
TRIANGULACIÓN METODOLÓGICA Y BIG DATA

TRIANGOLAZIONE METODOLOGICA E BIG DATA

TRIANGULAÇÃO METODOLÓGICA E BIG DATA

METHODOLOGICAL TRIANGULATION AND BIG DATA

Paolo Parra Saiani¹<https://orcid.org/0000-0003-0227-4411>**Juan Ignacio Piovani²**<https://orcid.org/0000-0002-8774-3258>

Recibido em: 11/09/2021

Aceito em: 21/09/2021

RESÚMEN: En este artículo se abordan las perspectivas de la triangulación metodológica en relación con los nuevos métodos computacionales para el análisis de los big data. En primer lugar, se revisa la apropiación del término triangulación por parte de las ciencias sociales, sus diferentes significados y su lugar en el debate entre métodos cuantitativos y cualitativos. Por otra parte, se reflexiona sobre el futuro de la triangulación metodológica a partir del auge de los big data y de la sociología computacional, y del debate entre métodos “viejos” y métodos “nuevos” al que ha dado lugar recientemente. Se toman como referencia aspectos críticos sobre la producción, recolección, interpretación y transparencia de los big data y se afirma que, más allá del curso que sigan los nuevos debates metodológicos y la investigación computacional con big data, nuevas formas de triangulación entre este tipo de investigación y los métodos cuantitativos y cualitativos clásicos de la sociología podrían ser importantes para enriquecer las perspectivas metodológicas futuras de las ciencias sociales.

Palabras clave: triangulación; big data; ciencias sociales; metodología

RIASSUNTO: Questo articolo affronta le prospettive della triangolazione metodologica in relazione ai nuovi metodi computazionali per l'analisi dei big data. In primo luogo si passa in rassegna l'appropriazione del termine triangolazione da parte delle scienze sociali, le sue diverse accezioni e la sua collocazione nel dibattito tra metodi quantitativi e qualitativi. Dall'altro, si riflette sul futuro della triangolazione metodologica basata sull'ascesa dei big data e della sociologia computazionale, e sul conseguente dibattito tra "vecchi" e

¹ Doctor en Sociología (Università di Trieste); Profesor Asociado en Sociología en el Departamento de Ciencias Políticas de la Universidad de Génova; Director del Programa en Ciencias Políticas del Doctorado en Ciencias Sociales; Coordinador de la Laurea Magistrale in Amministrazione e politiche pubbliche de la Universidad de Génova. Email: paolo.parra.saiani@unige.it

² Doctor en Metodología de las Ciencias Sociales (Sapienza Università di Roma), MSc. Advanced Social Research Methods and Statistics (City University London), Investigador Principal de CONICET, Profesor titular en la Facultad de Humanidades de la Universidad Nacional de La Plata, Coordinador de la Red Latinoamericana de Metodología de las Ciencias Sociales, Director del Doctorado en Estudios Sociales Interdisciplinarios de Europa y América Latina (Universidad Nacional de La Plata / Universität Rostock). Email: juan.piovani@presi.unlp.edu.ar

"nuovi" metodi che ha originato. Vengono presi a riferimento gli aspetti critici della produzione, raccolta, interpretazione e trasparenza dei big data e si afferma che, al di là dello sviluppo dei nuovi dibattiti metodologici e la ricerca computazionale con i big data, nuove forme di triangolazione tra questo tipo di ricerca e i metodi quantitativi e qualitativi classici della sociologia potrebbero essere importanti per arricchire le future prospettive metodologiche delle scienze sociali.

Parole Chiave: triangolazione; big data; scienze sociali; metodologia

RESUMO: O artigo aborda as perspectivas de triangulação metodológica em relação à os novos métodos computacionais para análise da Big Data. Em primeiro lugar, o artigo revisita a apropriação do termo triangulação pelas ciências sociais, seus diferentes significados e seu lugar no debate entre métodos quantitativos e qualitativos. Em segundo, reflete sobre o futuro da triangulação metodológica baseada no surgimento da Big Data e da sociologia computacional, e o debate entre “velhos” e “novos” métodos que recentemente apareceu. Aspectos críticos da produção, coleta, interpretação e transparência da Big Data são tomados como referência. A conclusão é que além do curso que os novos debates metodológicos e pesquisas computacionais com Big Data seguem, novas formas de triangulação entre este tipo de pesquisa e métodos quantitativos e qualitativos clássicos na Sociologia podem ser importantes para enriquecer as futuras perspectivas metodológicas nas Ciências Sociais.

Palavras-Chave: triangulação; Big Data; Ciências Sociais; Metodologia.

ABSTRACT: This article addresses the perspectives of methodological triangulation in relation to the new computational methods for the analysis of big data. In the first place, the appropriation of the term triangulation by the social sciences, its different meanings and its place in the debate between quantitative and qualitative methods is reviewed. On the other hand, it reflects on the future of methodological triangulation based on the rise of big data and computational sociology, and the debate between "old" and "new" methods that has recently given rise to. Critical aspects of the production, collection, interpretation and transparency of big data are taken as a reference. The conclusion is that, beyond the course that the new methodological debates and computational research with big data follow, new forms of triangulation between this type of research and classical quantitative and qualitative methods of sociology could be important to enrich the future methodological perspectives of the social sciences.

Keywords: triangulation; big data; social sciences; methodology

El término triangulación, de larga trayectoria en la geometría, la geodesia y otras disciplinas científicas, se introdujo en la literatura metodológica de las ciencias sociales en la década de 1960. En efecto, de acuerdo con los textos dedicados a la historia de los métodos de investigación social, este término se usó por primera vez en el ya célebre libro de *Webb, Campbell, Schwartz y Sechrest* titulado *Unobtrusive measures: Nonreactive research in the social sciences*, publicado en 1966.

Si en su campo científico de origen triangulación “refiere tanto a la subdivisión de una superficie o polígono plano en un conjunto de triángulos (...), como a la aplicación de la trigonometría de triángulos para determinar la posición de puntos, a través de métodos de intersección y resección, y para calcular distancias o superficies” (Piovani, 2018: 438), en su primera apropiación en las ciencias sociales, y por analogía, el término se utilizó para designar un procedimiento empleado en el proceso de medición, orientado a confirmar una

proposición mediante “dos o más mediciones independientes” (*Webb et al. 1966: 3*). De acuerdo con estos autores, la triangulación permitiría reducir la ambigüedad interpretativa del resultado de una medición.

Cabe señalar, sin embargo, que incluso antes de la introducción del término, el concepto de triangulación, aunque expresado con otras palabras, tenía ya cierta raigambre en las ciencias sociales. En este sentido, y tal como señalan Jick (1979) y Knafl y Breitmayer (1989) la apropiación del concepto antecede a la adopción del término. En la década de 1950, autores como Garner (1954); Campbell (1956); Garner, Hake y Eriksen (1956) y Campbell y Fiske (1959), habrían recurrido al concepto de triangulación, al menos en un sentido metafórico, cuando propusieron la convergencia de mediciones independientes con el fin de dar cuenta de un mismo constructo, retomando una vieja idea desarrollada a partir del siglo XVIII en el marco de la tradición astronómica de la *error theory* (Piovani, 2006). Pero en lugar de emplear el término triangulación, se refirieron a su propuesta metodológica con expresiones tales como delineación operacional (*operational delineation*), operacionismo múltiple (*multiple operationism*), validación convergente (*convergent validation*) y operacionismo convergente (*convergent operationism*).

Más allá de estas consideraciones históricas sobre el término triangulación, su introducción en las ciencias sociales y su significado originario, en la actualidad es inevitable entenderlo en el contexto del debate entre métodos cualitativos y cuantitativos. Esto se debe a que, como resultado de los habituales procesos de estiramiento semántico y cambio conceptual, el término fue adquiriendo diferentes significados y connotaciones a lo largo del tiempo y, de todos ellos, posiblemente el de triangulación metodológica –o más específicamente triangulación inter método– sea el que más pregnancia haya tenido, al punto de ser el más usado en la literatura actual de las ciencias sociales.

La idea de triangulación metodológica fue propuesta por Denzin (1970), quien sugirió que la triangulación no se limitaba a su sentido original de convergencia de datos y mediciones, sino que podía aludir también a la de investigadores, teorías y métodos. Para este autor, la triangulación metodológica abarca la articulación de métodos cuantitativos y cualitativos, o triangulación inter-método (between-method), y la triangulación intra-método (within-method), una categoría retomada por Jick (1979) para dar cuenta del uso conjunto de variantes de un mismo método –por ejemplo, entrevistas y observación participante en el caso cualitativo– en la misma investigación social.

Se puede argumentar que la triangulación metodológica cobra sentido y condiciones de posibilidad a la luz de lo que conoce como argumentos técnicos del debate cualitativo vs cuantitativo, que adquirieron preminencia luego de una primera fase caracterizada por la centralidad de los argumentos epistemológicos. En efecto, tal como señala Bryman (1988), este clásico debate metodológico atravesó distintas fases a lo largo del tiempo. Cuando comienza a configurarse, en el marco de la crisis del consenso ortodoxo de la sociología (Giddens, 1979), los principales argumentos que se esgrimían para justificar las diferencias entre ambos métodos eran de carácter epistemológico. Partiendo del concepto kuhneano de paradigma, la idea dominante en ese momento histórico con respecto a los debates metodológicos de las ciencias sociales, era que éstas se diferencian de las ciencias naturales por la coexistencia de paradigmas, y que ellos –a su vez– implican diversas formas de resolver cuestiones clave –acerca de la ciencia, la sociedad, etc.– en las que se fundan diversas formas de entender, producir, justificar y validar el conocimiento sociológico.

Esta perspectiva concebía a los métodos cualitativos y cuantitativos como antagónicos,

consecuencia lógica del supuesto según el cual ellos representan las alternativas metodológicas de paradigmas inconmensurables. No obstante, con el paso del tiempo estos argumentos fueron cediendo frente a los de carácter técnico, según los cuales la pertinencia de los métodos cuantitativos y cualitativos debe analizarse en relación con el tipo de problema de investigación que se postula y el tipo de respuesta cognoscitiva que se pretende alcanzar.

Una deriva de esta nueva mirada técnica sobre el debate cualitativo vs cuantitativo fue la transición, en el campo de la metodología de las ciencias sociales, desde lo que Pawson (1994) denomina “guerra paradigmática” hacia un “pluralismo pragmático”. Pero toda vez que los métodos ya no se conciben como antagónicos, y que a cada uno se le atribuyen ventajas y desventajas relativas diferentes, dicho pluralismo también abrió la puerta a las propuestas de investigación que apelan a la combinación de métodos, que en la literatura especializada suelen rotularse como métodos mixtos, y una de cuyas formas particulares es la triangulación inter metodológica.³ En esta línea, y tal como señalan Knafl y Breitmayer (1989) se ha propuesto la triangulación como forma de convergencia y de completitud, ya sea como medio de validación convergente o como una vía para alcanzar una comprensión más completa del fenómeno bajo investigación, que se aborda desde diferentes perspectivas metodológicas.

Entre finales del siglo XX y principios del XXI, e incluso en gran medida aun en la actualidad, el pluralismo pragmático y los criterios técnicos del debate cualitativo vs cuantitativo que le sirven de fundamento fueron moldeando los saberes metodológicos estándar que se reflejan en los manuales especializados y que se enseñan en los cursos universitarios de metodología. Los métodos mixtos y la triangulación, entendida como forma de articulación entre métodos cuantitativos y cualitativos, también forman parte del repertorio actual de estrategias y enfoques metodológicos ampliamente legitimados, a pesar de no haber estado exentos de críticas.

Pero en la última década, la irrupción de nuevos métodos de investigación computacionales basados en los big data han comenzado, para algunos autores, a alterar este panorama. Esto se debe a sus implicancias para las formas de concebir y practicar la investigación social, en general, y la triangulación metodológica y los métodos mixtos, en particular. En efecto, luego de 2009, con la publicación en la revista *Science* del célebre artículo de Lazer et al. *Life in the network: the coming age of Computational Social Science*, el fenómeno de los Big Data dio pie al desarrollo de una nueva ciencia social cuantitativa (Conte, 2016). En esta misma línea, Burrows y Savage (2014) puntualizan que a partir de 2011 hemos sido testigos de una verdadera explosión de interés en el tema.

En sus orígenes, el término big data daba cuenta de la inmensa cantidad de datos generados ininterrumpidamente gracias a las nuevas tecnologías digitales, y producidos en diversos ámbitos: interacciones en las redes sociales, comercio electrónico, transacciones financieras, trámites administrativos y gubernamentales, etc. (Piovani, 2018). En la actualidad, big data no solo alude a esta enorme masa de datos, sino también a ciertas cualidades vinculadas a las “huellas digitales” que dejan actividades rutinarias de la vida cotidiana (Burrows y Savage, 2014) y que han llamado la atención de los investigadores sociales porque permitirían acceder a las vidas y estilos de vida de millones de personas en tiempo real y a lo largo del tiempo (Halford y Savage, 2017)

El tema de los big data presenta algunos desafíos desde diferentes puntos de vista: desde el lado de la producción, de su recolección y organización, su interpretación y –finalmente– su

³ Para un análisis más detallado de los puntos de contacto y las diferencias entre métodos mixtos y triangulación metodológica véase Piovani (2018).

transparencia.

i) En relación con la *producción*, cabe subrayar que en la literatura se afirma que con los big data se superaría el problema de la escasa fiabilidad de las informaciones, ya que la información no es solicitada por el investigador. En cambio, se pueden utilizar las “huellas digitales” (Golder, Macy, 2014) que dejan los que utilizan redes sociales, tarjetas de crédito, teléfonos móviles, etc. Pero, si bien el uso de una tarjeta, la posición detectada por el GPS, o los datos de tráfico urbano se pueden considerar como datos fiables (no manipulables por parte del usuario), no se puede afirmar lo mismo de los tweets y –más en general– de los mensajes, fotografías y video que se comparten en Twitter, Facebook, TikTok y otras redes sociales. En cualquier lugar, ya sea físico o virtual⁴, las personas utilizan “máscaras”: ellas deciden qué publicar, eligiendo cuidadosamente qué mostrar y qué no mostrar (Martire y Pitrone, 2016). En lugar de “huellas” digitales deberíamos hablar de grafitis (Parra Saiani, 2017): signos dibujados deliberadamente por aquellos que quieren dejar un mensaje. Hay que añadir que existe una tendencia a estudiar lo que ya está disponible, lo “*ready-made*”, que da pie a investigaciones que ya no son *theory-driven*, sino *data-driven*. En este caso la elección del tipo de información, o el tipo de técnica a utilizar, no tiene un soporte teórico, sino que se basa en la disponibilidad de información sin costo.⁵

ii) En lo que concierne a la *recolección de los datos*, hay un enorme problema de representatividad: los usuarios de las redes sociales no representan cabalmente a la población general, e incluso entre dichos usuarios existen muchas diferencias con respecto, por ejemplo, al nivel educativo y la edad. El problema de la fiabilidad de la información es aún más relevante si tenemos en cuenta los cambios en los algoritmos de Google (86 cambios sólo entre junio y julio de 2012), o en las plataformas Twitter o Facebook, que a menudo hacen que sea imposible comparar la información obtenida tan solo un año después, por no hablar de la manipulación fraudulenta con fines comerciales o financieros (Lazer et Al. 2014:1204).

iii) En cuanto a la *interpretación*, los big data presentan el problema que Dormagen llama “riesgo de regresión sociológica” (2016: XVI): se ocultan las características sociales de la participación en las redes sociales ya que hay muy poca –¿ninguna?– información sobre quiénes son los productores de tweets (Bouillier, 2015: 810; Boyadjian, 2016: 41). Es imposible situar a los productores de opiniones digitales (numéricas) en el espacio social, ya que el análisis se focaliza en el flujo de información en detrimento de los productores de estas publicaciones, impidiendo cualquier análisis según criterios sociodemográficos. No obstante, en los últimos años se han desarrollado técnicas que permiten situar socialmente las opiniones que circulan en la red sin recurrir a las clásicas herramientas de la entrevista, y poder así aprovechar el valor agregado de la web: la observación de mensajes en tiempo real y en gran número.

iv) El último punto –la transparencia– es particularmente importante cuando hablamos de integración de las técnicas: la información no se construye directamente con las personas, hablando, entrevistando, observando, etc., sino que es detectada y organizada por parte de instituciones/empresas que la recogen para sus fines y que decidirán si, cómo y cuándo estará disponibles para fines de investigación. Estamos frente a verdaderos *gatekeepers*, que en lugar de filtrar la distribución del *output* (Hirsch 1972) tienen la facultad de regular el acceso a la

⁴ Es interesante como muchas veces ‘virtual’ está considerado como opuesto de ‘real’, como si las interacciones que tienen lugar en Facebook tuvieran lugar en un mundo diferente al nuestro.

⁵ Cabe agregar que, muy a menudo, se trata de una elección forzada por recortes en fondos públicos destinados a la investigación, que afectan selectivamente a las universidades, y que acentúan la brecha entre universidades ricas y pobres

información. boyd y Crawford (2012: 674-675) señalan que en el sector privado algunas empresas limitan el acceso a todos los potenciales usuarios externos, otras venden los derechos y otras ofrecen la opción de una búsqueda limitada para la investigación académica, lo que produce una injusticia sustancial en el sistema, vinculada a los desequilibrios ya existentes en la asignación de recursos a disposición de las diferentes universidades (Lazer et Al. 2009: 721; Conley et Al. 2015; Strong 2015: 134).

La mayoría de los big data no son públicos: una excepción es Twitter, que hace accesible la información a cualquier persona a través de la interfaz de los canales dedicados (*application programming interface*, API), que permite el acceso a una pequeña selección de tweets, a una muestra de 10% (garden-hose), o a todo el corpus (firehose)⁶. Pero, como se señaló, Twitter es una excepción en el paisaje de grandes volúmenes de datos celosamente guardados, y por lo tanto no debería sorprender que haya despertado tanto interés entre los científicos sociales: entre 2008 y 2012 se produjeron más de 110 publicaciones científicas sobre Twitter (Tinati et Al. 2014: 664).

Por supuesto, la posesión de información por parte de entidades que no tienen ninguna obligación de publicarlas hace aún más difícil y problemática la transparencia y la replicación de los estudios, uno de los pilares de la práctica científica. Si la ciencia es una actividad acumulativa, y si todos estamos “parados sobre hombros de gigantes”, necesitamos saber cuáles son las bases sobre las que se está construyendo, para no caer al suelo a causa de hombros que alguien hizo resbaladizos. Para contrarrestar la progresiva "privatización" de la información relevante han surgido muchas iniciativas, como el programa federal estadounidense *Open Government Initiative*, que prevé la obligación de compartir los datos como una cláusula obligatoria en muchas propuestas de financiación. Otras iniciativas del sector público y privado, revistas y fundaciones, han dado lugar a la disponibilidad sin precedentes de datos para el análisis secundario (Shaikh et al. 2014).

¿Cómo garantizar, entonces, un nivel mínimo de credibilidad para la investigación científica, para que no “promueva” intereses particulares, o se base en consideraciones de conveniencia política o económica? En el contexto de los problemas empírico-científicos, el control intersubjetivo protegería –al menos parcialmente– de distorsiones y prejuicios de naturaleza extra científica. De ahí la importancia de la transparencia de los procedimientos adoptados por el investigador y de la controlabilidad del proceso de formulación de los enunciados. Las ciencias naturales, las más formalizadas –como la física, la química, etc.– respondieron parcialmente a la necesidad de transparencia con el criterio de replicabilidad, es decir, con la posibilidad de repetir un experimento obteniendo los mismos resultados que el primero. Pero esto solo es posible bajo ciertas condiciones: posibilidad de controlar factores externos, posibilidad de manipular una propiedad y observar sus efectos sobre la propiedad experimental, posibilidad de realizar mediciones precisas, posibilidad de descartar razonablemente los efectos de propiedades no controladas en la situación experimental. Se asume una relación positiva entre replicación y científicidad, en el sentido de que los estudios con múltiples repeticiones que producen resultados idénticos (o muy similares) tienden a considerarse más robustos desde un

⁶ Cómo recuerdan Tinati et al., la elección de analizar un subconjunto así determinado del corpus disponible compromete las cualidades que hacen a los mismos big data tan interesantes, en particular, el tamaño y dinamismo: «Rendering Big Data manageable in this way overrides its nature as ‘big’ data, bypassing the scale of the data for its availability or imposing an external structure by sampling users or tweets according to a priori criteria, external to the data themselves» (2014: 665).

punto de vista científico. Por lo tanto, se considera que las investigaciones basadas en procedimientos impersonales son más fiables porque el resultado no dependerá de quién realice el experimento (Piovani, 2015: 222). La transparencia es la condición necesaria para la replicabilidad: solo un estudio adecuadamente descrito en sus procedimientos puede ser replicado por otro investigador. La replicabilidad ha sido definida como el "principio ordenador de la ciencia" (Collins, 1985: 2), la Corte Suprema del sistema científico (ibid.: 20), pero también –por este mismo autor– como un principio que se proclama con mayor frecuencia que lo que se lo practica. Son muy pocos los casos de investigaciones repetidas por el mismo investigador y, menos aún, por otros, ya que la ciencia reserva los máximos honores a quien hace algo primero, y la confirmación del trabajo de otro solo confirma que es el trabajo de otros (Collins, 1985: 19). Además, es difícil conseguir financiación por no descubrir nada nuevo y, aún más difícil, que las revistas publiquen el artículo resultante.

En las ciencias sociales la replicación de un estudio plantea aún más interrogantes, a menos que nos limitemos a un significado muy restringido de replicabilidad, entendida como la repetición del análisis de datos teniendo disponible la matriz original. Por un lado, las ciencias sociales comparten algunos problemas con las ciencias naturales, por ejemplo, el impulso a publicar tanto como sea posible, la inversión creciente de tiempo para procurar financiamiento, la cada vez mayor carga de trabajo administrativo y burocrático que quita tiempo a la investigación. Por otro lado, enfrentan la inevitable mutabilidad humana, lo que significa que una persona puede cambiar de opinión incluso varias veces durante el mismo día. Entonces, al replicar la administración de un cuestionario, o al repetir una entrevista...: ¿cómo podemos estar seguros de que el entrevistado no ha cambiado de opinión de un mes a otro, de una semana a otra, de un día a otro, de un momento a otro? Ante la imposibilidad de ajustarse en sentido estricto al principio de replicabilidad, las ciencias sociales se han apoyado en un principio de orden menor, la controlabilidad (Guba, 1978: 71) o inspeccionabilidad (Ricolfi, 1995: 394): la posibilidad, por parte de un tercero (un lector, otro investigador), de controlar las elecciones realizadas por el investigador durante su trabajo. Becker escribió: "Debemos usar nuestras técnicas de tal manera que se pueda probar que una opinión a la que estamos particularmente apegados no es cierta. Siempre debemos inspeccionar nuestro trabajo con suficiente cuidado para saber si nuestras técnicas y teorías están lo suficientemente abiertas para permitir esta posibilidad" (1967: 246).

Pero no toda la investigación social se funda en bases empíricas inspeccionables: algunos tipos pueden inspeccionarse en su totalidad (si se basan en un corpus textual) o en parte (si se basan en una matriz de datos o en la transcripción de entrevistas), pero para otros tipos los márgenes de inspección son extremadamente estrechos (por ejemplo, las notas de campo de la investigación etnográfica). Sin embargo, el criterio de la posibilidad de inspección o control como vía hacia una mayor formalización del proceso de investigación no es aceptado por unanimidad. Existen muchas propuestas alternativas que se han consolidado a lo largo del tiempo: fiabilidad y confirmabilidad (Lincoln y Guba, 1985); sinceridad, verdad subjetiva, verdad histórica y verdad ficticia (Denzin, 1989: 23); persuasión, coherencia, correspondencia y pragmatismo (Riessman, 1993: 64-9). No obstante, aunque la terminología, y a menudo también los supuestos epistemológicos varíen, "la naturaleza argumentativa de la ciencia" (Campelli, 1991) resulta transversal a todas estas perspectivas. Esto requiere no solo argumentar las elecciones del investigador sino, sobre todo, proporcionar a la comunidad de expertos herramientas de evaluación. Aquí surge la pregunta: ¿cómo se puede garantizar la naturaleza argu-

mentativa de la ciencia si la comunidad científica no puede acceder a la base de datos?

Más allá de estas cuestiones clave en relación con la producción, recolección, interpretación y transparencia de los big data, desde el punto de vista metodológico resulta relevante señalar que, al menos para algunos autores, con los big data las ciencias sociales están experimentando un renacimiento en el marco del cual las divisiones clásicas sobre el método están destinadas a desaparecer (King, 2014). Pero la supuesta obsolescencia de las divisiones metodológicas clásicas no implica necesariamente una cancelación de los debates metodológicos. Por el contrario, como argumentan Burrows y Savage (2014), la irrupción de los métodos computacionales en el escenario de las ciencias sociales está dando lugar a una renovada “política de los métodos” caracterizada por un nuevo debate, no ya entre los clásicos métodos cuantitativos y cualitativos, sino entre métodos “viejos” y métodos “nuevos”. En este sentido, como señalan Halford y Savage (2017), se observa una creciente tensión entre quienes producen conocimiento científico sobre la sociedad a partir cálculos computacionales y algoritmos basados en los big data, y los científicos sociales “tradicionales” escépticos acerca de su pertinencia y sus resultados.

Es interesante hacer notar que, en muchos sentidos, esta nueva disputa metodológica se organiza tanto en torno de los clásicos argumentos epistemológicos utilizados durante las fases iniciales y más combativas del debate cuantitativo vs. cualitativo, basados en las supuestas diferencias inconmensurables en las nociones de realidad social y conocimiento científico en las que se fundan los métodos, como en los argumentos técnicos que ponen el acento en las características específicas de cada método y, de ahí, en su mayor pertinencia para abordar distinto tipo de objetos (Piovani, 2018).

Reconocer los puntos de contacto entre los argumentos relacionados con el más tradicional debate metodológico del siglo XX –cualitativo-cuantitativo– y el actual debate entre métodos “nuevos” y métodos “viejos”, que se estaría configurando en la actualidad, permite especular acerca de posibles derivas hacia un fuerte antagonismo, análogas a las de la guerra paradigmática de la fase epistemológica de aquél debate, o hacia posicionamientos más pluralistas y pragmáticos que, además de aceptar la convivencia de lo “nuevo” y lo “viejo”, den paso a posibles formas de articulación metodológica entre ambos.

Todavía no hay elementos concluyentes para determinar el curso que tomará este novedoso y reciente debate metodológico. Podría consolidarse, e incluso erigirse como el debate rector en los próximos años en el campo de la metodología de la investigación social—desplazando al clásico clivaje cuantitativo-cualitativo, que para muchos parece ya obsoleto—, o podría perder fuerza y quedar relegado a una posición marginal luego de un breve y tal vez solo aparente momento de protagonismo. En cuanto a la triangulación, como se acabada de señalar más arriba, si bien hay algunos indicios que apuntan a posicionamientos irreductibles del tipo de los que caracterizaron a la guerra paradigmática, también se observan perspectivas más moderadas que, a su vez, dejan espacio para diversas formas de convivencia y de convergencia.

Las controversias que rodean al método y, por lo tanto, las oportunidades para combinar técnicas, aumentan si añadimos no solamente una técnica de recolección de informaciones más, sino una manera totalmente diferente respecto a la caja de herramientas del sociólogo. Lo que sí parece cierto, es que la importancia creciente de los big data conllevará la necesidad de repensar las competencias del científico social: no necesitamos convertir a todos los sociólogos a *computer scientists* —en una suerte de versión actualizada de lo que imaginaba Ogburn en la década de 1930, cuando planteaba un futuro en el que todos los sociólogos serían estadísti-

cos⁷— sino que sepan que los big data pueden ofrecer mucho a la investigación social, si se los utiliza en conjunción con otras técnicas, y que conozcan a fondo sus límites para poder matizar el entusiasmo desmesurado de los últimos años.

Más allá de estos debates, el énfasis en la cuestión metodológica se debe a que existe cierto consenso acerca de la necesidad de desarrollar nuevos métodos y técnicas que logren operacionalizar más cabalmente las altas expectativas que han generado los big data. Y, en este contexto, “los científicos sociales que asumen posiciones relativamente moderadas, es decir que no rechazan de plano las nuevas modalidades de investigación basadas en big data, pero tampoco las aceptan acríticamente, alertan sobre el riesgo de nuevas modas científicas que impliquen un total abandono de las formas de investigación más conocidas y consolidadas” (Piovani, 2018: 454). Entre quienes se identifican con este posicionamiento, frecuentemente se encuentran llamados a generar novedosas formas de articulación metodológica en línea con la triangulación y los métodos mixtos, pero que combinen en este caso big data con los small data producidos a través de métodos convencionales cuantitativos —como las encuestas— o cualitativos —como las entrevistas, las observaciones participantes o los grupos focales, etc.—. Por ejemplo, Mauceri (2016) sugiere estrategias para integrar la investigación convencional con el análisis de big data; desde su punto de vista, un eventual ocaso de los métodos tradicionales de producción de datos primarios sería ruinoso para las ciencias sociales. En el nivel aplicado, es decir, el de la investigación empírica, se pueden encontrar ejemplos de este tipo de integración: en el campo de los estudios sociopolíticos, por tomar solo un caso, Wells y Thorson (2015) y Di Giammaria y Faggiano (2016), entre muchos otros, emplearon estrategias mixtas basadas en el uso conjunto de métodos computacionales y encuestas tradicionales, combinando big y small data.

Siguiendo a Bennato (2016), podemos afirmar que el análisis de los big data abre oportunidades muy importantes para las ciencias sociales, pero sería muy importante evitar que su análisis, basado en métodos computacionales, conlleve un descuido del papel clave que la imaginación sociológica tiene en todo proceso de investigación e interpretación de la realidad social. Asimismo, deberíamos mantenernos alertas frente a la posibilidad de que se genere una “datafrenia” en las ciencias sociales, una renovada y actualizada variante de la “cuantofrenia” que caracterizó —y en alguna medida aún caracteriza— a muchas prácticas de investigación *mainstream*. No menos importante resulta recordar el papel central de la teoría en la investigación social, que mucho cultores de los big data, especialmente aquellos con menos formación sociológica, tienden a minimizar, desconocer o incluso negar.

Teniendo en cuenta este conjunto de advertencias y llamados a la cautela, no parece descabellado sugerir que la triangulación con métodos convencionales, ya sean cuantitativos o cualitativos, o con nuevos métodos y técnicas orientados a la producción y análisis de datos primarios sobre la vida social, podría ser una vía prometedora para potenciar a los métodos computacionales y algorítmicos elaborados con el fin de analizar big data en el campo de las ciencias sociales.

REFERENCIAS

- Becker H. S. (1967), *Whose Side Are We On*, “Social Problems”, 14, 3, pp. 239-247.
Bennato D. (2016), “Il rischio della datafrenia nell’analisi dei big data. Appunti a partire dal

⁷ “... all sociologists will be statisticians” (Ogburn, 1930:4).

- caso “Mafia capitale””, *Sociologia e ricerca sociale*, 109, pp.83-101.
- Boyd d., Crawford K. (2012), “Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon”, *Information, Communication & Society*, 15, 5, pp. 662-679, doi:10.1080/1369118X.2012.678878.
- Bouillier D. (2015), “Les sciences sociales face aux traces du big data société. Opinion ou vibrations? ”, *Revue française de science politique*, 65, 5, pp. 805-828.
- Boyadjian J. (2016), *Analyser les opinions politiques sur internet. Enjeux théoriques et défis méthodologiques*. Paris: Dalloz.
- Bryman A. (1988), *Quantity and quality in social research*. Londres: Routledge.
- Burrows R., Savage M. (2014), “After the crisis? Big data and the methodological challenges of empirical sociology”, *Big Data & Society*, 1, pp. 1-6.
- Campbell D.T. (1956), *Leadership and its effects upon the group*. Columbus: Ohio State University. Bureau of Business Research. Research Monograph, 83, xi, 92.
- Campbell D. T., Fiske D. W. (1959), “Convergent and discriminant validation by the multitrait-multimethod matrix”, *Psychological Bulletin*, 56, 2, pp. 81-105.
- Campelli E. (1991), *Il metodo e il suo contrario. Sul recupero della problematica del metodo in sociologia*. Milán: Franco Angeli.
- Collins H. M. (1985), *Changing order. Replication and induction in scientific practice*. Londres: Sage.
- Conley D. et Al. (2015), “Big Data. Big Obstacles”, *The Chronicle of Higher Education*, 2.
- Conte R. (2016), “Big data: un’opportunità per le scienze sociali?”, *Sociologia e ricerca sociale*, 109, pp. 18-27.
- Denzin N. K. (1970), *The Research Act in Sociology*. Chicago: Aldine.
- Denzin N. K. (1989), *Interpretive Biography*. Londres: Sage.
- Di Giammaria L., Faggiano M. P. (2016), “Tra big e small data: la partecipazione politica degli utenti Twitter”, *Sociologia e ricerca sociale*, 109, pp. 143-158.
- Garner W.R. (1954), “Context effects and the validity of loudness scales”, *Journal of Experimental Psychology*, 48, pp. 218-224.
- Garner W.R., Hake H.W. and Eriksen C.W. (1956), “Operationism and the concept of perception”, *Psychological Review*, 63, pp. 149-159.
- Giddens A. (1979), *Central Problems in Social Theory. Action, Structure and Contradiction in Sociological Analysis*. Londres: Macmillan.
- Golder S.A., Macy M.W. (2014), “Digital Footprints: Opportunities and Challenges for Online Social Research”, *Annual Review of Sociology*, 40, pp. 129-152, doi:10.1146/annurev-soc-071913-043145.
- Guba E. G. (1978), *Toward a Methodology of Naturalistic Inquiry in Educational Evaluation*. Los Angeles: Center for the Study of Evolution.
- Halford S., Savage, M. (2017), “Speaking sociologically with big data: symphonic social science and the future for big data research”, *Sociology*, 00, 0: pp. 1-17.
- Hirsch P.M. (1972), “Processing Fads and Fashions: An Organization-Set Analysis of Cultural Industry Systems”, *American Journal of Sociology*, 77, 4, pp. 639-659, doi:10.1086/225192.
- Jick T. (1979), “Mixing qualitative and quantitative methods: triangulation in action”, *Administrative Science Quarterly*, 24, 4, pp. 602-611.
- King G. (2014), “Restructuring the Social Sciences: Reflections from Harvard’s Institute for Quantitative Social Science”, in *PS: Political Science & Politics*, 47, 1, pp. 165-172.
- Knafl K. A., Breitmayer B. J. (1989), “Triangulation in qualitative research: Issues of concep-

- tual clarity and purpose”. En J. M. Morse (Ed.), *Qualitative nursing research: A contemporary dialogue*. Rockville, MD: Aspen.
- Lazer D. et Al. (2009), “Computational Social Science”, *Science*, 323, 6, pp. 721-723, 10.1126/science.1167742.
- Lazer D. et Al. (2014), “Big data. The parable of Google Flu: traps in big data analysis”, *Science*, 343, 14, pp. 1203-1205, doi:10.1126/science.1248506.
- Lincoln Y. S., Guba E. G. (1985), *Naturalistic inquiry*. Beverly Hills: Sage.
- Martire F., Pitrone M.C. (2016), *Lo studio dell'opinione pubblica al tempo dei big data. Una sfida per la ricerca sociale, Sociologia e ricerca sociale*, XXXVII, 109, pp. 102-115; doi: 10.3280/SR2016-109009
- Ogburn W. F. (1930), “The Folkways of a Scientific Sociology”, *Publications of the American Sociological Society*, 16, pp. 1-11;
- Mauceri S. (2016), “Contro la deriva scissoria. Le funzioni dei big data nelle strategie integrate di ricerca sociale”, *Sociologia e ricerca sociale*, 109, pp. 70-82.
- Parra Saiani P. (2001), “Triangolazione e processi valutativi”, *Rassegna Italiana di Valutazione*, VI, 24, pp. 49-65.
- Parra Saiani P. (2016), “Los gatekeepers y los recursos de la investigación. Viejos desafíos y nuevas perspectivas en el tiempo de los big data”, *Revista Colombiana de Sociología*, 39, 2, pp. 221-240. DOI: dx.doi.org/10.15446/rsc.v39n2.58973
- Parra Saiani P. (2016), “Le risorse e il controllo. I big data oltre il mito”, *Sociologia e Ricerca sociale*, XXXVII, 109, pp. 28-41, doi: 10.3280/SR2016-109004.
- Parra Saiani P. (2017), *Las ciencias sociales en tiempos del Big Data. Viejos desafíos y nuevas perspectivas*, Universidad Nacional de La Plata, Doctorado en Ciencias Sociales, 11 de diciembre 2017, mimeo.
- Pawson R. (1994), “Quality and Quantity, Agency and Structure, Mechanism and Context, Dons and Cons”. *World Congress of Sociology ISA*, Bielefeld.
- Piovani J. I. (2006), *Alle origini della statistica moderna*. Milán: Franco Angeli.
- Piovani J. I. (2015), “Falacias de la retórica científicista”. En A. Marradi (coord.), *Las ciencias sociales ¿seguirán imitando a las ciencias duras? Un simposio a distancia*. Antigua: Ciudad Autónoma de Buenos Aires, pp. 217-23.
- Piovani J. I. (2018), “Triangulación y métodos mixtos”. En A. Marradi, N. Archenti y J. I. Piovani, *Manual de Metodología de las ciencias sociales*. Ciudad Autónoma de Buenos Aires: Siglo Veintiuno.
- Ricolfi L. (1995), “La ricerca empirica nelle scienze sociali. Una tassonomia”, in *Rassegna Italiana di Sociologia*, XXXVI, 3, pp. 389-418.
- Riessman C. K. (1993), *Narrative Analysis*. Londres: Sage.
- Shaikh A.R. et Al. (2014), “Collaborative Biomedicine in the Age of Big Data: The Case of Cancer”, *Journal of Medical Internet Research*, 16, 4, e101; doi:10.2196/jmir.2496.
- Strong C. (2015), *Humanizing big data. Marketing at the meeting of data, social science and consumer insight*. Londres: Kogan Page.
- Tinati R. et Al. (2014), “Big Data: Methodological Challenges and Approaches for Sociological Analysis”, *Sociology*, 48, 4, pp. 663-681, doi:10.1177/0038038513511561.
- Webb E. J. et al. (1966), *Unobtrusive Measures: Nonreactive Measures in the Social Sciences*. Chicago: Rand McNally.
- Wells C. y Thorson K. (2017), “Combining Big Data and Survey Techniques to Model Effects of Political Content Flows in Facebook”, *Social Science Computer Review*, 35, 1, pp. 33-52.