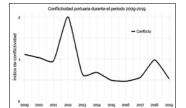
Intersecciones entre historia digital e historia social: un ejercicio de lectura distante sobre la conflictividad marítima en la historia argentina reciente*

Agustín Nieto

INHUS-CONICET/UNMDP (agustin.nieto77@gmail.com)

Data de recepció: 18 de desembre 2020 Data d'acceptació: 11 de gener de 2021 DOI: https://doi.org/10.51829/Drassana.28.650



■ RESUMEN

El giro digital es una revolución en curso y sus impactos en el escritorio de quienes se abocan la disciplina histórica es una realidad aún poco asumida y menos reflexionada. A partir de algunos ejercicios de lectura distante se busca reflexionar sobre estos cambios. Esta lectura permitirá presentar los resultados de una mirada panorámica en base a información extraída con técnicas de web scraping y analizada con técnicas de text mining. El corpus analizado está compuesto por el universo de notas publicadas por el portal de noticias portuarias Revista Puerto: la otra cara de la pesca. A partir de las técnicas mencionadas se procesará el corpus documental con el fin de listar la frecuencia y la forma del conflicto. Asimismo, se hará un ejercicio de georreferenciación de la conflictividad a lo largo de los años 2009-2019.

Palabras clave: historia social, historia digital, lectura distante, conflictividad marítima, Argentina.

Intersections between Digital History and Social History: a distant reading exercise on maritime conflict in recent Argentine history

■ ABSTRACT

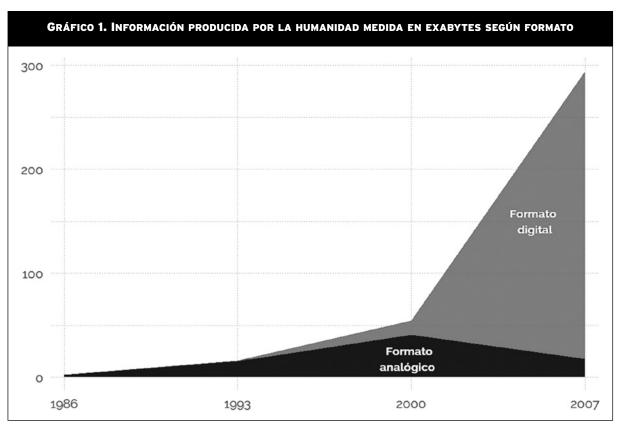
The digital turn is an ongoing revolution and its impact on the desktop of historians is a reality still hardly accepted and even less reflected on. This article reflects on these changes through distant reading exercises. This approach will allow us to present the results of a panoramic view based on information extracted with 'web scraping' techniques and analyzed with 'text mining' techniques. The analyzed corpus is composed by the notes published by the port news portal *Revista Puerto: la otra cara de la pesca*. The documentary corpus will be processed in order to list the conflict's frequency and form using the previously mentioned techniques. Likewise, a georeferencing exercise of the conflict will be carried out throughout the years 2009-2019.

Keywords: social history, digital history, distant reading, maritime conflict, Argentina.

■ HACER ARCHIVO EN LA ERA DIGITAL: NOTAS INTRODUCTORIAS

Según un cálculo realizado por Google, hasta el año 2003, la humanidad había generado cinco exabytes de información. Solo ocho años más tarde, la información ge-

nerada había alcanzado los 1.800 exabytes, gran parte de ella en formato digital. A partir del año 2000, la información en soporte analógico inició su declive¹ (ver gráfico 1). Hoy se producen 2,5 billones de bytes de datos diarios. Y aunque estamos muy lejos del "fin de la teoría" decretado por Chris Anderson² –afirmación que concitó duras críticas³–, el crecimiento exponencial de datos en formato digital nos desafía a interseccionar teoría, big data y lenguajes de programación.



Fuente: elaboración propia a partir de Martin Hilbert i Priscila López, "The World's Technological Capacity to Store, Communicate, and Compute Information", en Science, 332.6025 (2011): 60-65. https://doi.org/10.1126/science.1200970

Este tsunami de datos se enmarca en un giro digital que impacta en el quehacer cotidiano de quienes ejercemos el oficio de la historia⁴. Tecnologías como la mensajería electrónica, las cámaras fotográficas digitales, los archivos PDF (portable document format), los smartphones, las redes sociales, los blogs, los buscadores, entre muchas otras, introdujeron "capas de sentido" en las labores cotidianas⁵. Ya no es difícil percatarse de que en nuestro escritorio de trabajo los documentos físicos se desvanecen al ritmo del crecimiento de nuestros archivos digitales. El característico desorden material que adornaba nuestro espacio de trabajo dejó lugar al "desorden digital" que se traslada de la web a nuestra computadora personal⁶. El fichaje, que hasta hace unas décadas era analógico, al igual que las fuentes y la bibliografía fichadas, hoy deja el lugar a los procesadores de texto, las hojas de cálculo y las bases de datos digitales que acumulan anotaciones sobre fondos documentales digitalizados o nacidos digitales. Hace tiempo que los corpus heurísticos se miden en bytes y no en metros lineales. Bajo esta Iluvia de datos, adoptamos y usamos con cierta candidez algunas de estas innovaciones técnicas sin reparar demasiado en sus posibles impactos en el quehacer disciplinario. En cierto modo, actuamos como si el cambio en las formas fuera solamente eso, y no una profunda alteración de los contenidos7.

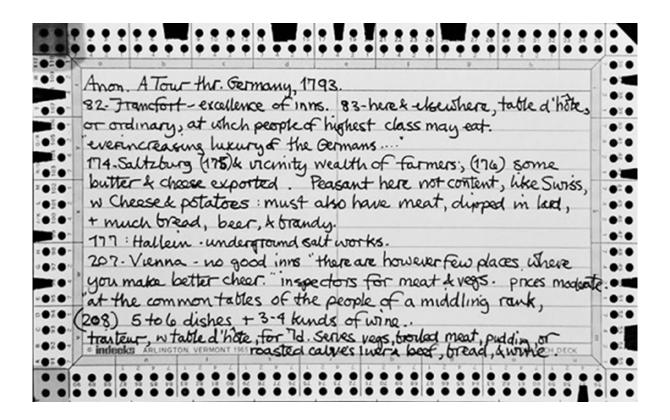
En este contexto, cuando los manuales sobre metodología de la historia citados en planes y programas de estudios abandonan el conservador apego al carácter artesanal del oficio y avanzan sobre las técnicas de investigación, se asemejan más a una fuente documental pertinente para un capítulo de la historia de la historiografía que a un recurso didáctico para enseñar el oficio. Las menciones e imágenes de tarjetas perforadas procesadas por "computadoras de habitación" dan cuenta de una brecha con el presente más amplia que su cronología. Estas tarjetas, nacidas al calor de la revolución industrial, fueron las semillas de los actuales formatos de almacenamiento, el procesamiento computacional de datos y, también, de la empresa International Business Machines (IBM), que lanzó en 1969 el Statistical Package for the Social Sciences (SPSS). Sin embargo, estos cambios técnicos fueron dosificados y, en el caso de la historia, focalizados (cliometría –New Economic History–), razón por la cual sus alcances e impactos no son equiparables a las consecuencias del actual giro digital. Hasta no hace tantos años se podía practicar la historia sin contacto con el mundo digital, desde entonces a esta parte ese vínculo se ha vuelto imprescindible.

Como sostiene Nicolás Quirogaº, esta es una época que nos invita a reflexionar sobre el trabajo de archivo con repositorios digitalizados y documentos nacidos digitales: ¿cómo pensarlos y procesarlos "desde y más allá de las herramientas tradicionales de la investigación social"? Esta es la cuestión.

Notas metodológicas: de la lectura cercana a la lectura distante

Aparte de mostrar su carácter arcaico, la reproducción de la imagen de la tarjeta perforada (ver imagen 1) busca llamarnos la atención sobre las protoformas de las "lecturas no cercanas". Formas que nunca estuvieron reñidas con las "lecturas atentas". Esta tarjeta perforada nos habla de una forma algorítmica supervisada para identificar tópicos al estilo del actual Atlas-ti. Sí, dije algorítmica, pues debemos saber que, después de todo, el halo de misterio que flota sobre la noción *algoritmo* no condice con su definición: "conjunto de instrucciones o reglas definidas y no ambiguas, ordenadas y finitas que permite, típicamente, solucionar un problema, realizar un cómputo, procesar datos y llevar a cabo otras tareas o actividades" 10. Y en esto consistía el trabajo con tarjetas perforadas.

En este apartado vamos a prescindir de entrar en el amplio, profundo y, probablemente, perpetuo debate sobre la conveniencia o inconveniencia del uso de la lectura distante en la investigación histórica". El desacuerdo en este campo es grande, pero no nos impide avanzar en la realización de ejercicios que impliquen alguna forma de lectura distante. En todo caso, debemos saber que, al igual que ocurre con los distintos tipos de fuentes históricas, todo método presenta ventajas y desventajas. En la historia digital y en la no digital siempre es deseable la triangulación de fuentes, datos, técnicas y métodos.



Antes de seguir, vale recordar que la "lectura distante" surge como iniciativa para "leer" un volumen de libros no abarcable desde una "lectura a ojo". Y esto en el marco de un análisis de historia cultural de la literatura universal que pudiera escapar al sesgo de las obras canónicas. En este sentido, los vínculos entre la vieja cliometría asociada a la historia económica y la lectura distante asociada a la historia de la literatura no son tan claros, pero sus distancias sí. Mientras la primera se basa en experiencias de "lectura cercana" de números, la segunda se basa en la ruptura de la experiencia de la "lectura cercana" de textos. Sin embargo, el enfoque de Moretti es reproducible por fuera de la historia de la literatura universal y podría dar lugar a una renovación y redefinición de la cliometría; quién sabe. En todo caso, aquí nos interesamos por otras mixturas y justificamos el ejercicio de lectura distante como una ventana a la datificación computacional en el campo de la historia social. Volvamos a Moretti.

¿Cuál fue el pacto que hizo Moretti con el diablo? iAprender a no leer textos! i¿De eso se trata la lectura distante, de una no lectura?! Es conveniente que no seamos tan literales en el tratamiento de la tesis de Moretti¹². En este sentido, la modulación hecha por Martin Mueller nos parece sugerente: una lectura escalable. ¿En qué consiste este tipo de lectura? Según Mueller:

Scalable reading, then, does not promise the transcendence of reading, – close or otherwise – by bigger or better things. Rather it draws attention to the fact that texts in digital form enable new and powerful ways of shuttling between 'text' and 'context'. Who could complain about tools that let you rapidly expand or contract your angle of vision?¹³.

Pensamos que no deberíamos quejarnos...

Ya con la decisión tomada, nos gustaría enunciar los enfoques analíticos prevalecientes en el ejercicio que vamos a encarar con un corpus documental reducido (small data). Usaremos un volumen de datos pequeño porque el presente artículo pretende ser un ejercicio de lectura distante controlado. Minaremos el corpus de textos con técnicas de procesamiento de lenguaje natural (interface

computacional para interactuar con el leguaje humano) para reconocer entidades (NER)¹⁴. Claro está que para realizar estos análisis necesitamos auxiliarnos con algún tipo de *software* y algún tipo de *hardware*.

Notas tecnológicas: sobre el software y el hardware

El proceso de recolección, estructuración, normalización, análisis y visualización de la información semiestructurada disponible en el sitio web de la *Revista Puerto* se realizó a través de programación computacional. Cada etapa de la investigación requirió un desarrollo específico de líneas de código en leguaje R Project¹5. Muchas de las ideas para desarrollar el código fueron producto del diálogo con los avances realizados en el marco de otras investigaciones¹6.

El trabajo de archivo requirió extraer información semiestructurada de forma automatizada de la Revista Puerto. La unidad de extracción de información fue la nota individual publicada en formato de marcado HTML y el recurso específico para llevar adelante esta tarea fue el paquete¹⁷ rvest¹⁸. Es un paquete para raspado y análisis web creado por Hadley Wickham e inspirado en el módulo Beautiful Soup de Python. Internamente utiliza funciones de otros paquetes como xml2. La información extraída se reorganiza en formato tabular, cada fila de la tabla refiere a una única nota publicada en la Revista Puerto, mientras que cada columna refiere a una parte de la nota como puede ser el título, o el subtítulo o entradilla, el cuerpo de la nota, la fecha, el link, la foto, etc. El formateo estructurado de la información extraída se logra con el uso del paquete tibble¹⁹. Los tibbles son data frames, pero con un formato más ágil que los clásicos marcos de datos²⁰. El análisis de los datos requiere de un conjunto de paquetes, muchos de ellos agrupados en tidyverse que está diseñado para facilitar la instalación y la carga de los paquetes principales en un solo comando gracias a compartir representaciones de datos comunes y diseño de API²¹. Finalmente, para la visualización de los datos hacemos uso del paquete ggplot2, un sistema para crear gráficos de forma declarativa²².

Ahora bien, el uso de lenguajes de programación como R Project requiere de un soporte físico para su ejecución. Lo interesante del caso es que podemos optar por algo distinto a la tradicional PC de escritorio. Un hardware reducido en tamaño y precio (Raspberry Pi) que solo necesita periféricos o acceso remoto para ser utilizado y que tiene una gran potencia de procesamiento. Su tamaño se asemeja al de una caja de naipes españoles. Si bien es verdad que su origen estuvo ligado al espacio educativo (Yamanoor y Yamanoor 2017), desde muy temprano fue utilizado como una PC alternativa, y en los últimos años cosechó adherentes en las humanidades digitales y las ciencias sociales computacionales²³.

Una breve presentación del corpus documental utilizado

Antes de avanzar sobre este punto, nos gustaría recordar que el presente artículo no busca dar una explicación exhaustiva de los sucesos de conflictividad en la pesca comercial argentina, su intención principal es presentar las posibilidades que brindan algunas técnicas computacionales para la recolección ("hacer archivo") y el análisis de corpus documentales nacidos digitales, en el marco de una investigación histórica, aunque no exclusivamente. Veamos. Nuestro corpus documental es el universo de notas publicadas por el portal de noticias portuarias Revista Puerto: la otra cara de la pesca (https://revistapuerto.com.ar/). Es una de las pocas publicaciones específicas del sector pesquero que, desde el año 2009, se publica solo en formato digital. Desde entonces, mantiene su cobertura periódicamente y cuenta con un archivo de acceso libre en línea (2009-continúa). Esta publicación es de carácter comercial y se define como "independiente". Cuenta con anuncios de distintas empresas del sector. A diferencia de los otros pocos portales sobre la actividad pesquera, la Revista Puerto brinda información detallada y abundante sobre eventos de conflictividad en el sector que suceden en las distintas provincias con litoral marítimo. Finalmente, su sesgo ideológico no es más problemático que el que presenta

cualquier medio de difusión o cualquier otro tipo de fuente documental²⁴.

■ LA CONFLICTIVIDAD EN LOS PUERTOS PESQUEROS ARGENTINOS

La historia de los puertos es la historia de quienes allí laboran cotidianamente. Y la historia de este colectivo laboral es la historia de sus luchas. Dentro del universo portuario, la historia de la masa obrera del pescado no es la excepción. Es imposible descifrar la actual configuración pesquera argentina sin dar cuenta de los procesos de lucha que la entretejieron y entretejen.

La mirada de sentido común considera que la conflictividad es un momento excepcional en la vida de las sociedades. El mismo sentido común inunda las miradas que se ciñen sobre las protestas en los puertos pesqueros. Evocan un pasado pacífico jalonado por irrupciones iracundas e irracionales de masas obreras en protesta abierta. Pero entre jalón y jalón existe una trama, invisible a estas miradas, que conecta los grandes eventos de protesta a través de una pléyade de pequeñas y cotidianas escaramuzas entre quienes solo portan capacidad de trabajo y quienes se adueñan del trabajo ajeno.

La conflictividad en los puertos pesqueros de Argentina está íntimamente entrelazada con el discurrir económico del sector. La última década (2009-2019) se enmarca en un proceso de mayor duración y que hunde sus raíces hasta los primeros años de la década de 1960, cuando la industria conservera relegó su lugar de preeminencia frente al fileteado de pescado fresco y el mercado internacional se impuso al mercado interno. Treinta años más tarde, en los años noventa, la actividad conoció su punto más alto en las exportaciones y el comienzo de una larga crisis que llega hasta nuestros días, producto de la sobrepesca y la precarización laboral²⁵.

En los siguientes párrafos nos abocamos a presentar una mirada panorámica (y de superficie) basada en información extraída con técnicas de *web scraping* y analizada con técnicas de *text mining*. Están presentes tanto los grandes eventos como la conflictividad cotidiana protagonizada por la fuerza de trabajo de la industria pesquera

argentina. No solo nos ocupamos de listar la frecuencia y la forma del conflicto. También presentamos su geografía y a sus protagonistas a lo largo de los años que van de 2009 a 2019.

Conflictividad social: una definición elemental

Aunque el eje del presente texto no sea indagar en profundidad cuestiones conceptuales en torno a la noción de conflictividad y sus implicancias en el campo de la historiografía, consideramos que es oportuno dedicar unas líneas a presentar lo que entendemos por conflictividad. En el marco de este artículo, la noción de conflictividad refiere a acciones colectivas dirigidas contra alguna expresión del estado de cosas existente, de modo tal que incluye acciones tan heterogéneas como una declaración o conferencia de prensa hasta una toma de fábrica o una huelga. Por otra parte, sus sujetos pueden ser diversos e incluye acciones emprendidas por personificaciones de relaciones de dominación y explotación como puede ser, por ejemplo, el *lockout* de una cámara empresarial del sector pesquero²⁶.

Breve estado del arte

El persistente perfil agropecuario de este país le granjeó el mote de "granero del mundo". En ese contexto, las comunidades marítimas y portuarias tuvieron (tienen) poca centralidad tanto en la agenda gubernamental como en la agenda historiográfica. Por esta razón, el historiador José Mateo²⁷ afirma que la sociedad argentina se desarrolló de espaldas al mar. Por su parte, Eduardo Pradas, en su estudio sobre la industria pesquera, sostiene que la sociedad argentina "no solo vive de espaldas al mar sino, lo que es mucho más grave, de espaldas a las luchas y conflictos sociales que le dan sentido a los tiempos actuales"²⁸. Esta situación generó una vacancia que fue cubierta solo parcialmente.

En relación con los primeros momentos de la conflictividad social protagonizada por comunidades marítimas y portuarias vinculadas a la actividad pesquera, encontramos tres referencias, una centrada en el papel del sindicalismo pesquero en los orígenes del peronismo²⁹; otra preocupada por el protagonismo femenino en el mundo laboral pesquero durante el período de industria-lización por sustitución de importaciones³⁰, y la tercera, interesada en la historia del colectivo obrero de la conserva de pescado desde sus orígenes hasta 1966³¹. Por otro lado, en los últimos años se desarrollaron novedosas investigaciones sobre la historia reciente de la conflictividad en los puertos pesqueros de la provincia de Chubut³² y de las ciudades bonaerenses de Necochea³³ y Mar del Plata³⁴.

Como balance, podemos decir que en la última década el conocimiento sobre la conflictividad en comunidades marítimas pesqueras se amplió y diversificó, y logró en algunos casos avanzar hacia estudios comparados³⁵. De todos modos, los espacios de vacancia siguen siendo amplios, en particular en lo que hace a la historia más reciente. Por otra parte, son inexistentes las investigaciones sobre esta temática que intentan interseccionar historia social e historia digital. A partir de este diagnóstico, en los párrafos siguientes arriesgaremos en forma temeraria un primer acercamiento computacional a la historia reciente de la conflictividad en los puertos pesqueros argentinos ocurrida durante la última década.

Notas en torno a la conflictividad general

Gracias a los aportes de la escasa bibliografía existente, sabemos que, luego del conflicto de 2007 protagonizado por quienes trabajan en la industria del fileteado de pescado de la ciudad de Mar del Plata, el 2012 emerge como el año con mayor conflictividad laboral en la constelación de puertos pesqueros argentinos. Pero poco sabemos sobre lo que ocurrió antes y después del conflictivo año 2012. Comencemos por ver, gracias a un ejercicio controlado de lectura distante, la frecuencia anual de conflictividad en la pesca según lo registrado por la prensa sectorial.

La silueta dibujada por el índice interanual de conflictividad ratifica lo que ya sabíamos: el 2012 es el año con mayor índice de conflictividad del período bajo análisis (ver gráfico 2). También nos advierte algo que podíamos intuir pero que no conocíamos: entre los años 2013 y 2019



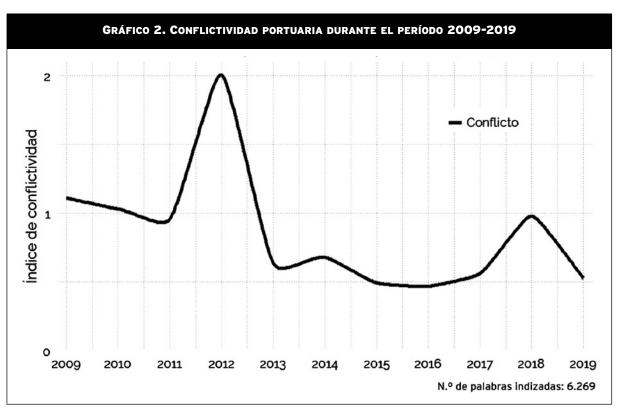
la media anual fue más baja que durante el período previo (2009-2012).

Indicios sobre los movimientos huelguísticos

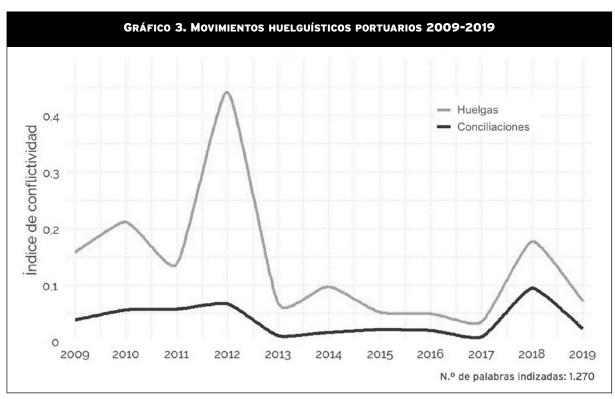
Cuando pasamos de la conflictividad general a los movimientos huelguísticos, notamos que la silueta muestra pequeñas variaciones perceptibles solo en los años 2009 y 2017 (ver gráfico 3). El carácter más pronunciado de los picos es un punto a destacar. Una razón puede deberse al rol dinamizador de la huelga en el marco de conflic-

tos de amplio alcance social y temporal. Asimismo, es notoria también cierta correlación entre movimiento huelguístico y dictado de conciliaciones obligatorias por parte de las carteras de trabajo provinciales y nacional.

Esta correlación es más marcada en el año 2018, dato que contrasta con la correlación del año 2012. Es esperable que esta diferencia se deba al cambio de orientación política en materia laboral mostrada por los ejecutivos provinciales y nacionales entre ambos periodos. Recordemos que en diciembre de 2015 se pasó de un gobierno de origen peronista (Partido Justicialista), crítico al neolibe-



Fuente: elaboración propia a partir de datos de la Revista Puerto.

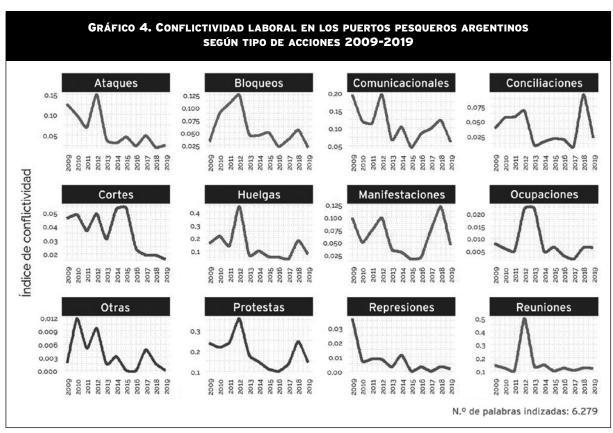


ralismo, a un gobierno antiperonista de origen liberal (Juntos por el Cambio), de alto contenido antisindical. Este perfil puede explicar el uso más reiterado de la conciliación obligatoria como recurso para disciplinar a los colectivos obreros en huelga.

La diversidad de las formas que adopta la lucha

La desagregación de la conflictividad por formato de acción habilita nuevas lecturas sobre la protesta obrera en los puertos pesqueros de nuestro país. La silueta inicial de la conflictividad general se despliega en un racimo de temporalidades divergentes (ver gráfico 4). Como primer aspecto a destacar emerge la pérdida de centralidad del año 2012.

Aunque el 2012 sigue mostrando preeminencia en formatos de acción como ataques, bloqueos, huelgas, protestas y reuniones, en otros aparece superado o igualado. Esto ocurre, por ejemplo, con las acciones estatales como conciliaciones y represiones, así como con otros formatos de protesta: manifestaciones, cortes y ocupaciones. Es sugestiva la caída abrupta y sostenida de los cortes de calles y rutas a partir del año 2016. Nuevamen-



te, esta caída coincide con el cambio de gobierno y, más interesante aún, converge con una política comunicacional de un gobierno que buscó decididamente deslegitimar este formato de lucha por la vía de la criminalización. Por tomar un ejemplo, durante esos años el intendente de Mar del Plata –también de Juntos por el Cambio– interpuso una acción de amparo para que la Justicia intervenga ante los cortes de calles frente a la Municipalidad de General Pueyrredon³⁶.

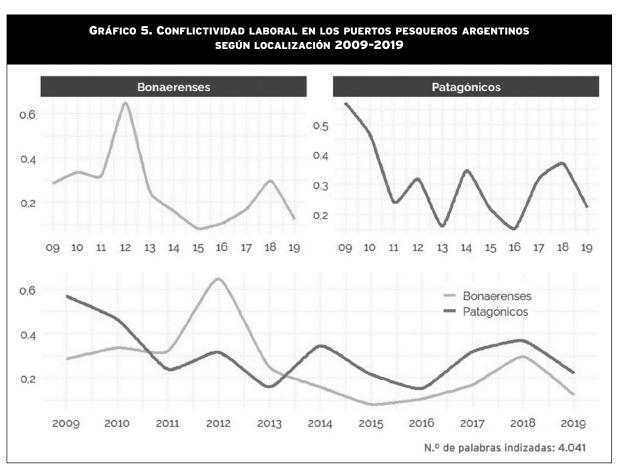
La geografía de la protesta en la industria pesquera

Las divergencias en las temporalidades no se acotan a los formatos de las acciones, también emergen si diferenciamos entre puertos patagónicos y puertos bonaerenses. Históricamente los puertos bonaerenses fueron los de mayor participación en los desembarques y en las exportaciones de pescado. Esta situación comenzó a cambiar hacia mediados de la década de 1970, cuando el gobierno

nacional incentivó, a través de políticas fiscales beneficiosas para las empresas del sector pesquero, su radicación en las provincias patagónicas. En un principio, los resultados no fueron los esperados, pero la situación comenzó a cambiar hacia los últimos años de la década de 1980 y se profundizó en los años noventa con el acuerdo firmado entre Argentina y la Comunidad Europea para la

incorporación de barcos españoles en el caladero argentino. Todas estas transformaciones no dejaron de impactar en las formas y los contenidos de la conflictividad laboral en el sector.

Mientras que en los puertos bonaerenses el año 2012 persiste en ser el de mayor conflictividad, en los puertos patagónicos el año 2009 emerge como el de mayor regis-

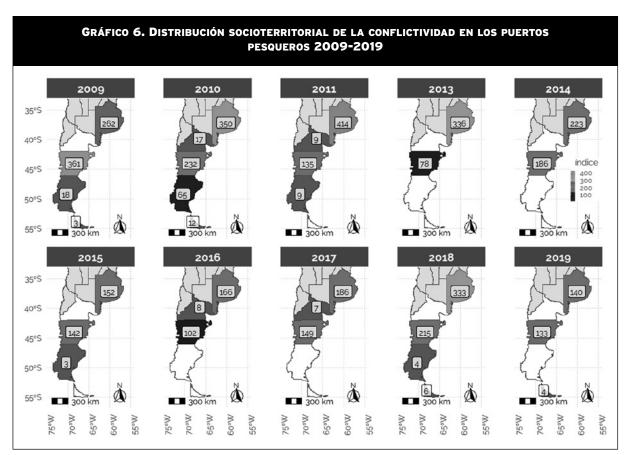


Fuente: elaboración propia a partir de datos de la Revista Puerto.

tro de conflictos. Esto no es un dato menor, ya que la mirada agregada sobre el conjunto de la conflictividad invisibilizaba la dinámica conflictual en los puertos patagónicos. De este modo, detectamos que 2009, 2010, 2014 y 2018 se presentan como años de alta conflictividad en los puertos al sur del paralelo 42º (ver gráfico 5).

Ahora bien, las siluetas divergentes entre Buenos Aires y Patagonia pueden desagregarse aún más en el territorio y que el contrapunto pase a ser entre provincias con puertos pesqueros.

El pasaje de la silueta del gráfico de líneas al mapa nos permite ver la frecu encia interanual según provincia. De esta forma, Chubut se presenta como la provincia pesquera más relevante de la Patagonia (ver gráfico 6). También se evidencia que 2009 es el único año en el que la provincia de Buenos Aires no encabeza la conflictividad



Nota: quitamos el año 2012 por su excepcionalidad.

sectorial: es el año en que Chubut lidera el ranking. Asimismo, el año 2010 destaca como el único año en el que la totalidad de las provincias con litoral marítimo participan de la conflictividad en el sector pesquero. Sin embargo, queda muy claro que las provincias más pesadas en desembarques y exportaciones son dos: Buenos Aires (Mar del Plata) y Chubut (Puerto Madryn). Finalmente, consideramos que el cruce entre esta información y la visualizada en la gráfico 5 permite identificar el pico represivo de 2009 con el pico de protesta ese mismo año en los puertos patagónicos. Es así que la impronta represiva del Estado sobre la protesta pesquera recae en el Gobierno de la provincia de Chubut.

Una aproximación a los sujetos de la protesta

En relación con las organizaciones participantes sucede algo similar a lo que ocurre con la geografía de la protesta. Su desagregación nos permite identificar temporalidades disímiles, lo que no es de extrañar si advertimos que los sindicatos de una y otra región difieren en su nomenclatura. Tanto el Sindicato Marítimo de Pescadores (SIMAPE) como el Sindicato Obrero de la Industria del Pescado (SOIP) son sindicatos con actuación acotada al Partido de General Pueyrredon, mientras que el Sindicato de Obreros Marítimos Unidos (SOMU) y el Sindicato de Trabajadores de Industrias de la Alimentación (STIA) son de alcance nacional, pero tienen mayor presencia en los puertos patagónicos. Por su parte, el Sindicato Argentino de Obreros Navales (SAON) y el Sindicato Unidos Portuarios Argentinos (SUPA) tienen incumbencia en ambas regiones, pero su peso específico en el entramado sindical es menor³⁷.

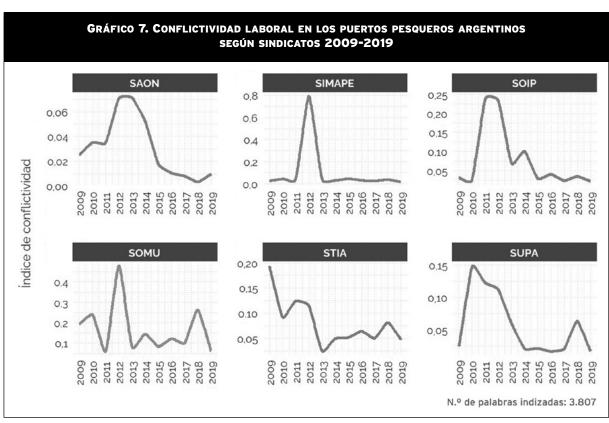
Las estelas que dibuja cada organización con su accionar no coinciden entre sí (ver gráfico 7). Como diferencia tajante emergen las relativas al STIA y al SUPA. Estas dos organizaciones sindicales posicionan a los años 2009 y 2010, respectivamente, como los de mayor conflictividad. También es destacable el pico del SOIP en el año 2011, que muy levemente le gana en intensidad al año 2012. Durante ese año, impulsada por el cuerpo de delegados y delegadas, el sindicato del fileteado de pescado protagonizó

la huelga más importante de la década. Sus reclamos se organizaron en torno a dos ejes: aumento salarial por encima del tope impuesto por el gobierno nacional y registración laboral bajo convenio colectivo de trabajo para el sector informal de las cooperativas fraudulentas de trabajo.

Un sondeo elemental de los motivos de las protestas

Los motivos de las protestas son diversos, en esta primera aproximación los agrupamos en cuatro categorías, tres positivas ("crisis y despidos", "laboral" y "salarial") y una residual ("otros"). Así distribuidos, los motivos dibujan siluetas levemente divergentes, con un punto alto común: el 2012. En términos de peso relativo, como era de esperar, las demandas más importantes son aquellas iniciadas por cuestiones salariales y laborales. Las acciones emprendidas por motivos de crisis y despidos, que tienen un peso relativo mucho más bajo, en su gran mayoría están asociadas con conflictos de menor alcance. Esto no se debe a que haya pocos despidos en el sector. La razón está en la vigencia, desde principios de los años noventa, de las cooperativas fraudulentas, que es una forma de contratación informal usada por los grandes grupos económicos del sector. Como el grueso de la fuerza de trabajo está contratada bajo esas condiciones, los despidos se invisibilizan, ya que las personas asociadas a estas cooperativas aparecen como "trabajadores por cuenta propia".

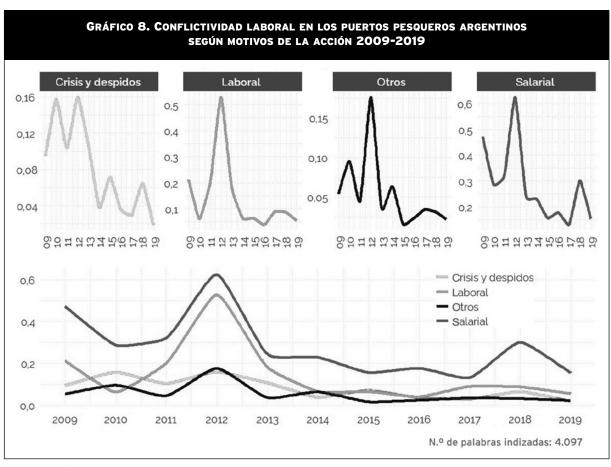
Pese a que las diferencias son leves, se puede ver cierta correspondencia entre los motivos laborales y salariales, por un lado, y entre los vinculados a situaciones de crisis, suspensiones y despidos y los motivos residuales, por otro lado (ver gráfico 8). Asimismo, en el tramo final de la silueta, 2016-2019, parece haber una mayor correspondencia entre "crisis y despidos" y "salarial", por un lado, y entre "laboral" y "otros", por otro lado. Finalmente, el leve pero perceptible aumento de la conflictividad en 2017 por motivos vinculados a reclamos laborales y a "otros" parece deberse al proceso de lucha iniciado por familiares de las víctimas del hundimiento de El Repunte.



tf-idf: otro enfoque para el análisis de la conflictividad

La métrica tf-idf (term frequency - inverse document frequency) es una técnica de procesado de lenguaje natural (PLN) que se usa para estimar el peso relativo de determinadas palabras dentro de un documento con relación a los otros documentos que conforman el corpus documental bajo análisis³⁸. En este apartado, usaremos esta métrica para detectar los tópicos más importantes de cada año. Para lograr este objetivo, agrupamos las notas

en doce documentos -uno por año-, estimamos el peso de las palabras y las listamos en orden decreciente para quedarnos con las primeras de la lista según la métrica tf-idf. De esta forma, nos aproximamos al contexto de las estimaciones realizadas en los apartados precedentes. Vale aclarar que las palabras no son preseleccionadas, no hay supervisión ni diccionarios, el algoritmo las ordena según su peso específico y solo intervenimos parta decidir cuál será el *top* de palabras que visualizaremos en el gráfico correspondiente.



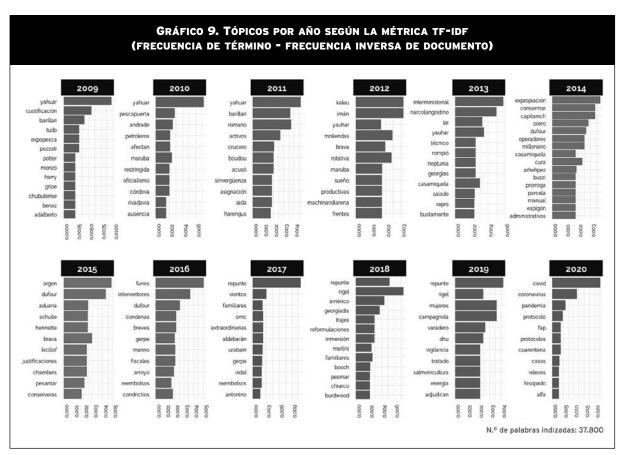
En primer lugar, aplicamos este procedimiento sobre el conjunto total de notas, sin diferenciar entre las notas que se refieren a conflicto y las que no. De esta forma, nos daremos una idea de cuáles fueron las temáticas noticiables más relevantes para la actividad pesquera argentina. En segundo lugar, replicaremos el mismo procedimiento, pero esta vez solo sobre el corpus referido a

conflictos. Este segundo ejercicio nos brindará información contextual sobre los conflictos más sobresalientes en cada año. En ambos casos, decidimos incluir el año 2020 para que la visualización sea más armoniosa.

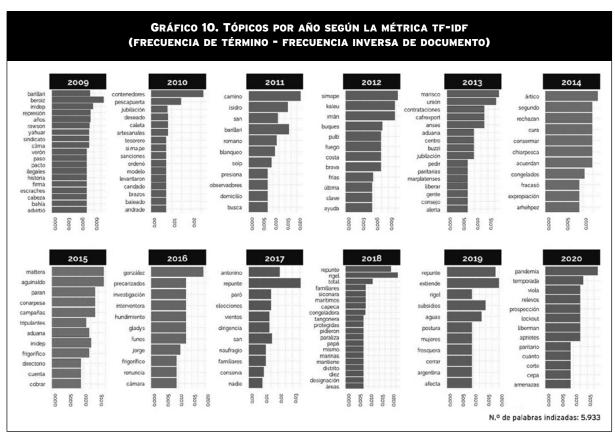
Como podrán comprobar en los siguientes gráficos (ver gráficos 9 y 10) los *rankings* están dominados por nombres propios, razón por la cual es posible identificar personas, lugares, organizaciones e instituciones con peso en los acontecimientos. También es posible identificar acciones y eventos de relevancia en cada uno de los años.

El gráfico nos permite contextualizar los conflictos más destacados en cada año a partir de la identificación de actores, lugares y motivos (ver gráfico 10). Por ejemplo, mientras que en 2011 y 2012 emergen, respectivamente, el SOIP y el SIMAPE como actores clave, entre 2017 y 2018 los actores claves fueron las mujeres y los familiares de la tripulación de los buques hundidos *El Repute, Rigel* y *San Antonino*.

Un enfoque un poco distinto, aunque seguimos dentro del PLN, es el referido a los n-gramas. Este enfoque permite trabajar con cadenas de palabras (dos o más). El ngrama más utilizado es el bigrama. Aquí lo usamos para



Fuente: elaboración propia a partir de datos de la Revista Puerto (todas las notas).



Fuente: elaboración propia a partir de datos de la Revista Puerto (solo notas referidas a conflictos).

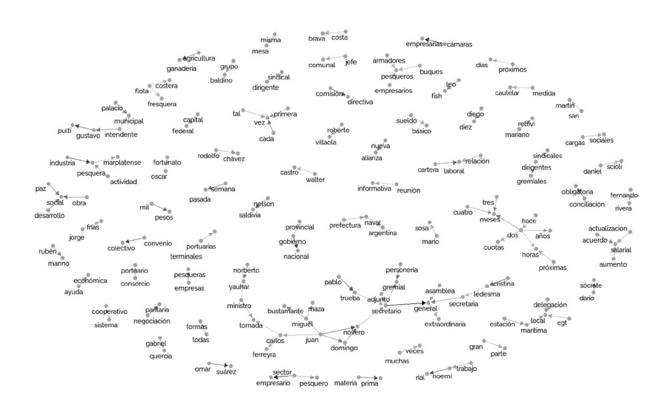
analizar con mayor detalle y profundidad el año más conflictivo de la serie: el 2012 (ver imagen 2).

La red de bigramas es una forma potente de complementar los análisis realizados con la métrica tf-idf. Al poner en relación dos o más palabras, logra una contextualización más acabada, como en la red visualizada arriba. Además de identificar, con este enfoque encontramos posibles relaciones ente sujetos y acciones en el mundo pesquero argentino de la última década.

Resultados y nuevas preguntas de investigación

Este primer avance sobre la conflictividad laboral en los puertos pesqueros de la Argentina desde un enfoque computacional deja un saldo más que prometedor. En primer lugar, pudimos reconstruir un indicador de conflictividad general para el período 2009-2019 que es posible ampliar hacia años anteriores y posteriores. Logramos identificar subperíodos con tasas medias de conflictivi-

Imagen 2. Red de bigramas para las noticias del año 2012 referidas a conflictos. (Fuente: elaboración propia a partir de datos de la Revista Puerto)



dad distintas y los años de mayor y menor actividad en términos de acciones de protesta. En segundo lugar, conseguimos desagregar y diferenciar la conflictividad general según tipo de acción emprendida. A partir de esta identificación, alcanzamos detectar temporalidades diversas. En tercer lugar, esta mirada fue complementada con una visualización de las geografías de la protesta. La georreferenciación de los índices de conflictividad nos permitió establecer dos polos de alta conflictividad, Puerto Madryn, en la Patagonia, y Mar del Plata, en la provincia de Buenos Aires. En cuarto lugar, la identificación de las organizaciones sindicales más importantes del sector nos habilitó una mirada atenta a las singularidades gremiales de cada región. Mientras que las acciones empren-

didas en los puertos patagónicos fueron dinamizadas por el STIA y el SOMU, los índices de conflictividad en los puertos bonaerenses fueron alimentados por el accionar del SIMAPE y el SOIP. En quinto lugar, los motivos de las acciones emprendidas desde distintos espacios pudieron ser identificados. Dentro de los motivos enunciados destacan los referidos a salarios y a cuestiones laborales por condiciones de trabajo en las cooperativas fraudulentas. En sexto lugar, nos acercamos cautelosamente a otros enfoques de lectura distante que permiten una contextualización más acabada, nos referimos a la métrica tf-idf y a la red de bigramas. Finalmente, si hay algo que estas notas lograron constatar es la extensa y constante conflictividad que dinamiza y da sentido a la trama profunda

de una actividad que perfuma su geografía con olor a pescado y a goma ardiendo...

Estos primeros resultados sobre la conflictividad laboral en los puertos pesqueros marítimos de la Argentina abren nuevos interrogantes: ¿a qué se debe la diferenciación en las dinámicas conflictuales entre los puertos patagónicos y los puertos bonaerenses? ¿Por qué algunas de las organizaciones sindicales, como el STIA, presentan una actividad conflictual divergente de la conflictividad general? ¿Cómo se articulan los movimientos huelguísticos con otros formatos de lucha como los cortes de calle v los bloqueos? ¿En qué momentos y de qué forma los colectivos obreros del sector lograron mayor unidad y rompieron las fronteras organizacionales y geográficas que los dividen? ¿Qué relación existe entre la estructura productiva de la industria pesquera y la conflictividad que la anima? ¿Quiénes son objetivo de las demandas iniciadas por los colectivos laborales en protesta? Estos interrogantes pretenden ser abordados en futuros avances de investigación que triangulen las técnicas computacionales usadas aquí con las técnicas tradicionales del oficio.

■ BALANCE PROVISORIO SOBRE LA LECTURA DISTANTE DEL CONFLICTO

En el aún poco explorado mundo de las humanidades digitales y las ciencias sociales computacionales las escasas incursiones sobre lectura distante para el análisis de la conflictividad social fueron encaradas casi con exclusividad desde una perspectiva sociológica. El enfoque histórico en ejercicios de lectura distante para el análisis de la conflictividad social es virtualmente inexistente. En parte, esta ausencia en la agenda historiográfica actual se debe al predominio del enfoque cualitativo en el tratamiento de los documentos históricos. También es verdad que la transitada dicotomía cualitativo/cuantitativo es cada vez más difícil de sostener y no aporta nada. Y es así porque son enfoques total y necesariamente complementarios en la investigación histórica, como lo son los enfoques clínicos y epidemiológicos en el ámbito de la salud. Números y palabras son de andar juntos, se mezclan porque se necesitan, y su separación solo se explica por nuestro irreductible afán clasificador. En cierta forma, el impacto del giro digital en nuestro oficio, con sus avanzados desarrollos informáticos y de programación, no solo no dificulta, sino que torna más fácil la intersección de enfoques, perspectivas, técnicas y fuentes.

Fue este estado de situación el que nos animó a arriesgar en solitario un ensayo en el cual buscamos interseccionar la historia social, entendida como la historia de las clases subalternas y sus luchas, con la historia digital, entendida como un campo en construcción que entrecruza cuestiones epistémicas propias de la historiografía con cuestiones tecnológicas propias del tratamiento computacional de grandes volúmenes de datos. Por suerte, en el mismo momento en que comencé a escribir este artículo, Jairo Melo desarrolló un seminario cuyo título está en completa sintonía con nuestra apuesta: "¿Historia Digital e Historia Social? Intersecciones posibles". El propósito del seminario era debatir las posibles intersecciones entre la historia digital y la historia social. Se preguntaba por las implicancias que la lectura distante puede tener sobre los fondos documentales de la historia social y por las posibilidades de datificar la historia social más allá de las experiencias cliométricas con datos económicos y demográficos. Los ejercicios desarrollados en estas páginas son un intento precario y exploratorio de responder los interrogantes compartidos por quienes se preocupan por ensayar este tipo de intersecciones.

Realizados los ejercicios, los primeros resultados asoman como alentadores. Por medio de la lectura distante de un corpus pequeño que no superaba las 8.000 notas periodísticas, encontramos regularidades, patrones, relaciones, similitudes, diferencias, "desvíos", "anomalías", "excepcionalidades" que nos ayudan a ensayar interpretaciones. En esa búsqueda, las técnicas de visualización de datos no son ni ornamentales ni ilustrativas, son una parte fundamental del análisis exploratorio³⁹.

Esperamos haber sido lo suficientemente convincentes como para que se animen a transitar estos cruces entre historia y programación.

■ NOTAS

* Agradecemos los comentarios y observaciones críticas realizadas por quienes hicieron las evaluaciones de la versión prelimi-



nar del presente artículo. También agradecemos los comentarios y las sugerencias que realizaron Nicolás Quiroga y Silvana Ferreyra.

- 1. Martin HILBERT i Priscila LÓPEZ, "The World's Technological Capacity to Store, Communicate, and Compute Information", en Science, 332.6025 (2011): 60-65. https://doi.org/10.1126/science. 1200970; Matthew J. SALGANIK, Bit by Bit: Social Research in the Digital Age (New Jersey, 2018).
- 2. Chris Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete", en *Wired* (2008). https://www.wired.com/2008/06/pb-theory/. (Consultado el 11 de diciembre de 2020).
- 3. Massimo PIGLIUCCI, "The End of Theory in Science?", en EMBO Reports, 10.6 (2009): 534. https://doi.org/10.1038/embor.2009.111.
- 4. Jairo MELO, "Lectura distante, fragmentada y colaborativa en el archivo infinito", en *Relaciones: Estudios de historia y sociedad*, 38.149 (2017): 169-89.
- 5. Nicolás QUIROGA, "Blogs de historia: usos y posibilidades", en *Historia Crítica*, 43 (2011): 62-80; Nicolás QUIROGA, "¿Qué tengo si no tengo papeles? Materialidad y juego en el trabajo de archivo", en *Población & Sociedad*, 25.2 (2019). https://cerac.unlpam.edu.ar/index.php/pys/article/view/3915. (Consultado el 22 de noviembre de 2020).
- 6. Anaclet Pons, El desorden digital: guía para historiadores y humanistas (Madrid, 2013).
- 7. Nicolás Quiroga, "El archivo y la toma de notas. El lugar del software en la interpretación histórica", en *Revista de Humanidades Digitales*, 2 (2018): 33-48. https://doi.org/10.5944/rhd.vol.2.2018.21936.
- 8. Montse HIDALGO PÉREZ, "Recetarios medievales y tarjetas perforadas: esta historiadora lleva 50 años creando el gran archivo culinario de Internet", en *La Nación* (Buenos Aires, 2020). https://www.lanacion.com.ar/tecnologia/recetarios-medievales-tarjetas-perforadas-esta-historiadora-lleva-nid2481360/. (Consultado el 12 de diciembre de 2020).
 - 9. Nicolás QUIROGA, "¿Qué tengo si no tengo papeles?".
- 10. COLABORADORES DE WIKIPEDIA, "Algoritmo", en Wikipedia, la enciclopedia libre, 2020. https://es.wikipedia.org/w/index.php?t itle=Algoritmo&oldid=131399938. (Consultado el 13 de diciembre de 2020). Vale aclarar que hacemos uso de Wikipedia porque, además de considerarla como una enciclopedia tan o más válida que otras, junto con el verbo googlear es una de las marcas más notables del giro digital.
- 11. Nan Z. DA, "The Computational Case against Computational Literary Studies", en *Critical Inquiry*, 45.3 (2019): 601-639. https://doi.org/10.1086/702594; Franco MORETTI, *Lectura distante* (Buenos Aires, 2016).
- 12. Franco Moretti, "Más conjeturas sobre la literatura mundial", en New Left Review, 20 (2003): 83-92.

- 13. Martin MUELLER, "Scalable Reading", en Scalablereading, 2012. https://scalablereading.northwestern.edu/?page_id=22. (Consultado el 13 de diciembre de 2020).
- 14. El NER "es una tarea de extracción de información que busca localizar y clasificar en categorías predefinidas, como personas, organizaciones, lugares, expresiones de tiempo y cantidades, las entidades nombradas encontradas en un texto". COLABORADORES DE WIKIPEDIA, "Reconocimiento de entidades nombradas", en Wikipedia, la enciclopedia libre, 2020. https://es.wikipedia.org/w/index.php?title=Reconocimiento_de_entidades_nombradas&oldid=128832539. (Consultado el 13 de diciembre de 2020).
- 15. R es un lenguaje y un entorno para gráficos y computación estadística. COLABORADORES DE WIKIPEDIA, "R (lenguaje de programación)", en Wikipedia, la enciclopedia libre, 2020. https://es.wikipedia.org/w/index.php?title=R_(lenguaje_de_programaci%C3%B3n)&oldid=128000324. (Consultado el 13 de diciembre de 2020).
- 16. Ramiro Albrieu i Gabriel Palazzo, "Categorización de conflictos sociales en el ámbito de los recursos naturales: un estudio de las actividades extractivas mediante la minería de textos", en *Revista CEPAL*, 131 (2020): 29-59; Roberto Franzosi, "What's in a Text? Bridging the Gap between Quality and Quantity in the Digital Era", en *Quality & Quantity* (2020). https://doi.org/10.1007/s11135-020-01067-6; Hai Liang i Jonathan J. H. Zhu, "Big Data, Collection of (Social Media, Harvesting)", en *The International Encyclopedia of Communication Research Methods* (New Jersey, 2017), 1-18. https://doi.org/10.1002/9781118901731.iecrm0015.
- 17. Un *package* (paquete) es un conjunto de funciones y métodos computacionales empaquetados.
- 18. Hadley WICKHAM, *Rvest: Easily Harvest (Scrape) Web Pages*, version 0.3.6, 2020. https://CRAN.R-project.org/package=rvest. (Consultado el 13 de diciembre de 2020).
- 19. Kirill MÜLLER i Hadley WICKHAM, *Tibble: Simple Data Frames*, version 3.0.4, 2020. https://CRAN.R-project.org/package=tibble. (Consultado el 13 de diciembre de 2020).
- 20. Garrett GROLEMUND i Hadley WICKHAM, *R para ciencia de datos* (github: R4DS, 2019). http://es.r4ds.hadley.nz/. (Consultado el 13 de diciembre de 2020).
- 21. Hadley WICKHAM, *Tidyverse: Easily Install and Load the 'Tidyverse'*, version 1.3.0, 2019. https://CRAN.R-project.org/package=tidyverse. (Consultado el 13 de diciembre de 2020).
- 22. Hadley WICKHAM et al., *Ggplot2*: Create Elegant Data Visualisations Using the Grammar of Graphics, version 3.3.2, 2020. https://CRAN.R-project.org/package=ggplot2. (Consultado el 13 de diciembre de 2020).
- 23. James SMITHIES, "Full Stack DH: Building a Virtual Research Environment on a Raspberry Pi", en *Making Things and Drawing Boundaries*. *Experiments in the Digital Humanities*, Jentery Sayers ed., (Minneapolis, 2017), 102-114. https://doi.org/10.5749/j.cttlpwt6wg.13.

- 24. Roberto Franzosi, "The Press as a Source of Socio-Historical Data: Issues in the Methodology of Data Collection from Newspapers", en *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 20.1 (1987): 5-16. https://doi.org/10.1080/01615440.1987.10594173.
- 25. Agustín NIETO, "Industria pesquera y mundo obrero", en *Historia de la provincia de Buenos Aires. Del primer peronismo a la crisis de 2001*, Osvaldo Barreneche ed. (La Plata, 2014).
- 26. SISMOS, "La conflictividad social en clave local. Un ejercicio de ponderación sobre los hechos de rebelión en la Mar del Plata actual (Argentina, 2011-2016)", en *Clivatge. Estudis i testimonis sobre el conflicte i el canvi socials*, 7 (2019): 228-285.
- 27. José MATEO "De espaldas al mar: la pesca en el Atlántico Sur (siglos XIX y XX)" (tesis doctoral, Universitat Pompeu Fabra, 2004). https://dialnet.unirioja.es/servlet/tesis?codigo=141556. (Consultado el 13 de diciembre de 2020).
- 28. Eduardo «Maro» PRADAS, Un acercamiento a la problemática pesquera marplatense (Buenos Aires, 2006): 26.
- 29. Elisa PASTORIZA, Los trabajadores de Mar del Plata en vísperas del peronismo (Buenos Aires, 1993).
- 30. Irene Delfina MOLINARI, "El rol del sindicato y las experiencias colectivas de las trabajadoras de la conserva de pescado. Mar del Plata, 1940/1983", en *Il Seminario Mercosur* (Mar del Plata, 1997).
- 31. Agustín NIETO, Entre anarquistas y peronistas. Historias obreras a ras del suelo (Buenos Aires, 2018).
- 32. Gonzalo PÉREZ ÁLVAREZ, "Aunque parezca, la red no está vacía. Luchas de los obreros pesqueros del noreste de Chubut, 1990-2005", en *Revista de Estudios Marítimos y Sociales*, 2 (2009): 171-183.
- 33. María Luciana NOGUEIRA, "El papel de las luchas previas en la trayectoria de una empresa pesquera recuperada por sus trabajadores (Necochea, 2004-2011)", en *Archivos de historia del movimiento obrero y la izquierda*, 15 (2019): 119-141. https://doi.org/10.46688/ahmoi.n15.82.
- 34. Guillermo Colombo, "Hasta que el recurso nos falló... Crisis de la merluza y protesta obrera. La dinámica de los enfrentamientos en el puerto marplatense (1997-2002)" (tesis de licenciatura, UNMdP, 2006); María Soledad Schulze, "Representaciones morales y grados de participación en el conflicto de 2007 en la industria pesquera de Mar del Plata, Buenos Aires, Argentina", en Jangwa Pana, 19.3 (2020). https://doi.org/10.21676/16574923.3518; Diego Adrián Solimeno i Gonzalo Julián Yurkievich, "Sobran cuchillos, falta el pescado. Auge del langostino patagónico y reactivación del conflicto social en Mar del Plata", en Geograficando, 16.2 (2020). e076-e076. https://doi.org/10.24215/2346898Xe076; Gonzalo Yurkievich i Agustín Nieto, "Organización y lucha obrera en la Argentina contemporánea: La experiencia del Sindicato Marítimo de Pescadores (2000-2012)", en Trabajo y Sociedad, 23 (2014): 19.

- 35. María Luciana NOGUEIRA i María Soledad SCHULZE, "Conflictividad sociolaboral y recuperación de empresas pesqueras en Argentina, Necochea/Quequén (2010-2012)", en Conflicto Social, 11.19 (2018): 188-215; Gonzalo PÉREZ ÁLVAREZ i María Soledad SCHULZE, "Hacia una comparación de la conflictividad pesquera en la costa bonaerense y patagónica en Argentina: de la huelga del 2005 en Chubut a la del 2007 en Mar del Plata", en Anuario de La Escuela de Historia, 33, (2020): 1-30.
- 36. LA CAPITAL, "Presentan acción de amparo contra los piquetes frente al municipio", en *Diario La Capital de Mar del Plata* (2017). https://www.lacapitalmdp.com/presentan-accion-de-amparocontra-los-piquetes-frente-al-municipio/. (Consultado el 16 de diciembre de 2020).
- 37. Agustín NIETO, "Industria Pesquera y Mundo Obrero".
- 38. COLABORADORES DE WIKIPEDIA, "Tf-idf", en Wikipedia, la enciclopedia libre, 2019. https://es.wikipedia.org/w/index.php?title=Tf-idf&oldid=121943707. (Consultado el 18 de diciembre de 2020).
- 39. Marcelo RAIMUNDO, "Visualización y animación de tablas históricas con R", en *The Programming Historian*. https://programminghistorian.org/ (en prensa).

