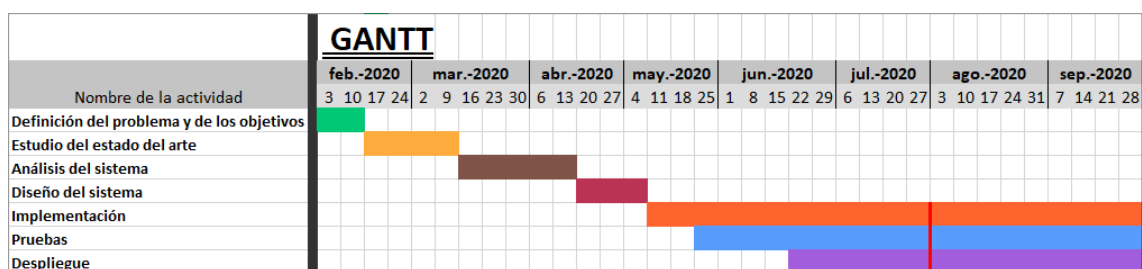


Fig. 7.1. Diagrama de Gantt del proyecto.

7.2. Presupuesto del proyecto realizado

Como presupuesto para este proyecto, se estimarán los costes reales en los que se tuvieron que incurrir a la hora de proporcionar personal y material para el desarrollo de la herramienta. Se omitirán los costes de software, puesto que en el entorno de la beca la utilización de las herramientas no ha supuesto coste adicional alguno. Todos los costes que aparecen en este apartado están en euros (€).

En la tabla 7.1 se reflejan los costes en los que se incurren por el personal. El autor trabajó en el proyecto en calidad de becario, con ayuda de algunos trabajadores de la cátedra, pero estos no se añadirán en el presupuesto para reducir ruido en el mismo. El sueldo percibido por el becario durante el proyecto ha sido 7,5 euros brutos por hora. En las semanas en las que se solapan tareas, se supondrá que se reparten equitativamente entre las tareas en cuestión.

Tarea	Coste por hora	Horas dedicadas	Coste total
Definición del problema y los objetivos	7,5 €	40	300 €
Estudio del estado del arte	7,5 €	80	600 €
Análisis del sistema	7,5 €	100	750 €
Diseño del sistema	7,5 €	60	450 €
Implementación	7,5 €	120	900 €
Pruebas	7,5 €	80	600 €
Despliegue	7,5 €	40	300 €
Total	-	-	3.900 €

Tabla 7.1. Desglose de los costes de personal para el proyecto.

Para estimar el coste del material utilizado (véase tabla 7.2), se listará el que usó el becario durante todo el proyecto. Las estimaciones están basadas en [70] y [71]. Se han depreciado el coste de los periféricos como teclado y ratón porque no se adquirieron nuevos dispositivos para la realización del proyecto. También cabe destacar que se estima una vida útil del material de 5 años.

Concepto	Coste unitario	Fracción imputable al proyecto	Cantidad	Coste total
Ordenador	518,01 €	1/10	1	51,80 €
Monitores	129,99 €	1/10	2	26,00 €
Total	-	-	-	77,8 €

Tabla 7.2. Desglose de los costes del material para el proyecto.

Respecto a los costes indirectos incurridos durante el proyecto, en los que se incluyen desplazamientos, reuniones, etc., se asumirán como un quince por ciento de los costes directos en lo que se haya incurrido.

Por último, para reflejar el presupuesto total (véase tabla 7.3), se aplicará un veintiuno por ciento de IVA, de acuerdo a [72].

Concepto	Total
Trabajadores	3.900 €
Material	77,8 €
Costes directos	3.977,8 €
Costes indirectos	551,67€
Total sin IVA	4.529,47 €
IVA (21 %)	888.19 €
Total	5417,66 €

Tabla 7.3. Desglose de los costes totales para el proyecto.

7.2.1. Presupuesto del proyecto a futuro

En este apartado se pretende aportar una breve descripción del presupuesto a futuro en el que se incurriría si se pretendiera seguir con el proyecto dos años más (a partir de septiembre de 2020) dentro de la cátedra por trabajadores cualificados. De esta forma se facilitaría la realización de presupuestos futuros y la toma de decisiones a aquellos encargados de decidirlo. Esto viene motivado por la posibilidad de futuras necesidades de actualización o revisión.

Se supondrán las intervenciones de un analista y dos programadores senior (con sus respectivos sueldos brutos por hora de 19,08 €y 21,62 €, de acuerdo a datos extraídos de *Glassdoor* [73]). Se propone repetir ciclos de seis meses con la misma estructura, para poder adecuarse a los cambios que se den en las necesidades del cliente, en la tecnología o en

los problemas encontrados. Esto supondría realizar el ciclo de seis meses cuatro veces. En la tablas 7.4, 7.5 y 7.6 se detallan los costes para el supuesto descrito de uno de los ciclos.

Tarea	Horas analista	Horas programadores	Coste total
Definición del problema y los objetivos	40 x 19,08 €	20 x 2 x 21,62 €	1.628 €
Estudio del estado del arte	60 x 19,08 €	40 x 2 x 21,62 €	2.874,4 €
Análisis del sistema	100 x 19,08 €	50 x 2 x 21,62 €	4.070 €
Diseño del sistema	40 x 19,08 €	60 x 2 x 21,62 €	3.357,6 €
Implementación	40 x 19,08 €	100 x 2 x 21,62 €	5.087,2 €
Pruebas	30 x 19,08 €	80 x 2 x 21,62 €	4.031,6 €
Despliegue	20 x 19,08 €	40 x 2 x 21,62 €	2.111,2 €
Total	6.296,4	16.863,6	23.160 €

Tabla 7.4. Desglose de los costes de personal para el supuesto de uno de los ciclos.

Para los costes de material de los cuatro ciclos, en la tabla 7.5 se supone de nuevo que la vida útil de los equipos es de cinco años.

Concepto	Coste unitario	Fracción imputable al proyecto	Cantidad	Coste total
Ordenador	518,01 €	2/5	3	612,61 €
Monitores	129,99 €	2/5	6	311,98 €
Total	-	-	-	924,59 €

Tabla 7.5. Desglose de los costes del material para los cuatro ciclos del supuesto.

Por último, el presupuesto completo de los dos años (es decir, cuatro ciclos de seis meses) se refleja en la tabla 7.6. Los costes indirectos se volverán a asumir como un quince por ciento de los directos y se aplicará un IVA del veintiuno por ciento [72].

Concepto	Total
Trabajadores (x4)	92.640 €
Material	924,59 €
Costes directos	93.564,59 €
Costes indirectos	14.034,69 €
Total sin IVA	107.599,28 €
IVA (21 %)	22595,85 €
Total	130.195,13 €

Tabla 7.6. Desglose de los costes totales para los cuatro ciclos del supuesto.

8. MARCO REGULADOR Y ENTORNO SOCIOECONÓMICO

8.1. Marco regulador

La legislación en España restringe y se aplica, en casos de distribución de software, licencias de uso y protección de la propiedad intelectual. No obstante, el software que se ha desarrollado y que se propone en este trabajo no entra dentro de ninguna de las categorías anteriores. El propósito de la herramienta mostrada sería para un uso interno, controlado y propio por parte de RTVE, sin que se distribuyera el software a terceros ni se creasen licencias de uso, y siendo los periodistas y trabajadores de la corporación los únicos usuarios de la misma. Tampoco se observa que la vida de personas vaya a depender en el correcto funcionamiento del software, con lo que no conllevará ninguna responsabilidad de esta índole.

Por otra parte, los datos que se usan en el desarrollo y durante el funcionamiento de esta herramienta son de dominio público (los resultados de los partidos, los equipos y sus componentes, la localización, etc.), lo que conlleva a afirmar que la Ley de Protección de Datos [74] no supondrá ninguna barrera a la hora de que la herramienta realice sus funciones.

Para legislación aplicable a RTVE, por favor, diríjase la sección 9.2.

8.2. Entorno socioeconómico

En todo el mundo la tecnología lleva avanzando los últimos años, y la generación de lenguaje natural no es ninguna excepción. Ya se ha mostrado en la sección 3 cómo el sector audiovisual y de medios a nivel global está adoptando cada vez más tecnologías que simplifiquen o realicen, de forma autónoma, tareas de las que los trabajadores antes eran encargados de llevar a cabo. Además, para que un medio tenga impacto en la actualidad, necesita de una fuerte presencia digital y la herramienta propuesta la fomenta.

La situación económica en España lleva un rumbo favorable en los últimos años, pero España y el mundo han sido testigos de una pandemia a nivel global. Esto ha supuesto el parón de muchos negocios y eventos de todo tipo (entre ellos los deportivos) y un golpe económico importante para muchas personas y empresas. No obstante, para el desarrollo de este trabajo se espera una nueva normalidad en la que los eventos deportivos sigan sucediendo, y que la población tenga un interés en ellos similar al actual.

Por otro lado, al ser esta una herramienta desarrollada para RTVE, y como la corporación está parcialmente financiada con fondos públicos, no se prevé problema alguno para el desarrollo, funcionamiento ni mantenimiento de la misma.

En el aspecto social, al ser una herramienta de uso interno, su impacto será limitado

y no se aprecian aspectos sociales que puedan interferir ni influir en el uso de la misma. Por otra parte, observando que la finalidad de la herramienta es la generación de textos para consumo a través de plataformas digitales, y la accesibilidad actual a estos medios, no se considera que ningún gran segmento de la población al que fueran dirigido estos contenidos tuviera problemas de acceso.

Además, un aspecto a tener en cuenta sobre el entorno social de esta herramienta son las *fake news*. Son de tal relevancia que ya hay periódicos que incluyen una sección especializada en sus parrillas informativas [75]. Mediante el uso de tecnología como la propuesta en este trabajo se podría asegurar la calidad de las fuentes de información, pues el sistema obtendría esta solo de las fuentes especificadas previamente.

Por otra parte, hay un tema social delicado que rodea a las herramientas como la que se presenta en este documento: el aspecto ético de reemplazar a una persona por una máquina. Desde este trabajo se mantiene que el uso del sistema propuesto sea el de una herramienta asistida por una persona y no el reemplazo completo de dicha persona. No obstante, como en otras industrias, la automatización sucede de forma inevitable en procesos compatibles, pero se mantiene el factor humano pues es indispensable.

Por último, se quieren aportar ejemplos de tecnología similar en el sector de los medios de comunicación en nuestro país. En el ámbito deportivo, el diario Sport lleva contando con los servicios de Narrativa desde hace un par de años para la generación de contenido [76]. Además de esto, la notable agencia EFE estableció un acuerdo también con la misma compañía para contar también con su tecnología [77].

Parte III

Mejora de procesos

9. PUNTO DE PARTIDA

9.1. Descripción de la empresa

Como ya se ha mencionado anteriormente, RTVE es la corporación encargada de ofrecer los servicios de radio y televisión en todo el ámbito español de forma pública. Sin embargo, no solo pone a disposición del espectador contenido a través de estos canales de televisión y radio. A continuación se describirán las áreas en las que opera RTVE [78]:

- **Televisión.** La división de RTVE encargada de ofrecer servicio de imagen y vídeo se llama Televisión Española (en adelante referida como TVE). Comenzó sus emisiones en 1956, siendo la más veterana en el panorama español. Los canales con los que cuenta son [79]:
 - **La 1.** El canal principal e insignia de TVE, con programación de interés general para todos los públicos.
 - **La 2.** Un canal más centrado en la cultura, cine, música, nuevas tecnologías y medio ambiente.
 - **Canal 24h.** Donde se ofrecen noticias nacionales e internacionales en español todos los días de la semana, las veinticuatro horas del día.
 - **Clan.** Espacio dedicado para el público infantil, en este canal se emiten contenidos educativos y de entretenimiento.
 - **Teledeporte.** Canal dedicado enteramente al deporte, que intenta también reflejar deportes que no son tan televisados (esquí, gimnasia, etc.).
 - **TVE Internacional.** El canal de emisión internacional centrado en los españoles residentes en el extranjero.
 - **Star TVE.** Canal internacional dedicado al público familiar americano.
- **Radio.** Radio Nacional de España (en adelante referida como RNE) es la división de RTVE encargada de la radiodifusión de contenidos. Siendo más longeva que su análoga televisiva, sus primeras emisiones se realizaron en 1937. Las emisoras que ofrece son [80]:
 - **Radio Nacional.** Emisora principal de RNE, con emisiones también de carácter general para todos los segmentos de la población.
 - **Radio Clásica.** La emisora especial centrada exclusivamente en la difusión de la cultura española y la música clásica.

- **Radio 3.** Emisora también dedicada a la cultura, pero esta vez más enfocada en contenido contemporáneo y joven. Se difunden distintos estilos de música y también cine, teatro y artes plásticas.
 - **Ràdio 4.** Emitiendo únicamente para Cataluña, es la primera emisora española que emitió en catalán. Se centra en la promoción de la cultura catalana.
 - **Radio 5 Todo Noticias.** Como su nombre indica, esta emisora está centrada en contenidos informativos.
 - **Radio Exterior de España.** La emisora en la que se emite contenido referente a España al mundo, en siete idiomas distintos.
- **Web.** RTVE cuenta con una web que desde 2008 [81] ofrece a los usuarios la posibilidad de acceder al contenido publicado por cada uno de sus otros canales y también permite dar visibilidad a la Orquesta y Coro de RTVE, así como a su Instituto.
 - **Formación.** El Instituto RTVE lleva desde 1975 [82] ofreciendo formación para personas que busquen un futuro en el sector de los medios audiovisuales.
 - **Música.** RTVE también cuenta con una Orquesta y un Coro. Estos tienen como propósito la difusión e interpretación de la música clásica, así como la composición de esta para cualquiera de las otras ramas de RTVE (sintonías de programas, música de series o películas de producción propia, etc.).
 - **Innovación y emprendimiento.** Impulsa Visión RTVE es la marca que tiene la corporación que apoya proyectos de relacionados con startups, empresas e investigación en el ámbito universitario relacionado con el sector audiovisual [83].
 - **Redes sociales.** Aunque no aparecen listadas en su página web, RTVE cuenta con perfiles presentes en las principales redes sociales (Twitter, Facebook, Instagram, etc.). Estos perfiles suponen otra fuente de difusión importante para RTVE.

Todas estas áreas de RTVE están alineadas en torno a los siguientes ejes que se pueden encontrar en su página web [84]:

- **Misión.** *Refleja la razón de ser de la empresa y cómo pretende llegar a cumplir su visión.*

RTVE responde a su carácter de empresa de comunicación de servicio público. Este la obliga a garantizar una información rigurosa, independiente y plural y un entretenimiento de calidad. A fomentar el debate, la innovación y la creación. A apoyar la difusión de las artes, la ciencia y la cultura. Todo ello bajo las premisas de cohesionar y dar cauce a la participación."

- **Visión.** *Indica el futuro hacia el que se quiere dirigir la empresa.*

"Desde su función de servicio público, RTVE mantiene la vocación de ser el medio de comunicación de referencia en España. Quiere estar presente en todos los acontecimientos sociales relevantes y ser un espejo en el que se vea reflejada en toda su variedad y riqueza la realidad nacional e internacional. Realizará su trabajo con criterio estrictamente profesional y atendiendo la encomienda de difundir valores constitucionales."

- **Valores.** *Son los principios morales y éticos por los que se rige el funcionamiento de la empresa.*

RTVE defiende y promociona en toda su programación los valores constitucionales, especialmente los de libertad, igualdad, pluralismo y tolerancia, sobre los que se asienta la convivencia democrática. Los valores de RTVE los marca la ley y se reflejan en lo que hacemos."

9.2. Análisis del negocio y su entorno

Durante este trabajo el foco va a estar situado en la web de RTVE y sus cuentas en las redes sociales. Ambos formatos publican contenido y en ambos está presente la redacción de textos. Las herramientas que se han decidido utilizar para este análisis son el Business Model Canvas [85] y el análisis PESTEL [86]. El primero nos aportará información sobre el negocio de la empresa y el segundo sobre su entorno macroeconómico. Se ha optado por no utilizar otras herramientas como las cinco fuerzas de Porter [87] o el análisis DAFO [88], ya que se considera que el la herramienta Business Model Canvas aporta suficiente información para la realización de este trabajo.

9.2.1. Business Model Canvas

Para dar un análisis del negocio en el área web y de las redes sociales de RTVE, se utilizará el Business Model Canvas de Osterwalder [85] (puede observarse el caso de RTVE en la figura 9.2 en la página 78). Según el autor, un modelo de negocio es una herramienta que "describe cómo una organización crea, provee y retiene valor" [85] y el propone un 'lienzo' sobre el que plasmar dicho modelo, siendo la mitad izquierda del lienzo centrada en la eficiencia y la mitad derecha centrada en el valor (obsérvese la figura 9.1). Esta representación se basa en nueve puntos del negocio que tiene en cuenta tanto aspectos puramente internos, como aspectos en los que el negocio o la compañía se relacionan con diversos actores y factores. La representación del negocio de RTVE mencionado puede observarse en la figura 9.2 (página 78).

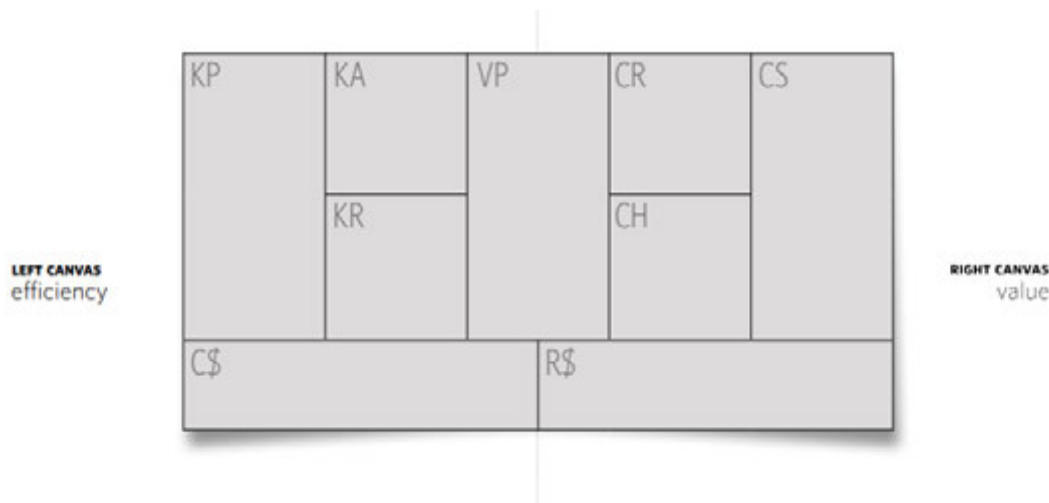


Fig. 9.1. División del Business Model Canvas en eficiencia y valor [85].

Segmentos de clientes

Del inglés *Customer Segments (CS)*. Aquí se deben identificar a quiénes quiere alcanzar y dar servicio la organización. Para ello, los clientes se dividen en grupos conocidos como segmentos y cada uno puede diferenciarse por: las necesidades que tienen, los canales mediante los que se les contacta, las relaciones que requieren, el beneficio que se puede obtener de ellos, o los aspectos del producto o servicio por los que están dispuestos a pagar. En función de los segmentos a los que se quieran alcanzar, el segmento de cliente que se persigue puede ser:

- **Mercado masivo.** Si este es el objetivo de la organización, no hay interés alguno en diferenciar entre los distintos segmentos y se busca alcanzar a un grupo grande y poco diferenciado.
- **Nicho de mercado.** Cuando el segmento al que se quiere dirigir una empresa es uno muy específico y especializado, se considera que se dirige a un nicho de mercado.
- **Segmentado.** A través de este foco, la organización pretende centrarse en segmentos poblacionales ligeramente distintos (por ejemplo, dos grupos tienen básicamente las mismas necesidades pero un poder adquisitivo distinto).
- **Diversificado.** Al contrario que el enfoque anterior, si la organización pretende diversificarse en este sentido, tratará de alcanzar a segmentos de la población que no tienen relación alguna y con distintas necesidades y problemas.
- **Plataforma multilateral.** Si se da el caso en el que la organización necesita de dos partes opuestas, muy posiblemente se trata de una plataforma multilateral. Un claro ejemplo de ello es una plataforma que ofrece contenido gratuito. Para que funcione, necesitará tanto atraer muchos consumidores para poder atraer así a patrocinadores

y agencias de publicidad y empresas que quieren darse a conocer, pero también deberá adecuar su contenido y formato a estos últimos para no perder atractivo.

Se ha identificado que RTVE está centrado en este tipo de clientes ya que ofrece contenido gratuito y su caso es prácticamente el descrito en el párrafo anterior (aunque cuente también con fondos públicos).

Propuesta de valor

Del inglés *Value Proposition (VP)*. En este apartado se debe identificar y definir la razón por la que los clientes deberían escoger a la organización y lo que les crea valor al hacerlo. El autor propone los siguientes tipos de propuesta de valor:

- **Novedad.** Este tipo de propuesta de valor se basa en productos o servicios que no existían anteriormente o que suponen un cambio radical a los actuales.
- **Rendimiento.** Si lo que la organización ofrece supone mejores prestaciones que otras alternativas de mercado, esta es su propuesta de valor.
- **Personalización.** Ofrecer al cliente algo específicamente adecuado a su perfil y sus necesidades se identifica con esta propuesta de valor.
- **Externalización del servicio.** Las empresas suelen encargar a terceros que pueden realizar tareas que o bien ellas no podrían realizar, o las podrían realizar de forma menos eficiente. Esta propuesta de valor se basa en este hecho.
- **Diseño.** La propuesta de valor se identifica con esta si la forma, el color, los detalles o cualquier característica del diseño del producto o servicio es lo que atrae al cliente.
- **Marca.** Hay grupos de clientes que quiere sentirse parte de un grupo selecto, y las marcas con una fuerte presencia son la propuesta de valor que permite atraerles.
- **Precio.** Cuando la organización es capaz de ofrecer prácticamente lo mismo que la competencia, pero a un precio menor, eso suele llamar la atención de los clientes. En ese principio se basa esta propuesta de valor.
- **Reducción de costes.** A diferencia de la propuesta de valor anterior, en esta los clientes consiguen ahorrar costes en sus procesos.
- **Reducción de riesgos.** Siguiendo el mismo principio que con la reducción de costes, pero esta vez con los riesgos se identifica esta propuesta de valor.
- **Accesibilidad.** Si productos o servicios comunes se hacen accesibles a otros segmentos de población que antes no podían llegar a ellos, se trata de esta propuesta de valor.

- **Comodidad/Conveniencia.** Se da cuando la característica diferencial del producto o servicio que ofrece la organización es la facilidad de uso o adquisición.

Teniendo en cuenta las propuestas anteriores, se ha advertido RTVE por las propuestas de valor de precio, marca y comodidad. La personalización porque los contenidos a los que acceden sus usuarios son exclusivos de RTVE (aunque haya alternativas en otros competidores). También son cómodos, ya que disponen de una gran variedad de ellos sin salir de su plataforma y son inmejorables en precio ya que no suponen ningún coste del que se pueda deshacer al cliente.

Canales

Del inglés *Channels (CH)*. Los canales son los medios a través de los cuales la organización se comunica y llega al cliente para hacerle llegar la propuesta de valor. Hay cinco fases que cada canal elegido por la organización puede llevar a cabo durante todo el proceso: comunicación, evaluación, compra, entrega y posventa. En la comunicación, el canal informa al cliente de su propuesta de valor y mediante la evaluación ayudan al cliente a valorar dicha propuesta. Durante la compra y la entrega, se establece cómo se permite al cliente adquirir los bienes o servicios y posteriormente, cómo se le entregan. Por último, la posventa engloba todos los esfuerzos que realiza la organización por atender al cliente en lo que necesite una vez realizada la venta.

Además de las fases descritas en el párrafo anterior, los canales se pueden clasificar según a quién pertenezcan (propios si pertenecen a la organización y de socios si pertenecen a terceros) y según el contacto que haya entre el cliente y la organización (directos e indirectos). A continuación se clasifican los distintos tipos de canales: fuerza de ventas (propio y directo), web propia (propio y directo), tiendas de la organización (propio e indirecto), tiendas o web de socios (de socios e indirecto) y mayoristas (de socios e indirecto).

En este caso, se considera que RTVE cuenta con un canal propio y directo, como es su página web y un canal de socios y directo, como las redes sociales (aunque el servicio lo ofrezca la red social, el contenido es directamente creado por RTVE). Ambos canales sirven a la organización para comunicarse con el cliente y permitir que le evalúe. También ofrecen el contenido y lo “entregan” de forma inmediata, además de ofrecerle al consumidor la posibilidad de contactar con RTVE, lo que supondría el servicio posventa.

Relaciones con clientes

Del inglés *Customer Relationships (CR)*. Aquí se define específicamente cómo interactúan los clientes con la organización. Cabe destacar que no son iguales a los canales que la organización haya definido, ya que la organización podría tener distintas relaciones

con un cliente en el mismo canal, o emplear la misma relación en dos canales distintos. Debajo se muestran los diferentes tipos de relaciones:

- **Asistencia personal.** Este tipo de relación se da cuando una persona de la organización interactúa con el cliente, pero no está disponible específicamente para ese cliente.
- **Asistencia personal dedicada.** Siendo similar a la anterior, este caso se da cuando se asigna a una persona de la organización o presente en el canal para dedicarle atención individual, seguimiento, etc. al cliente. Sirve para que el este tenga una sensación de atención mayor.
- **Autoservicio.** En estas relaciones, la organización no tiene ningún contacto directo con los clientes, ya que deja a su disposición todo lo que podría querer y necesitar para que ellos se sirvan de forma independiente.
- **Servicio automatizados.** Normalmente acompañan a relaciones de autoservicio y permiten realizar servicios o entregar productos específicos para cada cliente de forma automatizada. La personalización suele ir ligada a este tipo de relaciones.
- **Comunidades.** Mediante este tipo de relaciones, las organizaciones permiten al cliente también ponerse en contacto con otros que tengan sus mismos intereses o que hayan consumido los mismos productos o servicios.
- **Creación colaborativa.** Además de ofrecer comunidades a los clientes, hay organizaciones que ofrecen a los clientes aportar valor al proceso. Esto toma formas como las reseñas, participar en el diseño del producto (personalización), etc.

A la luz de este esquema de relaciones, se ha determinado que RTVE mantiene relaciones de autoservicio con sus clientes, puesto que la organización publica su contenido y son los clientes los que pueden decidir consumirlo cuando deseen. Además, ofrecen comunidades y creación colaborativa, ya que pueden interactuar tanto en su página web como en las redes sociales, y también dejar reseñas y participar en encuestas sobre contenido a través de los canales ya mencionados.

Fuentes de ingresos

Del inglés *Revenue Streams (RS)*. Una organización no puede funcionar si no recibe unos ingresos por sus actividades y es importante identificar de dónde proceden esos ingresos. Hay seis posibles fuentes:

- **Venta de activos.** Siendo esta la fuente más común y obvia de ingresos, esta se da cuando el cliente paga por la obtención de un producto físico.

- **Cuota de uso.** Esta es la fuente de ingresos cuando el cliente paga en función de la utilización de un servicio. Cuanto más lo usa, más paga.
- **Cuota de suscripción.** Siendo similar a la anterior, el cliente esta vez paga una cantidad fija por la utilización de un servicio. Esta cantidad fija tiene que ser abonada cada cierto tiempo para que el cliente pueda seguir disfrutando del servicio.
- **Préstamo/alquiler/leasing.** En estos casos, la organización permite al cliente la utilización exclusiva de un activo por una duración de tiempo determinada. Cuando es un préstamo, el cliente debe hacerse cargo de los gastos, pero en los dos últimos casos los gastos se reparten como decida la organización.
- **Licencias.** Este tipo de fuente de ingresos permite a la organización generar ingresos a partir de propiedad intelectual. El cliente paga por el derecho a utilizar dicha propiedad a cambio de una cantidad de dinero.
- **Comisiones por intermediación.** Cuando el cliente contrata a una organización para que medie por ella en un proceso y esta recibe una cantidad a cambio, es considerado una comisión por intermediación.
- **Publicidad.** La organización recibe ingresos de otras organizaciones por publicitar sus productos, servicios, marcas, etc.

En el caso de RTVE, se ha obviado la aportación pública a la organización ya que es ajena a ella, para centrarse en los ingresos sobre los que puede influir de forma directa. Teniendo en cuenta la fuentes anteriores y conociendo la naturaleza de RTVE, la fuente de ingresos que tiene la organización es la publicidad. Aunque por ley no está permitida la publicidad directa en sus plataformas, los patrocinios y otras formas de publicidad indirecta son actividades que aportan ingresos a RTVE.

Recursos clave

Del inglés *Key Resources (KR)*. Estos se definen como los activos importantes para la organización a la hora de que su modelo de negocio funcione. Los puede haber de cuatro tipos: físicos (normalmente activos), intelectuales (como marcas, patentes, etc.), humanos (cualquier trabajador de la organización) y financieros (líneas de crédito, garantías financieras, etc).

Los recursos de RTVE identificados han sido sus periodistas y su marca. Los periodistas son los que hace a RTVE lo que es (en sus distintos puestos y con sus respectivas funciones) y los que han dotado a la marca RTVE el valor que tiene. Sus periodistas son un recurso humano mientras que su marca es un recurso intelectual.

Actividades clave

Del inglés *Key Activities (KA)*. En la misma línea que el apartado anterior, estas actividades son aquellas cruciales para el correcto funcionamiento del modelo de negocio de la organización. Se pueden clasificar en tres grupos: las que producen un bien o servicio para vender al cliente, producción; las que resuelven problemas que tengan los clientes, resolución de problemas; y las que ofrecen los medios a los clientes para que realicen sus actividades o satisfagan sus necesidades, plataforma o red.

RTVE centra sus esfuerzos en dos actividades: la creación de contenido y la presencia virtual. La primera es una actividad de producción, puesto que sus periodistas deben generar ese contenido que va a ser ofrecido al cliente. Para que este servicio pueda ser ofrecido, RTVE tiene que mantenerse presente de forma virtual mediante su web (actividad de plataforma) y sus cuentas de redes sociales.

Socios clave

Del inglés *Key Partners (KP)*. Como los dos factores anteriores, una organización necesita una red de proveedores y socios para que su modelo de negocio funcione. Para identificarlos, el autor ofrece tres tipos de asociaciones con ellos:

- **Optimización y economías de escala.** Hay situaciones en las que las organizaciones comparten recursos con otras para que les salga más rentable su utilización.
- **Reducción de riesgos e incertidumbre.** Cuando nuevas inversiones están rodeadas por incertidumbre para las compañías, suelen formarse alianzas entre ellas para reducir los riesgos para paliar esa incertidumbre.
- **Adquisición de recursos o actividades.** Normalmente las organizaciones no están a cargo del cien por cien de todas las actividades detalladas en su modelo de negocio. En algunos casos, cuentan con terceros para que realicen algunas por ellas, ya que estos pueden tener mejores medios, más experiencia, etc.

En el caso de RTVE, las agencias de las que obtenga información y los proveedores de los servicios web que ofrece son sus principales socios claves a la hora de llevar a cabo su modelo de negocio.

Estructura de costes

Del inglés *Cost Structure (CS)*. Las diferentes organizaciones poseen distintas prioridades con respecto a los costes. Algunas se centran en minimizar los costes a los que incurrirán al máximo (aquellas enfocadas en los costes) y otras en que el valor de lo que produzcan o generen sea el mayor posible (aquellas centradas en el valor).

A parte de la distinción entre los focos de las compañías respecto a sus costes, las estructuras de costes tienen las siguientes características:

- **Costes fijos.** Estos costes son aquellos en los que la organización incurre sea cual sea su nivel de producción.
- **Costes variables.** Al contrario que los anteriores, este tipo de costes dependen directamente del nivel de producción al que se haya llegado.
- **Economías de escala.** A medida que el volumen que una empresa puede producir aumenta, suele poder reducir costes debido a la cantidad de productos o servicios que produce y necesita para producir.
- **Economías de alcance.** Si una organización puede aprovechar sus recursos para varias actividades distintas al mismo tiempo, se dan este tipo de economías.

RTVE está centrada en generar el máximo valor y no hay ningún indicio que confirme que pretende reducir sus costes al mínimo (todas las empresas quieren tener menores costes, pero no todas priorizan recortar costes antes que entregar valor al cliente). Los costes principales que deberá vigilar serán los relacionados con el desarrollo y mantenimiento de su web (dada la importancia observada dentro de su modelo de negocio) y los relacionados con la información que obtiene de proveedores, pues le permite generar contenido a partir de ella.

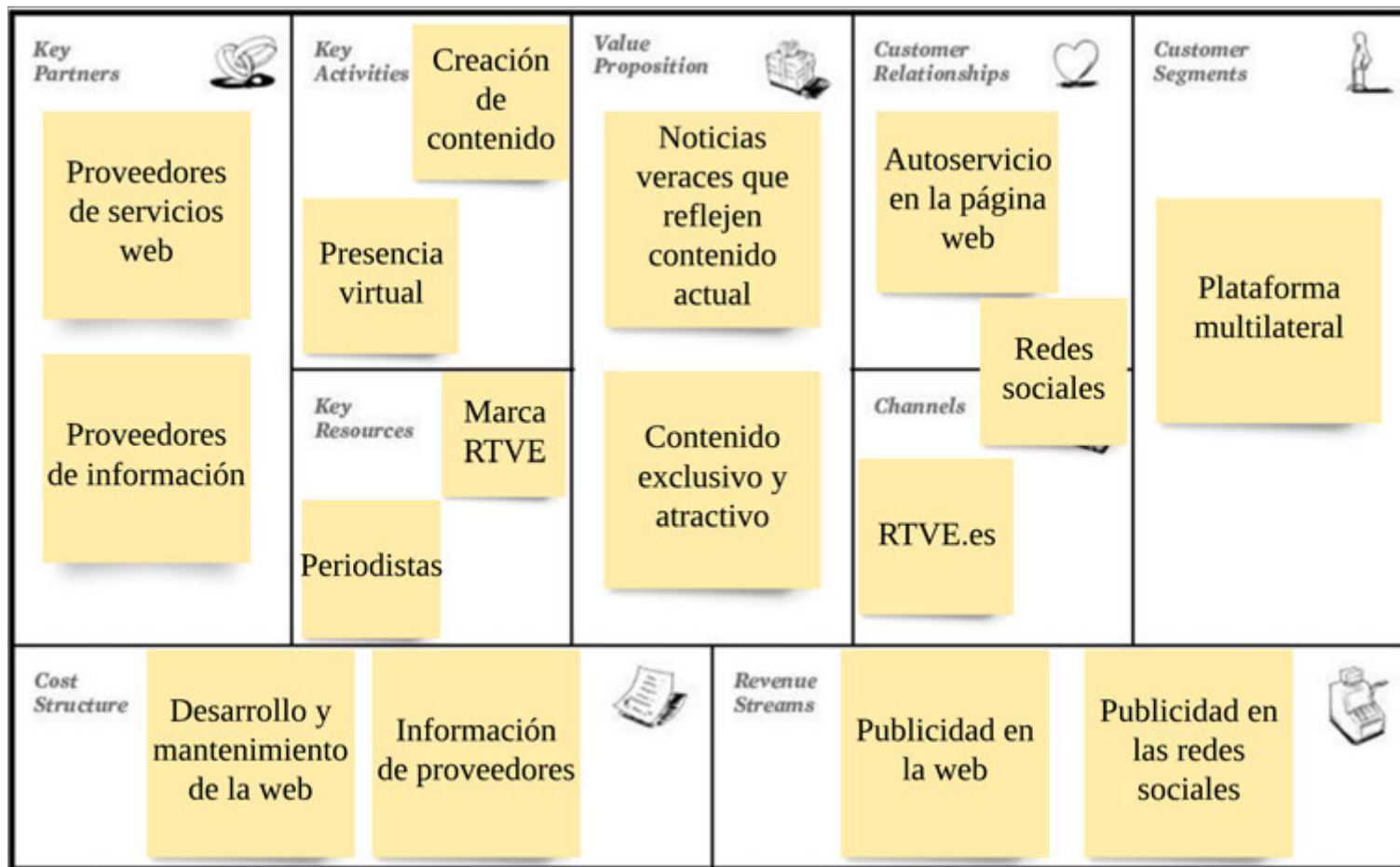


Fig. 9.2. Business Model Canvas del rea web y de redes sociales de RTVE.

9.2.2. Análisis PESTEL

Este análisis fue utilizado por primera vez [89] por Liam Fahey y V.K.Narayanan en 1986 en [86], siendo una ligera ampliación del modelo utilizado por Francisco Aguilar en 1967 en [90]. Para que el análisis no resulte tedioso ni redundante, se excluirá información que no afecte de forma directa a RTVE y se intentará enfocar hacia la misma.

Factores políticos

A pesar de que es una empresa dedicada a la divulgación de información veraz, al ser un ente público, RTVE siempre estará ligada en mayor o menor medida al gobierno que sea elegido en cada periodo. El gobierno actual (elegido en el año 2019) está formado por una coalición entre el PSOE (Partido Socialista Obrero Español) y Unidas Podemos, y no cambiará (salvo que se den circunstancias extraordinarias) hasta 2023. Esto supondrá de manera más que probable que si el transcurso de la legislatura sucede con normalidad, RTVE no se debería ver afectada por este aspecto. Además, el último cambio en la cúpula directiva de la corporación, por parte de El Gobierno, sucedió antes de las últimas elecciones [91], por lo que se estima que no habrá más cambios en este frente (los últimos cambios en la dirección sucedieron en julio de 2020 [92]).

Factores económicos

España estuvo sometida a una crisis económica desde aproximadamente 2008 hasta 2012. No obstante, como se puede comprobar a través de el gráfico (véase figura 9.3) que muestra el desempleo en España (uno de los indicadores de crecimiento económicos), hace años que España salió de esa situación.

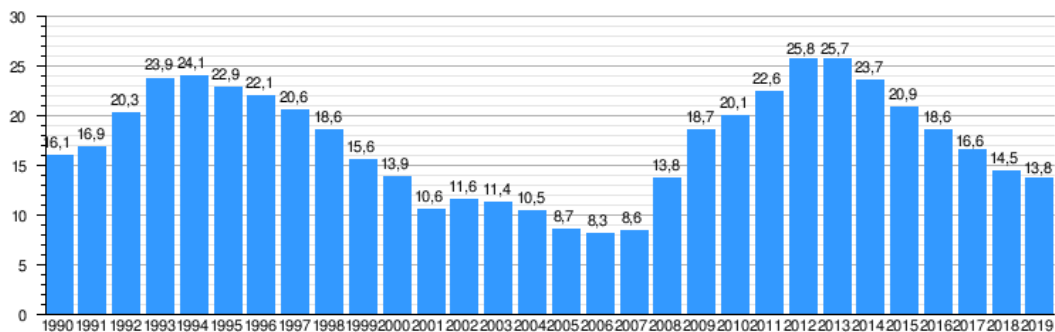


Fig. 9.3. Evolución de la tasa de desempleo en España [93].

Sin embargo, los datos del año actual aún no están disponibles al completo y muy probablemente pudiera observarse en ellos una cifra devastadora en referencia a él. No obstante, en la línea que se ha comentado en el apartado 8, se espera una evolución hacia

una normalidad similar a la previa a la pandemia sufrida, salvando las distancias pertinentes.

Por otro lado, RTVE cuenta con apoyo gubernamental lo que redundará en que siempre tendrá un respaldo económico por los fondos públicos con los que cuenta.

Factores sociales

La pirámide poblacional española actual (véase la figura 9.4) tiene una forma invertida. Esto quiere decir que hay más gente de edades más altas que gente joven. Esto, en lo relativo a la tecnología, podría haber supuesto en un pasado un problema, ya que para las generaciones más avanzadas en edad se hubiera esperado un rechazo o una menor adaptación a la tecnología. No obstante, hoy en día la tecnología se ha hecho accesible para un gran número de personas y cada vez más gente, lo cual no supone ningún problema para RTVE. Es más, se observa como un entorno social favorable.

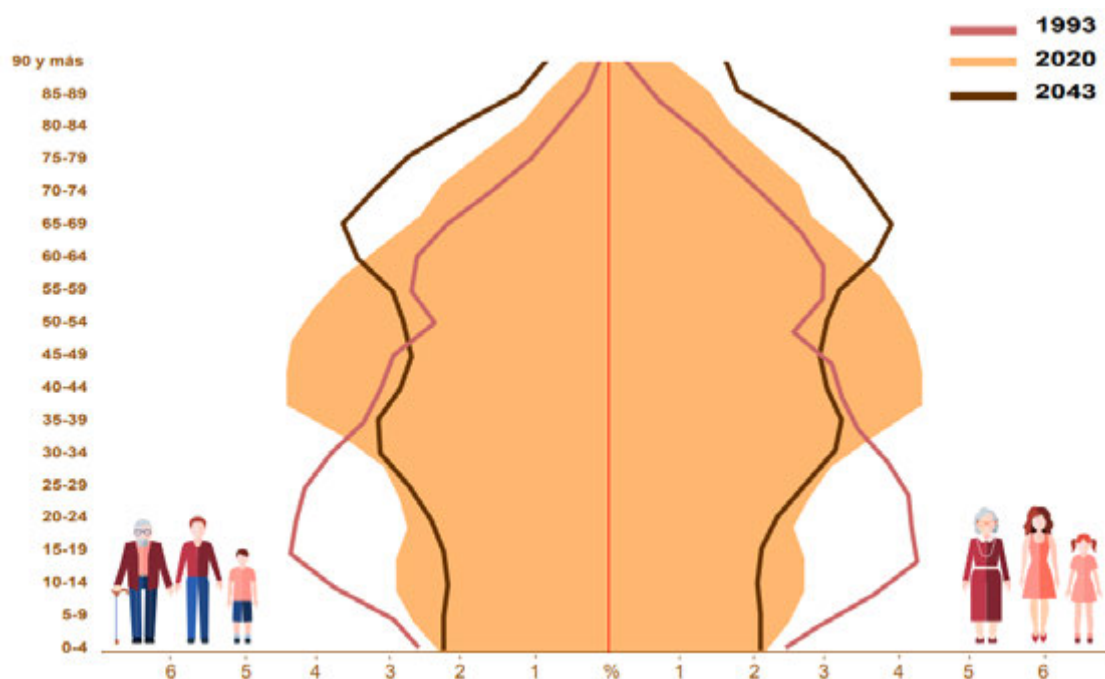


Fig. 9.4. Pirámide de población en España en el pasado, el presente (a 1 de enero de 2020) y el futuro [94]^a

^aEsta gráfica refleja la población y sus niveles previos a la pandemia. El autor de este trabajo es consciente de que un número devastador de vidas han sido perdidas en los últimos meses y que la pirámide actual bien podría diferir notablemente de la proporcionada. No obstante, nótese su uso meramente ilustrativo.

Factores tecnológicos

Como ha sido mencionado ya a lo largo de este trabajo, la tecnología ha avanzado muchísimo en los últimos años. De hecho, ha sido ese avance el que ha sentado los ci-

mientos sobre los que se basa este trabajo. La creación de contenido es un aspecto que ha estado siempre presente y ahora con el entorno digital actual aún más. Los cambios en las tecnologías de transmisión, en las formas de visualización, en las redes sociales podrían suponer un cambio en la forma en la que RTVE tuviera que enfocar su negocio, pero no se prevé ningún cambio que impida o dificulte que RTVE siga ejecutando el negocio descrito anteriormente.

Factores medioambientales

La generación de contenido y actividades en las redes sociales, bajo el prisma de este trabajo, no se ven sujetos a factores medioambientales evidentes ni notables.

Factores legales

Al ser un ente público, RTVE se ve sujeta a una fuerte legislación que debe cumplir [95]. El primer caso es la Ley 17/2006 [96] de la Radio y la Televisión de Titularidad Estatal, que regula el servicio público de estos formatos y establece su régimen jurídico. El segundo es el mandato-marco [97] a la Corporación RTVE previsto de la ley mencionada anteriormente cuyo objetivo es concretar los objetivos generales de la función de servicio público esencial. Por último, RTVE se ve sujeta a la Ley 8/2009 [98] de financiación de esta corporación en la que se pretende regular el sistema de financiación de RTVE y de sus filiales, además de que se establecen obligaciones adicionales que se exigen para la prestación de los servicios públicos encomendados.

9.3. Descripción del proceso

El proceso en el que se quiere profundizar durante este trabajo es la creación de contenido. En especial de noticias y de publicaciones en redes sociales. No obstante, no se dispone de información clara sobre el proceso exacto que se sigue en la creación de estos contenidos. Hay que tener en cuenta que la información es poder y más aún en el sector audiovisual. Revelar información sobre qué proceso siguen podría suponer que otros competidores se beneficiaran de ella, explotándola para su propio beneficio. Por ello se aportará un esquema general sobre el que poder basar el desarrollo de esta mejora. Para definirlo se utilizarán como base el siguiente manual de redacción de la Universitat Jaume I [99] y también el proceso ilustrado en [100]. Cabe destacar que este proceso se tendrá en cuenta tanto para publicaciones que se realicen de noticias en la web de RTVE como en sus cuentas de las redes sociales.

Durante el desglose del esquema, se atribuirá una duración temporal a cada fase y tarea en base a una duración total hipotética. Es decir, si para realizar todo el proceso el periodista tardase cien unidades de tiempo, a cada parte del proceso podría atribuírsele una fracción de ese total. Se ha optado por esta opción porque no se tiene un esquema fiel

proporcionado por los periodistas del proceso y sin él no hay forma de saber el tiempo que lleva redactar un texto (sería complicado incluso teniéndolo porque podría variar la duración del proceso completo en función de lo complicado que fuera investigar sobre el tema de la noticia, lo polémico que fuera, etc.). Además, utilizando esta representación para la duración del proceso, podremos cuantificar de forma general la mejora que se pueda producir puesto que hablaremos de qué porcentaje se ahorra el periodista en lugar de de qué cantidad de minutos que pudiera evitarse en el proceso.

A continuación se ofrece el esquema (ilustrado en la figura 9.5) que se considerará como referencia en este documento, en el que se ordenan cronológicamente las fases y tareas identificadas en él:

1. **Planificación.** [Duración: 60 de 100] Fase previa a la redacción en la que el periodista obtiene y ordena información sobre la que quiere escribir. También identifica y define a quién quiere dirigirla y puede entrevistar a personas de interés para el texto.
 - a) **Investigación.** [Duración: 15 de 100] Primera tarea a realizar en esta fase en la que el periodista recaba información acerca del tema sobre el que versará el texto.
 - b) **Identificación de hechos.** [Duración: 10 de 100] Después, hay que identificar de entre toda la información encontrada qué hechos son sobre los que se quiere informar en el texto.
 - c) **Diseño del esquema del texto.** [Duración: 5 de 100] Como tercera tarea, el periodista debe detallar el orden en el que presentará la información, así como en qué partes dividirá el texto.
 - d) **Localización de audiencia.** [Duración: 10 de 100] No refiriéndose solo a localización geográfica, en esta tarea el periodista elige o identifica a quién va a ir dirigido el texto.
 - e) **Toque personal.** [Duración: 5 de 100] Dependiendo del tipo de texto y el tipo de audiencia, el periodista podría querer acentuar su autoría en el texto añadiendo ciertas oraciones, frases y/o expresiones.
 - f) **Entrevistas.** [Duración: 15 de 100] No siendo siempre una tarea necesaria, la realización de esta puede aportar testimonios o fuentes de información que mencionar o sobre los que basarse en el texto.
2. **Redacción.** [Duración: 30 de 100] Fase en la que el periodista debe reflejar los datos de acuerdo al esquema que ha diseñado previamente y después redactarlos para que se dejen ver los hechos de forma clara, centrándose en la audiencia que desee y teniendo en cuenta la información que haya recibido y recogido en la fase anterior.

3. **Revisión.** [Duración: 10 de 100] Fase final en la que el periodista comprueba su la calidad de su trabajo y la adecuación de lo que ha producido a lo que se quería producir en un principio.

- a) **Revisión propia.** [Duración: 5 de 100] En esta tarea el periodista mismo repasa el texto que ha creado, buscando posibles fallos que corregir, datos que se hayan quedado sin reflejar o cambios que pudiera necesitar aplicar al texto.
- b) **Revisión externa.** [Duración: 5 de 100] Como última tarea, el periodista podría querer o necesitar de una segunda opinión. Puede darse el caso en el que sea obligatoria la supervisión de alguien superior en la jerarquía de la empresa.

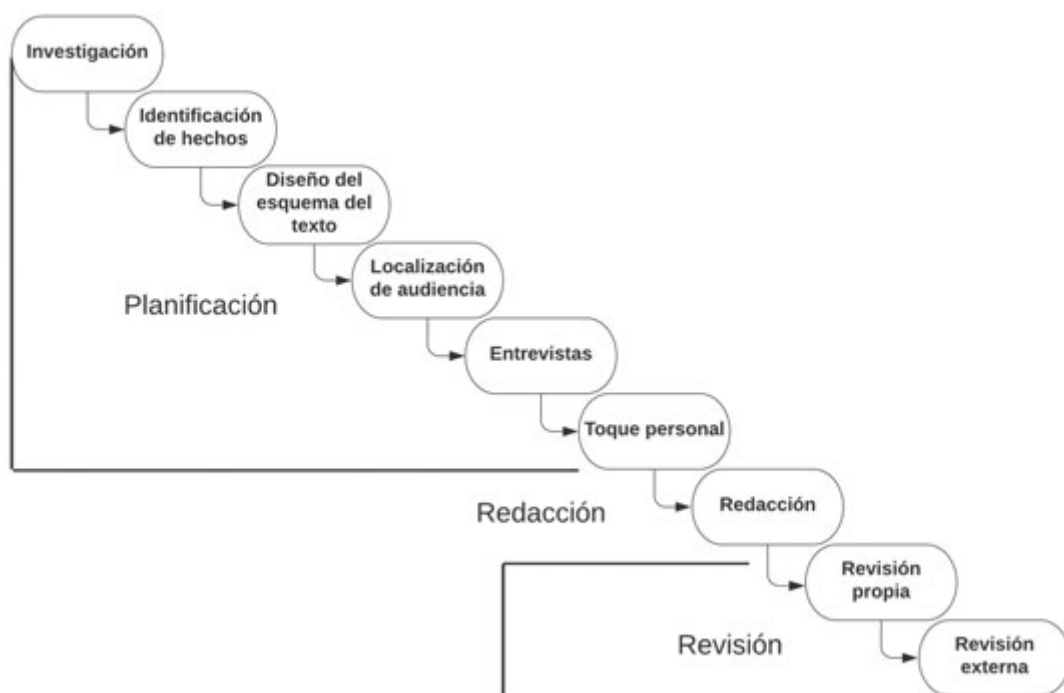


Fig. 9.5. Esquema del proceso de redacción propuesto.

9.4. Descripción del problema

Aunque no se puede decir que el proceso descrito sea o tenga un problema *per se*, se puede afirmar con total rotundidad que a día de hoy la tecnología es capaz de ahorrar tiempo en parte del proceso al periodista, llegando a poder realizar tareas por este de forma autónoma. Desde este trabajo quiere plantearse que el proceso en la forma descrita para los tiempos que corren y la tecnología es un problema para el periodista si tiene que realizar cada tarea de forma personal. Esto se basa en que el periodista podría aprovechar el tiempo que dedica en gran parte del proceso delegando tareas a una máquina, realizando otras tareas que solo él pueda hacer, ya sea porque posee los conocimientos necesarios

para ello, porque tiene un criterio humano intransferible a una máquina, porque siente y padece.

Dada la cantidad de fuentes públicas de información que hay en la actualidad y la existencia de software capaz de obtener información, sintetizarla y organizarla, el periodista podría dedicar el tiempo de investigación a investigar alguna parte del tema o alguna fuente que requiriese más conocimiento o presencia personal, por ejemplo. Este supondría que el periodista podría centrarse esas excepciones y supervisar, dejando el resto a una herramienta software.

La identificación de hechos también es un proceso que podría ser delegado en una herramienta software, ya que la tecnología actual permite a las computadoras la extracción de elementos importantes presentes en un texto (véase el ejemplo de The New York Times en 3.2).

Es más que probable que el diseño del esquema de los distintos textos ya sea un proceso maquinal y que se tengan guardadas plantillas, guías o esquemas sobre cómo redactar diferentes textos. No obstante, estos recursos podrían no estar estandarizados o actualizados, lo que también supondría un problema durante el proceso.

La localización de audiencia es algo que sin duda solo puede hacer el periodista si nos referimos a la elección e identificación de los lectores. Sin embargo, el criterio de elección o los distintos intereses de ciertas audiencias a las que se dirige con frecuencia, así como las expresiones que se usan para cada uno de los grupos son parámetros que el profesional podría introducir también en un sistema software.

Si suponemos que todo periodista deja su toque personal (en mayor o menor medida) en cada texto que crea, esto es algo que el profesional podría también dejar grabado en un sistema para que este lo replicase.

Las entrevistas son algo en lo que las máquinas no pueden sustituir a las personas. Ya sea por el trato humano, la información que una persona pueda extraer de otra que vaya más allá de las palabras que dice (el contexto, la comunicación no verbal, los sentimientos, etc.). Esta parte del proceso está reservada únicamente para el periodista y, aunque pudiera ser optimizada mediante uso de la tecnología en ciertos casos (comunicación remota en vez de presencial si el asunto lo acepta, etc.), no es un aspecto que se pueda ni se desee automatizar completa ni parcialmente.

La parte de redacción dentro del proceso, aunque no se hayan identificado tareas para ella, es una parte troncal y bastante costosa en lo que a tiempo se refiere. Dado los avances tecnológicos en los últimos años y la herramienta propuesta en este trabajo, el periodista podría, previa indicación de ciertos parámetros necesarios para la redacción, delegar esta parte de forma íntegra a una computadora.

Por último, la revisión es otra parte que también deben realizar personas. La revisión propia es esencial para que el periodista pueda corregir errores cometidos durante el proceso o, en el caso de que adoptase la alternativa tecnológica, errores o aspectos a mejorar

de un texto creado por una máquina. También, la revisión externa permite que (al menos) otro par de ojos lea el texto y ofrezcan su opinión, sin haber sido parte activa en el proceso de creación del mismo.

10. MARCO TEÓRICO

Primero se definirá lo que es un proceso, para después dar un contexto teórico a la mejora de procesos.

Definición de proceso Según se puede observar en [101], proceso es definido por [102] como una actividad o grupo de ellas que transforman unos elementos de entrada en otros de salida, ya sean productos o servicios.

Mejora de procesos Se considera mejora de procesos como el conjunto de tareas que estudian y entienden un proceso para poder optimizarlo y así aumentar su calidad, maximizar la satisfacción del cliente o reducir los costes que este supone para la organización [103].

Se han propuesto diferentes enfoques a la hora de proceder a la mejora de procesos como son la reingeniería de procesos y la mejora continua de los mismos.

10.1. Reingeniería de procesos

La reingeniería de procesos consiste en conseguir una mejora de los procesos actuales mediante una completa reinvención de los mismos. “Reingeniería es la revisión fundamental y el rediseño radical de procesos para alcanzar mejoras espectaculares en medidas críticas y contemporáneas de rendimiento, tales como costos, calidad, servicio y rapidez” es la definición que le dan a esta alternativa Hammer y Champy en [104, p. 34].

10.1.1. Principios clave

La reingeniería se basa en doce principios clave [105]:

1. Se debe contar con el apoyo gerente para realizar el proceso.
2. La reingeniería debe estar alineada con la estrategia empresarial.
3. El objetivo final es crear valor para el cliente.
4. Las funciones no son el foco de atención, lo son los procesos.
5. Se deben designar equipos de trabajo capaces e incentivados para realizar el proceso de reingeniería, y recompensarles por ello tras la exitosa realización del mismo.
6. Para conocer si se están cumpliendo los objetivos se deben observar las necesidades del cliente y comprobar su nivel de satisfacción de forma activa.

7. El plan para el proceso de reingeniería debe ser flexible y estar preparado para realizar cambios a raíz de las primeras evaluaciones.
8. Los planes de reingeniería son específicos para cada negocio y situación, por lo que no podrán reutilizarse.
9. Los sistemas de medición son cruciales para la comprobación de la consecución de objetivos y no deben limitarse solo a tener en cuenta el tiempo, deben elegir otras métricas.
10. El factor humano debe tenerse en cuenta para no fracasar y evitar o reducir en lo posible la resistencia al cambio.
11. La reingeniería se debe apreciar como un proceso continuo dentro de la organización.
12. Es esencial la comunicación entre actores internos y con factores externos durante todo el proceso de reingeniería.

10.1.2. Rasgos característicos

Además de cumplir estos requisitos, todos los procesos de reingeniería cuentan con las siguientes características [105]:

- **Unificación de tareas.** Al estar los procesos por delante de las tareas, en pos de logran una optimización de los primeros, muchas tareas suelen verse eliminadas o combinadas con otras.
- **Participación de los trabajadores en la toma de decisiones.** Se busca una reducción de plazos y costes al hacer que las estructuras horizontales y verticales dentro de la organización se compriman, mediante el empoderamiento de los trabajadores encargados de esos procesos.
- **Cambio del orden secuencial por el natural en los procesos.** El cómo se ha realizado el proceso hasta ahora pasa a un segundo plano mientras que la forma que beneficia a los procesos se antepone al resto.
- **Realización de diferentes versiones del mismo producto.** A fin de llegar a una mayor adecuación de los productos y servicios a los gustos y las necesidades del cliente, se promueve la rotura de la estandarización y la generación de distintos modelos.
- **Reducción de las comprobaciones y los controles.** Se pretende flexibilizar y agilizar la supervisión y el control mediante un plan de evaluación que tenga en cuenta solo aquellos elementos con sentido económico.

- **Papel protagonista del responsable del proceso.** Al ser el único punto de contacto, se permite que aumente la eficiencia.
- **Operaciones híbridas.** La reingeniería utiliza muchas herramientas infotecnológicas y eso le permite que los equipos y las operaciones que se realizan cuenten con una conveniente naturaleza dual: son centralizadas y descentralizadas al mismo tiempo. Pretenden beneficiarse de la agilidad de la descentralización, sin perder las ventajas de formar parte de la organización.

10.1.3. Herramientas utilizadas

Las herramientas comúnmente utilizadas en la reingeniería de procesos de acuerdo a [105] son:

- **Visualización de procesos.** Consiste en un diseño meticuloso en el que se plasman todas las partes del proceso objetivo, anotando tareas, plazos y costes.
- **Investigación operativa.** Se basa en cinco elementos básicos: identificación del problema, selección de la estrategia de solución, prueba de la estrategia y evaluación, difusión de los resultados y utilización de los mismos.
- **Gestión el cambio.** Los cambios drásticos suponen reacciones en las personas que los sufren y no tratados de la manera correcta, pueden llevar a problemas. Por ello, en los procesos de reingeniería se le presta especial atención para evitar el fracaso.
- **Benchmarking.** Mediante esta técnica se estudian a los líderes del mercado para obtener información y aprender de sus éxitos y sus fracasos.
- **Infotecnología.** Se considera la herramienta básica de la reingeniería, ya que permite mantener la naturaleza híbrida de las operaciones.

10.2. Mejora continua de los procesos

Al contrario que la propuesta de la reingeniería, en la que se proponen cambios radicales sustanciales, la mejora continua de los procesos se centra en desarrollar mecanismos sistemáticos para mejorar los procesos desde el punto de vista del desempeño y así aumentar la satisfacción de los stakeholders [103]. ‘Kaizen’ es otro nombre con el que referirse a la mejora continua [106].

10.2.1. Requisitos

Para que la mejora continua se pueda llevar a cabo de manera satisfactoria, se deben cumplir los siguientes requisitos [107]:

- **Apoyo de la dirección.** Es fundamental que la cúpula de la organización respalde la mejora de los procesos para que esta pueda completarse satisfactoriamente.
- **Compromiso a largo plazo.** Para que los cambios que se produzcan sean notable, se debe superar la barrera de los problemas que surgen en el corto plazo para poder obtener resultados.
- **Metodología disciplinada y unificada.** Todas las personas que participen en la mejora de procesos deben tener clara la misma metodología y cumplirla, pues solo así se conseguirá finalizar con éxito.
- **Propietario de cada proceso.** Es necesario que haya una cabeza visible para cada proceso.
- **Sistemas de evaluación y retroalimentación.** Conocer cómo se está haciendo en todo momento y poder determinar en qué áreas se debe mejorar es crucial en la mejora de procesos.
- **Foco en los procesos y en los clientes.** El proceso tiene mayor importancia que cualquier trabajo individual y este debe estar centrado en satisfacer las necesidades y deseos de los clientes.

A continuación se ofrecerán algunos ejemplos que prácticos de la mejora continua de los procesos como son el Ciclo de Deming, Seis Sigma y el Ciclo de Mejora de Procesos de negocios.

10.2.2. Ciclo de Deming

Este modelo fue ideado por en 1951 por W. Edwards Deming, basándose en el trabajo publicado por el doctor Walter A. Shewhart en 1939 [108]. A día de hoy sigue siendo una de las principales herramientas que usan las empresas y es la base de muchos otros

métodos. Su simplicidad y universalidad permite su aplicación en multitud de campos y situaciones. Está compuesto por por cuatro pasos (véase figura 10.1) [107]:

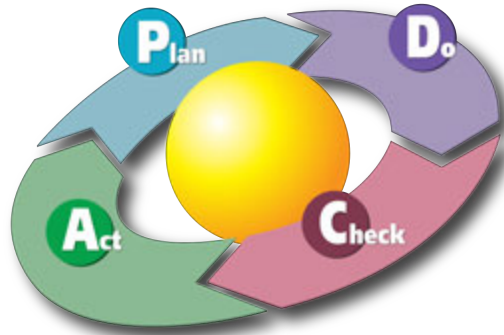


Fig. 10.1. Esquema del ciclo PDCA [109].

- **Planificar (Plan).** En este paso se debe reconocer dónde y cómo se puede mejorar, y también se debe conseguir el apoyo de la dirección. A parte, hay que realizar un análisis de los problemas y los procesos implicados, y realizar la planificación de su implantación.
- **Hacer (Do).** Consta de la puesta en marcha de la planificación anterior en pequeña escala para probar su funcionamiento.
- **Verificar (Check).** Aquí se controla que los resultados de esa puesta en marcha reducida son los esperados. En caso negativo, se vuelve al paso de planificar.
- **Actuar (Act).** Último paso del proceso en el que se aplica de forma definitiva los cambios planteados en el paso de planificar. De nuevo, si los resultados no fueran los esperados, se vuelve al principio del ciclo.

10.2.3. Seis Sigma

Otro de los modelos más conocidos de mejora continua de los procesos es el Seis Sigma. Este modelo fue utilizado por el ingeniero Bill Smith en Motorola, en el año 1988. Su nombre viene de que el método utiliza herramientas estadísticas para analizar los procesos, y a través de ellos se pretende alcanzar un número de fallos por el millón de unidades igual a 3,4 (en una distribución normal este es el valor que se alcanzaría si la desviación típica, sigma, tuviera un valor de 6). Para producir esos resultados, se caracteriza por una organización del personal basada en las artes marciales según las siguientes funciones [110]:

- **Líder ejecutivo.** Grupo formado por los ejecutivos de alto nivel.
- **Campeones.** Nivel superior de gestión encargado de la aplicación de Seis Sigma y, normalmente, dueños de los proyectos.

- **Maestro cinturón negro.** Designados por los campeones, su responsabilidad es la de transmitir el método dentro de la organización como mentores.
- **Cinturón negro.** Expertos centrados en el mando medio. Colaboran en la mejora de productos, servicios y procesos.
- **Cinturón verde.** También expertos, pero centrados en ayudar a los que hayan sido designados como cinturón negro.

Una vez creada la estructura personal descrita, se sigue el modelo descrito en la figura 10.2 [107]:



Fig. 10.2. Esquema DMAIC para el método Seis Sigma [111].

- **Definir (*Define*).** Fase en la que los proyectos son identificados y evaluados por la dirección. Se identifica el problema, se detalla la mejora y se establece el plan a seguir.
- **Medir (*Measure*).** Se prepara a los trabajadores para que tomen las medidas correctas, se ponen a punto las herramientas de medida y se toman datos asegurando su integridad.
- **Analizar (*Analyze*).** Con la información conseguida en la fase anterior, se desarrollan las hipótesis de las causas que más afectan al proceso.
- **Mejorar (*Improve*).** A raíz de los resultados del análisis, se acota el rango operacional de los cambios y diseñan y ejecutan las mejoras.
- **Controlar (*Control*).** Última fase cuyo propósito es perpetuar las mejoras obtenidas y corregir aquellos aspectos a mejorar. En el caso de que sea necesario, se repetiría el modelo.

10.2.4. Ciclo de mejora de los procesos de negocios

El Ciclo de Mejora de los Procesos de negocios (más conocido como CMP y de ahora en adelante referido como tal) es un método desarrollado por José María Berenguer Peña y Juan Antonio Ramos-Yzquierdo de la Universidad de Navarra [112]. El esquema de este método se basa en el DMAIC descrito en el apartado anterior, pero ampliando ciertas áreas. Consta de la siguiente estructura (figura 10.3) [112]:



Fig. 10.3. Representación del Ciclo de Mejora de Procesos de negocios [112].

- 1. Sensibilización.** Aquí se prepara el terreno para el cambio que se va a producir en la organización. Consta de las siguientes actividades:
 - **Apreciar.** Se debe conseguir la aceptación y la comprensión por parte de la alta dirección sobre el cambio, delimitándolo en alcance.
 - **Auscultar.** Se debe analizar el marco de referencia de la organización y determinar la propuesta de valor y los stakeholders.
 - **Comprometer.** Se debe conseguir el compromiso de la dirección y difundirlo en la organización.
- 2. Identificación.** En esta fase se identifican los procesos a mejorar. Consta de las siguientes actividades:
 - **Listar.** Se listan los posibles procesos de la organización.
 - **Valorar.** Se valoran cuales necesitan mejora.

- **Seleccionar.** Se obtiene aprobación de la dirección una vez seleccionados los procesos a mejorar.
3. **Planificación.** Supone una fase esencial y en ella se definen los parámetros del proyecto y la comunicación del mismo. Consta de las siguientes actividades:
- **Exponer.** Se da a conocer a todos los grupos implicados la visión de lo que se pretende conseguir.
 - **Proyectar.** Se define el proyecto, junto con los objetivos, los beneficios y los responsables.
 - **Comunicar.** Se divulga la necesidad del cambio para conseguir la aceptación de la misma en toda la organización.
4. **Análisis.** En esta fase sucede el estudio del proceso. Consta de las siguientes actividades:
- **Confirmar.** Se corrobora que el flujo de valor del proceso está alineado con la estrategia de la organización.
 - **Modelar.** Se determinan las actividades que no aportan valor al cliente.
 - **Identificar.** Se señalan las ineficiencias en el proceso.
 - **Parametrizar.** Se indican los indicadores de desempeño.
 - **Medir.** Se establece la medición del rendimiento de los procesos.
 - **Sustanciar.** Se priorizan las mejoras propuestas.
5. **Diseño.** Esta fase está caracterizada por aportar la estructura de la solución o soluciones que se van a llevar a cabo. Consta de las siguientes actividades:
- **Alinear.** Se comprueba la satisfacción de expectativas internas y externas de la organización.
 - **Especificar.** Se concretan las soluciones al problema.
 - **Evaluar.** Se escoge la solución más idónea.
 - **Estructurar.** Se determinan los cambios en la estructura y los sistemas de información.
6. **Implantación.** En esta fase se procede al asentamiento del resultado de las fases anteriores. Consta de las siguientes actividades:
- **Digitalizar.** Se procede a la automatización de las partes pertinentes del proceso.
 - **Reorganizar.** Se acomoda la organización a los ajustes del nuevo proceso.

- **Formar.** Se aporta información a todos los implicados por el cambio, tanto de forma directa como indirecta.
 - **Probar.** Se pasa a un periodo piloto.
 - **Arrancar.** Se ejecuta el despliegue de forma definitiva.
7. **Control.** Aquí se monitoriza el rendimiento del proceso. Consta de las siguientes actividades:
- **Mantener.** Se debe conservar la calidad del proceso, documentando cambios y controles.
 - **Responsabilizar.** Se identifica a los responsables y se les concienta de los cambios a asumir.
 - **Aprender.** Se revisan y se analizan las mejoras.
8. **Capacitación.** Por último, se intenta aprovechar la experiencia adquirida por la aplicación del ciclo. Consta de las siguientes actividades:
- **Adjudicar.** Se determina el nivel de madurez de la estructura de la organización.
 - **Ejercitar.** Se estructura a los implicados para fomentar la mejora continua.

10.3. Selección del método.

Para la realización de este trabajo, se ha optado por utilización del CMP porque se considera que se adecúa más al propósito del trabajo y porque tiene en cuenta la progresión de la organización después del cambio en el proceso. Además, es un método que afecta a el área directiva, operacional y de recursos humanos, y ofrece una secuencia coherente entre el proceso en las tres [112].

Se ha descartado la reingeniería de procesos porque supone un cambio total y notable, y en este trabajo se busca un cambio menor en la estructura y el proceso como tal. Además, suele conllevar muchos riesgos por la naturaleza radical del método.

Dentro de la mejora continua de procesos, se han descartado el Ciclo de Deming y el método Seis Sigma por tener en cuenta menos aspectos que el CMP, lo cual no sugiere que no sean aptos para otras mejoras continuas de procesos.

11. SOLUCIÓN

Ya que se ha elegido el método CMP para mejorar procesos en este caso, se va proceder a detallar el proceso descrito anteriormente, omitiendo las fases que no sean relevantes o aplicables en este caso particular. Este es un caso en el que se ha identificado el proceso a mejorar antes de realizar el ciclo, y es un aspecto a tener en cuenta. Antes de empezar, también cabe destacar que este trabajo es una propuesta y no una memoria del proceso, por lo que hay secciones sobre las que no se tiene información suficiente.

El cambio que se propone a partir de aquí es la utilización de una herramienta de software de redacción de textos por parte de los periodistas y redactores que tengan que publicar contenido en la web y en las cuentas de las redes sociales de RTVE. No como un sustituto de estos trabajadores, si no como una herramienta que puedan usar para reducir su carga de trabajo y emplear mejor su tiempo. En la primera parte de este trabajo (véase la parte II) se muestra el desarrollo de una herramienta software similar, aunque en dicho desarrollo no se haya llegado a obtener un producto completo.

11.1. Propuesta de aplicación del CMP

A continuación se hará un desglose de las fases enumeradas en la sección 10.2.4 de forma concreta para el problema y el cambio planteados:

11.1.1. Sensibilización

En esta fase se ha de empezar por explicar el cambio a la alta dirección de forma clara y acotada para que comprendan de qué se les intenta hablar y lo acepten. En este caso, el cambio está muy bien localizado en las áreas de creación de contenido en formato texto, ya sea para publicar en la web o en redes sociales. El cambio viene motivado por una mejora en el rendimiento de los trabajadores y una mejor inversión del tiempo de estos, en tareas que solo ellos puedan realizar.

Acto seguido, hay que tener en cuenta el marco de referencia de RTVE (misión, visión y valores; véase la sección 9.1), para asegurarse de que el cambio que se propone está alineado con estos factores. Tal y como se han descrito estos, no se interpreta que el cambio propuesto contradiga ninguno de ellos. No obstante, en la aplicación real del cambio debería contrastarse con la organización para que no haya lugar a confusiones.

La propuesta de valor, ya descrita en el Business Canvas Model (véase sección 9.2.1), consta de aportar noticias veraces que reflejen contenido actual y ofrecer contenido exclusivo y atractivo a los usuarios. Tampoco se aprecia una falta de alineamiento entre las propuestas de valor mencionadas y el cambio que se pretende realizar.

Por otro lado, los stakeholders en este caso son la organización y todos los que la rodean, pues con este cambio se beneficia la organización por contar con mayor eficiencia; los trabajadores por disponer con más tiempo; los proveedores porque uno de sus clientes está teniendo buenos resultados; y los clientes porque se les está ofreciendo un servicio de calidad.

Por último, se debe conseguir el compromiso por parte de la dirección. En esta actividad se deberá enfatizar a la dirección (y más adelante a los empleados) que el cambio propuesto no supone un cambio radical ni negativo para nadie, y que solo es la inclusión de una herramienta que pretende facilitar sus labores a los trabajadores para así conseguir mejores resultados con esfuerzos mejor aprovechados.

11.1.2. Identificación

En esta segunda fase no sería necesario realizar las actividades que se mencionan, puesto que el problema que se quiere solucionar está definido desde el principio. Además, para la realización de este apartado no se cuenta con información sobre los procesos internos de RTVE.

Sin embargo, si en el caso real se requiriese la identificación de los procesos en lo que se requiere la mejora, deberán identificarse cuáles son esenciales en el funcionamiento de la organización. Posteriormente deberán valorarse con respecto a unas métricas pactadas con la organización, priorizando respecto a dichas métricas. Más adelante, habrá que seleccionar los que quisieran implantarse en el corto plazo para afectar positivamente a los resultados de la empresa, y en el largo para contribuir a la cultura de gestión de procesos. Por último, justificar la elección a la dirección para obtener su aprobación.

11.1.3. Planificación

En un caso real, en esta fase debería realizarse un Project Charter para poder cumplir con las actividades de las que esta consta. Un Project Charter es un documento en el que se detallan: la definición del problema, la declaración de objetivos, las medidas o hechos que ilustran el problema, los recursos y las limitaciones del proceso, las reglas de actuación del equipo, los miembros del equipo, los riesgos y el plan preliminar del proyecto, así como las de realización y de revisión por parte de la organización. Este documento entonces se vuelve la carta magna del proceso, en la que se expone, se proyecta y se comunica el cambio.

De nuevo por la falta de información, se ha optado por redactar el problema, los objetivos, los afectados y los riesgos para que, en el caso de que se realizase el método, se pudiera utilizar esta información para completar el mencionado documento:

Problema En el caso descrito a lo largo del documento, el problema que se observa en el proceso de redacción de textos que serán publicados para la plataforma web y las redes sociales es la falta de eficiencia por poco aprovechamiento de los recursos tecnológicos actuales.

Objetivos Los objetivos de este cambio son la mejora de eficiencia en el proceso, llevando a los trabajadores a poder dedicar más tiempo a tareas que solo ellos puedan realizar y aligerando sus responsabilidades para que gocen de tiempo que emplear en la mejora del producto final.

Afectados Los afectados por este cambio sería todos los periodistas y redactores que trabajen en las áreas web y de redes sociales que se dediquen al proceso en cuestión, así como sus supervisores.

Riesgos Queda ilustrado en estudios como [113] y [114] que la aceptación de la tecnología no es un tema banal dentro de las organizaciones. Por miedo, falta de formación, falta de una explicación y puesta en contexto clara; importancia de las costumbre y diversas razones; cambios como los propuestos en este trabajo pueden trabajar de manera estrepitosa llevando, además, a favorecer el rechazo a la tecnología.

11.1.4. Análisis

En esta fase el trabajo se basa sobre el mapa As-Is del proceso. Este mapa describe con detalle el estado actual del proceso sin ningún cambio aplicado. Cabe destacar que el mapa de este trabajo se ha realizado a modo ilustrativo y no se dispone de información detallada para el proceso. No obstante, se considera que será de ayuda para ejemplificar y delinear el caso.

Observando la figura 11.1, se pueden apreciar todas las fases y tareas con sus duraciones entre paréntesis (previamente descritas en la sección 9.3). A raíz de su estudio no se percibe falta de alineamiento con la estrategia de la organización ni se distinguen actividades que no aporten valor al cliente. Sin embargo, se observan ineficiencias dado el contexto de este trabajo. Las tareas de investigación, identificación de hechos, diseño del esquema del texto, toque personal y redacción podría delegarse (a veces en su totalidad) a una herramienta software para reducir la carga del trabajador liberarle de dicha tarea.

En este caso en particular, se propone para comprobar el funcionamiento del cambio el tiempo total de redacción de una noticia, así como el tiempo en el que se finaliza cada tarea. También cabría realizar encuestas en las primeras fases para comprobar si el cliente percibe algún tipo de cambio en el contenido y qué tipo de cambio (positivo o negativo).

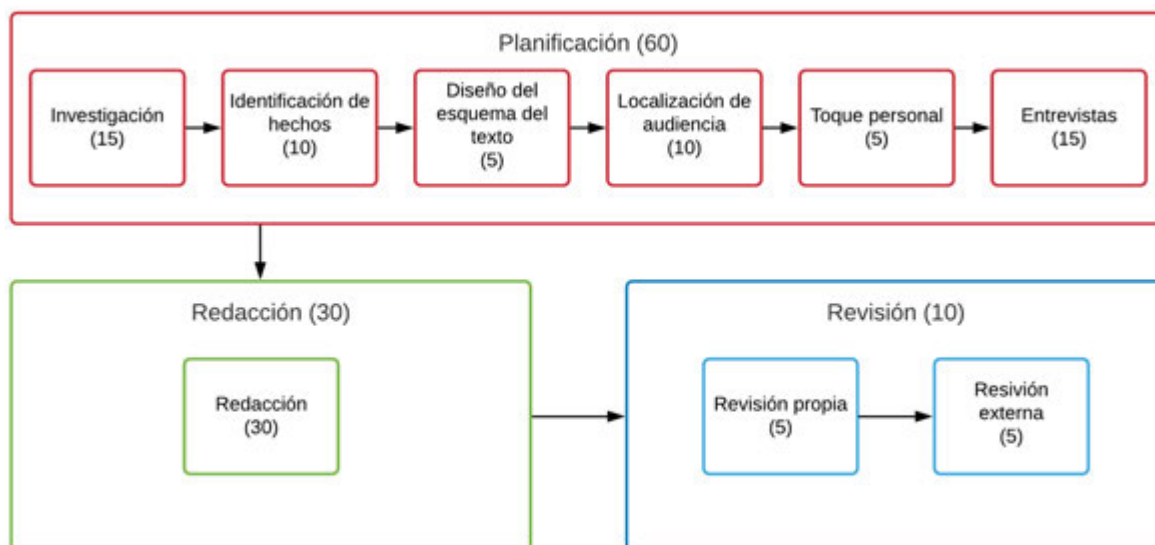


Fig. 11.1. Mapa del proceso de redacción de textos As-Is.

11.1.5. Diseño

De forma similar que en la fase anterior, en esta también toma protagonismo un mapa del proceso, el mapa To-Be. Este mapa pretende reflejar de manera fiel cómo será el proceso una vez se haya realizado los cambios a los problemas detectados en el análisis. Al igual que el mapa anterior, el siguiente tiene fines ilustrativos y no se basa en un reflejo fiel del proceso actual.

El mapa propuesto (véase la figura 11.2) sigue la misma estructura que el anterior, con la diferencia de que las tareas que serían total o parcialmente automatizadas aparecen en amarillo. El primer paso después de observar este mapa sería comprobar que se alinean con las especulativas internas y externas.

Una vez comprobado, se determinaría las soluciones. En este caso son:

- Reducir o eliminar el tiempo de investigación por parte del periodista/redactor. La herramienta recibe la información de las fuentes especificadas, puesto que no hay necesidad aparente de la investigación por parte del trabajador. No obstante, se contempla que, por cualquier motivo, el periodista necesitara evaluar una fuente el mismo, de ahí que se contemple una reducción parcial. **Se determina que la duración de la tarea se reduciría de 15 unidades de tiempo a 5.**
- Eliminar el tiempo de identificación de hechos por parte del periodista/redactor. La herramienta, al recibir la información, subrayaría los hechos que están contenidos en ella, permitiendo al trabajador quedar libre de la tarea. **Se determina que la duración de la tarea se reduciría de 10 unidades de tiempo a 0.**
- Eliminar el tiempo de diseño del esquema del texto por parte del periodista/redactor.

El esquema que se quiere seguir en la redacción del texto se introduciría como parámetro a la herramienta, liberando de nuevo al trabajador de la tarea en cuestión. **Se determina que la duración de la tarea se reduciría de 5 unidades de tiempo a 0.**

- Eliminar el tiempo de añadir toque personal al texto por parte del periodista/redactor. De igual manera que el anterior, el cómo identificarse en el texto podría ser introducido en la herramienta, ahorrando tiempo al trabajador. **Se determina que la duración de la tarea se reduciría de 5 unidades de tiempo a 0.**
- Eliminar el tiempo de redacción por parte del periodista/redactor. Con toda la información introducida y la correcta configuración, el trabajador podría olvidarse de la redacción para centrarse en la revisión u otros aspectos de su trabajo. **Se determina que la duración de la tarea se reduciría de 30 unidades de tiempo a 0.**

Dado que no hay distintas soluciones al problema, no tiene cabida la evaluación ni elección de las mismas.

Por último, con el uso de el esquema proporcionado y siguiendo los cambios descritos se espera una reducción teórica del **sesenta por ciento en el tiempo que los periodistas y redactores tienen que dedicar al proceso** de redacción descrito. Esto supone un cambio positivo para tanto trabajadores, como directivos pues genera eficiencia en el proceso y aporta bienestar a los trabajadores.

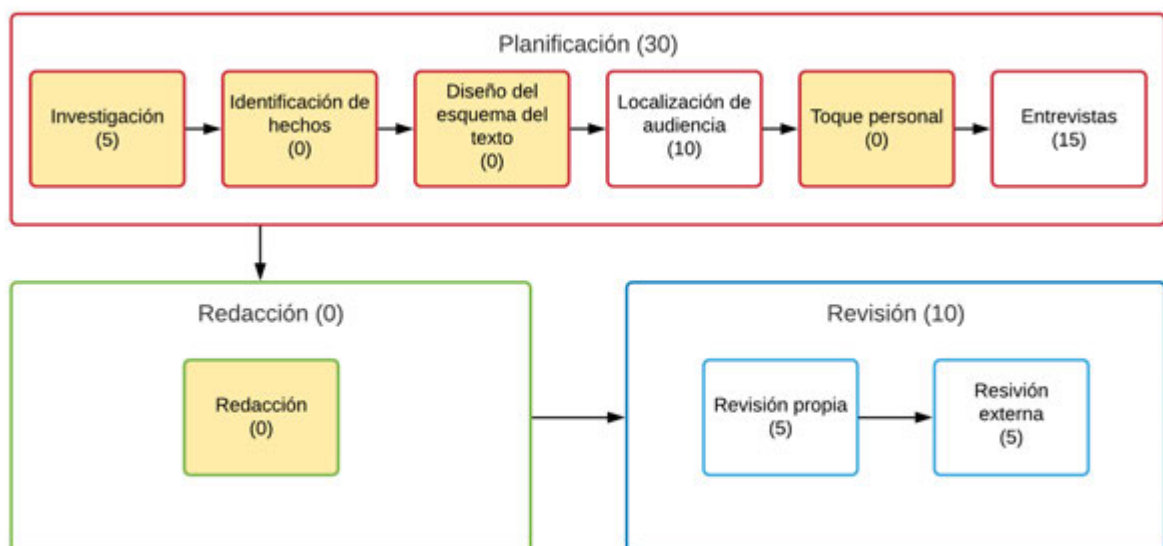


Fig. 11.2. Mapa del proceso de redacción de textos To-Be.

11.1.6. Implantación

Para implantar los cambios descritos sería necesario dotar a los redactores y periodistas de equipos informáticos para la utilización de la herramienta. Se considera que tendrán

equipos competentes, pues un equipo informático normal y actualizado debería poder soportarla, pero se contempla que la organización tuviera que hacer una pequeña inversión en caso de que no fuera así.

No se tiene que reorganizar la estructura de la corporación al cambio, ya que no supone ninguna alteración en la misma.

Se recomienda un periodo de formación de un mes para que los profesionales tengan tiempo de adaptarse a esta nueva herramienta, aprender su funcionamiento, conocer sus funciones y opinar sobre ella. Después, se sugiere otro periodo de un mes en el que se fueran seleccionando a grupos de profesionales por semanas para que fueran utilizando de manera piloto la herramienta.

Una vez pasados los dos meses, si no han surgido problemas que requieran que el cambio se parase o revirtiese, podría escalarse al resto de profesionales afectados por el cambio.

Cabe reseñar la importancia de la implantación. Si no se consigue formar de manera completa y satisfactoria a los empleados podría, en vez de conseguirse un aumento de productividad y una aceptación de la herramienta, un rechazo y una negativa al proceso descrito en este trabajo.

11.1.7. Control

A modo de asegurar que la calidad se mantiene en las publicaciones, se recomienda llevar un seguimiento constante de las opiniones y la afluencia de la web y las redes sociales. El usuario podría notar los textos demasiado robóticos, o al cabo del tiempo muy repetitivos. Es clave realizar el seguimiento y actualizar de forma periódica la herramienta para que esto no suceda.

Como método para responsabilizar a los trabajadores pertinentes, se podría instaurar un sistema de incentivos a los encargados de los trabajadores y a los trabajadores en cuestión, para que se vean motivados a revisar de forma concienzuda los productos generados por la herramienta y así mantener la calidad de los servicios.

A raíz de la monitorización mencionada, podría modificarse el software de observarse mejoras necesarias o modificar el proceso descrito si hubiera cabida de nuevas tareas después de que se asentasen los cambios.

11.1.8. Capacitación

La última fase de este ciclo queda a discreción de la organización, pues permite que se aproveche la experiencia obtenida de la consecución del desarrollo de estas fases y promueve la futura utilización de métodos de mejora continua.

11.2. Plan de Marketing hipotético

La solución propuesta en este trabajo no se plantea de forma que admita un plan de marketing, pues RTVE es quien la aplicaría y lo haría dentro de su misma organización. No obstante, se van a delinear brevemente los aspectos a tener en cuenta en el hipotético caso de que lo descrito en páginas anteriores quisiera venderse, por si resultase de interés.

Primero deberían definirse los objetivos de marketing que tuviera la empresa con la intención de comercializar la solución. Eso es algo de lo que no se puede disponer en este trabajo. Sin embargo, es crucial a la hora de definir dichos objetivos tener en cuenta el marco de referencia (misión, visión y valores) de la empresa que lo comercialice. De esta forma, se evitarían las incoherencias de cara al cliente y el mercado.

Además, cabría identificar los segmentos a los que dirigir dicha propuesta. A raíz de la realización de este trabajo se identifican dos principales segmentos: medios de comunicación con procesos de redacción sin automatizar y redactores independientes que pretendan aumentar la demanda que puedan cubrir. De todos modos, en el momento de la puesta en marcha de este plan habría que comprobar si los ejemplos aquí propuestos siguen siendo factibles y si han aparecido nuevos a raíz de necesidades no conocidas en el momento de realización de este trabajo.

En cuanto al posicionamiento que debería realizar depende, de nuevo, de la empresa que realice la comercialización. Tendría que asegurarse de mantener la campaña en línea con su imagen y los principios que le representan. En este proceso se recomienda la utilización del modelo de marketing mix de las 7 Ps (*Place, Price, Promotion, Physical evidence, People, Product y Process*) [115] en vez del modelo de las 4 Ps (*Place, Product, Promotion y Price*) [116] ya que resulta más completo, pues el modelo de las 4 Ps se considera que deja fuera de sus directrices aspectos muy relevantes del marketing [117].

Otro aspecto muy importante a tener en cuenta a la hora de comercializar la solución propuesta es la resistencia al cambio que supone en las organizaciones que se ha comentado con anterioridad. Al suplantar al trabajador por una máquina, este puede tener una reacción exagerada y perjudicial si no comprende en que situación se haya con respecto a la herramienta. Por ello, se pretende enfatizar en que a lo largo de todo el plan de marketing se debería hacer especial hincapié en que todos los elementos implicados (no solo los directivos en el caso de una organización) comprendieran la utilidad de la solución propuesta y que el software que se emplearía sería solo una herramienta y no un sustituto del trabajador. Este aspecto no solo permitiría el contento de las organizaciones en las que se lograra implantar, sino también una excelente publicidad de cara a otras organizaciones.

Parte IV

Conclusiones

Durante este trabajo se ha expuesto el desarrollo de un módulo de generación de noticias deportivas con sentimientos (añadidos durante la generación) y se ha ilustrado el proceso, en líneas generales, que debería una gran organización como RTVE para incluir un sistema basado en dicho módulo dentro de sus procesos de redacción habituales.

Se ha mostrado que la tecnología actual permite ayudar a los redactores y periodistas en sus procesos de creación de contenido, pues si se le asigna todo el proceso de recogida de información, detección de temas y redacción a la herramienta propuesta, las personas previamente encargadas de esas tareas gozarían de una cantidad de tiempo muy considerable. Cantidad de tiempo que, por otra parte, estos beneficiados podrían emplear en investigación robusta y exhaustiva de fuentes que las máquinas no puedan tratar, entrevistar a personas para recoger información con una profundidad que los ordenadores y sistemas no pueden ni rozar.

En este trabajo no se pretende mostrar cómo las herramientas de software y la tecnología pueden sustituir a las personas para dejarlas en un segundo plano. Todo lo contrario. Durante las páginas de este documento se apunta hacia un futuro en el que los humanos puedan delegar tareas perfectamente automatizables, de menor o mayor calibre, en sistemas que puedan asumirlas, para que los primeros puedan disponer de un tiempo añadido para ocupar en tareas únicamente humanas. En el caso particular de este trabajo, se ofrece una solución para que los periodistas tengan que, sin renunciar a cubrir noticias de gran importancia, puedan seguir cubriendo (con esfuerzo reducido) noticias de menor impacto como las referentes a los partidos de 2ª división B en España.

Además, aunque pueda resultar un comentario un poco obvio, los sistemas como el descrito en este trabajo no sufren el cansancio. Esto no les supone problema a la hora de estar trabajando de forma continua, lo que abre nuevas posibilidades. Posibilidades como el cubrir tareas que no se podían realizar por falta de recursos (humanos) y tiempo, y la inclusión de nuevas tareas en los procesos actuales que realizan las organizaciones, pues hay tiempo que, a raíz de la automatización, ha quedado liberado.

Trabajo futuro

Sobre la herramienta de generación desarrollada. Como aparece reflejado en el trabajo, el funcionamiento deseado de la herramienta no ha sido alcanzado. No obstante, no se pierde de vista el trabajo realizado y la experiencia adquirida durante el proceso de desarrollo. Esto sirve perfectamente como base para que el sistema genere noticias completas sin problema. Además, su rango de actuación podría ampliarse a otras noticias de mayor importancia si se quisiera y también a otros idiomas.

Por otra parte, el aprendizaje automático es algo que se ha dejado de lado en este trabajo. Estando tan a la alza en la actualidad y siendo usado por gigantes del sector audiovisual y de los medios de comunicación como se ha ilustrado en estas páginas, no cabe duda de que podría mejorar la solución actual. Se propone desde estas líneas un

enfoque híbrido, que consiguiera información para la redacción y, posiblemente, poblase la base de conocimiento del sistema mediante aprendizaje automático (llegando a poder detectar plantillas incluso), y que contase con una versión completa y mejorada del motor de generación para realizar la creación de textos.

Sobre la mejora de procesos propuesta. Como propuesta para continuar con el trabajo descrito siguiendo el CMP, se contempla la posibilidad de que la mejora de procesos se llevase a cabo de manera real. No solo aportaría información que ha estado ausente durante el desarrollo, si no que serviría como una fuente de aprendizaje notable.

Una vez se hubieran implantado los cambios descritos, sería de interés estudiar las sensaciones de los empleados a la utilización de la herramienta, pues sería útil para futuros cambios tecnológicos en la organización. También se contempla como información relevante los datos que se puedan obtener de los clientes, pues esto serviría a la corporación para que pudiera acercarse más a sus intereses y necesidades.

Además, a medida que la herramienta evolucionara, su utilidad dentro de la organización podría aumentar, pudiendo aplicarla en otras tareas e, incluso, en otros procesos.

Por último, si la implantación resulta exitosa, se urge a la corporación a contemplar cambios de índole similar en otros procesos que se lleven a cabo en la organización. La tecnología avanza a pasos de gigante y parte de ese avance se centra en facilitar tareas a las personas y ocuparse de aquellas en las que no sea necesaria la intervención de una persona, pues de esta forma se libera tiempo de las mismas. Aquellas organizaciones que puedan maximizar la utilidad y el rendimiento de su capital humano serán aquellas que destaquen por encima del resto en un futuro.

Limitaciones

Como cierre a este trabajo se quieren mostrar las limitaciones que se han percibido durante la realización del mismo. Referente a la parte tecnológica, la falta de información y el alcance inicial tan reducido del proyecto suponen que el resultado del trabajo necesita esfuerzos posteriores a la hora de tener una aplicación efectiva y satisfactoria en una organización real. Respecto a la mejora de procesos, también la ausencia de información de los procesos de RTVE, así como el carácter hipotético de la mejora suponen que lo escrito en las páginas de este trabajo, sin que carezca de valor, necesita de una dosis de realidad empresarial para alcanzar la madurez. Por último y de forma general a todo el trabajo realizado, la aparición de la COVID-19 no estaba contemplada a inicios de este trabajo. Tampoco está contemplada en muchas de las fuentes utilizadas a lo largo de todo el desarrollo presente. Esto puede suponer una falta de información notable a la hora de haber realizado, dependiendo de la realidad a la que depara el futuro. No obstante, sin perder de vista la tragedia que ha sido y sigue siendo, desde este trabajo se proyecta optimismo para esa realidad en lo que atañe a este trabajo y en general.

REFERENCIAS

- [1] *RTVE*, es, Page Version ID: 128426981, ago. de 2020. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=RTVE&oldid=128426981> (Acceso: 17-08-2020).
- [2] A. Gatt y E. Krahmer, “Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation”, *arXiv:1703.09902 [cs]*, ene. de 2018, arXiv: 1703.09902. [En línea]. Disponible en: <http://arxiv.org/abs/1703.09902> (Acceso: 29-07-2020).
- [3] *Natural-language generation*, en, Page Version ID: 969537795, jul. de 2020. [En línea]. Disponible en: <https://en.wikipedia.org/w/index.php?title=Natural-language-generation&oldid=969537795> (Acceso: 29-07-2020).
- [4] *Natural language generation (NLG)*, en. [En línea]. Disponible en: <https://securities.bnpparibas.com/insights/natural-language-generation.html> (Acceso: 02-09-2020).
- [5] R. Nuseibeh, *NLP, NLU and NLG Chatbots explained*. en, Library Catalog: medium.com, nov. de 2018. [En línea]. Disponible en: <https://medium.com/botique-ai/nlp-nlu-and-nlg-chatbots-explained-64820bc32ad> (Acceso: 03-08-2020).
- [6] E. Reiter y R. Dale, “Building applied natural language generation systems”, en, *Natural Language Engineering*, vol. 3, n.º 1, pp. 57-87, mar. de 1997. doi: [10.1017/S1351324997001502](https://doi.org/10.1017/S1351324997001502). [En línea]. Disponible en: http://www.journals.cambridge.org/abstract_S1351324997001502 (Acceso: 03-08-2020).
- [7] F. J. Och y H. Ney, “A Systematic Comparison of Various Statistical Alignment Models”, *Computational Linguistics*, vol. 29, n.º 1, pp. 19-51, mar. de 2003, Publisher: MIT Press. doi: [10.1162/089120103321337421](https://doi.org/10.1162/089120103321337421). [En línea]. Disponible en: <https://doi.org/10.1162/089120103321337421> (Acceso: 15-08-2020).
- [8] J. Clarke y M. Lapata, “Discourse Constraints for Document Compression”, *Computational Linguistics*, vol. 36, n.º 3, pp. 411-441, jul. de 2010, Publisher: MIT Press. doi: [10.1162/coli_a_00004](https://doi.org/10.1162/coli_a_00004). [En línea]. Disponible en: https://doi.org/10.1162/coli_a_00004 (Acceso: 15-08-2020).
- [9] I. Macdonald y A. Siddharthan, “Summarising News Stories for Children”, en *Proceedings of the 9th International Conference on Natural Language Generation (INLG 2016)*, Edinburgh: ACL, sep. de 2016, pp. 1-10. [En línea]. Disponible en: <https://www.aclweb.org/anthology/W/W16/W16-6601.pdf> (Acceso: 15-08-2020).

- [10] R. Dale, I. Anisimoff y G. Narroway, “HOO 2012: A Report on the Preposition and Determiner Error Correction Shared Task”, en *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, Montréal, Canada: Association for Computational Linguistics, jun. de 2012, pp. 54-62. [En línea]. Disponible en: <https://www.aclweb.org/anthology/W12-2006> (Acceso: 15-08-2020).
- [11] A. Bartoli, A. De Lorenzo, E. Medvet y F. Tarlao, “Your Paper has been Accepted, Rejected, or Whatever: Automatic Generation of Scientific Paper Reviews”, en, en *Availability, Reliability, and Security in Information Systems*, F. Buccafurri, A. Holzinger, P. Kieseberg, A. M. Tjoa y E. Weippl, eds., ép. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2016, pp. 19-28. doi: [10.1007/978-3-319-45507-5_2](https://doi.org/10.1007/978-3-319-45507-5_2).
- [12] D. Kauchak y R. Barzilay, “Paraphrasing for Automatic Evaluation”, en *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, New York City, USA: Association for Computational Linguistics, jun. de 2006, pp. 455-462. [En línea]. Disponible en: <https://www.aclweb.org/anthology/N06-1058> (Acceso: 15-08-2020).
- [13] V. Rus et al., “The First Question Generation Shared Task Evaluation Challenge”, en, Num Pages: 7, Trim Castle, Ireland, jul. de 2010. [En línea]. Disponible en: <http://www.scss.tcd.ie/conferences/INLG2010/> (Acceso: 15-08-2020).
- [14] M. Theune, E. Klabbers, J. Odijk, J. Pijper y E. Kraemer, “From Data to Speech: A General Approach”, *Natural Language Engineering*, vol. 7, jun. de 2001. doi: [10.1017/S1351324901002625](https://doi.org/10.1017/S1351324901002625).
- [15] D. L. Chen y R. J. Mooney, “Learning to sportscast: a test of grounded language acquisition”, en, en *Proceedings of the 25th international conference on Machine learning - ICML '08*, Helsinki, Finland: ACM Press, 2008, pp. 128-135. doi: [10.1145/1390156.1390173](https://doi.org/10.1145/1390156.1390173). [En línea]. Disponible en: <http://portal.acm.org/citation.cfm?doid=1390156.1390173> (Acceso: 04-08-2020).
- [16] M. Molina, A. Stent y E. Parodi, “Generating Automated News to Explain the Meaning of Sensor Data”, vol. 7014, 2011, pp. 282-293. doi: [10.1007/978-3-642-24800-9_27](https://doi.org/10.1007/978-3-642-24800-9_27).
- [17] L. Leppänen, M. Munezero, M. Granroth-Wilding y H. Toivonen, “Data-Driven News Generation for Automated Journalism”, en *Proceedings of the 10th International Conference on Natural Language Generation*, Santiago de Compostela, Spain: Association for Computational Linguistics, sep. de 2017, pp. 188-197. doi: [10.18653/v1/W17-3528](https://doi.org/10.18653/v1/W17-3528). [En línea]. Disponible en: <https://www.aclweb.org/anthology/W17-3528> (Acceso: 04-08-2020).

- [18] R. v. d. Wal, N. Sharma, C. Mellish, A. Robinson y A. Siddharthan, “The role of automated feedback in training and retaining biological recorders for citizen science”, es, *Conservation Biology*, vol. 30, n.º 3, pp. 550-561, 2016, _eprint: <https://conbio.onlinelibrary.wiley.com/doi/pdf/10.1111/cobi.12705>. doi: [10.1111/cobi.12705](https://doi.org/10.1111/cobi.12705). [En línea]. Disponible en: <https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/cobi.12705> (Acceso: 15-08-2020).
- [19] V. Plachouras et al., “Interacting with Financial Data using Natural Language”, en *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, ép. SIGIR '16, New York, NY, USA: Association for Computing Machinery, jul. de 2016, pp. 1121-1124. doi: [10.1145/2911451.2911457](https://doi.org/10.1145/2911451.2911457). [En línea]. Disponible en: <https://doi.org/10.1145/2911451.2911457> (Acceso: 15-08-2020).
- [20] H. Banaee, M. U. Ahmed y A. Loutfi, “Towards NLG for Physiological Data Monitoring with Body Area Networks”, eng, 2013, pp. 193-197. [En línea]. Disponible en: <http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-30257> (Acceso: 15-08-2020).
- [21] O. Stock et al., “Adaptive, intelligent presentation of information for the museum visitor in PEACH”, en *User Modeling and User-Adapted Interaction*, vol. 17, n.º 3, pp. 257-304, jul. de 2007. doi: [10.1007/s11257-007-9029-6](https://doi.org/10.1007/s11257-007-9029-6). [En línea]. Disponible en: <https://doi.org/10.1007/s11257-007-9029-6> (Acceso: 15-08-2020).
- [22] G. Carenini y J. D. Moore, “Generating and evaluating evaluative arguments”, en *Artificial Intelligence*, vol. 170, n.º 11, pp. 925-952, ago. de 2006. doi: [10.1016/j.artint.2006.05.003](http://www.sciencedirect.com/science/article/pii/S000437020600066X). [En línea]. Disponible en: <http://www.sciencedirect.com/science/article/pii/S000437020600066X> (Acceso: 15-08-2020).
- [23] E. Reiter, R. Robertson y L. M. Osman, “Lessons from a failure: Generating tailored smoking cessation letters”, en *Artificial Intelligence*, vol. 144, n.º 1, pp. 41-58, mar. de 2003. doi: [10.1016/S0004-3702\(02\)00370-3](http://www.sciencedirect.com/science/article/pii/S0004370202003703). [En línea]. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0004370202003703> (Acceso: 15-08-2020).
- [24] H. Venigalla y B. Di Eugenio, “UIC-CSC: The Content Selection Challenge Entry from the University of Illinois at Chicago”, en *Proceedings of the 14th European Workshop on Natural Language Generation*, Sofia, Bulgaria: Association for Computational Linguistics, ago. de 2013, pp. 210-211. [En línea]. Disponible en: <https://www.aclweb.org/anthology/W13-2134> (Acceso: 15-08-2020).
- [25] P. A. Duboue y K. McKeown, “Statistical Acquisition of Content Selection Rules for Natural Language Generation”, en, 2003. doi: [10.7916/D81J9JND](https://doi.org/10.7916/D81J9JND). [En línea]. Disponible en: <https://doi.org/10.7916/D81J9JND> (Acceso: 15-08-2020).

- [26] J. Shaw, “Clause Aggregation Using Linguistic Knowledge”, en *Natural Language Generation*, 1998. [En línea]. Disponible en: <https://www.aclweb.org/anthology/W98-1415> (Acceso: 15-08-2020).
- [27] R. Barzilay y M. Lapata, “Aggregation via Set Partitioning for Natural Language Generation”, en *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, New York City, USA: Association for Computational Linguistics, jun. de 2006, pp. 359-366. [En línea]. Disponible en: <https://www.aclweb.org/anthology/N06-1046> (Acceso: 15-08-2020).
- [28] G. Angeli, C. Manning y D. Jurafsky, “Parsing Time: Learning to Interpret Time Expressions”, en *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Montréal, Canada: Association for Computational Linguistics, jun. de 2012, pp. 446-455. [En línea]. Disponible en: <https://www.aclweb.org/anthology/N12-1049> (Acceso: 05-08-2020).
- [29] K. v. Deemter, M. Theune y E. Krahmer, “Real versus Template-Based Natural Language Generation: A False Opposition?”, en *Computational Linguistics*, vol. 31, n.º 1, pp. 15-24, mar. de 2005. doi: [10.1162/0891201053630291](https://doi.org/10.1162/0891201053630291). [En línea]. Disponible en: <https://www.mitpressjournals.org/doi/abs/10.1162/0891201053630291> (Acceso: 14-08-2020).
- [30] L. Sanby, I. Todd y C. M. Keet, “Comparing the template-based approach to GF: the case of Afrikaans”, en *Num Pages: 50-53*, Association for Computational Linguistics, 2016. [En línea]. Disponible en: <http://pubs.cs.uct.ac.za/id/eprint/1104/> (Acceso: 14-08-2020).
- [31] C. van der Lee, E. Krahmer y S. Wubben, “PASS: A Dutch data-to-text system for soccer, targeted towards specific audiences”, ene. de 2017, pp. 95-104. doi: [10.18653/v1/W17-3513](https://doi.org/10.18653/v1/W17-3513).
- [32] A. Mazzei, C. Battaglino y C. Bosco, “SimpleNLG-IT: adapting SimpleNLG to Italian”, en *Proceedings of the 9th International Natural Language Generation conference*, Edinburgh, UK: Association for Computational Linguistics, sep. de 2016, pp. 184-192. doi: [10.18653/v1/W16-6630](https://doi.org/10.18653/v1/W16-6630). [En línea]. Disponible en: <https://www.aclweb.org/anthology/W16-6630> (Acceso: 12-08-2020).
- [33] *TreeBank*, es, Page Version ID: 117324320, jul. de 2019. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=TreeBank&oldid=117324320> (Acceso: 13-08-2020).
- [34] C. Underwood, *Automated Journalism - AI Applications at New York Times, Reuters, and Other Media Giants*, en-US. [En línea]. Disponible en: <https://emerj.com/ai-sector-overviews/automated-journalism-applications/> (Acceso: 17-08-2020).

- [35] *Google Code Archive - Long-term storage for Google Code Project Hosting*. [En línea]. Disponible en: <https://code.google.com/archive/p/word2vec/> (Acceso: 17-08-2020).
- [36] *Perspective*. [En línea]. Disponible en: <https://www.perspectiveapi.com/#/home> (Acceso: 17-08-2020).
- [37] *Improving Conversations using the Perspective API*, en, ago. de 2017. [En línea]. Disponible en: <https://css-tricks.com/improving-conversations-using-perspective-api/> (Acceso: 02-09-2020).
- [38] *The Juicer*, en. [En línea]. Disponible en: <https://bbcnewslabs.co.uk/projects/juicer/> (Acceso: 17-08-2020).
- [39] WashPostPR, “The Washington Post to use artificial intelligence to cover nearly 500 races on Election Day”, en-US, *Washington Post*, [En línea]. Disponible en: <https://www.washingtonpost.com/pr/wp/2016/10/19/the-washington-post-uses-artificial-intelligence-to-cover-nearly-500-races-on-election-day/> (Acceso: 17-08-2020).
- [40] *Heliograf, el robot del Washington Post, crea 850 artículos en un año | Marketing Directo*. [En línea]. Disponible en: <https://www.marketingdirecto.com/anunciantes-general/medios/heliograf-robot-del-washington> (Acceso: 02-09-2020).
- [41] *Yahoo!*, en-US. [En línea]. Disponible en: <https://automatedinsights.com/customer-stories/yahoo/> (Acceso: 18-08-2020).
- [42] *The company behind the AP’s ‘robot journalists’ is opening up its technology for everyone - The Verge*. [En línea]. Disponible en: <https://www.theverge.com/2015/10/20/9572975/automated-insights-wordsmith-natural-language> (Acceso: 02-09-2020).
- [43] *Social Media Monitoring - Spike*, en-US. [En línea]. Disponible en: <https://www.newswhip.com/spike-social-media-monitoring/> (Acceso: 02-09-2020).
- [44] *Turing-NLG: A 17-billion-parameter language model by Microsoft*, en-US, feb. de 2020. [En línea]. Disponible en: <https://www.microsoft.com/en-us/research/blog/turing-nlg-a-17-billion-parameter-language-model-by-microsoft/> (Acceso: 06-09-2020).
- [45] T. B. Brown et al., *Language Models are Few-Shot Learners*. 2020, _eprint: 2005.14165.
- [46] N. Braun, M. Goudbeek y E. Kraemer, “The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch”, en *Proceedings of the 9th International Natural Language Generation conference*, Edinburgh, UK: Association for Computational Linguistics, sep. de 2016, pp. 74-78. doi: [10.18653/v1/W16-6612](https://doi.org/10.18653/v1/W16-6612). [En línea]. Dis-

- ponible en: <https://www.aclweb.org/anthology/W16-6612> (Acceso: 20-08-2020).
- [47] L. Gatti, C. van der Lee y M. Theune, “Template-based multilingual football reports generation using Wikidata as a knowledge base”, en *Proceedings of the 11th International Conference on Natural Language Generation*, Tilburg University, The Netherlands: Association for Computational Linguistics, nov. de 2018, pp. 183-188. doi: [10.18653/v1/W18-6523](https://doi.org/10.18653/v1/W18-6523). [En línea]. Disponible en: <https://www.aclweb.org/anthology/W18-6523> (Acceso: 15-08-2020).
- [48] *Wikidata*. [En línea]. Disponible en: https://www.wikidata.org/wiki/Wikidata:Main_Page (Acceso: 20-08-2020).
- [49] *Git*. [En línea]. Disponible en: <https://git-scm.com/> (Acceso: 05-09-2020).
- [50] *Build software better, together*, en. [En línea]. Disponible en: <https://github.com> (Acceso: 05-09-2020).
- [51] BillWagner, *Documentos de C#: inicio, tutoriales y referencias*. es-es. [En línea]. Disponible en: <https://docs.microsoft.com/es-es/dotnet/csharp/> (Acceso: 05-09-2020).
- [52] *Angular*. [En línea]. Disponible en: <https://angular.io/> (Acceso: 05-09-2020).
- [53] J. Deacon, “Model-view-controller (mvc) architecture”, *Online* [Citado em: 10 de março de 2006.] <http://www.jdl.co.uk/briefings/MVC.pdf>, 2009.
- [54] *MySQL*. [En línea]. Disponible en: <https://www.mysql.com/> (Acceso: 05-09-2020).
- [55] *MySQL :: MySQL Workbench*. [En línea]. Disponible en: <https://www.mysql.com/products/workbench/> (Acceso: 05-09-2020).
- [56] *IDE de Visual Studio, editor de código, Azure DevOps y App Center*, es-ES. [En línea]. Disponible en: <https://visualstudio.microsoft.com/es/> (Acceso: 05-09-2020).
- [57] *Visual Studio Code - Code Editing. Redefined*, en. [En línea]. Disponible en: <https://code.visualstudio.com/> (Acceso: 05-09-2020).
- [58] *Apache Subversion*. [En línea]. Disponible en: <https://subversion.apache.org/> (Acceso: 05-09-2020).
- [59] P. B. Kruchten, “The 4+ 1 view model of architecture”, *IEEE software*, vol. 12, n.º 6, pp. 42-50, 1995, Publisher: IEEE.
- [60] D. D. McDonald, “Issues in the choice of a source for natural language generation”, *Computational Linguistics*, vol. 19, n.º 1, pp. 191-197, 1993.
- [61] C. Mellish y R. Dale, “Evaluation in the context of natural language generation”, *Computer Speech & Language*, vol. 12, n.º 4, pp. 349-373, 1998, Publisher: Elsevier.

- [62] E. Reiter y S. Sripada, “Should corpora texts be gold standards for nlg?”, en *Proceedings of the International Natural Language Generation Conference*, 2002, pp. 97-104.
- [63] C. B. Callaway y J. C. Lester, “Narrative prose generation”, *Artificial Intelligence*, vol. 139, n.º 2, pp. 213-252, 2002, Publisher: Elsevier.
- [64] M. Lapata, “Automatic evaluation of information ordering: Kendall’s tau”, *Computational Linguistics*, vol. 32, n.º 4, pp. 471-484, 2006, Publisher: MIT Press.
- [65] J. Lester y B. Porter, “Developing and empirically evaluating robust explanation generators: The KNIGHT experiments”, *Computational Linguistics*, vol. 23, n.º 1, pp. 65-101, 1997.
- [66] S. Sripada, E. Reiter y L. Hawizy, “Evaluation of an NLG system using post-edit data: Lessons learnt”, en *Proceedings of the Tenth European Workshop on Natural Language Generation (ENLG-05)*, 2005.
- [67] M. White, R. Rajkumar y S. Martin, “Towards broad coverage surface realization with CCG”, en *Proceedings of the Workshop on Using Corpora for NLG: Language Generation and Machine Translation (UCNLG+ MT)*, 2007, pp. 267-276.
- [68] K. Binsted y G. Ritchie, “Computational rules for generating punning riddles.”, *Humor: International Journal of Humor Research*, 1997, Publisher: Walter de Gruyter.
- [69] A. Gatt y A. Belz, *Introducing shared tasks to NLG: the TUNA shared task evaluation challenges, Empirical methods in natural language generation: data-oriented methods and empirical evaluation*. Springer-Verlag, Berlin, Heidelberg, 2010.
- [70] *Medion Akoya M80 PC888 P66033 Intel Core i3-8100/8GB/1TB+128GB SSD | PcComponentes.com*. [En línea]. Disponible en: <https://www.pccomponentes.com/medion-akoya-m80-pc888-p66033-intel-core-i3-8100-8gb-1tb-128gb-ssd> (Acceso: 02-09-2020).
- [71] *Philips V Line 272V8A/00 27"LED FullHD | PcComponentes.com*. [En línea]. Disponible en: <https://www.pccomponentes.com/philips-v-line-272v8a-00-27-led-fullhd> (Acceso: 02-09-2020).
- [72] *IVA - Agencia Tributaria*. [En línea]. Disponible en: https://www.agenciatributaria.es/AEAT.internet/Inicio/_Segmentos_/Empresas_y_profesionales/Empresarios_individuales_y_profesionales/IVA/IVA.shtml (Acceso: 02-09-2020).
- [73] *Búsqueda de empleo en Glassdoor | Encuentra el empleo perfecto para ti*. [En línea]. Disponible en: <https://www.glassdoor.es/index.htm> (Acceso: 06-09-2020).
- [74] *BOE.es - Documento consolidado BOE-A-2018-16673*. [En línea]. Disponible en: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673> (Acceso: 02-09-2020).

- [75] E. E. PAÍS, *Fake news en EL PAÍS*, es. [En línea]. Disponible en: <https://elpais.com/noticias/noticias-falsas> (Acceso: 06-09-2020).
- [76] david.llorente, *Narrativa generates automatic news for Diario Sport*, en-US, nov. de 2018. [En línea]. Disponible en: <https://demo.narrativa.com/spanish-narrativa-generacion-automatica-de-noticias-para-el-diario-sportnarrativa-generacion-automatica-de-noticias-para-el-diario-sport/> (Acceso: 06-09-2020).
- [77] *EFE y Narrativa acuerdan potenciar contenidos basados en inteligencia artificial*, es, jul. de 2019. [En línea]. Disponible en: <https://www.agenciaefe.es/efe-narrativa-firman-acuerdo-potenciar-los-contenidos-basados-inteligencia-artificial/> (Acceso: 06-09-2020).
- [78] *Sobre Nosotros - Áreas de actividad - Corporación*, es. [En línea]. Disponible en: <https://www.rtve.es/corporacion/sobre-nosotros/index.shtml> (Acceso: 29-08-2020).
- [79] RTVE, *Áreas de actividad | TVE*, es, Section: RTVE, mayo de 2014. [En línea]. Disponible en: <https://www.rtve.es/rtve/20140522/areas-actividad-tve/942142.shtml> (Acceso: 29-08-2020).
- [80] —, *Áreas de actividad | RNE*, es, Section: RTVE, mayo de 2014. [En línea]. Disponible en: <https://www.rtve.es/rtve/20140522/areas-actividad-rne/942144.shtml> (Acceso: 29-08-2020).
- [81] —, *Áreas de actividad | Área Digital*, es, Section: RTVE, mayo de 2014. [En línea]. Disponible en: <https://www.rtve.es/rtve/20140522/areas-actividad-rtvees/942160.shtml> (Acceso: 29-08-2020).
- [82] —, *Áreas de actividad | Instituto RTVE*, es, Section: RTVE, mayo de 2014. [En línea]. Disponible en: <https://www.rtve.es/rtve/20140522/areas-actividad-instituto-rtve/942161.shtml> (Acceso: 29-08-2020).
- [83] *¿Qué es Impulsa Visión RTVE?*, es, Section: RTVE, sep. de 2018. [En línea]. Disponible en: <https://www.rtve.es/rtve/20180928/impulsa-vision-rtve/1813101.shtml> (Acceso: 29-08-2020).
- [84] *Innovación - RTVE.es*. [En línea]. Disponible en: <https://www.rtve.es/rtve/20140526/rc-compromisos/943260.shtml> (Acceso: 27-08-2020).
- [85] A. Osterwalder e Y. Pigneur, *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, en. John Wiley & Sons, jul. de 2010, Google-Books-ID: UzuTAAwAAQBAJ.
- [86] L. Fahey y V. Narayanan, “Análisis macro-ambiental en gestión estratégica”, *St. Paul*, 1986.
- [87] M. Porter, “E., 1979. How Competitive Forces Shape Strategy”, *Harvard Business Review*, vol. 57, n.º 2, pp. 137-145, 1979.

- [88] A. S. Humphrey, “SRI Alumni Association Newsletter”, *SWOT Analysis for Management Consulting*, 2005.
- [89] *Historia del Análisis PEST-PESTEL*. [En línea]. Disponible en: <https://foda-dafo.com/historia-del-analisis-pest-pestel/> (Acceso: 01-09-2020).
- [90] F. Aguilar, “Análisis del entorno empresarial”, *Sevilla: Universidad de Sevilla*, 1967.
- [91] *El Gobierno cambia la cúpula de RTVE en vísperas de unas posibles elecciones*, es, Section: Noticias, sep. de 2019. [En línea]. Disponible en: https://www.abc.es/play/television/noticias/abci-gobierno-prepara-cambios-inminentes-cupula-directiva-rtve-201909041024_noticia.html (Acceso: 01-09-2020).
- [92] *Nueva crisis en los informativos de RTVE: ceses en la alta dirección de RNE y dimisiones en TVE*, es, Section: Series, jun. de 2020. [En línea]. Disponible en: https://www.abc.es/play/series/noticias/abci-nueva-crisis-informativos-rtve-ceses-alta-direccion-y-dimisiones-202006302223_noticia.html (Acceso: 01-09-2020).
- [93] *Economía de España*, es, Page Version ID: 128155360, jul. de 2020. [En línea]. Disponible en: https://es.wikipedia.org/w/index.php?title=Econom%C3%ADa_de_Espa%C3%B1a&oldid=128155360 (Acceso: 01-09-2020).
- [94] *INE. Instituto Nacional de Estadística*, es. [En línea]. Disponible en: <https://ine.es/> (Acceso: 01-09-2020).
- [95] *Corporación RTVE - Web Oficial*, es. [En línea]. Disponible en: <https://www.rtve.es//corporacion/index.shtml> (Acceso: 01-09-2020).
- [96] “Ley 17/2006, de 5 de junio, de la radio y la televisión de titularidad estatal.”, es, p. 22,
- [97] L. Ley, “PRIMER MANDATO-MARCO A LA CORPORACIÓN RTVE PREVISTO EN EL ARTÍCULO 4 DE LA LEY 17/2006, DE 5 DE JUNIO, DE LA RADIO Y LA TELEVISIÓN DE TITULARIDAD ESTATAL APROBADO POR EL PLENO DEL CONGRESO DE LOS DIPUTADOS EN SU SESIÓN DEL DÍA 11 DE DICIEMBRE DE 2007 Y POR EL PLENO DEL SENADO EN SU SESIÓN DEL DÍA 12 DE DICIEMBRE DE 2007”, es, p. 16,
- [98] “Ley 8/2009, de 28 de agosto, de financiación de la Corporación de Radio y Televisión Española.”, es, p. 12,
- [99] A. Iranzo y T. Latorre, *La noticia: manual práctico de redacción*, es, 1.^a ed. Universitat Jaume I, 2019. doi: [10.6035/Sapientia149](https://doi.org/10.6035/Sapientia149). [En línea]. Disponible en: <http://hdl.handle.net/10234/181827> (Acceso: 27-08-2020).
- [100] *Cómo redactar una noticia*, es. [En línea]. Disponible en: <https://es.wikihow.com/redactar-una-noticia> (Acceso: 27-08-2020).

- [101] F. Rios y G. Edith, “Mejora de procesos en una empresa textil exportadora mediante la metodología Six Sigma”, spa, *Repositorio de Tesis - UNMSM*, 2017, Accepted: 2017-12-15T16:25:02Z Publisher: Universidad Nacional Mayor de San Marcos. [En línea]. Disponible en: <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/6732> (Acceso: 02-09-2020).
- [102] D. A. Collier y J. R. Evans, *Administración de operaciones*. Cengage Learning, 2016.
- [103] L. J. Krajewski, L. P. Ritzman y M. K. Malhotra, *Administración de operaciones: procesos y cadenas de valor*. Pearson Educación México, 2008.
- [104] M. Hammer y J. Champy, *reingeniería*. Editorial Norma, 1994.
- [105] G. de la Calidad Total, “F. Sáez Vacas, O. García, J. Palao y P. Rojo”, *Innovación Tecnológica en las Empresas [en línea]*. Departamento de Ingeniería de Sistemas Telemáticos. Escuela Técnica Superior de Ingenieros de Telecomunicación. Universidad Politécnica de Madrid,
- [106] M. Imai, *Kaizen*. Random House Business Division New York, 1986, vol. 201.
- [107] O. A. M. Ceballos, “PROYECTO DE MEJORA Y ESTANDARIZACIÓN DEL PROCESO DE COBROS EN INDUSTRIAS PROMAR S. A. ARMENIA - QUINDÍO - COLOMBIA”, es, p. 87,
- [108] R. Moen y C. Norman, *Evolution of the PDCA cycle*. Citeseer, 2006.
- [109] K.-b. K. G. Bulsuk, *A diagram showing the PDCA Cycle*, nov. de 2008. [En línea]. Disponible en: https://commons.wikimedia.org/wiki/File:PDCA_Cycle.svg (Acceso: 04-09-2020).
- [110] E. Bonilla-Pastor-de-Céspedes, B. Díaz-Garay, F. Kleeberg-Hidalgo y M.-T. Noriega-Aranibar, *Mejora continua de los procesos: herramientas y técnicas*, spa. Universidad de Lima. Fondo Editorial, 2010, Accepted: 2020-04-28T16:50:28Z Publication Title: Repositorio Institucional - Ulima. [En línea]. Disponible en: <http://repositorio.ulima.edu.pe/handle/ulima/10832> (Acceso: 03-09-2020).
- [111] NatiPeña, *Español: Six Sigma consta de cinco fases de mejora continua, las cuales permiten cumplir con los objetivos de la metodología*. Mar. de 2016. [En línea]. Disponible en: https://commons.wikimedia.org/wiki/File:Fases_seis_sigma.png (Acceso: 04-09-2020).
- [112] J. M. B. Peña y J. A. Ramos-Yzquierdo, *Dirección de procesos digitales: crear organizaciones más eficientes*. Ediciones Universidad de Navarra, 2004.
- [113] S. Talwar, M. Talwar, P. Kaur y A. Dhir, “Consumers’ resistance to digital innovations: A systematic review and framework development”, *Australasian Marketing Journal (AMJ)*, 2020, Publisher: Elsevier.

- [114] P.-T. Chen y S.-C. Kuo, "Innovation resistance and strategic implications of enterprise social media websites in Taiwan through knowledge sharing perspective", en, *Technological Forecasting and Social Change*, vol. 118, pp. 55-69, mayo de 2017. doi: [10.1016/j.techfore.2017.02.002](https://doi.org/10.1016/j.techfore.2017.02.002). [En línea]. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0040162517301555> (Acceso: 03-09-2020).
- [115] B. Booms, "Marketing strategies and organizational structures for service firms", *Marketing of services*, 1981, Publisher: American Marketing Association.
- [116] E. J. McCarthy y W. D. Perreault Jr, "Basic Marketing, Richard D", *Irwin, Homewood, IL*, vol. 181, 1964.
- [117] C. L. Goi, "A Review of Marketing Mix: 4Ps or More?", en, *International Journal of Marketing Studies*, vol. 1, n.º 1, p. 14, 2009.