

Dirección

Clara Martínez
Cantón

Gimena del Río
Riande

Francisco Barrón

Editor Asociado

Rubén Iñiguez
Pérez

I+D+i COVIDTECA. Hemeroteca de Humanidades sobre el impacto socio-económico y la incidencia cultural de la pandemia de coronavirus

<https://covidteca.org/>

Reseña realizada por:
Natàlia RIBÉS
natalia-ribes@pm.me

COVIDTECA es un proyecto llevado a cabo por la Unidad de Excelencia IberLab de la Universidad de Granada, financiado por los Fondos FEDER y la Junta de Andalucía y dirigido por Ana Gallego Cuiñas. Se trata de una hemeroteca de humanidades donde se recopilan materiales producidos en español relativos a la pandemia del COVID-19, en el ámbito de Iberoamérica y Estados Unidos. El objetivo del proyecto es realizar una cartografía de los relatos y reflexiones acerca de la pandemia desde distintos ámbitos del pensamiento, como la filosofía, la política, el arte o la literatura.

El proyecto se realizó en tres fases: en la primera se recopilaron los materiales aparecidos en plataformas, blogs, revistas, etc.; en la segunda se extrajeron los textos de los recursos recopilados para procesarlos; y, finalmente, en la tercera fase se analizaron los recursos combinando técnicas de aprendizaje automático y procesamiento de lenguaje natural, como *topic modeling* y *topic classification*, para categorizar estos recursos en base a una serie de temáticas.

A partir de este enfoque cuantitativo, un equipo de expertos en diversas disciplinas humanísticas, principalmente procedentes de los campos de la filosofía, la lingüística y la literatura, realizaron diversos análisis interpretativos de los datos, que fueron recogidos en el libro editado por Ana Gallego Cuiñas y José Antonio Pérez Tapias *Pensamiento, pandemia y Big Data. El impacto sociocultural del coronavirus en el espacio iberoamericano*¹.

RHD 9 (2024)

ISSN

2531-1786

¹ Gallego Cuiñas, A., Pérez Tapias, J.A. (Eds.). (2022). *Pensamiento, pandemia y Big Data. El impacto sociocultural del coronavirus en el espacio iberoamericano*. De Gruyter. <https://doi.org/10.1515/9783110693928>.



Los materiales se recogieron entre marzo de 2020 y abril de 2021 y abarcan, según la página web de COVIDTECA, 7107 recursos analizados (artículos, entrevistas, material de audio, vídeo, imágenes, libros, etc.) de dieciséis países de habla hispana (incluyendo EE. UU.) y clasificados en doce temáticas y siete tipos de materiales.

En la web destaca la hemeroteca, que ofrece un sistema de navegación facetada en el que se puede añadir un término de búsqueda y luego ir refinando a través de diversos filtros, tales como fecha, país, material y temática.

La categoría país muestra un desplegable con los dieciséis países analizados. Por su parte, la categoría materiales muestra siete tipos de materiales: artículo, audio, entrevista, imagen, miscelánea, revista y vídeo. La última categoría muestra las doce temáticas clasificadas a partir del material recopilado, utilizando el modelado de datos. Estas doce temáticas son arte, bioética, biopolítica, editoriales, estudios sensoriales y afectos, género, historia, poscolonialismo, redes sociales, retórica y ficción, teatro y espectáculo y universidad.

Al seleccionar un resultado obtenemos la información básica del recurso (referencia, título, autor, país, material, medio, fecha y URL), además de las distintas temáticas sobre las que trata, un gráfico con los porcentajes en los que aparecen dichas temáticas y los resultados de esa entrada, en los que se muestran las palabras asociadas a cada uno de los temas que aparecen en el recurso.

Otro elemento fundamental de la COVIDTECA es la sección de resultados en la que, al igual que en la hemeroteca, se puede filtrar por país, material y temática, con el mismo sistema de facetas. Al realizar una búsqueda combinando estos tres tipos de filtros, se obtiene un gráfico de torta. Si queremos combinar las distintas opciones de búsqueda, podremos observar las relaciones entre las distintas variables (filtros) que hemos añadido. Además, un botón en la parte inferior nos ofrece la posibilidad de consultar los recursos asociados a nuestra búsqueda.

Es importante señalar que el tema con mayor presencia en los materiales recogidos es con mucha distancia la biopolítica (70, 57 %) y que el pico de mayor presencia de este tema en los materiales recogidos fue de marzo a mayo de 2020. Estos dos datos tienen una gran relevancia en la posterior interpretación de los materiales.

Por otro lado, de los dieciséis países de los que se recogían datos, España aporta un 22,75 %, seguida de México, que aporta un 18,45 %, y Argentina, con un 11,28 %. El tipo de material más frecuente son los artículos, seguidos de revistas, entrevistas y miscelánea.

En el apartado de divulgación aparece una serie de publicaciones realizadas a partir de los resultados de la COVIDTECA, en su mayoría artículos aparecidos en medios de tipo periodístico, así como de un congreso realizado en diciembre de 2021.

Además, el libro, publicado a partir de los datos recopilados y analizados, nos presenta diversas interpretaciones a través de los datos obtenidos y está organizado en dos grandes bloques. La primera parte, "Pandemia, Filosofía y Política", se centra en el tratamiento de la

pandemia en cuanto a acontecimiento que irrumpió en la vida de las personas y en la organización de la sociedad, y en cómo la biopolítica arrebató el protagonismo en el discurso público. La segunda parte, “Cultura y Sociedad”, se enfoca en cómo la pandemia fue narrada en la literatura y en las artes y en cuáles fueron sus consecuencias en distintos ámbitos sociales como la educación, el género o la “nueva normalidad”.

El proyecto se presentó en la primera conferencia de *Humanidades y Big Data* en Iberoamérica que tuvo lugar en la Universidad de Granada en diciembre de 2021. Además, en esa misma universidad se realizó el Congreso Internacional COVIDTECA en marzo de 2022.

Desde el punto de vista de las Humanidades Digitales, COVIDTECA resulta un ejemplo muy interesante para analizar, ya que se trata de un corpus de materiales creados y difundidos a través de internet en una época de producción masiva de contenidos. Además, esta producción masiva se da en un contexto de polarización política y de ruido informativo. Por si esto no fuera suficiente, el tema que nos ocupa, la pandemia del COVID-19, se trata de la primera “enfermedad posverdadera”, según las palabras de Francisco J. Alcalá en el propio libro, lo que añade una capa extra de dificultad interpretativa a la sobreinformación que caracteriza a nuestra época.

Dadas estas circunstancias, la lectura distante, que Moretti aplicaba a la crítica literaria, se nos presenta en este caso como una oportunidad para dar un paso atrás, respirar y tomar perspectiva a la hora de enfrentarnos a esta enorme cantidad de datos sobre un fenómeno polarizante en el que estamos todos inmersos. Las posibilidades técnicas de los ordenadores nos ofrecen una mirada más global que nos facilita realizar una interpretación subjetiva pero rigurosa de estos acontecimientos desde las humanidades, sin dejarnos llevar por la vorágine de la actualidad. Parece útil a priori aplicar el método científico a la hora de realizar este macroanálisis. ¿Pero esto es realmente así?

A partir del proyecto COVIDTECA me gustaría plantear unas preguntas generales de carácter epistemológico. La primera pregunta que me surge es si toda esa cantidad de datos recopilados, procesados y presentados en gráficos, que conforman el proyecto, son realmente relevantes para una mejor comprensión y un análisis más riguroso, desde las humanidades, del fenómeno de la COVID, o si, por el contrario, esa “obsesión” por los datos es una excusa para contar relatos que ya teníamos elaborados previamente.

El medio digital obliga a los humanistas a situarse en su tiempo con una mirada adaptada y un nuevo repertorio de competencias técnicas y analíticas. Una vez superada la primera y la segunda ola de las Humanidades Digitales, en las que el foco estaba puesto en generar, primero las infraestructuras, más tarde las herramientas y las técnicas necesarias para desarrollar la disciplina, en esta tercera ola en la que nos encontramos, resulta imprescindible reflexionar sobre la ideología que llevan consigo los materiales digitales que analizamos. En este punto me gustaría lanzar otras preguntas. La aspiración de aplicar análisis cuantitativos a disciplinas humanísticas, ¿no tiene ya de por sí una fuerte carga ideológica? Además, si esta duda se nos plantea cuando hablamos

de materiales digitalizados, procedentes de disciplinas históricas, ¿no será más cierto aun cuando se trata de un corpus de textos “nativos digitales” relativos a un tema de actualidad que nos afecta en nuestro día a día y donde todos tenemos “nuestra” opinión?

Para responder a estas preguntas, me gustaría remitirme al concepto *Cultural Analytics*, acuñado por Lev Manovich. Manovich siente la necesidad de establecer un paradigma de investigación diferenciado para la detección y análisis de patrones en la cultura digital contemporánea, que tiene unas características distintas a las disciplinas clásicas por su “escala, rapidez, diversidad y conectividad”². El autor plantea que los *big data* pueden ayudarnos a cuestionar nuestros estereotipos y concepciones previas sobre el objeto de estudio. Además, propone un marco teórico apropiado para estudiar la cultura digital con todas sus especificidades, resaltando que los métodos de análisis de datos y las visualizaciones nos aportan un lenguaje que no teníamos para describir fenómenos (culturales) contemporáneos, pero también reflexiona sobre los límites de la cuantificación, destacando la importancia del análisis cualitativo y la interpretación del humano, que es capaz de reconocer los detalles significativos y poner la información en su contexto.

Desde este punto de vista, COVIDTECA resulta una propuesta muy sugerente, ya que se centra en un fenómeno en el cual la producción de materiales se desarrolló a una escala gigantesca en un período corto de tiempo, con una enorme rapidez, diversidad y conectividad. Por lo que el análisis de los grandes datos realmente nos puede servir para combatir las ideas preconcebidas que tenemos sobre este evento y a poner palabras sobre acontecimientos que nos pueden resultar desbordantes. Además, los materiales recopilados, que en definitiva son productos culturales contemporáneos, han sido analizados por un grupo de investigadores expertos, que aportan la labor interpretativa a partir de los marcos teóricos de sus disciplinas.

Por otro lado, cabe plantear también ciertas reflexiones sobre la metodología llevada a cabo y, sobre todo, acerca de la transparencia con la cual se comunica esa metodología. Una de las primeras preguntas que surgen respecto al enfoque metodológico, dado que no queda claro en la documentación disponible, es el criterio utilizado para recopilar los materiales. Por lo que se da entender, la selección de los más de 7000 materiales se ha realizado de forma manual por un equipo de cinco personas. ¿Qué criterios han aplicado entonces para seleccionar los materiales y qué métodos han empleado para evitar sesgos a la hora de incluir o descartar dichos materiales? Por ejemplo, la selección del material tiene una clara sobrerrepresentación de recursos procedentes de España (México es un país mucho mayor en número de habitantes y tiene una menor representación). La directora del proyecto lo justifica por la disponibilidad de las fuentes, pero cabe argumentar que en internet hay fuentes disponibles de todos los países, por lo que sería interesante que se expusiera la metodología para la selección de la muestra.

También resultaría muy interesante que se explicara con mayor detalle el proceso de *topic modeling* y *topic classification* que se ha llevado a cabo para analizar los datos. ¿Qué *software* se

² Manovich, L. (2020). *Cultural Analytics* (p. 9). The MIT Press.

ha utilizado? ¿Se ha llevado a cabo un tipo de aprendizaje automático no supervisado o había hipótesis previas que han sido validadas? ¿Bajo qué criterios se han realizado las categorizaciones y cómo se definen? A nivel técnico, la única información disponible es que el *software* ha sido desarrollado por Everywhere Technologies, una empresa que se ha formado en la Universidad de Granada y que ha colaborado con ella en diversos proyectos.

Con esto no se pretende poner en cuestionamiento la validez de los datos, ni su utilidad para realizar los análisis. Los análisis llevados a cabo por los investigadores aportan resultados realmente interesantes. Por ejemplo, el hecho de que la biopolítica sea el tema con mayor presencia en los materiales recopilados, con más de un 70 %, es un dato rotundo, que da pie a reflexiones muy relevantes a lo largo del libro. Los expertos que participan en el proyecto ponen estos datos en contexto sacando a la luz discursos que son de gran importancia para interpretar lo que ha sido (o está siendo) la pandemia, como la vulnerabilidad, las narrativas acerca de la realidad o la diferencia entre biopolítica y necropolítica. Tan solo se plantea que resultaría interesante disponer de más información sobre cómo han sido obtenidos y procesados porque es importante tematizar los posibles sesgos que hay en todo proceso de selección, agregación y categorización.

En definitiva, la COVIDTECA se presenta como una hemeroteca virtual de humanidades en la que se recogen distintos tipos de materiales que tratan sobre la pandemia del COVID-19 y además se puede acceder a visualizaciones de los análisis de esos materiales.

Este proyecto tiene un gran valor por varios motivos. El primero es técnico, ya que supone un ejemplo de cómo se puede usar grandes cantidades de datos para sistematizar, jerarquizar y visualizar material cuantitativo y darle una interpretación humanística a un fenómeno contemporáneo, con todo lo que ello implica. A partir de un acontecimiento de gran relevancia histórica, que afecta a la totalidad de la población mundial, y sobre el cual, de un día para otro, se empiezan a producir enormes cantidades de material, realizar una lectura distante nos permite adoptar una perspectiva que se aleje del ruido. Este macroanálisis sirve de base a los investigadores para realizar interpretaciones y reflexiones con un enfoque propio. Sin embargo, se echa en falta una mayor transparencia a la hora de comunicar el procedimiento metodológico, particularmente en la selección del material analizado y en los detalles sobre cómo se ha realizado el modelado, categorización y jerarquización de los temas.

El segundo motivo por el que este proyecto resulta relevante es porque pone en valor la importancia del pensamiento filosófico, la ética, la crítica, la literatura y las artes visuales a la hora de dar sentido a la realidad. En una época en la que las humanidades se ven reducidas en los temarios de los colegios, y el discurso público está acaparado por los avances de la tecnología, la polarización política y los devenires del mercado, resulta más necesario que nunca que el pensamiento humanista tome la palabra y nos proporcione ciertas líneas interpretativas desde las cuales dar sentido al mundo que nos rodea.

Por último, el tercer motivo interesante que aporta este proyecto es el valor testimonial. Quizá en el momento en el que escribo estas líneas ya nadie tiene ganas de pensar en la pandemia. Probablemente, después de tres años duros y disruptivos, poca gente tenga ganas de curiosear entre los filtros de la hemeroteca del COVID y ponerse a revisar relatos sobre el confinamiento o artículos sobre la distancia social. No obstante, el material recogido y los análisis realizados quedarán para quienes, en un futuro, deseen revisar el fenómeno que marcó el comienzo de nuestra década. Los historiadores del futuro, los filósofos, los críticos o los humanistas digitales tendrán a su disposición todo un corpus catalogado con el que podrán examinar los relatos realizados en nuestra época y darles sentido desde una nueva perspectiva.