

Diferencias en el abordaje de la realidad entre el muestreo estadístico, el mapeo del análisis de redes sociales y los modelos de simulación

Introducción

Uno de los problemas básicos en todo tipo de investigación es el recorte que se establece del fenómeno que se quiere estudiar. Desde el momento en que no es posible abarcar toda la realidad, ya que ésta puede ser descompuesta en infinitas variables, se torna necesario operar recortando aquellos factores que se consideren más relevantes. Toda referencia a la realidad, aún aquella más primaria como es la que resulta del uso del lenguaje natural, exige una selección de elementos, imposibilitando un abordaje completo. Esta verdad de perogrullo, no siempre es reconocida en el ámbito de las ciencias sociales y muchas veces se escuchan voces que critican el uso de modelos, argumentando que con esta estrategia explícita se dejan de lado importantes factores de la realidad en estudio.

Consideramos que es mucho más importante para lograr un buen trabajo, ser conscientes de los recortes que se van a efectuar; pues de ese modo es posible considerar nuevos enfoques que o bien incorporen las variables que han sido dejadas fuera o bien encarar un nuevo modelo tomando en cuenta los factores que fueron dejados de lado.

Sin ánimos de pretender ser exhaustivos vamos a dar un panorama de tres modalidades de investigación en donde se practica un recorte explícito, cada uno con sus reglas, sus límites y potencialidades. En primer lugar trataremos acerca de lo que se denomina Muestreo y que es utilizado en Estadística y en lo que podemos considerar la investigación tradicional en sociológica. En segundo lugar nos adentraremos en la modalidad que utiliza el Análisis de Redes Sociales para el armado de los mapas y sociogramas. Por último nos concentraremos en el enfoque de los modelos de simulación y en los objetivos que se persiguen con este tipo de abordajes.

Generalmente, desde la sociología que podríamos denominar clásica al menos desde Durkheim, el estudio de los fenómenos sociales se aborda con el auxilio de las técnicas estadísticas. Si bien existe todo un conjunto de modalidades que enfocan la investigación sociológica incorporando información cualitativa, en las taxonomías más tradicionales sobre las ciencias sociales suele adjudicársele el rol estadístico. Con la utilización de esta modalidad metodológica, es que se extraen los patrones y

regularidades de los fenómenos sociales. Como en general no es posible, por ejemplo, realizar encuestas a todo el conjunto de la población que se asume como universo, en función básicamente de dificultades de orden práctico, es que la teoría de muestreo se presenta como una solución óptima al problema del total poblacional.

El análisis de redes sociales es una modalidad de investigación social que tuvo su origen en la década del '30 en círculos relacionados con la psicología social, la antropología y la sociología. Básicamente el enfoque se separa de las metodologías tradicionales en que el objeto de estudio no son los atributos de los individuos aislados, sino que lo trascendente son las relaciones que vinculan a los actores. Y estos lazos, no son meras intuiciones o sospechas acerca del grado de unión que enlaza a los sujetos sino que son registrados y medidos, y en torno a ellos es que se construyen las hipótesis y se revelan las estructuras subyacentes. Este abordaje excede el marco de una metodología, ya que el propio enfoque obliga a generar determinado tipo de hipótesis por lo que el impacto es también claramente en el dominio de la teoría.

Los modelos de simulación son relativamente nuevos en las ciencias sociales. El auge de las computadoras personales permitió, dentro de los límites impuestos por el escaso apego a la formalización de las humanidades, su relativa difusión entre los investigadores inquietos. El objetivo detrás de esta clase de modelos es por un lado acceder a un conocimiento más acabado del fenómeno en cuestión, ya que es posible reproducir dinámicamente los componentes aislados, brindando una visión completa del problema. Por otra parte, con los modelos de simulación, es posible manipular el objeto y hacerlo correr en un tiempo manejable por el investigador, a diferencia de lo que ocurre cuando se está investigando el fenómeno real, en donde lo único que puede hacerse es o reconstruir arqueológica o históricamente la situación o apelar a lo que se denomina experimentos naturales (i.e. poblaciones aisladas o sometidas a determinadas condiciones particulares).

Somos conscientes que las modalidades expuestas no son las más utilizadas dentro del ámbito particular de la antropología. Sin embargo consideramos que, más allá de la utilidad metodológica de las tres modalidades, la elección también responde a la forma en que se brinda una solución al problema esbozado. Lo más importante es el grado de consciencia del recorte efectuado sobre el fenómeno empírico, en donde es el investigador el que escoge explícitamente que variables y constantes incorporar y cuales dejar de lado.

Al igual que ocurre con las discusiones acerca de la objetividad y subjetividad a la hora de ejercer la ciencia, es preferible siempre tener en claro que es lo que se pierde y que es lo que se gana con el enfoque elegido. Dejar librado al azar, o como diría con equivalencia un einsteniano, a la buena de Dios, el material sobre el que se trabajará en la investigación es casi como un suicidio metodológico. Un error común es pensar que como es imposible alcanzar una objetividad absoluta, entonces es factible proponer cualquier interpretación y que en función de que toda observación es realizada por un ser humano, cualquier punto de vista es válido.

Si bien es imposible acceder a una instancia de objetividad, es factible y de hecho la ciencia así trabaja, basarse en la contrastación intersubjetiva para ajustar las observaciones y lograr, con un grado de rigor aceptable, un tipo de conocimiento que no dependa únicamente de la intuición del investigador, sino que permita a otros científicos llegar a las mismas conclusiones siguiendo los mismos pasos.

Del mismo modo que se pueden tomar recaudos epistemológicos, metodológicos y teóricos para evitar caer en la peligrosa modalidad de igualar todo tipo de interpretación, con las previsible consecuencias morales y políticas que en última instancia igualan a opresores con oprimidos, es factible reconocer los límites de los enfoques que se emplean para lidiar con la compleja realidad.

Como una forma de ilustrar lo anterior exponemos en los siguientes acápites el modo en que las tres modalidades mencionadas abordan la problemática, motivo central del presente artículo.

Teoría de Muestreo

Uno de los problemas con los que se enfrentaron aquellos que comenzaron a utilizar las técnicas estadísticas fue la imposibilidad de encontrar información relativa a toda la población en estudio. Específicamente aquellos que trabajaban en cuestiones concernientes al control de calidad se vieron en la imposibilidad de realizar censos, entendidos como los estudios de una población completa, ya que la propia prueba implicaba la destrucción del objeto que se sometía al control.

Un ejemplo famoso fue el que llevó a cabo William Gosset, más conocido como “Student”, en la fábrica de cerveza Guinness en la que el control de calidad implicaba la apertura de la botella para corroborar sus condiciones. Evidentemente no era posible abrir todas las botellas, por lo que fue necesario desarrollar un método que permitiera

extrapolar los valores obtenidos al universo en su conjunto. Surge así el muestreo en el que las botellas sometidas a verificación son sorteadas y merced a un procedimiento basado en la teoría de las probabilidades y a una serie de supuestos que deben ser cumplidos, es posible no sólo extrapolar los valores a todo el universo sino también conocer de forma precisa el error en el que se está incurriendo. Esta teoría fue ampliándose para adaptarse a los distintos tipos de universo a los que se enfrenta el investigador en donde no siempre es conveniente la utilización del muestreo aleatorio simple.

Los diferentes tipos de muestreos son: muestreo aleatorio simple, muestreo sistemático, muestreo estratificado y muestreo por conglomerados. Todos ellos conllevan, en algún momento del procedimiento, al muestreo aleatorio simple; si bien se adaptan y contrapesan los diferentes sesgos que puedan ser identificados a priori dentro del universo a estudiar. De este modo se reduce la posibilidad de encontrar sobreestimaciones o subestimaciones en las muestras escogidas y se abre la puerta para el uso de los métodos inferenciales estadísticos.

Un muestreo aleatorio simple es aquel en el que cada posible muestra de la medida escogida tiene la misma probabilidad de ser seleccionada. El método requiere de una lista de todos los integrantes del universo, de un tamaño de muestra escogido apriorísticamente y de algún procedimiento que permita la elección al azar de los miembros del universo. En general suelen utilizarse tablas de números aleatorios para la selección, aunque puede usarse cualquier otro sistema que garantice la aleatoriedad.

El muestreo aleatorio sistemático es más sencillo que el anterior aunque igualmente válido para la aplicación de la estadística inferencial. Para su realización es necesario contar con una lista de los integrantes del universo, de un tamaño de muestra apriori, y del cálculo de la razón entre el número de los integrantes del universo sobre el tamaño de la muestra. El procedimiento que se emplea es elegir de la lista a los sujetos con un intervalo igual al número resultado de la división.

Con el muestreo estratificado es necesario primeramente dividir a la población en grupos separados, denominados estratos. Luego, dentro de cada categoría, se aplica un muestreo aleatorio simple. Cuando las proporciones de los estratos son las mismas que las de la población, el muestreo se proporcional; cuando no, es desproporcional. La estratificación desproporcionada es útil cuando la medida de algún estrato, en la población, es relativamente pequeña.

El método de muestreo por conglomerados divide a la población en agrupamientos y dentro de esos grupos se escogerán a las unidades elementales. A los sujetos puede elegírselos por un muestreo simple al azar o tomando la totalidad del conglomerado. Cuando dentro del agrupamiento se utiliza el azar, el método se denomina bietápico. Uno de los requerimientos de este tipo de muestreo es que la variabilidad de la población se encuentre presente en forma natural en los agrupamientos.

En general suele hablarse de esta clase de muestras como representativas, pero esto es un claro error semántico. La única garantía que ofrece este método está relacionada con la probabilidad de ocurrencia del elemento del universo dentro de la muestra. Esto implica que todos los individuos que son definidos para ese universo pueden aparecer con la misma chance dentro de la muestra. Pero ello de ningún modo implica una “representatividad” ya que no es que exista un elemento que defina cada subgrupo del universo dentro de la muestra, sino que simplemente la rueda de la fortuna puede beneficiarlos o condenarlos al ostracismo muestral. Veamos un pequeño ejemplo: supongamos que tenemos en una bolsa 5 clases de bolitas de los siguientes colores: rojo, verde, azul, naranja y amarillo. El muestreo al azar no garantiza de ningún modo que en la muestra estén representados todos los colores, si bien es cierto que si existe la misma cantidad de bolitas de cada color, la mayor probabilidad hará que se cumpla la premisa, sino que, justamente, por tratarse de un procedimiento en el que se involucra al azar, existen chances de que, por ejemplo, todas las bolitas que se extraigan sean del mismo color. Aunque sea extraño, cada tanto la gente gana fortunas en la lotería.

Otro de los inconvenientes que trae aparejado el uso del concepto de representatividad de la muestra está relacionado con el hecho de que, en general, el universo o población es desconocido. Para poder establecer que un fenómeno cualquiera es representativo, en cualquier ámbito en el que nos plantemos, es necesario conocer el referente del cual se desea realizar la representación. La mayor parte de las veces, cuando se utilizan los métodos de la estadística, las características del universo son desconocidas, por lo tanto es necesario contar con una metodología que nos permita extrapolar los atributos observados en la muestra al total poblacional. La única forma eficiente para realizar esa tarea, no exenta de error, es la utilización de alguno de los métodos de muestreo que en alguna parte de su desarrollo involucran al azar. De este modo, aunque el error persista, es posible no sólo realizar inferencias, sino, más importante aún, manejar y controlar el error que se está cometiendo.

Una de las lecciones metodológicas que deben extraerse del uso de los muestreos que autorizan el uso de la estadística inferencial, es que el tamaño de la muestra no es lo que verdaderamente importa. En general suele creerse que a medida que aumenta el tamaño de la muestra, disminuye la posibilidad de error. Esto es cierto sí y solo sí la forma de escoger los elementos de la muestra (sean cuales sean las unidades preferidas) involucre a algunos de los métodos de muestreo que permitan la utilización de las inferencias estadísticas. Sin este supuesto, el tamaño de la muestra pierde todo sentido estadístico. Existen varios ejemplos famosos de exceso de confianza en los tamaños de las muestras que culminaron en estrepitosos fracasos. Tal vez el caso más famoso fue el que sucedió durante las elecciones presidenciales de 1936 en los Estados Unidos. El Literary Digest realizó una encuesta entre sus lectores, recolectando cerca de dos millones de respuestas. Según los datos obtenidos, el candidato del partido republicano, Alf Landon, debía ser el ganador. Al mismo tiempo George Gallup realizó un sondeo de opinión con la misma temática, pero con una muestra mucho más pequeña, tres mil casos, y sus resultados dieron como ganador a Franklin Delano Roosevelt. La historia le dio la razón a Gallup, Roosevelt gobernó los Estados Unidos y merced al New Deal, sacó al país de la depresión económica. El literary Digest, luego del estrepitoso fracaso, entró en bancarrota. Cuál fue la diferencia en ambos enfoques? Básicamente Gallup utilizó un muestreo aleatorio, que le permitió aplicar las herramientas de la estadística inferencial con un margen de error conocido y aceptable; el Literary Digest obtuvo una muestra muy sesgada (la mayoría de sus lectores, sujetos de la encuesta, eran conservadores) y por más que el n era inmenso comparado con el de Gallup, la falta de azar impidió que los datos obtenidos puedan ser extrapolados. Con muchísimo menos dinero gastado en encuestas, pero con un diseño científico, la incipiente consultora, logró un rotundo éxito como predictor de la opinión pública.

Resumiendo, podemos decir que mediante los diferentes tipos de muestreo, podemos reducir el objeto de estudio en forma confiable, manejando una menor cantidad de elementos, pero sin perder efectividad a la hora de extrapolar los valores al total poblacional.

Mapeo de redes en el ARS

El análisis de redes sociales (ARS) es un conjunto de metodologías, con una fuerte carga teórica, que enfatiza las relaciones entre los actores, en vez de focalizar en los

atributos individuales. El cambio del punto de vista implica, no sólo la utilización de técnicas de recolección y análisis de los datos diferentes a las de la sociología tradicional, sino básicamente un enfoque distinto de los problemas que tratan las ciencias sociales.

Los fundamentos del ARS se basan en la teoría de grafos y en el álgebra de matrices. Los grafos, que cuando se aplican a fenómenos sociales se denominan sociogramas, nos brindan una forma icónica de representación de la red, con sus nodos y sus lazos. El álgebra nos permite realizar los cálculos pertinentes que brindan los resúmenes de datos que se utilizarán como insumo de la interpretación. Toda la información dibujada en el grafo, puede ser traducida sin pérdida en una matriz, con lo cual la podemos considerar equivalente tanto en una representación como en la otra. Sin embargo, el uso de los algoritmos aplicables a las matrices permiten su tratamiento en forma más sencilla, esto último siempre y cuando se cuente con una computadora que sea la encargada de realizar los procesos pertinentes.

Los nodos que conforman los sujetos no se encuentran restringidos a individuos concretos, sino que pueden ser cualquier entidad que el investigador decida, tanto personas, como agrupamientos de personas, o inclusive atributos. Las relaciones que utiliza el ARS pueden ser de cualquier naturaleza y sólo están limitadas por la imaginación del científico, así como por el rigor necesario de toda investigación. Estos lazos pueden ser de varias clases y en función de ellos, variará la información que pueda extraerse. Las relaciones pueden poseer dos valores que indiquen presencia o ausencia o tener una gradación que manifieste los diferentes pesos de la misma. Los vínculos pueden ser no direccionados, en donde no importa el sentido, o direccionados, en donde se establece una diferencia entre la entrada y la salida al nodo; dentro de los direccionados los lazos pueden ser simétricos, cuando los nodos poseen una relación de entrada y salida en forma simultánea o en una sola dirección.

Una diferencia que puede encontrarse entre la forma de representación matricial de la sociología tradicional y la del ARS es que en el primer caso, la matriz está formada por las filas, que son los casos u observaciones y las columnas, que son las variables a observar. En el tratamiento del ARS suelen utilizarse matrices cuadradas, es decir que poseen el mismo número de columnas que de filas, en donde tanto las filas como las columnas representan a los casos u observaciones y la intersección entre la fila y la columna es la que expresa la relación.

El recorte del objeto de estudio también implica un cambio de perspectiva tanto si se trata del ARS como de la sociología tradicional. En los estudios sociológicos el universo a estudiar está, usualmente, fuera del alcance de los investigadores. De allí que la técnica más utilizada sea la del muestreo, tal como se vio en el apartado anterior.

Dentro del análisis de redes sociales, los analistas no suelen utilizar muestreos, sino que se investigan redes completas, es decir se identifican las poblaciones y se establecen censos para el mapeo. Por la clase de información que se genera con el ARS es imposible, una vez determinada la red, establecer un muestreo dentro de la propia red, ya que por la propia estructura reticular, cualquier subconjunto de nodos y vínculos que se extraiga no contará con un patrón similar al de la red completa. Esto sucede debido a que usualmente las redes no poseen una distribución gaussiana en cuanto a sus nodos y a sus vínculos, es decir no todos los nodos poseen aproximadamente la misma cantidad de vínculos. Pero aún si tuvieran una distribución aleatoria de sus lazos, las conexiones en sí no son completas (como puede suceder por ejemplo en una red neuronal en donde cada neurona de un nivel se encuentra conectada con todas las neuronas del siguiente nivel, lo que de algún modo facilita que si se pierden algunas conexiones no se afecta la conectividad total), sino que responden a las preferencias o inclinaciones de los elementos tomados como nodos (sean estos individuos, conglomerados de individuos, atributos, etc.). Esto implica que cuando se toma una muestra de una red armada, la subred que se selecciona no tiene ninguna similaridad con la red original.

Los métodos de la sociología tradicional suelen utilizarse en el análisis de redes sociales para escoger la población sobre la cual se mapeará la red. Pero una vez escogida la población objetivo, ya no se puede aplicar el muestreo clásico y la red se registrará en forma completa. Si se tomaran sólo algunos de los actores de lo que constituye nuestro universo, será necesario tomar en cuenta todas las relaciones que posea ese actor, ya que eliminar algunas falsearía la red de relaciones con la que ese individuo actúa.

El punto central del análisis de redes es la definición de los límites del universo a mapear; es decir los niveles de relaciones que se tomarán durante el procesamiento de la información. No existen recetas previas que garanticen que los límites elegidos sean los necesarios y precisos para el tipo de investigación que se desea realizar. Lo que queda claro es que una vez seleccionados los límites, el mapeo de la red deberá realizarse hasta su confluencia.

La arbitrariedad del recorte no impide que éste sea preciso y que cumpla con todas las exigencias de la investigación. En muchos casos la red es tan evidente que los límites se

observan muy fácilmente. Hanneman define dos tipos de límites para el armado de la red. El primero es el que se impone por la propia acción de los individuos, los ejemplos que nos brinda el autor están relacionados con los socios de un club, los alumnos de un aula, los miembros de una organización, etc. El segundo asume un abordaje más naturalista (demográfico o ecológico según la definición de Hanneman), en donde los límites pueden estar dados por el lugar en el que vive la población o por compartir algún tipo de atributo que se considere importante a priori (contactar a la gente que se encuentra en un área geográfica delimitada o los que poseen algún nivel predefinido de ingresos en los ejemplos de Hanneman). El primero de los casos abordados está en relación con lo que podríamos considerar, desde la antropología, como una visión EMIC de la red; el segundo estaría más relacionado con una visión ETIC.

Las técnicas de recolección de la información utilizados en el ARS son: el método de redes completas, el método de bola de nieve y las redes egocéntricas. El primero requiere que se observen todas las relaciones de los actores con todos los demás que se definieron como límite de la red. Este tipo de enfoques brinda un tipo de información muy completa, ya que las descripciones son totales. Por contrapartida, obtener este tipo de datos suele ser muy costoso, sobre todo si la definición de la red abarca una gran cantidad de nodos. Cuando los grupos son pequeños, éste es el mejor método para conseguir la mejor información. El segundo método empieza considerando a un actor o grupo pequeño de actores y a cada uno de ellos se les va preguntando por los lazos que poseen. El proceso se reitera por cada actor que aparezca en el árbol que se genera hasta llegar a la totalidad de la población estudiada o hasta que el investigador lo decida. Uno de los problemas de este método es que deja afuera a aquellos actores que no se encuentran conectados, el otro inconveniente es que no hay forma de saber si el comienzo del árbol es el correcto, en el sentido de que nos va a permitir alcanzar la totalidad de la red estudiada. El tercero implica la localización de determinados nodos, considerados focales, e identificar al resto de los actores con los que se encuentran enlazados. En general este tipo de abordaje se utiliza en combinación con un análisis de atributos y es el que más cerca está del enfoque tradicional ya que puede utilizarse junto a un muestreo aleatorio de los actores y sus propiedades. Ciertos tipos de análisis, clásicos dentro del ARS, no pueden ser utilizados si la red se mapea de esta forma, tales como la distancia geodésica, la centralidad o los distintos tipos de equivalencia posicional.

Modelos de simulación

Los modelos de simulación son estructuras formales, escritas generalmente en algún lenguaje de programación, que se utilizan para representar en forma dinámica determinada clase de fenómenos. Si bien tanto la noción de modelo como la de la simulación no son una novedad en la historia del pensamiento humano, la aplicación dentro de un ámbito computacional sí puede considerarse como una consecuencia de la aparición de la informática y de su veloz y masiva difusión.

El concepto de modelo involucra la posibilidad de representar, en alguna forma, un aspecto de la realidad. Ya sea mediante el uso del lenguaje natural, mediante el uso de lenguajes formales o mediante el uso de alguna representación iconográfica, la modelización, como una forma de pensamiento, es una de las herramientas más importantes de la comunicación humana.

Las simulaciones tampoco hacen su aparición con las computadoras. En sentido estricto, una maqueta, como las construídas por los arquitectos, son también una simulación. Inclusive existen simulaciones dinámicas que no dependen de los bytes para su implementación. Un ejemplo clásico son las representaciones de batallas, hechas en los areneros, en las que los generales, ponen a prueba sus estrategias y vislumbran sus posibilidades en el campo de Marte.

En general las simulaciones computarizadas, dentro de la disciplina científica que fuere, permiten realizar experimentaciones o recrear situaciones en las que es muy difícil trabajar con el objeto real. La mayoría de las veces, las simulaciones involucran una serie de estados o pasos a lo largo del tiempo. Este devenir puede ser continuo o discreto. En las ciencias sociales las trayectorias simuladas son siempre discretas, esto es, cambian en un punto particular del tiempo. Esta clase de modelos permiten tanto la retrodicción, evaluación de sucesos pasados, como la predicción.

Una de las ventajas que poseen las simulaciones, para las ciencias sociales, es la posibilidad de someter a prueba un grupo de hipótesis, que muchas veces, dada la complejidad del ámbito empírico, con su carga de tiempo real y la confluencia de múltiples variables, se hace en la práctica imposible.

Entre las desventajas más apreciables se encuentran: 1) La desconfianza con que en general, los científicos sociales observan aquellos métodos que involucren algún tipo de formalización. Luego del auge del paradigma posmoderno, al menos dentro de la antropología social, quedó instalada la idea de que la mejor forma de ejercer la ciencia

social es confiando en la intuición y teniendo una gran sensibilidad literaria. 2) La falta de competencia de la mayoría de los científicos sociales ya formados, en las artes de la programación de software. Esta ignorancia, como toda ignorancia, trae aparejada una gran desconfianza en el método y en muchos casos la voluntaria incapacidad para comprender las nuevas modalidades de la ciencia. Por ventura, esta situación empieza a revertirse gracias a las nuevas generaciones, para quienes la computadora y sus secretos son parte de su vida cotidiana. Confiamos en que en el futuro, esta clase de prejuicios sea superada, no tanto por un ejercicio reflexivo sino más bien por una suerte de acostumbramiento informático.

Dentro del área de los modelos de simulación aplicados a las ciencias sociales, es central el recorte que se realice para simular el fenómeno de investigación. Los modelos de simulación suelen estar caracterizados por tres atributos: referencia, propósito y relación costo – beneficio (Rothenberg 1989). El primero de ellos está íntimamente ligado con el soporte (sea real o inventado) que sirve de base al modelo de simulación. La referencia, que insistimos, no necesariamente tiene porqué ser un fenómeno real, suele poseer mayor cantidad de atributos que el modelo a realizar. De allí que lo primero que se realice sea un recorte del objeto a investigar. Es necesario identificar que factores son los que se consideran centrales y que factores se consideran superfluos para aplicar con coherencia epistémica el recorte. El armado del modelo obliga a la reflexión, aunque aquí otra vez no tengamos garantía de haber escogido la opción correcta tal como sucede en el ARS, por lo que el diseño del modelo suele ser la etapa más difícil de su armado. El propósito implica el motivo por el cual se desea armar un modelo de simulación. Nunca el objetivo del modelo puede ser la simulación “miembro a miembro” del referente, ya que se incurriría en un error de tipo lógico, en sentido batesoniano, en cuanto a la confusión entre el mapa y el territorio. Este tipo de confusiones puede llevar a paradojas y contradicciones que invaliden todo uso del modelo. Por otra parte establecer un modelo que sea exactamente igual al soporte que le sirve de base nos lleva a una situación como la descrita, magistralmente, por Jorge Luis Borges en su relato sobre la ciencia de la cartografía, en donde las ruinas del reino se confundían con las ruinas del mapa. La relación costo – beneficio implica que existe una ventaja en la utilización del modelo con respecto al fenómeno a representar, ventajas que pueden ser tanto de manipulación experimental, como de claridad conceptual, como de imposibilidad física o simplemente económicas.

La diferencia que existe entre un modelo y un modelo de simulación es la dinámica que caracteriza al segundo, en donde las variables escogidas se ponen en juego y, como se denomina en la jerga computacional, el programa se hecha a correr. En general el uso de esta clase de abordajes está relacionado con un escaso conocimiento del sistema en su conjunto, es decir se conocen sus partes y sus resultados, pero no se tiene una visión cabal de cómo funciona en su conjunto.

El modelo de simulación pretende erigirse en una representación arbitraria del fenómeno a estudiar. Es el investigador quien decide que factores, variables y constantes, son las que considera relevantes, necesarias y suficientes, para obtener un conocimiento más acabado del sistema en estudio. En general los elementos que sirven como base para el armado de un modelo de simulación son situación empíricas, aunque existen casos de modelos de simulación que toman como base a una realidad inventada. En cualquiera de los casos debe aplicarse un recorte, pero este recorte a diferencia de los tratados anteriormente, debe ser lo suficientemente fiel, como para reproducir aquellas instancias que se están investigando.

Aquí encontramos grandes diferencias entre los enfoques planteados anteriormente, el muestreo y el mapeo de redes y el expresado en este apartado. Los modelos de simulación sí pretenden representar, dentro de las elecciones metodológicas del caso, una porción de la realidad con una fidelidad coherente con los objetivos de la investigación. Sin temor a equivocarnos, podemos hablar claramente de emular la realidad. La dificultad radica en que variables y constantes son las escogidas para el diseño del modelo. Y no existen recetas o criterios aplicables en todos los casos. Más allá de lo formalizados que son los simuladores, por necesidad del soporte, ya que la supina ignorancia de las computadoras obliga a una especificación exhaustiva de todos los elementos intervinientes, la única fórmula para el éxito epistémico de un modelo de simulación es la contrastabilidad permanente y la verificación empírica constante.

Tal vez por primera vez dentro de las ciencias sociales, encontramos una herramienta que permita manipular su objeto de estudio, reproduciéndolo en todos los aspectos que se consideren necesarios, aún si el fenómeno implicara una enorme cantidad de años reales en realizar su recorrido. La representación del tiempo, casi como categoría kantiana en cuanto a su fundamento, puede ser manipulada por el investigador de acuerdo a sus intereses y los años pueden fácilmente ser convertidos en ciclos de procesador.

No se conocen, dentro de la historia de la antropología, críticas demoledoras que invaliden el uso de esta clase de formalismos. Sólo su utilización consciente debe ser el parámetro con que se midan tanto sus éxitos como sus fracasos.

Conclusiones

Cada vez que se encara una investigación, se aplica un recorte, denominado en el lenguaje epistemológico “objeto de estudio”. Muchas veces en las ciencias sociales surge una confusión que se manifiesta, pensando que aquellos que hablan de objeto de estudio, están cosificando a las personas que forman parte de la investigación. Claramente desde nuestra postura metodológica, el objeto de estudio, es el recorte que se aplica al fenómeno empírico que se pretende estudiar. Consideramos que confundir a la gente (fuente de información, sujetos de su propio cambio) con los datos que se obtienen es un error de tipo lógico, en el que los niveles se entremezclan, a veces ingenuamente y otras veces sin tanta ingenuidad. Uno puede enamorarse de una nativa (imaginemosla como una informante clave), pero, creemos, esperamos, nadie puede enamorarse de las notas de campo o de la etnografía o de la historia de vida.

En la búsqueda del conocimiento y de la verdad, camino que, desde nuestra perspectiva, es el que recorre la ciencia, más allá de que ésta nunca se alcance en forma plena, es necesario conocer los límites desde los que partimos.

Hemos recorrido tres formas en que diferentes modalidades metodológicas escogen su objeto de estudio.

La teoría del muestreo con su base en el azar y en las probabilidades, que garantiza que la muestra escogida permita extrapolar los resultados al universo en cuestión. Si bien existe un margen de error, lo interesante del caso es que ese error es manejable y conocido, con lo que no se corre el riesgo de sesgo, siempre y cuando se cumplan los procedimientos de rigor. El muestreo usado en estadística posee algunas limitaciones, propias del método, tales como por ejemplo, asumir la independencia de los sujetos entre sí (no con sí mismos, uno puede tomar los recaudos correspondientes y realizar mediciones en dos tiempos diferentes), cierta escasez de riqueza en el procesamiento de los datos (de acuerdo a los niveles de la variables serán los cálculos que pueden realizarse) y cierta distancia entre lo que sugiere la teoría y lo que puede realizarse en la práctica (muchísimas veces no se cumplen todos los supuestos que exige la teoría y aún así se aplican los métodos con el consiguiente problema que ello puede implicar).

El mapeo en el ARS exige delimitaciones muy marcadas del objeto de estudio. Estas demarcaciones no siempre están del todo claras y muchas veces el investigador debe realizar algunos artificios para cerrar la red. De algún modo la red funciona como un sistema cerrado, ya que si bien existe conectividad al interior de la misma, y diríamos es la condición básica, las relaciones del entramado con el exterior, son eliminadas adrede. Esto no es un obstáculo para que se pueda extraer información sumamente valiosa acerca del funcionamiento y estructura del agrupamiento, pero no deja de ser una limitación clara, ya que no existe alguna forma ad-hoc de incorporar esos datos, a menos que se expanda la propia red. Otra inconveniente que puede surgir con esta clase de estrategia de investigación, es la sincronía del ARS. Si bien ésta es una elección metodológica absolutamente válida, hay que ser conscientes de la pérdida que esto puede implicar. Si bien existen salidas al sincronismo, que permitan observar la trayectoria de la red, no son aún lo suficientemente masivas como para ser encontradas entre las opciones más comunes del investigador.

Los modelos de simulación eligen las variables de acuerdo al criterio del científico, sin que existan recetas previas que guíen su selección. No hay modos correctos de construir un modelo de simulación, ya que la realidad puede ser descompuesta en infinitas variables o como reza el viejo refrán de la teoría de conjuntos “existen más clases que elementos”. El único criterio que puede aplicarse, en cuanto a su verificación semántica, es la puesta a prueba del modelo, su correlación con el problema de investigación que quiere investigarse. Desde un punto de vista formal la adecuación del modelo se realiza en función de lo que se denomina la prueba de máquina, es decir la coherencia y consistencia necesarias que todo software debe poseer. Aquí el refrán que viene en nuestra ayuda es aquel que dice “garbage in, garbage out”, “basura entra, basura sale”. Las limitaciones más importantes en este tipo de implementación tienen que ver por un lado con la falta de competencia informática de los científicos sociales, que en general cuando pretenden realizar simulaciones, se rodean de expertos en computación que probablemente conozcan poco del problema empírico. Por el otro lado se corre el riesgo de ajustar el software a los preconceptos del investigador y caer en la profecía autocumplida. La única forma de evitar esta situación es con una rigurosidad y una honestidad intelectual que permita asumir, sin correr riesgos emocionales, que la tarea que uno ha emprendido haya sido absolutamente vana.

Somos conscientes que hemos dejado de lado una gran cantidad de otras modalidades de investigación en ciencias sociales, que establecen el recorte basándose en parámetros

diferentes. No hemos abordado la investigación propiamente cualitativa, aunque tanto el muestreo, como el ARS, como los modelos de simulación, se alimentan de esa clase de datos. Básicamente la recolección de la información en un clásico trabajo de campo antropológico quedó fuera del alcance del presente artículo. Nuestra intención no era abarcar toda la gama posible de estrategias de investigación, tanto en sus fases de recolección como de análisis, sino simplemente brindar un panorama que permita vislumbrar como cada vez que se realiza una tarea científica, se exige un recorte de la realidad que es la que va a conformar el objeto de estudio.

Bibliografía

Agresti, A

(1997) *Statistical methods for the social sciences*. Prentice Hall. New Jersey

Díaz, D.

(2003) *Modelos de simulación en antropología y arqueología*. Tesis de licenciatura (inédita)

Hanneman, R.

(2005) *Introduction to social networks methods*. University of California. Riverside

Lohr, S.

(1999) *Muestreo: diseño y análisis*. International Thomson Editores. México

Martin, D.,

(1983) "Lenguajes Formales y sus autómatas", en *Ciencias de la Computación Vol. II*, Editorial Limusa: México.

Reynoso, C.

(1998) *Corrientes en antropología contemporánea*. Editorial Biblos. Buenos Aires

Reynoso, C.

(2006) *Complejidad y caos, una exploración antropológica*. Editorial SB.
Buenos Aires

Rothenberg, J.

(1989) "The nature of modeling" en *Artificial intelligence, simulation and modeling*, Editorial Wiley, New York

Samaja, J.

(1999) *Epistemología y metodología. Elementos para una teoría de la investigación científica*, Editorial EUDEBA, Buenos Aires

Wasserman, S. y Faust, K.

(1994) *Social networks analysis: methods and applications*, Cambridge University Press, New York