ISSN: 2683-3247

HUMANITAS

REVISTA DE TEORÍA, CRÍTICA Y ESTUDIOS LITERARIOS

Vol. 4 Núm. 8 Enero-Junio 2025





INIVERSIDAD AUTÓNOMA DE

Humanitas

Revista de Teoría, Crítica y Estudios Literarios

Para una lectura distante de *Sábado*, suplemento cultural

For a distant reading of *Sábado*, cultural supplement

Jonathan Gutiérrez Hibler Universidad Autónoma de Nuevo León Monterrey, México orcid.org/0000-0002-7686-7896

Fecha entrega: 19-10-2024 Fecha aceptación: 21-01-2025

Editor: Víctor Barrera Enderle. Universidad Autónoma de Nuevo León, Centro de Estudios Humanísticos, Monterrey, Nuevo León, México.

Copyright: © 2025, Gutiérrez Hibler, Jonathan. This is an openaccess article distributed under the terms of Creative Commons Attribution License [CC BY 4.0], which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



DOI: https://doi.org/10.29105/revistahumanitas4.8-117

Email: jonathan.gutierrezhr@uanl.edu.mx

Para una lectura distante de *Sábado*, suplemento cultural

For a distant reading of *Sábado*, cultural supplementl

Jonathan Gutiérrez Hibler Universidad Autónoma de Nuevo León Monterrey, México

jonathan.gutierrezhr@uanl.edu.mx

Resumen. Uno de los grandes problemas en la lectura del periodismo cultural es la resistencia contra el tiempo. Esta resistencia va desde lo delicado del material hemerográfico, en su preservación física, hasta los contenidos que pierden vigencia frente a las necesidades de un presente con otra visión de los valores culturales. Dentro de este espectro entra una segunda cuestión para el caso de los suplementos longevos: el registro cuantitativo, es decir, lo que se publicó, la creación de índices o bases de datos para crear mapas que visualicen las constantes de una dirección o las directrices según la coyuntura donde se lleva a cabo la labor de dicho medio. Este artículo intenta realizar dicho análisis en el suplemento cultural Sábado.

Palabras clave: lectura distante, análisis de datos, periodismo cultural.

Abstract. One of the major problems in reading cultural journalism is its resistance to time. This resistance ranges from the delicate nature of the newspaper material, in its physical preservation, to the content that

loses its validity in the face of the needs of a present with a different vision of cultural values. Within this spectrum lies a second issue for the case of long-lived supplements: the quantitative record, that is, what was published, the creation of indexes or databases to create maps that visualize the constants of a direction or the guidelines according to the situation in which the work of said medium is carried out. This article attempts to carry out this analysis in the cultural supplement Sábado.

Keywords: distant reading, data analysis, cultural journalism.

Sábado inicia su aparición en el año de 1977 y transcurre con diferentes épocas editoriales donde resaltan los nombres de Fernando Benítez (hasta su salida a La Jornada Semanal) y Huberto Batis. Desde una lectura cercana pueden identificarse diferentes líneas de investigación para este suplemento cultural; sin embargo, la aplicación de una lectura distante (Moretti, 2015) puede identificar de manera concreta, a partir de las materialidades del presente medio estudiado, diferentes líneas de investigación. Esto es posible siempre y cuando exista una conceptualización de datos que permita a las herramientas digitales realizar dichas conexiones a partir de coincidencias: traducciones, ediciones, libros, poesía, entrevistas, reseñas, cuentos, fragmentos de novelas u obras de teatro, fotografía, dibujo, correo de sus lectores, recomendaciones de lectura, traducciones, entre otras tantas que pueden aparecer en este ejercicio.

El propósito de este artículo consiste en relatar la experiencia desde la Facultad de Filosofía y Letras, en conjunto con la Capilla Alfonsina Biblioteca Universitaria, en el proceso de digitalización de una colección privada con vías de análisis de datos. Para efectos de este trabajo, primero se definirá qué son las humanidades digitales y el concepto de lectura distante proporcionado por Moretti. En el apartado siguiente, se relatará el proceso de digitalización realizado dentro de una materia optativa llamada "Proyectos de investigación en entornos digitales", parte del currículum de la licenciatura en letras hispánicas, y el estado en el que se encuentra actualmente. Finalmente, se presentará un apartado, a manera de muestra con un número del suplemento, de las herramientas que se pueden utilizar y su aplicación para plantear una lectura distante en el resto del suplemento; esto tiene como propósito ampliar este proyecto o que su experiencia sea adaptada para otros casos en la lectura y

visualización del periodismo cultural y su preservación en diferentes repositorios.

Humanidades digitales y lectura distante

Para Moretti (2015) la lectura distante es una condición del conocimiento. El autor realiza una oposición entre close y distant, términos traducidos originalmente como cercana y distante, respectivamente, aunque en las notas de Lilia Mosconi se menciona que en otras traducciones se usa "atenta" para el primer caso. Moretti parte de esta necesidad metodológica y teórica al plantear lo infinitamente rica que es la realidad en contraste con lo limitado que suelen ser los conceptos al momento de abarcarla. Es decir, siempre hay una pérdida cuando se busca abarcar la totalidad.

Sin embargo, en lugar de entender lo anterior como un impedimento, esto le da la oportunidad de plantear una flexibilidad: su pobreza permite que sean más fácil de manejar y, en consecuencia, con menos se puede hacer más. El interés del autor respecto a esto recae en polémicas en torno a la mundialización de la literatura y el debate sobre lo moderno en algunas de ellas, por ejemplo, las lecturas de Frederic Jameson en torno a la literatura japonesa. Para Franco Moretti las ideas planteadas por la crítica parecen interesantes, pero no dejan de ser ideas. Por esta razón sugiere una lectura a gran escala para encontrar constelaciones que las reafirmen.

Su preocupación parte de que un crítico siempre se equivoca. ¿Cómo sabemos que nos equivocamos aun y cuando seguimos a otro crítico? La respuesta es sencilla: no hay corroboraciones, pero sobre todo el corpus del cual partimos en una investigación siempre es limitado desde una conceptualización que tal vez oculte el resto del universo estudiado, sin mencionar que una ley o principio planteado

en literatura puede dejar de encontrarse con el tiempo, lo cual trae problemas en la aproximación al objeto de estudio.

Por esta razón, a partir de la conceptualización de datos y la determinación de unidades de análisis, Moretti ve pertinente una forma de visualización para aclarar el comportamiento y creación de patrones a partir de las materialidades del texto. La visualización es una forma importante para la jerarquización de datos porque ésta permite corroborar los juicios del crítico por medio de materialidades, sobre todo en un mundo donde cada vez es más grande la cantidad que se maneja en la investigación de las humanidades digitales y área similares. Una de las elecciones de Moretti para lograr este fin es el árbol:

Al principio lo usé como una mera visualización abreviada, pero después de un tiempo me di cuenta de que el árbol era más que eso: funcionaba como una metáfora cognitiva que permitía ver—casi literalmente—la historia literaria de manera nueva. (2015: 94)

Moretti se concentra en qué es lo que eleva a ese árbol, lo que le da sus dimensiones. En realidad, no son los textos los que le dan dicha altura, sino los indicios. Los textos se distribuyen en las diferentes ramas, pero son los indicios los que, por presencia/ausencia, necesidad o visibilidad, le dan tamaño a esta metáfora visual por medio de recursos que podemos leer como unidades pequeñas dentro de diferentes textos. El autor de *Lectura distante* toma como ejemplo el caso del género policial para ejemplificar cómo recursos y géneros pueden ser dos unidades formales. Es en ese punto donde Moretti plantea un supuesto radical: los textos no son objetos de conocimiento, solamente son objetos reales. Esto, por supuesto, lo plantea desde una lectura de la crítica literaria y no de la crítica editorial, aunque pueden encontrarse puentes en la conceptualización de unidades.

Dentro de esta postura, con respecto a la visualización de datos en literatura, nos encontramos con la relación que adquiere frente al Big Data. María Peña Pimentel y Fernando Sancho Caparrini (2018) señalan la importancia de las metodologías en estos procesos. En este aspecto es necesario ser conscientes de que no debe existir un desapego referente a ellas, sobre todo respecto en el uso de mediciones para los datos obtenidos y filtrados desde y en el mundo digital. Datos, metodologías y teorías no deben ser ignorados al momento de acercarse a un objeto de estudio. Por eso es necesario establecer un vínculo entre los datos y los procesos mediante los cuales se obtienen a través de las computadoras y quienes trabajan con ellas. Entre las diferentes aportaciones al campo de las humanidades digitales, se han creado una variedad de algoritmos e inteligencias artificiales con el fin de resolver problemas en la manera de abordar grandes cantidades de datos; además de lo anterior, se han creado nuevas formas de almacenarlos, compartirlos, transformarlos en información con el fin de convertirlos en conocimiento, todo esto de forma rápida y masiva.

En la figura anterior, dentro del texto de Pimentel y Capirri, podemos observar la manera en la que se conforma la estructuración de la información. Tomemos como ejemplo el caso de *Sábado*. Como objeto real parte de un mundo donde una colección privada se encuentra en físico frente a las diferentes limitantes y riesgos: no es una colección a la que tenga acceso todo mundo, en todo el tiempo, incluso su digitalización fuera de fines académicos presenta problemas de derechos de autor; por otro lado, el material del que está hecho es sumamente delicado, en cuanto a lo hemerográfico, sobre todo en las condiciones reales en las que se encuentre la colección, ya sea privada o en un repositorio público con sus respectivos protocolos. En lo que se refiere al proceso de captura debe tomarse en cuenta, para el caso

de la digitalización, que no basta en la actualidad con la sola imagen, sino que es necesario añadir el reconocimiento óptico de caracteres (OCR). En un escenario ideal, la captura OCR es clara; sin embargo, dependiendo la calidad del programa o de la imagen (dependiendo la preservación), será necesario otro paso en la reestructuración de los datos: la limpieza del texto. No obstante, esto es sólo parte de un proceso. En un segundo, aunque no el único posible, los datos pueden estructurarse aparte en una base de datos para la creación de metadatos, dando la posibilidad del uso de otras herramientas para su reestructuración en busca de relaciones que en una lectura atenta serían más lentas de encontrar.

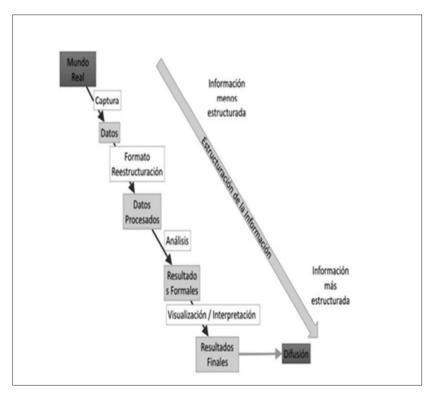


Fig. 1 Tomada de Pimentel y Caparrini (2018) en su texto "Big data y humanidades digitales"

En este punto vale la pena resaltar las observaciones de Lance Strate (2021) en "Humanidades digitales en contexto". Para este autor la digitalización transforma los mensajes y sus archivos en tipo de unidades mucho más discretas, capaces de compartirse de manera separada. Por ejemplo, una copia de un artículo científico puede ser descargada en diferentes dispositivos, incluso tenerlo abierto al mismo tiempo en dos o más. Se trata de un paso a la descentralización; por esta razón es importante estructurar la manera en que se ordenan estos datos para convertirlos en información y ser compartidos con la comunidad, tanto académica como periodística cultural, además del público interesado en el tema en cuestión. La programación, en el caso de estos medios, será importante para plantear metodologías que permitan observar lo que una lectura cercana no puede alcanzar, pero esto no descarta su aportación una vez que se procesen rumbo al análisis. Strate continúa:

La escritura, por otro lado, separa al cognoscente de lo conocido (Havelock, 1963); transformando un proceso, un verbo, un saber en conocimiento. Se vuelve algo que encuentras en un libro, que buscas cuando necesitas saber sobre algo. Mientras que la sabiduría implica entender relaciones; es decir, opera a nivel de relación o de medio; el conocimiento se traduce en contenido y trabaja en este nivel.

El interés de Strate radica en cómo podemos encontrar la información porque esta aparece de forma esporádica e incluso desestructurada o fuera de contexto. Esto quiere decir que la misma estructuración, su jerarquía que aumenta en complejidad, puede traer consecuencias. La máquina, por ejemplo, puede observar lo incómodo que una lectura cercana no puede explorar, pero esto no se logra de manera automática, debe pasar por una conceptualización

de los datos a través de una teoría o métodos, según sea el caso y los objetivos de investigación.

Al retomar las posturas de Pimentel y Caparrini (2018), esto toma otra dimensión en la digitalización de suplementos culturales, por ejemplo, el caso de *Sábado*. Los textos son parte del paradigma de la información no estructurada, esto desde el punto de vista de la Big Data y no el de la teoría literaria. Acerca de los textos, podemos realizar análisis dependiendo el método y el objetivo de investigación: si acaso es una cuestión de estructura física, como la llaman Pimentel y Caparrini, nos concentraremos en sus caracteres, palabras, oraciones, grupos de oraciones, párrafos, grupos de párrafos, capítulos, según sea el caso en específico. De manera similar, aunque con otro enfoque, si partimos de un interés sintáctico las condiciones de análisis serán otras las que recaen sobre el texto. Lo mismo sucederá con un contenido semántico con base en los elementos relacionales que haya entre sí en él.

Para el caso de *Sábado*, todos sus artículos, reseñas, secciones, aunque tengan una estructura editorial, son en realidad datos e información no estructuradas desde un horizonte amplio, es decir, el de la totalidad de sus números publicados. Lo que obtendríamos, al digitalizar, solamente serían datos sin estructura aunque estén en un formato para visualizarse, que es uno de los grandes retos actuales de los repositorios. Hay un cambio de soporte y de interfaz, pero no una manera distinta en lo que se refiere a lectura cercana. De esta manera es de vital importancia definir hacia dónde quiere ir un proyecto de este tipo antes de realizar posibles líneas de investigación. La implementación de herramientas digitales, ya sean programas de visualización o la aplicación de lenguajes de programación como Python o R, dependerán de las preguntas que

sean planteadas por seres humanos para su posible estructuración con base en las necesidades de su comunidad:

El proceso de análisis y reestructuración de sus partes es lo que permite que la información sea procesada de forma algorítmica por medio de un ordenador. Para que el proceso ejecutado por la máquina proporcione algún resultado interesante hay un requisito fundamental: que la 'estructura' que se busca en la información no estructurada sea suficientemente rica para obtener algún resultado. (Pimentel y Caparrini, 2018)

Los avances en la tecnología han sido muchos hasta el punto, como señalan los autores de "Big data y humanidades digitales", en que los análisis sintácticos ya están automatizados, y solamente requirieron una etapa posterior de supervisión. Para el momento de la publicación de su texto, los autores señalan que todavía está lejos de automatizarse el análisis semántico. En estos casos, la intervención humana sigue siendo importante. Esto puede ocurrir en otros puntos de un proceso de estructuración de datos. Tómese el caso del OCR donde algunas ediciones digitales siguen teniendo problemas en el traslado de un formato a otro, como el caso de una de las citas de este artículo (la de Strate) requirió limpieza por parte de quien ahora escribe, así como el formato indicado, según sea la publicación. La selección de algoritmos, para el caso del uso de un lenguaje de programación, como sucede con Python, requerirá un punto claro de partida en el interés de crear diferentes mapas interactivos para un suplemento cultural con una gran cantidad de datos sin estructurar.

El proceso de digitalización de la colección K

El objetivo de este proyecto consiste en la digitalización para uso interno y de investigación del suplemento cultural *Sábado*. Así mismo,

como segundo objetivo, éste busca crear una base de datos de acceso abierto para ser actualizada, previo consenso y verificación, y usada para diferentes propuestas de visualizaciones por medio de lenguaje de programación o programas especializados en esto, ya sea *Python* por medio de Pandas y otras librerías o de *Voyant Tools*, entre otros. La finalidad es plantear nuevas líneas de investigación con base en herramientas que permitan ampliar una lectura atenta o cercana, en los términos de Moretti, a través de una distante, y así emularlo para otros casos en la historia de la literatura mexicana del siglo XX y lo que va del XXI en cuanto a periodismo cultural, ejemplo, por mencionar uno de los tantos, el *Aquí vamos*, suplemento del periódico *El Porvenir*.

El presente trabajo partió de una materia optativa dentro del mapa curricular de la carrera Licenciado en Letras Hispánicas, del plan 420 de la Universidad Autónoma de Nuevo León, titulada "Proyectos de investigación en entornos digitales". Se contó con el apoyo de dos grupos, diferentes turnos, matutino y nocturno. El proceso de digitalización tuvo varias etapas hasta el momento del corte de este artículo: 1) evaluación de la colección privada, 2) inventario de los volúmenes proporcionados por el dueño de la colección, 3) cotejo de la colección privada con el de la Capilla Alfonsina Biblioteca Universitaria, 4) entrenamiento del equipo de digitalización por parte del equipo de la CABU, 5) digitalización (captura, formato, ocr), 6) creación de copias de seguridad y 7) control de calidad (en proceso, al momento de la redacción). Como parte complementaria, se dio inicio a la creación de una base de datos donde se capturaron los metadatos de cada número dentro de los 8 volúmenes de la colección. De esta manera, en un primer acercamiento se pueden hacer tablas dinámicas para después visualizarlo y plantear líneas de investigación. Cuando se finalice esta parte del proyecto, se vaciará todo en una sola base datos debido a que al inicio de la investigación se trabajó en diferentes equipos.

lo. Suplement		Director	No. páginas		Notas	Estado
87	14 de agosto de 1979	Fernando Benítez	16	12	Ligera rasgadura (página 1 y 2); página 9-10 está rota y	m Completo
	1 de septiembre de 1979		16	15	Error en página 4; Rojo en una palabra de la página 7	Completo
115	19 de enero de 1980	Fernando Benítez	16	14	Roto en las orillas todas las págs. del número.	Completo
116	26 de enero de 1980	Fernando Benítez	16	15		Completo
119	16 de febrero de 1980	Fernando Benítez	16	18		Completo
120	23 de de febrero de 1980	Fernando Benítez	16	11		Completo
121	1 de marzo de 1980	Fernando Benítez	16	10		Completo
122	8 de marzo de 1980	Fernando Benítez	16	14		Completo
123		Fernando Benítez	16	14		Completo
124	22 de marzo de 1980	Fernando Benítez	16	13		Completo
125	29 de marzo de 1980	Fernando Benítez	16	15		Completo
126	5 de abril de 1980	Fernando Benítez	16	14		Completo
127	12 de abril de 1980	Fernando Benítez	16	16		Completo
128	19 de abril de 1980	Fernando Benítez	16	13		Completo
129	26 de abril de 1980	Fernando Benítez	16	13		Completo
130	3 de mayo de 1980	Fernando Benítez	32	19		Completo
131	10 de mayo de 1980	Fernando Benítez	32	23		Completo
132	17 de mayo de 1980	Fernando Benítez	24	21		Completo
133	24 de mayo de 1980	Fernando Benítez	16	15		Completo
134	31 de mayo de 1980	Fernando Benítez	24	21		Completo
>	Volumen 1 Volumen 2	Volumen 3 V	olumen 4 Vol	umen 5 V	olumen 6 Volumen 7 Volumen 8 +	

Fig. 2 Captura de pantalla del inventario de la colección privada

Para efectos de este proyecto de digitalización, se decidió nombrar a la colección como "K" en honor de su dueño, el Dr. Roberto Kaput González Santos, quien heredó de su padre ocho volúmenes encuadernados, todos con leyendas que permitían catalogar a primera vista los años y su contenido. Sin embargo, entre el relato sobre la colección y el primer inventario hecho en la encuadernación original, fue necesario realizar un segundo para identificar el contenido real, tanto en los números como en su estado físico, de todos los volúmenes. Como puede observarse en el ejemplo de la figura 2, los equipos registraron el número del suplemento, su año de publicación, el director, el número de páginas totales, el número de textos, notas sobre algunas curiosidades del número y sobre el estado del material previo a su digitalización.

Esta etapa fue importante porque los grupos pudieron identificar la importancia de evaluar el material con el que se cuenta previo a la digitalización, además de iniciar un proceso de discusión al momento de plantear las etiquetas (su valor semántico), aspecto que se le conoce como conceptualización de datos. Por ejemplo, el debate fue intenso en el concepto de texto para efectos de esta investigación. Quedó como pendiente, debido al perfil de la carrera, integrar a la base de datos los de tipo visual, aunque desde la semiótica de la cultura (Lotman, 1996) los cartones (caricaturas) y fotografías entren dentro de una definición de texto. La investigación se centró solamente en textos verbales escritos, pero éste no fue el único punto de polémica en esta etapa. Al momento de entrar en la lectura del suplemento, hay títulos que tienen dentro más de un texto, como el caso de la poesía; por lo tanto, se tomó como criterio contar los poemas como textos individuales, aunque fueran parte de una sección entera. En el caso de la columna marcada en verde, ésta sólo funcionó como una forma de control en los números que fueron digitalizados.

Al observar este inventario, es evidente que los volúmenes no contaban con todos los números de los años señalados en sus leyendas de sus respectivas portadas. Esto es importante en el proceso de digitalización para crear una edición más completa, que agregue los datos faltantes en el diseño de un futuro mapa que registre desde la lectura distante las diferentes relaciones en el suplemento cultural *Sábado*. Posteriormente, un equipo de trabajo realizó el cotejo para constatar qué números faltantes se encontraban en la CABU. Esto se hizo con dos propósitos: 1) registrar los números disponibles para mejorar la edición digital (todavía en proceso) con base en el cotejo realizado con la colección de la hemeroteca de la CABU, y 2) capturar las copias que estén en mejores condiciones en comparación con los

datos de origen o por el desgaste en los volúmenes.

De esta manera, se registraron los números faltantes en la colección K que están disponibles en la biblioteca universitaria: a) 160, 154, 197, 195, 194, 193, 183, 176, 174, 173, 172, 170, 168, 166 (números sueltos); b) 101, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 117, 118, 219, 234, 236, 245, 250, 252, 261, 262, 263, 264, 270, 281, 285, 286, 288, 290, 294, 295, 296, 297, 299, 301, 302, 303, 305, 306, 309, 312, 313, 314, 315, 321, 332, 335, 336, 344, 345, 347, 350 (números en tomos encuadernados). Como puede observarse, la aplicación de una lectura distante para la visualización por medio de mapas y/árboles corre un riesgo importante en el planteamiento de preguntas de investigación e hipótesis si no se cuenta con datos suficientes.

Es atractivo, como ejercicio escolar, pero impreciso si la muestra no comprende una parte más amplia y representativa del universo con el fin de mostrar una evolución en este tipo de periodismo cultural. Cabe mencionar que este cotejo se realizó con base en los tomos existentes de la colección K. Un proyecto más ambicioso y con los recursos necesarios puede realizar un análisis de datos mucho más amplio para abarcar las dos épocas importantes, tanto la de Benítez como la de Batis, y observar en una escala más grande los cambios de temas, enfoques, géneros, entre otros tantos fenómenos presentes en dicha cantidad enorme de textos. Sin embargo, pueden plantearse lecturas interesantes a partir de una muestra de 179 números con aproximadamente 1495 textos¹, esto sin contar sus caracteres ni tokens al momento de dicha fase previa al de digitalización.

¹ En el control de calidad, se observó que en la hoja de cálculo del volumen 8 no se capturó la cantidad de textos para los números 337 y 346, con fechas 14 de abril de 1984 y 14 de junio del mismo año.

/olumen	Nùmero	Fecha	Autor	Titulo	Nacionalidad	Género	Descripción			Traductor	Género (\$)
- 1	1 2	84 16 de abril de 1983	Cioran, E. M.	Desgarraduras	Francesa	Ensayo	Fragmento del libro traducido y publicado por e	Existencialismo, Indecisión, Historia, IS	Sí I	María Dolores Aguilera	Hombre
	1 2	84 16 de abril de 1983	Montherlant, Henry de	Pasifae	Francesa	Teatro	Obra dramática donde se retoma al personaje m	Monólogo, Drama, Mitología, Grecia, 15	Si I	Raúl Falcó	Hombre
- 1		184 16 de abril de 1983	Sciascia, Leonardo	Recortes del diario de Sciascia	Italiana	Diario	Fragmentos del diario de Leonardo Sciascia.	Diario, Filosofia, Critica, André Malrau S	Si I	Héctor Abad	Hombre
- 1	1 2	184 16 de abril de 1983	Bedford, William	La pintura	Estadounidense	Poesía	Poema sobre la relación de una voz poética con	Pintura, Locus amoenus, Muerte, Tien S	Sí I	Dario Galicia	Hombre
- 7	1 2	84 16 de abril de 1983	Montiel, Edgar	¿Conformismo o subversión creadora? I	Mexicana	Ensayo	Texto sobre el debate de la emancipación fillosó	Filosofía, Imitación, Creación, Gabriel I	No I	NA.	Hombre
7	1 2	84 16 de abril de 1983	Benedetti, Mario	La comarca y el juglar	Uruguaya	Ensayo	Trata una polémica sobre la validez de los canta	Lluis Llach, Luis Eduardo Aute, Raimor I	No I	NA.	Hombre
		84 16 de abril de 1983	Aguilar Rivero, Mariflor	La filosofia, su historia y su sujeto	Mexicana	Ensayo	La filósofa mexicana argumenta la importancia o	Althusser, Foucault, Ideología, Discurs I	No I	NA.	Hombre
	1 2	184 16 de abril de 1983	Ruy Sánchez, Alberto	Dama de corazones de Xavier Villaurrutia	Mexicana	Ensayo	El autor toma un texto narrativo de Xavier Villau	Narrativa, Xavier Villaurrutia, Critica, (1	No I	NA.	Hombre
	1 2	184 16 de abril de 1983	Schavelzon, Guillermo	Los líbros en el mundo	Argentina	Reseña	El autor presenta tres noticias diferentes en el r	El País, Periodismo, Feria del Libro, Lit 1	No I	NA.	Hombre
- 1	1 2	84 16 de abril de 1983	Volkow, Verónica	Solo en la piel existe	Mexicana	Poesía	Poema de la escritora mexicana Verónica Volko	Piel, Versos, Noche, Visceral, Oscurid 1	No I	NA.	Mujer
7	1 2	84 16 de abril de 1983	Vera, Luis Roberto	Toledo dibuja un laberinto sobre una tortug-	Chilena	Poesía	Poema del escritor chileno donde una pintura d	Pintura, Mar, Dibujo, Muerte, Oscurid I	No I	NA.	Hombre
1		84 16 de abril de 1983	Lima, Bernardo	Socialización o muerte en el arte	Mexicana	Reseña	Reseña sobre el libro de Adolfo Sánchez Vázque	Marxismo, Praxis, Muerte, Arte, Ideol I	No I	NA.	Hombre
1	1 2	84 16 de abril de 1983	Verlichak, Victoria	Vidas ejemplares. Soledades insondables	Argentina	Reseña	Reseña sobre el libro de Mempo Giardinelli, titu	Ediciones del Norte, Novela, Novela M	No I	NA.	Mujer
		84 16 de abril de 1983	Falcó, Raúl	Prolegómeno	Mexicana	Poesía	Texto poética donde el Yo dialoga con la muerte	Minos, Laberinto, Muerte, Pasifae, Min	otauro, Linaj	je, Genealogía.	Hombre
- 1	1 2	184 16 de abril de 1983	Vallarino, Roberto	Cuatro bocetos de Carlos Chimal	Mexicana	Reseña	Reseña sobre el libro de Carlos Chimal titulado	Cuento, José Emilio Pacheco, Gustavo I	No I	NA.	Hombre
- 1	1 2	84 16 de abril de 1983	Patán, Federico	Marilyn Monroe y yo	Mexicana	Reseña	Reseña sobre el libro de Carlos Sampietro sobre	Marilyn Monroe, Poesia, Vacio, Tristez I	No I	NA.	Hombre
1	1 2	184 16 de abril de 1983	Blanco Figueroa, Francisco	La torre inclinade de Virginia	Mexicana	Reseña	Reseña sobre el libro "La torre inclinada", de Vir	Bertrand Russell, Keynes, M. Forster, II	No I	NA.	Hombre
- 1	1 2	84 16 de abril de 1983	Romo, Mariel	Poema	Mexicana	Poesía	Poema con dibujos de Francisco Toledo. En él ap	Francisco Toledo, Oscuridad, Muerte, 1	No I	NA.	Hombre
- 1	1 2	84 16 de abril de 1983	Navarro, Ornelas	El concepto de universidad	Mexicana	Reseña	Reseña sobre el libro de Jaime Castrejón Diez "	Jaime Castrejón Diez, Universidad, Co I	No I	NA.	Hombre
		184 16 de abril de 1983	Castel, Juan Pablo	Marx y América Latina	NA	Reseña	Reseña sobre el libro de José Arico, "Carlos Mar	Karl Marx, América Latina, Ensayo, Div I	No I	NA.	Hombre
7	1 2	84 16 de abril de 1983	Gómez Montero, Sergio	El partido de la revolución/II	Mexicana	Reseña	Reseña sobre el libro de Luis Javier Garrido "El p	PRI, Revolución, PNR, Luis Javier Garri I	No I	NA.	Hombre
-		84 16 de abril de 1983	Tiffenberg, Ernesto	Antología de Constantini	Argentina	Reseña	"Cuestiones con la vida". Editorial Katún.	Humberto Constantini, Funes, Bandec I	No I	NA.	Hombre
		184 16 de abril de 1983		Traza. Entrevista con Juan Pablo Castel	Mexicana	Entrevista	Entrevista con el Secretario de Redacción de la R	Juan Pablo Castel, Arquitectura, Critic I	No I	NA.	Mujer
		84 16 de abril de 1983	Valdéz Medellin, Gonzalo	Hablar de música/I	Mexicana	Entrevista	Entrevista con el músico José Antonio Alcaráz. S	Jospe Antonio Alcaraz, Música, México	No I	NA.	Hombre
		184 16 de abril de 1983	Cervantes, Francisco	Refleios o desdoblamientos en los dibuios d	Mexicana	Ensavo	Ensavo crítico sobe la obra de Bia Wouk. Se disci	Bia Wouk, Pintura, Dibujo, Arte, Critic I	No I	NA.	Hombre
		184 16 de abril de 1983	Angier, Natalie	Margaret Mead/Naturaleza VS cultura	Estadounidense	Ensayo	Ensayo sobre la figura de Margaret Mead, antro	Margaret Mead, Ensayo, Antropología S	Sí I	Mercedes Benet	Mujer
		84 16 de abril de 1983	García, Gustavo	La hojita perroquial	Mexicana	Ensayo	Ensayo crítico sobre la importancia de la difusión			NA.	Hombre

Figura 3. Metadatos del número 284 de Sábado

Al mismo tiempo se inició de manera transversal una captura de metadatos en ocho archivos separados para agilizar el trabajo en equipo. Se estableció en juntas previas qué etiquetas serían importantes para el proyecto antes del uso de herramientas digitales para diseñar mapas o árboles de relación. Cada número contenido en el volumen tendrá una hoja de cálculo para al final juntarlo todo en un solo archivo. En el ejemplo de la figura 3 puede observarse todos los textos, salvo las recomendaciones de Batis², con su número, fecha, autor(a), título, nacionalidad de su autor(a), género literario, descripción, temas a manera de palabras clave, si el texto es una traducción o no, nombre de su traductor(a) y sexo de su autor(a).

Pensando esto para una cantidad de casi 1,500 textos, el conjunto de metadatos aumentaría concretamente para el caso de los temas, lo cual dará como resultado un panorama interesante en la dinámica evolutiva de este suplemento. Es necesario mencionar en este punto que algunos de los datos tuvieron que obtenerse por medio del cotejo de otras fuentes, algunas de ellas virtuales,

² Pendiente para otra línea de investigación con la pregunta '¿qué recomendó *Sábado* para leer?', para el caso de esta sección y no el de las reseñas.

como enciclopedias en línea (literatura mexicana) o repositorios con trabajos de titulación relacionados con la época o la obra de quien participó dentro de *Sábado*. Algunos datos fueron imposibles, al momento de este corte, debido a nombres que posiblemente funcionaron como pseudónimos³.

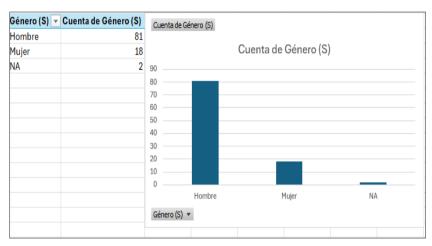


Figura 4. Datos del sexo de las personas involucradas en cuatro números de *Sábado*

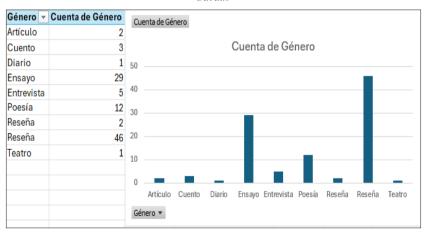


Figura 5. Datos de los géneros literarios presentes en cuatro números de Sábado

³ En etapa de verificación, por ejemplo, está el caso de Juan Pablo Castel.

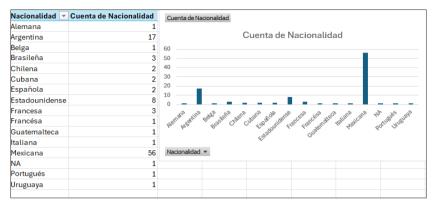


Figura 6. Datos de las nacionalidades de autoras y autores presentes en cuatro números de *Sábado*

En las figuras anteriores se puede apreciar un ejercicio previo con las herramientas de Excel, limitadas en comparación a las que se usarán en la siguiente etapa del proyecto. Sin embargo, lo anterior sirve como muestra, a partir de cuatro hojas de cálculo terminadas en su captura (números 280, 282, 283 y 284, del Vol. 8 de la colección K), de cómo puede cuantificarse los datos de forma ordenada a partir de un primer paso en la lectura distante. Si esta base de datos, al final de su elaboración, logra su objetivo de ser un recurso abierto para quien tenga el interés de plantear una línea de investigación, una persona con acceso a ella podrá realizar en principio tablas dinámicas como la siguiente:



Figuras 7 y 8. Ejemplos de tablas dinámicas en Excel.

La herramienta de tablas dinámicas permite la creación de columnas o filas en una nueva hoja de cálculo, pero entre sus opciones está el utilizar una etiqueta a manera de filtro; es decir, partiendo de los cuatro números capturados como ejemplo, es posible filtrar solamente autores y autoras que publicaron poesía en *Sábado* o seleccionar una de las nacionalidades que participaron en este suplemento cultural. No es el único filtro, pues se cuenta con la etiqueta de título y éste puede ser añadido en una nueva tabla dinámica junto a la fecha de publicación como puede verse a continuación:

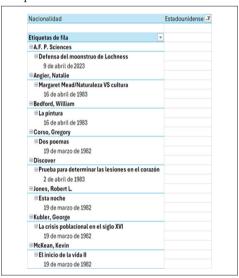


Figura 9. Imagen de una tabla dinámica

La ventaja de las tablas dinámicas es que pueden reutilizarse en tiempo real. Por ejemplo, podríamos cambiar la nacionalidad a la etiqueta mexicana y obtener todas las personas de ésta que participaron en el suplemento cultural. Esta base de datos se está llevando a cabo con la finalidad de proporcionar a diferentes profesionales de la investigación a jugar con los datos y plantear posibles líneas de investigación con base en un registro detallado de lo que hay dentro del suplemento, por ejemplo, "La aportación

estadounidense en *Sábado*", "La poesía escrita por mujeres en *Sábado*", "La traducción de poesía en *Sábado*", entre otras tantas que puedan surgir a partir de esta conceptualización de datos.

La versión digital del suplemento no estaría disponible debido a derechos de autor, pero una hoja de cálculo con etiquetas útiles puede ser de gran importante para realizar el cotejo con las diferentes ediciones que tengan a la mano cerca de sus instituciones o repositorios especializados en la materia. Al plantear la posibilidad de acceso abierto, esto permitiría al proyecto alcanzar los metadatos pendientes que no existan en la colección K y la de la CABU. Es importante resaltar que, hasta este punto del proyecto, la visualización es inexacta con respecto a todo este universo que denominaremos *Sábado*. Mientras más datos se registren en la evolución de éste, mejor será la imagen que se tendrá del suplemento desde la aplicación de una lectura distante no sólo a través de un Excel, sino en la aplicación de lenguaje de programación como *Python* o herramientas útiles en el campo, por ejemplo, *Voyant Tools*.

Digitalización y primera fase de la lectura distante

El proceso de digitalización tomó cerca de tres meses, dos sesiones de una hora y media por semana. Antes de iniciar el proceso, después de la realización del inventario, fue necesaria la asesoría de la gente encargada del taller de restauración de la CABU debido a detalles de origen en el encuadernado de la colección K. Una de las dificultades, además del estado de algunas de las impresiones, fue que varias páginas estaban cosidas de tal manera que la captura sería difícil en algunas partes.

La primera opción fue consultar qué tan viable era deshacer la encuadernación para continuar con su posterior restauración. Sin embargo, gracias a las observaciones de Reyna Vera, especialista en restauración, se optó por desechar esta opción debido a que los cortes no permitirían un nuevo proceso. La segunda opción que se tomó fue capturar las páginas de tal manera que el ojo humano pudiera leer su contenido y reconstruir los caracteres faltantes en el OCR por medio de archivos que funcionaran como una edición paralela al de la imagen.

Por esta razón en el proceso de vaciado se agregó un proceso más: la limpieza manual por parte del equipo de trabajo de cada texto al momento de ser trasladado a otro formato. Una de las herramientas alternativas a los programas de *Adobe Acrobat Reader* o *PDF 24 Tools* que se utilizó, en caso de largos apartados donde el reconocimiento óptico por computadora tuviera problemas, el *Google Keep* por su gran capacidad para realizar la captura de caracteres donde los programas populares o de acceso abierto más conocidos no pueden llegar. En los casos más difíciles, según el semáforo de evaluación física que se realizó en el inventario, la reconstrucción tomó criterios de cotejo y de reconstrucción morfológica de acuerdo con el contexto de cada texto. El número que se toma como punto de partida y ejemplo de las herramientas que se pueden utilizar para las siguientes fases del proyecto *Sábado* es el 188, publicado el 13 de junio de 1981. El índice de textos es el siguiente:

- El tapado y yo, de Manuel Maples Arce.
- Proyecto Nacional: Estado y Sociedad Civil, de Carlos Pereyra.
- Un drama de caza, de Anton Chéjov (traducción de Sergio Pitol).
- Entrevista a Antoni Peyrí: el centro de la obsesión irradia la diferencia, de Tununa Mercado.
- Exposición de Antoni Peyrí, de Mariana Frenk-Westheim.
- Joseph Grau Garriga: una revolución del tapiz, de Lelia Driben.

- Helarte de la errata, de Rafael Vargas.
- Las artes plásticas en el IX festival cervantino, de Graciela Kartofel.
- El santo contra la proletarización de la derecha, de Gustavo García.
- La música que no tiene futuro, de Carlos Chimal.
- Desde la tierra de nunca jamás a veces, de Guillermo Sheridan.
- Orquesta filarmónica de la UNAM, de Graciela Phillips.
- Entrevista a Miguel Ángel Granados Chapa. El periodismo: búsqueda de las claves políticas, de Otto Granados Roldán.
- Japón prepara la gran ofensiva científica, sección a cargo de Víctor Miguel Lozano (traducción y colaboración de Mercedes Benet).
- Los libros partisanos de la editorial Maspero. Libros en el mundo, sección a cargo de Guillermo Schavelzon.
- El hombre cuestionado de Eugene Ionesco, de Huberto Batis.
- La iglesia como grupo de presión, de José Salvador Arronte.
- Entrevista con Bernardo Ruiz. La rosa de los vientos, de Margarita Pinto.
- Las semillas de la granada. El libro eterno, de Ida Vitale.
- Sábado, domingo y feria, de Henrique González Casanova.

Es necesario aclarar que se omitió una sección de *Sábado* dedicada a la recomendación de libros debido a que se planteó en el curso de "Proyecto de investigación en entornos digitales" como una línea aparte de investigación dentro de este suplemento cultural. Tampoco se incluyen anuncios, que son importantes para un futuro estudio sobre la publicidad en los años ochenta, sobre todo el contraste entre los sectores público y privado presentes en diferentes números. Tampoco se incluyen caricaturas, que no son constantes, pero sí aparecen esporádicamente en otras semanas de publicación.

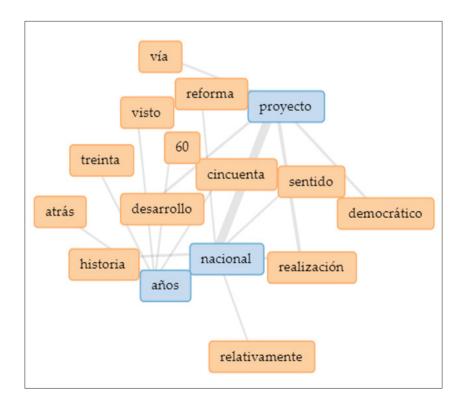
Como puede observarse, se trata de una cantidad considerable de textos, con diferentes géneros, extensión y enfoques. Dos son las formas en que puede abordarse cualquier número de *Sábado*, en principio, la primera es analizar todo el número como un conjunto frente a los otros números; la segunda consiste en realizar el análisis texto por texto para observar su comportamiento de forma mucho más particular. Por ejemplo, al aplicar la herramienta de análisis de datos *Voyant Tools* se obtiene el siguiente panorama de todo el número:

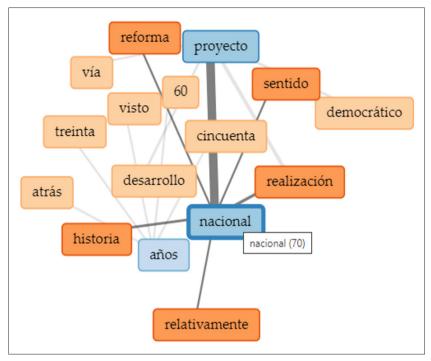




Figuras 10 y 11. Diferentes escalas de los términos más repetidos en todo el número 188

$\overline{}$	_	_		
\oplus		1	nacional	70
+		2	proyecto	44
+		3	años	36
+		4	conde	34
+		5	política	30
+		6	social	29
+		7	político	29
+		8	sociales	26
Η		9	tapiz	24
\pm		10	país	24
+		11	sociedad	23
+		12	sentido	23
+		13	caso	23
+		14	vida	22
+		15	crecimiento	22
+		16	arte	22
+		17	producción	21





Figuras 12, 13 y 14. Frecuencia de términos y relaciones entre ellos

Se trata de más de treinta y ocho mil palabras en todo el corpus que comprende el número 188, además de ser casi doscientos mil caracteres. La importancia de este tipo de análisis de datos y de proyectos de digitalización radica en un principio de interfaces. La cantidad de datos es muy grande, inmensa en otros medios como las redes sociales, y el visualizarlos por medio de estas herramientas permite la oportunidad de plantear preguntas de investigación sostenidas en una materialidad que difícilmente puede registrarse con la misma velocidad.

Es importante señalar una de las limitaciones principales del Voyant Tools al momento del corte de esta etapa de la investigación: su lista de palabras vacías en español es muy corta en comparación con el de las comunidades digitales, como el caso del lenguaje natural en Python, es decir, las llamadas "stop words". El consejo principal que se le extiende a quienes decidan usar esta herramienta es actualizar la lista que se tiene por default en la página voyant-tools.org.

	Término	Contexto del término	Contar (contexto)
	proyecto	nacional	38
	proyecto	realización	6
	tapiz	tradicional	6
	tapiz	contemporáneo	6
	social	dominante	5
	nacional	realización	4
	conde	amigo	4
	nacional	historia	3
	proyecto	desarrollo	3
	proyecto	democrático	3
Pank	∨ ? 1,106 contexto • 1,106 contexto		

Figura 15. Ejemplo de los términos en sus contextos

A partir de lo anterior, además de las líneas que puedan surgir en una lectura cercana, es posible plantear preguntas de investigación con base en una materialidad. Por ejemplo, ¿cuál es el sentido y percepción de lo nacional en Sábado, suplemento cultural? ¿Existe una evolución en las relaciones del término? ¿Quiénes son los autores o autoras y cuál es su continuidad y ruptura a lo largo de las diferentes épocas de dicho espacio cultural? Tal como se ve desde un enfoque del número 188 a manera de un todo, se observa el peso que tiene lo nacional con respecto a diferentes proyectos. Sin embargo, esto sería erróneo si se asevera que es la línea discursiva de todo el número, pues esta frecuencia de términos se debe a que el texto de Pereyra es uno de los más largos, a lado de la entrevista a Miguel Ángel Granados Chapa. Por esta razón es importante incluir la siguiente gráfica:

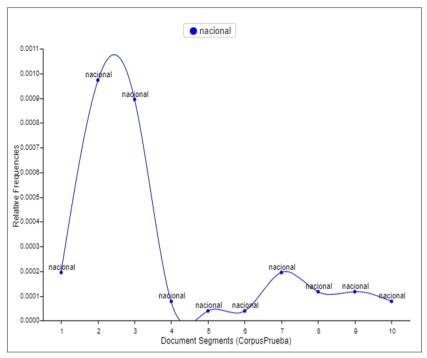


Figura 16. Evolución del término "nacional" en el número 188

Aunque son útiles las visualizaciones de una gran cantidad de datos, es importante contar con diferentes interfaces para contar con una imagen más precisa del vector en cuestión. En este caso, el término "nacional" no impera de manera uniforme en todo el número 188 de *Sábado*; sin embargo, es interesante ver cómo se mantiene a flote en el resto de los segmentos del corpus, en textos de diferente índole relacionados con la música clásica y el jazz y no con una crítica a las políticas económicas de los años ochenta, como en el caso de Pereyra. Por esta razón, aunque sea atractivo un criterio de lectura global por medio de estas herramientas, es necesario también partir de criterios metodológicos específicos con el fin de lograr una vista mucho más precisa. Tómese como ejemplo las siguientes visualizaciones:











Figuras 17, 18, 19, 20 y 21. Visualización de un corpus de cuatro entrevistas en el número 188 de *Sábado*

En las figuras anteriores es posible observar las siguientes entrevistas de izquierda a derecha, arriba a abajo: 1) Entrevista a Antoni Peyrí (fig. 17), 2) Entrevista a Miguel Ángel Granados Chapa, 3) Entrevista a Bernardo Ruiz, 4) Entrevista a Joseph Grau Garriga y 5) la visualización global de las cuatro entrevistas. En la gráfica que se presenta a continuación, mediante la Voyant Tools, es posible identificar vectores en común entre un grupo de textos. Para el caso que

se ilustra, tomando en cuenta solamente el género de entrevista (para un número), puede identificarse un punto de encuentro por medio de las coincidencias con una raíz: polit-. Se trata de un corpus pequeño donde se observa una frecuencia entre la entrevista 2 y 4. Esto podría aumentar, sin duda, con un corpus mucho más amplio, es decir, el total del proyecto que todavía sigue en curso. Sin embargo, es importante resaltar cómo esta herramienta puede ayudar en la búsqueda del contexto de un vector determinado para un corpus como el de *Sábado*.

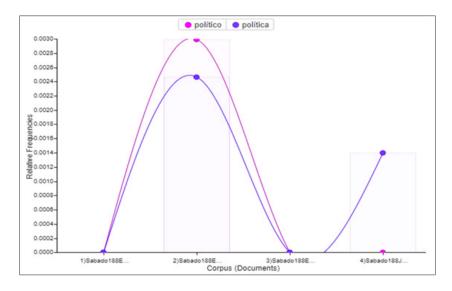


Figura 22. Ejemplo de gráfico en *Voyant* tools a partir de un término recurrente seleccionado

Además de la herramienta anterior, es posible aplicar lenguaje de programación como Python, con el fin de tokennizar un corpus por medio de librerías pre entrenadas en la clasificación de textos. Algo muy interesante que permite el lenguaje de programación, con respecto a los programas existentes, es la experimentación o el agregar una serie de pasos que permitan 'limpiar' el texto con mayor precisión que un programa abierto en una página Web. Hay dos

desventajas en la aplicación del *Voyant Tools* para el español, como lengua base del corpus: 1) la lista de palabras vacías en la aplicación Web no está actualizada, lo dará serias imprecisiones; 2) al tokenizar o lematizar, se obtiene un mejor manejo en la identificación de términos recurrentes. Como ejemplo de estas diferencias, se tomó una entrevista de Tununa Mercado realizada a Antoni Pyerí donde se realizaron los siguientes pasos de limpieza:

```
TununaNLP.py > ...
     Texto = "Nunca supe por qué la reja de entrada a la galería de arte Sloane R
   #Convertir el objeto en minúculas
    print (Texto.lower())
     #Segundo objeto: texto en minúsculas
    Textomin = (Texto.lower())
    #Limpieza puntos, comas, dos puntos, punto y coma, guiones
     print (Textomin.replace(".", "").replace(",", "").replace(":", "").replace("
     TextoLimpio = (Textomin.replace(".", "").replace(",", "").replace(":", "").r
     #Librería de lenguaje natural para tokenizar
     import nltk
    nltk.download('stopwords')
     import nltk
     nltk.download('punkt')
     #Importación de corpus de palabras vacías y tokens
     from nltk.corpus import stopwords
     from nltk.tokenize import word tokenize
     #Definición de palabras vacías en español
     stop_words = set(stopwords.words('spanish'))
    print(stop_words)
    tokens = word tokenize(TextoLimpio)
    print(tokens)
    #Texto limpio
     texto_filtrado = [word for word in tokens if not word in stop_words]
    print(texto filtrado)
```

Figura 22 y 23. Ejemplo de creación de objetos en Python y limpieza de un texto

Las imágenes anteriores muestran paso a paso cómo "limpiar" un texto en el sentido del contexto de programación. El ejemplo es de lo más simple y cambia cuando se convierte todo un archivo en una string. No obstante, lo anterior sirve como un punto de partida para observar los cambios que ocurren con textos pequeños y cómo esto podría afectar un corpus más grande. La ventaja actual del Visual Code Studio es que cada elemento del lenguaje de programación es identificado por un color. Las oraciones que se encuentran en verde, con un signo de gato al inicio, son comentarios que el programa VCS identifica como indicaciones que se pueden compartir en las diferentes comunidades de programación, pues Python es una herramienta de acceso abierto. Como se puede ver en las imágenes hay diferentes objetos para ilustrar la creación de los mismos en un programa, cosa que no es necesaria y puede simplificarse por medio de la práctica. Según el propósito que tenga la persona dedicada a la investigación, será necesario contar con las librerías adecuadas (el lenguaje pre entrenado). Para el ejemplo presentado anteriormente, se trabajó con base en el Natural Language Tool Kit (NLTK), el cual permite identificar que los datos con los que se están trabajando son parte de lo que llamamos lenguaje natural. Para ayudar a la computadora a procesarlos es necesario realizar otro tipo de limpieza, distinta a la que se realizó con el OCR de este proyecto. Los datos deben estar todos en minúsculas, por ejemplo, eliminar signos que puedan alterar la frecuencia de datos; se quitan las palabras vacías (stop words)y, por último, se tokenizan para darle un valor numérico a cada término. Al representarlo visualmente, ocurre lo siguiente:

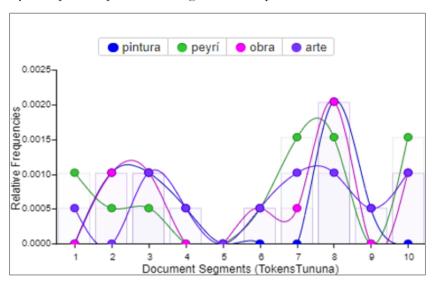


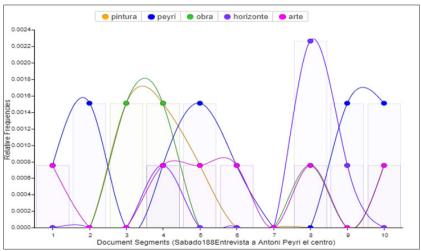


Figura 24 y 25. Del lado izquierdo no se aplicó la limpieza por medio de Python; del lado derecho, sí.

Parece no cambiar mucho en apariencia la anterior visualización por medio de un globo de palabras con los términos recurrentes, para la misma escala. Sin embargo, es importante señalar lo siguiente: 1) "Peyrí", "entrevista" y "paisaje" adquieren menos relevancia en comparación al primer caso; 2) los términos "obra" y "arte" adquieren mayor presencia. Parece ser algo intrascendente un cambio numérico de poco margen, pero al aumentar el tamaño del corpus esto puede llevar a etiquetas erróneas y una clasificación de textos menos precisa cuando se trata de algo limitado a una época, piénsese, por mencionar un caso, la de Fernando Benítez en

comparación con la de Huberto Batis. Una vez hecha la limpieza en Python puede apreciarse la siguiente comparativa.





Figuras 26 y 27. Del lado izquierdo, una gráfica con el análisis del texto completo; del derecho, solamente con los tokens

Es importante señalar el cambio que ocurre en la interfaz. Aunque la herramienta divide en diez segmentos, no se trata del mismo objeto. En el primer caso está toda la entrevista de Tununa Mercado a Antoní Pyerí, y en el segundo, solamente los tokens. Ambos casos permiten realizar minería de textos desde diferentes enfoques y necesidades; mientras el primero sirve para conocer el contexto de los términos, en el segundo se agiliza el modelaje en cuanto a clasificación de textos, no sólo por su etiquetado sustantivo, sino por su valoración; lo anterior será útil para identificar dentro de un gran corpus, como el de *Sábado*, los tipos de reseñas, no solamente por sus temáticas, artistas, sino por la valoración que realizan.

Conclusiones

La lectura distante es otra de las condiciones de conocimiento a la que se ha llegado en la actualidad. Sin embargo, a pesar de las grandes ventajas que ofrece la tecnología y la ciencia del Big Data o el apoyo de las inteligencias artificiales, es necesario resaltar la importancia de la creación de repositorios y bases de datos. Uno de los principales retos de las humanidades digitales radica en reconocer 1) la abundancia de datos en una lengua en proporción a otras, por ejemplo el inglés frente al español o el español con respecto a las lenguas originarias en diferentes países de Latinoamérica, y 2) que la asignación de recursos condiciona la memoria digital y, en consecuencia, su análisis, y esto puede reflejarse tanto en qué textos obtienen estos espacios o la ventaja comparativa entre una institución y otra.

Con respecto a *Sábado*, el presente proyecto continúa su siguiente etapa para explorar con mayor precisión un mapa de relaciones a lo largo de su producción y requerirá de más tiempo y recursos antes de entrar a una tercera. Los resultados presentados hasta este momento son preliminares e irán cambiando en su

ramificación, a manera de árbol o mapa conforme se agreguen el resto de los números al corpus de análisis y a la base de datos actual. Además de la importancia de los libros y de las revistas culturales, los suplementos son importantes para el caso del periodismo cultural por el territorio que crearon mediante la palabra; no obstante, ese territorio evoluciona, sobre todo si se trata de material impreso en un mundo donde la digitalización requiere de un trabajo de transición y preservación de otro tipo. Una de las claves principales para esto es la accesibilidad, pero esto involucra una cuestión de interfases cuando se trata de una cantidad enorme de datos.

Finalmente, para una aplicación de la lectura distante de un suplemento cultural como *Sábado* se requiere también crear espacios de oportunidad para que más gente pueda acercarse a su estudio desde este enfoque no encuentre obstáculos en la especialización de las herramientas. Se necesitan más recursos abiertos y simplificados en su proceso para convertir la lectura digital y el análisis de datos en un campo más accesible⁴.

Referencias bibliográficas

Lotman, I. (1996). La semiótica de la cultura y el concepto de texto. La semiófera I. Semiótica de la cultura y del texto (trad. Desiderio Navarro). Frónesis Cátedra /Universitat de Valéncia.

⁴ Muchas gracias al equipo de digitalización por su colaboración (captura y OCR) y esfuerzo en esta primera etapa del proyecto: Karla Daniela Lucio Silva, Fabián Arturo Macoy Ríos, Aldeny Ortega Martínez, Yazmín Madeleine Polanco Rodríguez, Francisco Javier Rodríguez Olivares, Debanhi Guadalupe Zapata Regis, Claudia Rebeca Barbosa Garza, Víctor Alejandro Flores Limón, Luis Gerardo Medina Morales, Marisol Yazmín Navarro Hernández, Yesenia Judith Navarro Hernández, Diego Jesús Torres Almaguer, Samantha Guadalupe Torres Rocha y Azael Contreras.

- Mercado, T. (1981). Entrevista a Antoni Peyrí: el centro de la obsesión irradia la diferencia. *Sábado* (suplemento cultural). 13 de junio (188).
- Moretti, F. (2015). *Lectura distante* (trad. Lilia Mosconi). Fondo de Cultura Económica.
- Peña Pimentel, M. y Sancho Caparrini, F. (2018). Big Data y Humanidades Digitales. *Humanidades Digitales*. *Lengua, texto, patrimonio y datos* (coord. Isabel Galina Russell et al). Bonilla Artigas/Red de Humanidades Digitales.
- Strate, L. (2021). Humanidades digitales en contexto (trad. Laura Meneses Trujillo y Marianna Montserrat Saldaña López). *Humanidades digitales en contexto* (coords. Laura Trujillo Liñán y José Rolando Islas Rivero). McGraw Hill.