

Inteligencia artificial aplicada al periodismo: traducción automática y recomendación de contenidos en el proyecto “A European Perspective” (UER)

Artificial intelligence in journalism: Automatic translation and recommendation system in the project “A European Perspective” (EBU)

João Canavilhas. Universidade da Beira Interior. Portugal.

jc@ubi.pt

[CV]     

Financiado por Labcom-Comunicação e Artes, un laboratório de Investigación apoyado por la Fundação para a Ciência e Tecnologia (UIDB/00661/2020)

Cómo citar este artículo / Referencia normalizada

Canavilhas, J. (2022). Inteligencia artificial aplicada al periodismo: traducción automática y recomendación de contenidos en el proyecto “A European Perspective” (UER). *Revista Latina de Comunicación Social*, 80, 1-13. <https://www.doi.org/10.4185/RLCS-2022-1534>

RESUMEN. Introducción: El 1 de julio de 2021, varios organismos públicos de radiodifusión europeos lanzaron oficialmente el proyecto “A European Perspective”. Se trata de un servicio de intercambio de noticias en línea desarrollado dentro del ámbito de la UER - Unión Europea de Radiodifusión, que reunió a diez empresas de radio y televisión públicas de nueve países. Este proyecto utiliza la Inteligencia Artificial para la traducción y recomendación de contenidos producidos por los participantes y que pueden ser utilizados libremente en sus sitios web. Entre los objetivos del proyecto se encuentran la lucha contra la desinformación y la afirmación de los valores europeos en un contexto social en el que las redes sociales han ido ganando cada vez más influencia entre la opinión pública. **Metodología:** Este trabajo es un estudio de caso y, además de explicar cómo funciona el proyecto, hace un análisis de los contenidos compartidos en los dos primeros meses de funcionamiento en el sitio web de la televisión pública portuguesa RTP- Rádio Televisão Portuguesa. **Resultados y Conclusiones:** El proyecto está cumpliendo los objetivos inicialmente definidos por la UER, especialmente en cuanto a ofrecer diferentes perspectivas sobre temas que son objeto de procesos de desinformación, como la pandemia, y solidificando una visión europea basada en el rigor y la transparencia de la información producida por los operadores de radio y televisión pública.

PALABRAS CLAVE: Inteligencia Artificial; Periodismo; Televisión; UER; Traducción automática; Sistemas de recomendación; Una mirada europea.

ABSTRACT: Introduction: On July 1st, 2021, several European public broadcasters officially launched the project “A European Perspective”. It is an online news exchange service developed within the EBU - European Broadcasting Union, which brought together ten public broadcast companies from nine countries. This project uses Artificial Intelligence for the translation and recommendation of content produced by the participants and that can be freely used on their websites. The main objectives of the project are the fight against misinformation and the affirmation of the European values in a social context in which social networks have been gaining more and more influence among public opinion. **Methodology:** This paper, a case study, explains how the project works and analyses the type of content shared in the first two months of operation on the website of RTP- Rádio Televisão Portuguesa.

Results and Conclusions The project is meeting the objectives initially defined by the EBU, especially in terms of offering different perspectives on issues that are subject to disinformation processes, such as the pandemic, and solidifying a European vision based on rigor and transparency of the news produced by public European broadcasters.

KEYWORDS: Artificial Intelligence; Journalism; Television; EBU; Automatic translation; Recommendation system; A European perspective.

CONTENIDO

1. Introducción. 2. Objetivos 3. Metodología. 4. Resultados y Discusión. 4.1 Traducción automática y Sistema de recomendaciones. 4.2 Ubicación del espacio en el sitio web. 4.3 Orígenes y Temas. 4.4 Original vs versión traducida. 5. Conclusiones. 6. Referencias. 7. Autor.

CONTENT

1. Introduction. 2. Objectives 3. Methodology. 4. Results and Discussion. 4.1 Machine translation and recommendation system 4.2 Location of space on the website. 4.3 Origins and Themes. 4.4 Original vs translated version. 5. Conclusions. 6. References. 7. Author.

1.Introducción

La inteligencia artificial (IA) sigue ganando terreno en la sociedad. Si hasta cierto tiempo solo se utilizaba en determinados nichos, hoy en día el uso de la IA se ha generalizado y nos acompaña a diario, incluso en objetos personales, como los teléfonos móviles. Aun así, la literatura sobre el tema sigue siendo algo escasa, centrándose más en la parte informática y menos en los impactos de la IA en la sociedad, con lo cual los estudios como este son importantes para analizar de qué forma los diferentes sectores están usando la inteligencia artificial.

El concepto se discutió por primera vez en 1956, en un congreso organizado por John McCarthy, en Dartmouth College. En aquel momento, Inteligencia Artificial (IA) se entendía como la posibilidad de que las máquinas pudieran imitar algunos aspectos de la inteligencia humana (Russell & Norvig 2016). A lo largo de las décadas, los enfoques conceptuales han evolucionado en paralelo a su propia aplicación, lo que ha permitido distinguir diferentes enfoques epistemológicos de la Inteligencia Artificial. El más prevalente defiende a la IA como un agente racional, es decir, su propósito es ofrecer más eficiencia, con o sin imitación del pensamiento/comportamiento humano. En esta perspectiva, la IA se define como “el estudio de los agentes que existen en un entorno y perciben y actúan” (Russell & Norvig, 2016, p. 7). En esta formulación, el concepto se aleja de perspectivas más cercanas a la IA como mimetismo del comportamiento humano, es decir, “el arte de crear máquinas que realizan funciones que requieren inteligencia cuando las realizan las personas” (Kurzweil, 1990, p.117).

Con una presencia importante en diversos sectores de actividad, la Inteligencia Artificial también ha llegado al campo del periodismo, concretamente a través del procesamiento del lenguaje natural (PLN) aplicado a la transformación de datos en noticias (Silverman, 2013; Harcup, 2014). Una de sus aplicaciones más conocidas es la producción automática de textos (Silverman, 2013; Carlson, 2015; Lokot & Diakopoulos, 2016) con recurso a bots, situación que ha conducido a la propuesta de términos como “robotperiodismo” (Carlson, 2015). Esta opción ha sido una respuesta a los problemas económicos del sector (Papadimitriou, 2016), uno de los cuales es la necesidad de mantener o incrementar la producción de noticias en un escenario de despidos (Graefe, 2016). Sin embargo, esta situación da lugar a nuevos problemas relacionados con la calidad del producto final (Manfredi Sánchez & Ufarte Ruiz, 2019) o con la ética (Ufarte Ruiz, Calvo Rubio & Murcia Verdú, 2021).

Los primeros experimentos con producción automática de noticias tienen ya más de una década. En 2010, The Big Ten Network, una asociación entre Fox Networks y Big Ten Conference, lanzó un servicio automatizado de producción de noticias deportivas (Lohr, 2011), utilizando software desarrollado por Narrative Science. No fue la primera experiencia en la producción automática de textos periodísticos, pero fue una de las primeras en utilizar Inteligencia Artificial, y permitió una mejora significativa en el producto final.

Pero el uso frecuente de la Inteligencia Artificial en este campo se remonta a 2014, cuando Los Angeles Times lanzó el primer bot (Flores-Vivar, 2018). Poco después, Associated Press también automatizó

la producción de noticias relacionadas con la economía (Brandom, 2014; Graefe, 2016; Lichterman, 2017), situación que dio mayor visibilidad a esta forma de producción. El hecho de que la sección de Economía fuera una de las primeras en ser testeadas no es casualidad: es una especialidad en la que hay mucha información online y donde los datos son imprescindibles, lo que facilita el recurso al PLN para convertirlos en textos. Lo mismo ocurre con la información deportiva y, por lo tanto, estas son las dos especialidades periodísticas donde mejor se hace la transformación de datos en narrativas textuales con poca o ninguna intervención humana (Carlson, 2015).

La Inteligencia Artificial también puede facilitar algunas actividades específicas del proceso de producción de noticias, como la búsqueda de más objetividad para evitar errores y ambigüedades (Essenfelder et al, 2019), la detección de tendencias informativas (Steiner, 2014) para construir la agenda, la recolección y organización de la información (Diakopoulos 2019), la traducción automática y desarrollo de sistemas de recomendación de noticias (Helberger, 2019; Túñez-López et al., 2021), o en la lucha contra la desinformación (Flew, 2012; Silverman, 2013; Manfredi Sánchez & Ufarte Ruiz, 2020). Estas últimas razones están directamente relacionadas con las posibilidades técnicas del proyecto estudiado en este trabajo y coinciden con algunos de los diez elementos del periodismo identificados por Kovach & Rosenstiel (2014)

Al analizar un proyecto de carácter periodístico, este trabajo cruza las dos dimensiones mencionadas anteriormente, buscando verificar cómo la IA permite profundizar en el compromiso con la verdad, hacer más interesante y relevante la información pertinente, contrastando fuentes, ofreciendo proporcionalidad entre hechos y noticias y, finalmente, cómo puede contribuir al fortalecimiento del periodismo como foro de discusión. Cumpliendo con estos principios enunciados por Kovach & Rosenstiel (2014), el proyecto “A European Perspective” puede ayudar a las emisoras públicas a reforzar su credibilidad (Túñez-López, et al, 2019) y a fortalecer su importante papel en las sociedades democráticas (Balkin, 2018), en este caso a través de su presencia informativa en la Web.

2. Objetivos

El proyecto “A European Perspective” se describe como “un servicio de noticias colaborativo que permite a las audiencias acceder a contenido confiable de otros países a través de un canal en línea seleccionado” (Unión Europea de Radiodifusión [UER], 2021). En este punto, surge una preocupación inicial - proporcionar información confiable - objetivo que recuerda el primer elemento identificado por Kovach & Rosenstiel (2014) - “la primera obligación del periodismo es con la verdad” (p.27) - porque la confianza está umbilicalmente conectada a la veracidad de los hechos denunciados. De forma complementaria, se menciona que el acceso a información de otros países en la lengua materna permite al lector comparar su realidad sociopolítica con la de otros países, lo que “actuará como un contrapeso a las campañas de desinformación dirigidas a explotar las fallas entre comunidades.” (UER, 2021). La pandemia de la COVID-19, el cambio climático y la migración se presentan como ejemplos de problemas transversales que han sido objeto de manipulación y, por ello, merecen una atención especial del proyecto una vez que la información de calidad es una de las más importantes armas para luchar contra la desinformación (Larrondo-Ureta, et al, 2021).

En este contexto surge un primer objetivo del proyecto: contribuir a la creación de una esfera pública europea compartiendo historias arraigadas en valores europeos fundamentales, utilizando contenidos que ayuden a los ciudadanos a contextualizar su realidad local en el contexto europeo. En este caso, podemos establecer una conexión con el séptimo elemento propuesto por Kovach & Rosenstiel (2014) – “Debe esforzarse por hacer que lo significativo sea interesante y relevante” (p.27) - ya que busca transformar la información de una determinada región en contenido que despierta la atención de los usuarios de toda Europa. Como señala Arendt (2005), “la realidad de la esfera pública tiene su raíz en la presencia simultánea de innumerables perspectivas y aspectos en los que se presenta el mundo común y para los que no conviene inventar el denominador común” (p. 66). La publicación de estas diversas perspectivas acaba constituyendo un foro de confrontación crítica de temas que interesan a Europa, un espacio de debate coincidente que también invoca el sexto elemento característico del periodismo: “Debe ser un foro de crítica y compromiso público” (Kovach & Rosenstiel, 2014, p. 27). Emerge así una primera cuestión de investigación:

¿los temas de las noticias publicadas se corresponden con los que más preocupan a los europeos en el momento de la muestra?

Dentro del proyecto, la parte que funciona para responder a esta cuestión es el sistema de recomendación (PEACH). Como bien nota (Helberger, 2019), los sistemas algorítmicos de recomendación, en sí mismos, nos son buenos o malos para la democracia: La forma como se utilizan son lo que hace la diferencia. En el caso presente, las propuestas del sistema son complementadas con la intervención del editor, constituyéndose como un equipo híbrido. Confirmándose la publicación de noticias sobre temas de fractura en la sociedad, y presentándose perspectivas originarias de diferentes países, el sistema estará contribuyendo para la construcción de una esfera pública europea.

El proyecto tiene igualmente como objetivo luchar contra la desinformación, ofreciendo a los consumidores información veraz y con posibilidad de contrastar fuentes. Como bien notan Manfredi Sánchez y Ufarte Ruiz (2020), “la veracidad es una cualidad que requiere una serie de competencias y habilidades profesionales como la práctica de la doble revisión de las fuentes, la calidad en la redacción y el manejo del idioma” (p. 54). En este caso, además de que anteriormente el periodista cumplió este recorrido profesional, el sistema permite al usuario hacerlo por sí mismo. Es decir, el usuario tiene la posibilidad de hacer una doble revisión de fuentes (leer noticias sobre el tema que le interesa publicadas en otros medios) y, además, hacerlo en su propio idioma porque el sistema hace la traducción.

Otra forma de combatir la desinformación es ofrecer información contextualizada, lo que recuerda el octavo elemento enumerado por Kovach y Rosenstiel (2014): “Debe presentar la noticia de una manera integral y proporcional” (p. 27). También en este caso, la oferta de otras perspectivas funciona como una segunda narración que hace más comprensible el hecho y otórgale un mayor valor social en función del énfasis que se le atribuye. Así que la segunda cuestión de investigación es la siguiente:

¿Las noticias publicadas en el espacio “Um Olhar Europeu” (RTP) cumplen su objetivo de luchar contra la desinformación a través de la publicación de contenidos sobre un mismo tema originarios de radiotelevisión de diferentes países?

En este caso no se trata de la aplicación directa de la IA porque no hablamos de un sistema de revisión automática de fuentes que presente un producto final, sino de ofrecer diversas fuentes sobre un mismo tema para que el usuario haga su propio fact-checking. En ese sentido, la Inteligencia Artificial actúa en la traducción automática de las noticias originales (EUROVOX) y en el sistema de recomendación (PEACH), con el objetivo de que los usuarios puedan ellos mismo confirmar la veracidad de información previamente recibida por otro canal.

Finalmente, entre los principios del proyecto se destaca la necesidad de dar atención a la “diversidad e inclusión, asegurándose de que el contenido de las noticias promueva las voces, las opiniones y los testimonios personales de todo el panorama social de Europa, incluidos los de los grupos vulnerables o marginados” (UER, 2021). Este es otro principio que recuerda un elemento identificado por Kovach & Rosenstiel (2014), en este caso el sexto, que apunta a la necesidad de “Vigilar el poder y poner voz al que no la tiene” (p.27). Así que la tercera cuestión de investigación en este trabajo es la opinión minoritarias y más frágiles?

La respuesta a esta cuestión representa una oportunidad para que el periodismo pueda corregir una limitación de los sistemas de IA, sobre todo los algoritmos de aprendizaje automático, que tienden a replicar las discriminaciones de las comunidades marginadas (Broussard et al, 2019). Al combinar un sistema de recomendaciones basado en IA (PEACH) con la posibilidad del editor local adjuntar sus conocimientos, se da “un enfoque híbrido, en el que tratamos a los seres humanos como expertos y utilizamos la inteligencia artificial para ayudarlos y aumentarlos, en lugar de suplantarlos” (Broussard et al, 2019, p. 683)

3. Metodología

El 1 de julio de 2021, RTP lanzó oficialmente el espacio “Um Olhar Europeu”. El proyecto involucra a diez empresas de radio y televisión públicas y se basa en Eurovision News Monitoring Tool, una herramienta de la Unión Europea de Radiodifusión (UER). Las empresas son RTP (Portugal), RTBF (Bélgica), YLE (Finlandia), France Télévisions (Francia), BR / ARD (Alemania), RTE (Irlanda), RAI (Italia), RTVE (España), SWI (Suiza) y ART (Francia/Alemania). Recientemente, la BBC se unió al proyecto, pero en esta etapa solo proporciona contenido que puede ser consultado, pero no republicado por los participantes. Este estudio podría tener al menos diez enfoques, tantos como el número de medios involucrados en el proyecto, pero en este caso se lleva a cabo desde la perspectiva de las pu-

blicaciones puestas a disposición por RTP - Radio e Televisão de Portugal, el representante portugués en este proyecto.

Este trabajo es un estudio mixto con recurso de dos técnicas: un análisis de contenido y entrevistas. Todo empezó con una toma de datos, en este caso la recolección de todas las noticias publicadas en la página web de RTP durante las primeras ocho semanas iniciales del proyecto. Se decidió iniciar el análisis solamente en la tercera semana considerando que los primeros días se suelen utilizar para realizar pequeñas mejoras en el sistema. Así, el período de recolección empezó el 20 de julio y terminó el 15 de septiembre de 2021. Durante este período se publicaron 54 noticias de todas las estaciones participantes en el proyecto.

Se utilizó el análisis de contenido, siguiendo los tres pasos recomendados por Bardin (1977): el preanálisis, que precedió a la planificación del trabajo, es decir, el período de recolección de la muestra, la definición de la tipología de datos a analizar y su organización en indicadores; la recolección y análisis de la muestra; y, finalmente, el procesamiento de los datos y su respectiva interpretación.

Consciente de que algunas conclusiones requerían un comentario de RTP, se utilizó igualmente una entrevista semiestructurada, en este caso enviando las preguntas por correo electrónico al responsable del proyecto, Alexandre Brito (RTP). El primer correo (28 de julio) para saber informaciones sobre la participación de RTP en el proyecto y esclarecer algunas dudas, y el segundo (22 de septiembre) con preguntas para interpretar algunos resultados. La elección de esta técnica está relacionada con la espontaneidad asociada a ella y porque permite una mejor decodificación de algunos resultados sin una interpretación especulativa del investigador.

4. Resultados y Discusión

En términos tecnológicos, la herramienta de seguimiento de noticias de Eurovisión se basa en dos instrumentos - EUROVOX y PEACH - sistemas desarrollados por la propia UER en estrecha colaboración con los editores y periodistas de las empresas participantes. Según Sebastien Noir (2021), responsable del departamento de ingeniería de software de la UER, se intentó desarrollar un sistema flexible y sencillo porque “el objetivo es una herramienta fácil de usar, que no requiera formación y que proporcione valor inmediato al usuario” (p.11).

Los contenidos ingresan al servidor de dos formas: compartidos por la empresa que los produjo o a través del propio sistema, que busca nuevos contenidos en periodos de tiempo previamente definidos. Después de cargar estos contenidos en el servidor, cada unidad de información se convierte a un formato estándar que identifica los diferentes elementos: título, subtítulo, lead, párrafos, imágenes, gráficos, vídeos, sonidos, citas, etc.

La herramienta EUROVOX transcribe y traduce automáticamente los contenidos mediante Inteligencia Artificial. A continuación, entra en acción PEACH, un sistema de recomendaciones que se complementa con el trabajo de los editores en la selección de noticias. Los editores pueden elegir el contenido buscando por categoría/sección de noticias, por fuente o por fecha. Una vez hecha la elección, el proceso entra en la parte final del proceso: publicar en sitios web y aplicaciones.

Cada empresa participante eligió un espacio específico para poner a disposición estos contenidos. En el caso de RTP (PT), RTE (IE) y RAI (IT) por ejemplo, se encuentra en la subpágina “Noticias” ya que son canales generalistas; por otro lado, en Swissinfo y FranceInfo, el espacio aparece en la página de inicio por ser canales de información. En otros casos, como RTVE, resulta difícil encontrar el espacio y solamente una busca en el sitio permite encontrarla.



Figura 1: Noticias destacadas en el sitio web da RTP
Fuente: RTP

Si bien el diseño e identificación del espacio son los mismos para todos los participantes (widget), la cantidad de noticias disponibles puede variar. En uno de los días en que se realizó el análisis (22 de septiembre de 2021), RTP ofrecía 6 noticias, RTVE tenía 7, Swissinfo 62 (mes actual y mes anterior), Franceinfo 24 y RTBF presentaba 70.

Una nota final para mencionar que en la mayoría de los casos la noticia se presenta cronológicamente (la más reciente a la izquierda), situación que no ocurre en el espacio de RTP: el operador portugués optó por publicar en “modo aleatorio” para transmitir la idea de renovación, comenta el responsable. Además de estas dos opciones, y dentro del espíritu abierto del proyecto, también existe la posibilidad de que cada operador haga una elección manual.

4.1 Traducción automática y Sistema de recomendaciones

EUROVOX está compuesto por dos partes: una API, que permite la integración de los diferentes sistemas de producción de los participantes y una interfaz, que realiza las traducciones mediante herramientas online (Microsoft Azure, Google, Speechmatics, DeepL, etc.) para mejorar el resultado final. Esta integración permite ahorrar recursos porque normaliza los contenidos y genera un flujo noticioso que puede proveer información en diferentes lenguas presionando un botón. Debe destacarse que, a pesar de la automatización, los editores pueden hacer correcciones, situación que ocurrió en varias ocasiones durante el período analizado.

Una de las grandes innovaciones del sistema es la traducción de contenidos audiovisuales. Estas traducciones se realizan de forma automática, pero también se pueden mejorar en términos de texto, sonido y tiempo, utilizando en las traducciones una voz-off sintética. El sistema está preparado para producir automáticamente píldoras de vídeo para las redes sociales, resumiendo noticias y traducíendolas a otras lenguas.

Las ventajas de este sistema son su flexibilidad y facilidad de uso en los procesos de traducción, transcripción y preparación de contenidos en varios idiomas. Los contenidos están etiquetados (tags) en varios idiomas para facilitar su identificación en los sistemas de recomendación.

Por su parte, PEACH - Personalisation and Recommendation Ecosystem, se puede resumir en la frase utilizada por UER (2021): “El contenido adecuado en el momento adecuado para la persona adecuada en el dispositivo adecuado” (s.f). Mediante Inteligencia Artificial, este sistema de recomendaciones, que ha sido desarrollado colaborativamente, tiene código abierto y busca adecuar la oferta informativa a los intereses del usuario, combinando las preferencias de la audiencia con las tendencias informativas de cada momento. Existen algoritmos para todo tipo de contenido (texto, sonido, vídeo, texto, etc.), con alrededor de 20 millones de elementos que se recopilan por día. Cada contenido se convierte en una representación vectorial utilizando modelos previamente entrenados de machine learning, buscando asociar contenido similar incluso si sus descriptores usan palabras diferentes. Según la UER (2021), estos vectores pueden tener más de cien dimensiones para mejorar la aproximación semántica entre artículos y hacer más eficaz el sistema de recomendación. Finalmente, las propias preferencias del usuario, en este caso el editor o el periodista, son consideradas por el sistema de recomendación.

Una gran ventaja de este sistema es, una vez más, su flexibilidad, adaptándose fácilmente a todas las plataformas / sistemas y facilitando una rápida implementación. Según la UER, la herramienta aún está en desarrollo y se espera que en el futuro pueda presentar tendencias informativas en tiempo real.

Tanto EUROVOX como PEACH utilizan la Inteligencia Artificial para cumplir con los objetivos del proyecto, sin embargo, siempre ha existido la preocupación por desarrollar herramientas que permitan a los editores y periodistas interferir en los procesos de traducción y recomendación, evitando así escenarios más pesimistas que auguran la sustitución de periodistas por ordenadores (Wölker & Powell, 2021). Se pueden introducir cambios en los diversos elementos de la noticia y en los descriptores de contenido mismos, siguiendo un conjunto de principios editoriales construidos entre los miembros de la UER que participan en el proyecto.

4.2 Ubicación del espacio en el sitio web

No es fácil encontrar el espacio “Um Olhar Europeu” en el sitio web de RTP porque no hay ninguna

referencia en la página de entrada. Debes ingresar al menú “Noticias” y desplazarte hasta encontrarlo después de varias noticias del día, un espacio promocional para los programas informativos de RTP, una solicitud de suscripción al boletín y otro espacio llamado Vídeos y Audios. La ubicación de este espacio en la página de noticias es común a todos los canales generalistas, pero en algunos de ellos aparece en una posición superior. En el caso de RTP, que tiene un sitio web organizado en dos columnas, sería recomendable tener al menos un destacado en la columna de la derecha, donde se promociona la App del canal, redes sociales, etc. Una situación que podrá cambiar tras este período experimental porque, según la RTP, el espacio “Um Olhar Europeu” será promocionado en otras partes del sitio web.

4.3 Orígenes y Temas

En cuanto a las fuentes de noticias, la más elegida por los editores de RTP fue SwissInfo (SWI), con un 20% de las noticias publicadas, la mitad de las cuales abordan temas relacionados con la pandemia. Le siguen la irlandesa RTE y Eurovision NewsWire, ambos con un 17%. Es importante explicar que Eurovision Social Newswire es una plataforma que comparte contenido extraído de las redes sociales por un proceso de filtrado y confirmación, y que habitualmente es compuesto de un vídeo y un pequeño texto descriptivo. Según RTP, estas preferencias pueden estar relacionadas con la mayor y más variada oferta de contenidos de estos operadores, ya que no existen indicaciones editoriales sobre fuentes prioritarias. “Depende mucho de la oferta. Es un proceso continuo, ya que todos intentan comprender lo que los demás buscan y por qué, con diferentes puntos de vista. De todos modos, en los últimos tiempos hemos tratado de diversificar lo más posible entre los distintos países” (Alexandre Brito, comunicación personal, 6 de octubre)

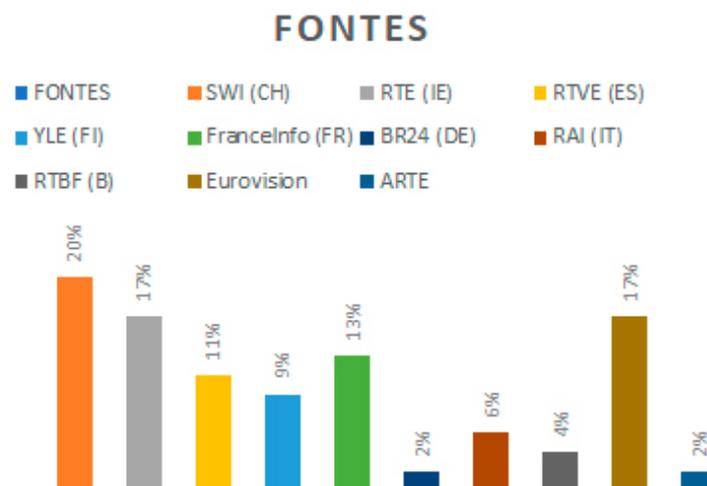


Figura 2: Empresa donde ha sido publicada la noticia original

Fuente: Elaboración del autor

La presencia de información de todos los operadores que integran el proyecto cumple con el objetivo de ofrecer una variedad de puntos de vista sobre determinados temas, lo que resulta en una forma de verificación (Kovach y Rosenstiel, 2014). La diversidad de operadores públicos identificados funciona como fuente para que el usuario pueda hacer su propio fact-checking en su lengua materna, contrastando información que consideró dudosa con fuentes prestigiosas como son los operadores públicos de radiotelevisión. Esta idea es reforzada con el análisis de las temáticas de las noticias publicadas en este período, que coinciden con los temas más discutidos en la sociedad europea en ese momento. El tema “Pandemia” destaca del resto con el 33% de las noticias publicadas, seguido de “Política” (26%), “Fait-divers” (18%) y “Clima” (11%). Un análisis más detallado revela que el 61% de las noticias incluidas en el tema “Pandemia” se refieren a temas de salud, el 17% está relacionado con la educación y el 11% vincula el tema a la “Política”, es decir, a las protestas sociales contra la vacunación.

En el caso de “Política”, el tema dominante es la situación en Afganistán (50%), el resto está relacionado con asuntos locales o geopolítica. El tema “Fait-divers” es muy variado, con mayor presencia de noticias relacionadas con fenómenos naturales no asociados al cambio climático (astronomía y erupciones volcánicas) o nuevas tecnologías (hackers, empresas tecnológicas multadas). “Educación” (6%) y “Cultura” (4%) son los temas menos discutidos, pero es importante señalar que la educación se

aborda también marginalmente en varias noticias sobre la pandemia.

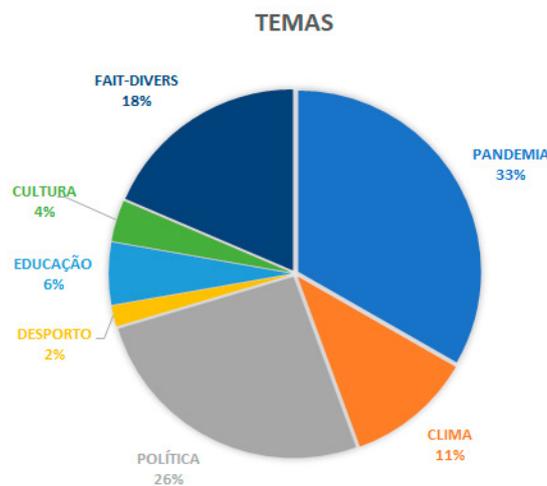


Figura 3: Temas de las noticias

Fuente: Elaboración propia

Este análisis nos permite verificar que los temas se presentan de manera proporcional (Kovach & Rosenstiel, 2014) con relación a los temas de actualidad (pandemia y Afganistán) y con puntos de vista de varios países, como se comentó anteriormente. En el caso de la pandemia, por ejemplo, hay noticias de siete de las nueve estaciones de televisión participantes (la décima es el canal Arte). Uno de los objetivos del proyecto - luchar contra agentes de desinformación que buscan explotar las fracturas entre comunidades en este tipo de temas – se concretiza por la oferta de puntos de vista originarios de las distintas geografías europeas. Cabe destacar que estos dos temas figuran entre los que más han sido afectados por la información manipulada. La necesidad de un entorno de confianza informativa en este tipo de temas es creado por la diversidad de fuentes que facilitan al usuario un contraste de informaciones, situación que combate la desinformación y, en simultáneo, crea la pretendida esfera pública basada en información con el sello de calidad de los operadores públicos de radiotelevisión.

Por otro lado, el análisis y contenido de las temáticas ayuda a responder a la tercera cuestión de investigación. Se confirma que el objetivo de la diversidad de temas está presente, como se ha visto anteriormente, pero en este período las corrientes de opinión minoritarias y/o los grupos más frágiles no han tenido una presencia significativa, con solamente 15% de las noticias dispersas en otros grandes temas. Dentro de las corrientes de opinión minoritarias se han identificado noticias relacionadas a un grupo muy específico dentro del pandemia, los antivacunas, que están presentes en 5% de las noticias, y los refugiados de Afganistán, igualmente con 5%. Hay igualmente una noticia sobre desigualdad de género y otra sobre un negocio de una adolescente de un grupo subrepresentado en las noticias, con lo cual se puede concluir que esta parte del objetivo del proyecto no se ha cumplido el periodo de análisis.

4.4 Original vs versión traducida

En general, podemos decir que el sistema de traducción funciona según los requisitos técnicos definidos por el proyecto, pero no siempre toda la información va desde el original a la versión traducida. Solamente en una situación el editor hizo un cambio significativo. Ocurrió el 23 de julio cuando se agregó el cuantificador existencial “algunos” al título “Cómo algunos escépticos de la vacunación ponen en peligro a sus hijos con lejía clorada” (BR24). Según RTP, “el medio que publica es responsable de lo que publica y, como tal, tiene derecho a realizar cambios fácticos muy leves que a veces se pierden en la traducción”. En situaciones como esta, los editores pueden intercambiar mensajes a través del propio sistema para aclarar dudas” (Alexandre Brito, comunicación personal, 23 de septiembre). Esta posibilidad de intervención humana es una forma de hibridación que perfecciona el sistema (Broussard et al, 2019).

Durante el período analizado, se registraron algunas traducciones que requirieron intervención humana (por ejemplo, traducción de “patients en réanimation” por “pacientes em terapia intensiva” cuando la expresión habitual en Portugal es “cuidados intensivos”) (FranceInfo, 15/09). Aunque las palabras

utilizadas no cambian el significado del texto, el uso de diferentes designaciones puede confundir a los lectores y por ello se recomienda la corrección editorial. Según RTP, la situación está identificada, pero en esta etapa del proyecto no ha sido posible destinar los recursos humanos necesarios al proyecto para este tipo de corrección. La situación puede cambiar con la contratación de periodistas dedicados exclusivamente al proyecto.

Como se menciona en la descripción de las herramientas, el software traduce por vectores, pero se encontró que algunos elementos de la noticia original desaparecieron en la versión RTP, a saber, fotos (20%), vídeo (17%), gráficos (6%) e, incluso, bloques de texto (12%). En el caso de las fotos, la situación ocurrió principalmente con RTE y RAI, donde las fotos fueron reemplazadas por una imagen con el nombre de la emisora que originalmente publicó la noticia. Según RTP, “las imágenes siempre son un problema porque las redacciones normalmente no están formateadas para tener sus propios fotografías y por lo tanto no tienen fotos originales. Cuando no hay derechos para la foto, se da esta indicación y una fallback image entra automáticamente” (Alexandre Brito, comunicación personal, 23 de septiembre). En este caso, no se trata de un error del sistema, sino de una dificultad relacionada con los derechos de autor, situación que se repite en el caso de los gráficos y algunos vídeos (tomados de Twitter, por ejemplo), ya que la televisión que produce el contenido tiene derechos para reproducción solo en sus plataformas.

En otros casos se realizó la traducción, pero se perdió el formato original. Lo que inicialmente era un cuadro de información que requería interacción para abrirse, aparece en la traducción como un texto continuo, como se puede observar en la imagen siguiente.



Figura 4: Original SWI (izquierda) y traducción RTP (derecha)
Fuente: SwissInfo y RTP

También con las citas resaltadas ocurre algo similar: en el original son un detalle de edición, pero en la traducción pasan a párrafos sueltos que a veces quiebran la lógica del texto.



Figura 5: Original FranceInfo (izquierda) y traducción RTP (derecha)
Fuentes: FranceInfo y RTP

Estas situaciones muestran que la traducción por bloques informativos funciona bien para el contenido en texto, pero se pierde en el diseño, con lo cual aún hay puntos para mejorar en el sistema. De todas formas, estos detalles no interfieren con los objetivos del proyecto y tan solo han sido enumerados para memoria futura.

5. Conclusiones

En el proyecto “A European Perspective”, la Inteligencia Artificial interviene en dos momentos - traducción automática de contenidos y recomendación de noticias - buscando alcanzar un conjunto de objetivos periodísticos descritos en los principios editoriales. Basado en estos objetivos han nacido las tres preguntas de investigación.

La primera cuestión de investigación buscaba saber si las noticias publicadas abordan temas que atraviesan varios países, sociedades o colectivos, buscando así crear una esfera pública europea. El análisis permitió constatar que las noticias publicadas en la web de RTP cumplen este objetivo, ya que los temas más discutidos (pandemia y conflicto en Afganistán) eran en ese momento de crucial importancia para el espacio europeo. En el primer caso, porque buscaba restablecer la libre circulación interrumpida por la pandemia (pasaporte verde y código de color), elemento fundamental para la recuperación de la economía en Europa. En el segundo caso, porque se prevé una nueva crisis de refugiados y Europa es un espacio geopolítico de referencia para su acogida. En este caso, al presentar un sistema que combina una recomendación de contenidos (PEACH) con la posibilidad de intervención del periodista, el proyecto facilita el enfoque híbrido donde los expertos humanos se apoyan en la inteligencia artificial (Broussard et al, 2019) para lograr el objetivo. Este modelo híbrido tiene dos ventajas: 1) es más eficaz porque, al análisis de sentimientos y de comportamientos fundados en los datos recogidos por la IA, este modelo añade factores que solamente un humano puede recoger en su vida social, algo importante cuando hablamos de opinión pública; 2) aparta las visiones más pesimistas de que los periodistas serán remplazados por ordenadores al comprobar que el trabajo conjunto aumenta la calidad del producto final.

La segunda cuestión se relaciona con el objetivo del proyecto luchar contra la desinformación al proveer perspectivas originarias de diferentes países sobre un determinado tema polémico. Se confirma que en los temas más conflictivos se buscó la variedad de fuentes informativas para ofrecer a los lectores la posibilidad de hacer su fact-checking, creándose un ambiente informativo de confianza. Una vez más tenemos un enfoque híbrido, porque hay dos niveles de actuación: en el caso de la IA, inicialmente actúa la traducción automática facilitando el acceso a varias fuentes informativas en la lengua materna del usuario, y, de seguida, funciona el sistema de recomendación, apoyado en criterios predefinidos; se añade el criterio periodístico del editor, que busca ofrecer variedad de fuentes y contexto, algo permitido porque la UER ha desarrollado una herramienta abierta. En este caso, el objetivo de combatir la desinformación puede cumplirse por la oferta de varias fuentes informativas que permiten al usuario un contraste de fuentes y la contextualización de la información que había recibido antes. Más que un mecanismo de combate directo a los procesos de desinformación, el proyecto simplifica los modelos para guiar usuarios en la validación de informaciones (Nguyen et al (2019), es decir, en hacer una autoverificación de hechos (user fact-checking).

Por fin, la tercera cuestión está conectada a la obligación de ofrecer contenidos diversificados e inclusivos, con el objetivo de dar visibilidad a las más variadas corrientes de opinión y los colectivos más frágiles. Los datos confirman que RTP apenas cumple parcialmente estos objetivos. El canal elige noticias provenientes de todos los socios y de todos los temas, pero en materia de temas dirigidos a las minorías no hay una presencia significativa. Los grupos representados están conectados con los dos temas más importantes en el momento, con lo cual se considera que este objetivo no se ha logrado en este periodo.

Concluimos que en los primeros pasos del espacio “Um Olhar Europeu” hay marcas muy positivas de un trabajo híbrido entre la inteligencia artificial y los periodistas en la labor de ofrecer información de calidad. Como bien nota Thurman et al (2019), los efectos de cada tecnología dependen de los humanos, en este caso de la forma como editores y periodistas decidan utilizarla. Al ofrecer perspectivas originarias de diferentes geografías sobre los temas más importantes a cada momento, el proyecto de UER posibilita al usuario hacer su propio fact-checking y, de esa forma, combatir la desinformación.

Además, esta variada oferta temática de noticias contribuye para la construcción de una esfera pública europea, un elemento fundamental en la estabilización de una identidad común en este espacio geográfico.

Por último, es importante señalar que estos objetivos se cumplen basados en un trabajo conjunto de UER, no exigiendo abultadas inversiones de cada medio en recursos humanos o tecnológicos. Esta situación confirma la idea de Manfredi Sanchez y Ufarte Ruiz, (2020) al defender que la IA aplicada al periodismo puede ser rentable.

6. Referencias

Arendt, H. (2005) [1958]. *La condición humana*. Paidós.

Balkin, J. M. (2018). *Free speech in the algorithmic society: Big data, private governance, and new school speech regulation*. *UC Davis Law Review*, (51), 1149-1210. https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf

Bardin, L. (1977). *Análise de conteúdo*. Edições 70.

Brandom, R. (2014). *How to teach a robot to write. Automated writing programs are coming to journalism. Is it good news?* *The Verge*. Sitio: <https://cutt.ly/ctiXGBn>

Broussard, M.; Diakopoulos, N.; Guzman, A. L.; Abebe, R.; Dupagne, M. & Chuan, C.H. (2019). *Artificial intelligence and journalism*. *Journalism & Mass Communication Quarterly*, 96 (3), 673-695. <https://doi.org/10.1177/1077699019859901>

Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism*, 3 (3), 416-431. <https://doi.org/10.1080/21670811.2014.976412>

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press

Essenfelder, R., Canavilhas, J., Maia, H. C. y Pinto, R. J. (2019). Automatización de textos periodísticos en la televisión brasileña: Estudio de caso del sistema AIDA (Globo-Brasil). *Doxa Comunicación*, (29), 255-274. <https://doi.org/10.31921/doxacom.n29a13>

Flew, T., Spurgeon, C., Daniel, A. y Swift, A. (2012). The promise of computational journalism. *Journalism Practice*, 6(2), 157-171. <https://www.doi.org/10.1080/17512786.2011.616655>

Flores Vivar, J. M. (2018). Algoritmos, aplicaciones y big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos. *Revista de Comunicación*, 17(2), 268-291. <http://www.doi.org/10.26441/RC17.2-2018-A12>

Graefe, A. (2016). *Guide to automated journalism*. Columbia Journalism School. Tow Center for Digital Journalism. <https://bit.ly/3FW1Uua>

Harcup, T. (2014). *Oxford dictionary of journalism*. Oxford University Press.

Helberger, N. (2019). On the Democratic Role of News Recommenders, *Digital Journalism*, 7 (8), 993-1012. <https://doi.org/10.1080/21670811.2019.1623700>

Kovach, B. & Rosenstiel, T. (2014). *The elements of journalism: What newspeople should know and the public should expect* (3rd Edition). Three Rivers Press

Kurzweil, R. (1990). *The age of intelligent machines*. MIT Press.

Larrondo-Ureta, A., Fernández, S.-P., y Morales-i-Gras, J. (2021). Desinformación, vacunas y Co-vid-19. Análisis de la infodemia y la conversación digital en Twitter. *Revista Latina De Comunicación Social*, 79, 1-18. <https://doi.org/10.4185/RLCS-2021-1504>

Lichterman, J. (2017). Want to bring automation to your newsroom? A new AP report details best practices. *NiemanLab*. <https://cutt.ly/ktiVwfj>

Lohr, S. (2011). *In case you wondered, a real human wrote this column*. The News York Times. <https://nyti.ms/3tG1RgK>

Lokot, T. & Diakopoulos, N. (2016). News bots: *Automating news and information dissemination on Twitter*. *Digital Journalism*, 4(6), 682-699. <https://www.doi.org/10.1080/21670811.2015.1081822>

Manfredi Sánchez, J.L. & Ufarte Ruiz, M.J. (2020). Inteligencia artificial y periodismo: una herramienta contra la desinformación. *Revista CIDOB d'Afers Internacionals*, 124, 49-72. <http://doi.org/10.24241/rci.2020.124.1.49>

Nguyen, T. T., Weidlich, M., Yin, H., Zheng, B., Nguyen, Q. V. H., & Stantic, B. (2019). User guidance for efficient fact checking. *Proceedings of the VLDB Endowment*, 12 (8), 850–863. <https://doi.org/10.14778/3324301.3324303>

Noir, S. (2021). Eurovision news monitoring tool: What's under the hood? *Tech-i*, (March 21). 11

Papadimitriou, A. (2016). *The future of communication: Artificial intelligence and social networks* [tesis inédita de maestría, Malmö University]. Repositorio Institucional de Malmö University. <http://bit.ly/379xa7O>

Russell, S. & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th Edition). Pearson Education

Silverman, C. (2013). *5 ways robots can improve accuracy, journalism quality*. Poynter Institute. <https://bit.ly/2V4MDFk>

Steiner, T. (2014). *Telling breaking news stories from Wikipedia with social multimedia: a case study of the 2014 winter Olympics*. <https://arxiv.org/abs/1403.4289>

Thurman, N., Lewis, S. & Kunert, J. (2019) Algorithms, automation, and news. *Digital Journalism*, 7:8, pp.980-992. <http://10.1080/21670811.2019.1685395>

Túñez-López, M., Toural-Bran, C., & Valdiviezo-Abad, C. (2019). Automatización, bots y algoritmos en la redacción de noticias. Impacto y calidad del periodismo artificial. *Revista Latina De Comunicación Social*, 74, 1411-1433. <https://doi.org/10.4185/RLCS-2019-1391>

Túñez-López, J. M., Fieiras Ceide, C. & Vaz-Álvarez, M. (2021). Impact of artificial intelligence on journalism: Transformations in the company, products, contents and professional profile. *Communication & Society*, 34(1), 177-193. <https://www.doi.org/10.15581/003.34.1.177-193>

Ufarte Ruiz, M. J. & Manfredi Sánchez, J. L. (2019). Algoritmos y bots aplicados al periodismo. El caso de Narrativa Inteligencia Artificial: estructura, producción y calidad informativa. *Doxa Comunicación*, 29, 213-233. <https://doi.org/10.31921/doxacom.n29a11>

Ufarte Ruiz, M. J., Túñez López, J. M. & Vaz Álvarez, M. (2019). La aplicación del periodismo artificial en el ámbito internacional: retos y desafíos. In J. L. Manfredi Sánchez, M. J. Ufarte Ruiz & J. M. Herranz de la Casa (Eds), *Periodismo y Ciberseguridad* (pp.67-88). Comunicación Social

Ufarte Ruiz, M.J.; Calvo Rubio, L.M., & Murcia Verdú, F.J. (2021). Los desafíos éticos del periodismo en la era de la inteligencia artificial. *Estudios sobre el Mensaje Periodístico*, 27(2), 673-684. <https://dx.doi.org/10.5209/esmp.69708>

UER – Union Europea de Radiodifusión (2021). *Providing a european perspective. Public service media allied to offer na innovative news sharing model across the continente.* <https://bit.ly/3l0imRd>

UER – Union Europea de Radiodifusión (s.f). EUROVOX: *An open toolbox to reduce the cost and complexity of transcription and translation.* <https://tech.ebu.ch/eurovox>

Vaz-Álvarez, M., Fieiras-Ceide, C., & Túnnez-López, M. (2021). Experiencias de co-creación en Me-dios de Servicio Público Europeos: Visión y Tendencias. *AdComunica*, 21, 71-84. <https://doi.org/10.6035/2174-0992.2021.21.5>

7. AUTOR

João Canavilhas
Universidade da Beira Interior

Profesor en la Universidade da Beira Interior (Covilhã – Portugal), es licenciado en Comunicação Social por esta universidad y doctor por la Universidad de Salamanca (Comunicación, Cultura y Educación). Es investigador y elemento de la dirección de la unidad de investigación Labcom – Communi-cation, and Arts evaluada por la Fundação para a Ciência e tecnologia com Muito Bom). Es consultor en proyectos de investigación en Portugal, España y Brasil.y autor/coautor de 10 libros, 45 capítulos de libro y 52 artículos en revistas científicas. Su trabajo de investigación se enfoca en el campo de la Comunicación y de las Nuevas Tecnologías, particularmente en la influencia de Internet y de los dispositivos móviles en el periodismo y en la política.

jc@ubi.pt

Índice H: 27 (Google Scholar); 5 (Scopus)

Orcid ID: <https://orcid.org/0000-0002-2394-5264>

Google Scholar: <https://scholar.google.com.br/citations?hl=pt-PT&user=ynI0OzkAAAAJ>

ResearchGate: <https://www.researchgate.net/profile/Joao-Canavilhas>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=37060317500>

Academia.edu: <https://ubi.academia.edu/JoãoCanavilhas>

Dialnet: <https://dialnet.unirioja.es/servlet/autor?codigo=341943>