Recolección de datos y modelización basadas en datos etnográficos: algunos aportes al debate

JORGE E. MICELI[1]

"La teoría científica antigua enseñaba – y lo sigue enseñando- que el punto de partida de la ciencia es nuestra percepción sensorial o la observación sensible. Esto suena, a primera vista, razonable y convincente, pero es absolutamente falso. Se puede mostrar fácilmente por medio de la siguiente tesis: sin problema, ninguna observación. Si les pido "Por favor, observen!", deberían preguntarme, con arreglo al uso del lenguaje: "Si, ¿pero qué? ¿qué debo observar?" (Karl Popper 1998)

El vínculo entre recolección de datos y la definición de un "problema" de interés ha sido considerado básico, en cualquier investigación tradicional, para formular preguntas que organicen la información de base. Estas preguntas, lejos de dejarse para después, son las que permiten, según las perspectivas epistemológicas más modernas y consensuadas, el propio registro de la realidad que nos interesa estudiar.

Sin embargo, para la gran mayoría de los antropólogos profesionales, una cosa es problematizar y otra muy distinta es modelizar. A todas luces, la modelización ha sido, tiempo atrás, expulsada del paraíso de los objetivos más a la mano para el trabajo etnográfico, y su mención parecería evocar, en la actualidad, abstrusos formalismos y criptográficas fórmulas aptas para iniciados.

En tiempos muy recientes, la programación de computadoras y el surgimiento del software han generado un cambio cultural paralelo a la consolidación de estas certezas disciplinares, generando dos curiosas realidades que conviven; mientras que por fuera de las ciencias sociales la modelización alimenta la ciencia básica y hasta la industria del entretenimiento, en las ciencias sociales

[1] jorgemiceli@hotmail.com

es considerada una sofisticación innecesaria para la mayoría de las investigaciones.

La idea de que algunos juegos simulados implican una modelización de procesos reales no es, tampoco, algo que se tenga muy presente, y quizás esta actividad ocupa, en la actualidad, el lugar de una hija bastarda a la que no muchos se animan a reconocerle su verdadero mérito intelectual.

Me animan, en este espacio y en cuanto a la modelización, no sólo el objetivo de corroborar este estado de la cuestión, sino agregar algunas reflexiones sobre lo que su uso analítico implica.

Aunque el epistemólogo Joshua Epstein subraye las ventajas de la etapa modelizadora recurriendo a elocuentes motivos para hacerlo (Epstein 2008), entre los cuales están la explicación, la recolección de datos, la revelación de las dinámicas profundas, etc., seguimos pensando en esta etapa como una necesidad ligada exclusivamente a la programación de las sociedades artificiales (modelos de simulación centrados en la representación abstracta de situaciones sociales), y no a la investigación standard que hacen los científicos sociales

Una de las consecuencias más engañosas de esta visión reificada de la modelización computacional es la creencia de que el pasaje de una etapa a otra es algo definitivo y no sujeto a reconsideraciones, dudas y decisiones retroactivas que ponen en tela de juicio la validez de los procesos ya realizados.

Otro efecto que quisiera cuestionar es el de una visión perspectiva discontinua de los vínculos existentes entre la modelización tradicional, propia de cualquier investigación en ciencias sociales, y aquella destinada a ser soporte de una simulación informática.

El recorrido planteado aquí no será exhaustivo ni erudito. Por cuestiones de espacio y formato descarto la exhaustividad; respecto de la erudición puedo decir que los conceptos que aquí anticipo podrían ser, sin duda, objeto de puntualizaciones mucho más técnicas y preciosistas, pero no me propongo detallar ideas que fueron mucho mejor desarrolladas por los expertos de cada

campo[2] sino constituir, en la medida de lo posible, enclaves de reflexividad interdisciplinaria.

La recolección de datos: yendo del territorio al mapa de modo controlado

Según el médico y neurólogo W. Ross Ashby *"el concepto fundamental en cibernética es el de diferencia", sea entre cosas evidentemente diferentes, sea entre dos estados de una misma cosa que ha cambiado en el transcurso del tiempo"* (Ashby 1960: 21). Aunque parezca obvio en algunas de sus aristas, podemos remitir todo proceso investigativo a la captación, organización y explicación minuciosa de las diferencias que se producen en un proceso empírico en el curso del tiempo. Estas diferencias permiten distinguir entidades a nivel perceptivo e intelectual, pero también posibilitan la distinción de los sistemas en el curso del tiempo.

Como antropólogos sociales, la reivindicación de los métodos de la etnografía clásica ha comportado, sin duda, la certeza de que las cosas no son tan sencillas cuando uno va al campo a hacer observaciones y recoger datos, como tener las categorías claras y ubicarlas, con cierta transparencia analítica, en el interjuego empírico que el investigador releva situacionalmente. Algunos antropólogos, como la reconocida antropóloga Elsie Rockwell, por ejemplo, han postulado que la claridad categorial es incluso un obstáculo al momento de formular inferencias válidas que partan directamente del proceso de recolección de datos:

> *"Al inicio, podemos tener cosas claras, categorías seguras; se trata de descubrir cuáles son y volverlas un poco más confusas, para poder empezar a reconocer lo que no conocemos. Esto último es clave: reconocer lo que uno mismo no comprende" (Rockwell 2009)*

[2] Quien quiera introducirse de modo técnico a los fundamentos de teoría de los sistemas desde un punto de vista cibernético, puede recurrir a la excelente "Introducción a la cibernética" de W. Ross Ashby (Ashby 1960) Como veremos, aquí recurriré a muchas de sus ideas, pero nada equivale a una inmersión cuidadosa en este clásico que data de la década del 50 pero que tiene una actualidad imprescindible.

Tal vez la fuerza evocativa de afirmaciones como esta sea estimulante, pero no puedo dejar de señalar que su interpretación literal implica consecuencias muy problemáticas. Si lo productivo, epistemológicamente hablando, es tornar las categorías confusas para forzar procesos reflexivos, entonces el papel de la modelización como instancia clarificadora también se reduce notablemente, porque la univocidad semántica de una construcción analítica es una de las claves de su replicabilidad y uso.

Probablemente la demanda de confusión categorial parcial sea la expresión apresurada de la necesidad de reexaminar, a cada instante, el campo referencial de nuestras afirmaciones, pero llevado a semejante extremo, huelga decir que todo proyecto de esta índole fracasa.

Sin embargo, no pretendo quedarme con este reclamo, sino reflexionar sobre sus implicancias directas. Tal vez lo que Rockwell y otros investigadores quieran decir es que la etnografía es un proceso iterativo que implica reformulaciones categoriales, pero, además de lo que caracterizo como una descripción impropia de esa implicación, pienso que la iteratividad, como concepto clave, no es sólo una propiedad del proceso reflexivo de construcción categorial de la etnografía, sino un rasgo de cualquier proceso de modelización que parta de datos reales, genere descripciones analíticas y pretenda construir explicaciones basadas en ellos.

Pero ¿cómo podemos definir a la recolección de datos como proceso general? Propongo la siguiente respuesta tentativa: es posible definir a la recolección de datos como el *proceso capaz de abstraer un conjunto de objetos y comportamientos considerados a priori como relevantes de acuerdo a un marco teórico más o menos explicitado.*

Esta abstracción no es algo que se desarrolle de un momento para otro y no contemple la vuelta atrás en su desarrollo. Por lo contrario, implica una necesaria y permanente reconsideración de las referencias empíricas en base a las que se construye. Lejos de desplegarse en un único sentido, mapeando elementos reales en objetos formales de manera unívoca, la recolección de

datos implica una reflexividad muy prematura sobre la conveniencia de considerar a los elementos reales de un modo específico.

¿Hay algo que diferencie, ontológicamente, a la colecta etnográfica de datos de cualquier otro proceso de recolección de información? Contra lo que recomienda la tradición antropológica clásica o la defensa más general de la llamada "metodología cualitativa", sostengo que no hay diferencias formales sustanciales entre el producto de colectar datos etnográficamente y hacerlo de un modo que resulte útil para la confección, por ejemplo, de un sistema informático. Tanto si se recogen datos de orden comportamental o cognitivo, las diferencias radican más en la cantidad de secuencias autocorrectivas de toma de información y en problemas de escala o de generalización que en la naturaleza final del dato que se recoge.

La complejidad relativa del trabajo etnográfico

Aunque parezca contraintuitivo, sesgado y difícil de digerir, la etnografía, como método de recolección, sólo se diferencia de los métodos no etnográficos en cuestiones de complejidad algorítmica, ya que, como sostiene el antropólogo Michael Agar:

> *"(....) el punto aquí no es medir la complejidad etnográfica de un modo preciso. El punto es mostrar qué diferente es comparada con la investigación social tradicional. Una etnografía será siempre más alta en complejidad algorítmica. Cuando se estudia metodología, dentro de la ciencia social tradicional, ¿qué es lo que se aprende? Se aprenden algunos algoritmos. La idea es que se pueden implementar esos algoritmos para investigar en un número equis de áreas.*
>
> *Así, por ejemplo, se aprende cómo diseñar y validar encuestas. Se toman cursos de estadística que en estos días son conceptos adosados a paquetes computacionales. Se toman cursos en los que se aprende la diferencia entre un diseño de muestreo aleatorio genuino y uno estratificado. Se aprenden estos algoritmos para poder ejecutarlos en secuencia y llevar adelante un estudio completo. No es ni muy no-lineal ni muy dinámico, pero tampoco se supone que lo sea. Se supone que*

debe ser lineal y predecible pero sobre todo, controlado. Tomemos un problema de investigación, encajémoslo en la secuencia correcta de algoritmos, y bingo, ya tenemos un estudio. No tiene importancia quién lo realiza, o cuándo o dónde lo realizó, siempre y cuando se sigan los algoritmos.

La investigación social tradicional está más abajo en la escala de la complejidad algorítmica cuando se la compara con la etnografía. Por supuesto que un etnógrafo aprende algunos algoritmos en la universidad. Y algunos de estos, como los algoritmos aprendidos por los científicos sociales más tradicionales, tienen que ver con cosas como realizar preguntas, darse cuenta de con quién deben hablar luego, y comparar diferentes clases de datos mediante el análisis. Sin embargo, las versiones etnográficas serán mucho más variadas y laxas." (Agar inédito: 4)

Si se trata sobre todo de aspectos cognitivos, el etnógrafo tiene que ir y venir más veces para relevar su dato, porque no le basta con una encuesta para considerarlo válido y lo suficientemente profundo para dar cuenta de un estado de cosas no evidente.

Este ir y venir no es trivial ni muchísimo menos, pero no genera, persé, ninguna diferencia intrínseca en la naturaleza formal del dato final relevado. El antropólogo debe repreguntar, contradecir estratégicamente al entrevistado para corroborar la fidelidad de un punto de vista, quizás disimular el rechazo cuando una opinión no es de su agrado, pero eso no hace que su dato, de algún modo, sea "más construido" que el relevado por otro investigador, al menos que consideremos esta complejidad algorítmica de la que habla Agar como medida de la complejidad del dato y no de su proceso de obtención.

Una vez que el dato ha sido obtenido (¿Cuáles son las creencias sobre la muerte de un grupo étnico determinado? ¿Cómo se representa su propia identidad ese mismo grupo?) su proceso de inclusión en un modelo formal no tiene diferencias respecto de información obtenida de otro modo.

Si pensamos en un esquema de desarrollo progresivo de estos procedimientos en el curso de una misma investigación, tal vez sea útil concebir a esta primera

etapa de recolección como iniciada por el ingreso de información en forma de percepciones y sensaciones desorganizadas y culminada, al menos provisionalmente, con la elaboración de lo que podemos llamar un *objeto de estudio sin modelizar, es decir, un conjunto de entidades, propiedades e interacciones generales registradas sin propósitos de modelización computacional.*

Aunque no queramos o no podamos reconocerlo, esta recolección implica modos preconcebidos y bastante complejos de considerar los datos perceptivos. No es necesario elaborar una definición operativa de cada concepto para ir al campo; llevamos nuestro fondo de conocimientos a él en un proceso de explicitación imperfecto y discontinuo en su formulación.

A un nivel más básico, y prescindiendo del marco teórico al cual adhiramos, las nociones mismas de religión, de identidad, de grupo social implican, por ejemplo, deslindes analíticos mucho más sofisticados de los que sospechamos. A pesar del diferente origen de los conceptos usados en uno y otro caso, es evidente que cualquier modelización informática recorre estas mismas etapas formales, e incluso, podríamos decir, con una conciencia de sus dificultades que la asemeja al mapeo etnográfico.

Datos y Metadatos

En realidad, desde el punto de vista informático, la recolección de datos en ciencias sociales define un doble standard de posibilidades y requerimientos; por un lado, lo que se conoce como *metadatos*[3] especifica el carácter formal de los

[3] Según una definición estándar "El término **metadatos** no tiene una definición única. Según la definición más difundida metadatos son «**datos sobre datos**». Debido a que muchas veces no se tiene en cuenta la diferencia entre datos e informaciones también hay muchas declaraciones como «informaciones sobre datos», «datos sobre informaciones» y «informaciones sobre informaciones». Otra clase de definiciones trata de precisar el término como «descripciones estructuradas y opcionales que están disponibles de forma pública para ayudar a localizar objetos» o «datos estructurados y codificadas que describen características de instancias conteniendo informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas». Esta clase de definiciones hace mayor hincapié en los metadatos en relación con la **recuperación de información**, y surgió de la crítica de que las declaraciones más simples son tan difusas y generales que dificultarán la tarea de acordarse de estándares, pero estas definiciones no son muy comunes. Los metadatos pueden describir colecciones de objetos y también los procesos en los que están involucrados, describiendo cada uno de los eventos, sus componentes y cada una de las restricciones que se les aplican. Los metadatos definen las relaciones entre los objetos, como las tuplas en una base de datos o clases en orientación a objetos, generando estructuras". Extractado de http://sistemasavanzadosderecuperaciondeinformacion.iespana.es/ el 2 de Junio de 2010.

datos a colectar. Si pienso en el concepto “clase social” voy a entrevistar, por ejemplo, individuos que satisfagan determinados requerimientos de ubicación socioeconómica en la estructura poblacional. A la vez, la definición de estos *metadatos*, de esta forma de colectar información, implica o habilita la recolección de nueva información, de nueva *data* que satisfaga los mismos requerimientos.

En definitiva ¿Qué es lo que se define en el metadato y lo que se releva como dato en una investigación? Básicamente, lo que ingresa en estas categorías son objetos, propiedades e interacciones de cualquier tipo. Pensemos, a efectos de ejemplificar con un caso ideal el funcionamiento de estas categorías, que somos antropólogos y vamos a una aldea para hacer un relevamiento etnográfico que mapee parte de la vida económica y algunos aspectos de la dinámica reproductiva de su población.

Los objetos pueden ser, dado el caso, la misma aldea, sus sub-grupos internos, las unidades domésticas y los individuos que conforman a estas unidades domésticas. Las propiedades, por ejemplo, pueden ser la pertenencia social de cada individuo, el tamaño de un grupo social o el papel productivo de cada unidad doméstica. Cada propiedad, aplicada a un objeto, adquiere a su vez un valor, que es el modo en que esa propiedad se *instancia* en el objeto. Las interacciones, a su vez, son secuencias de eventos, generalmente regulares, que vinculan a los objetos entre sí, produciendo modificaciones de valores de propiedades en los objetos. Si los eventos no son regulares tienen el carácter de acontecimiento, pero los acontecimientos, por su excepcionalidad, son menos susceptibles de incorporarse a cualquier modelo. Las interacciones, a su vez, pueden vincular a objetos del mismo nivel o de niveles diferentes.

Ejemplificando, los individuos pueden interactuar entre sí casándose, y eso puede tener como resultado una modificación en el estado marital de cada individuo. Los grupos sociales pueden comerciar entre sí, modificando el valor de sus posesiones respectivas. Las unidades domésticas, finalmente, pueden

intercambiar prestaciones reciprocitarias de trabajo, generando cambios temporales en la disponibilidad de mano de obra en su interior.

En principio, seguramente no tenemos en claro qué interacciones ligan a los objetos de nuestra recolección, pero progresivamente obtendremos información más precisa sobre estos aspectos. Lo importante de las interacciones es que, al contrario de los objetos y las propiedades, presuponen la temporalidad procesual. Si se modifican los valores de propiedades en los objetos que interactúan, tenemos, al menos, dos instancias temporales, una predecesora y otra sucesora de cada interacción (Tiempo 1 predecesor y Tiempo 2 sucesor).

¿Cómo opera la recolección de datos en cada tipo de entidades? En el caso de los objetos, la recolección actúa asignando valores de propiedades a cada uno de ellas, concebidas como variables.

Como sostiene el etnógrafo y metodólogo Russell Bernard:

> *"A variable is something that can take more than one value. The* values can be words or numbers*. If you ask a woman how old she was at her first pregnancy, the answer will be a number (16 or 40, or whatever), but if you ask her about her religion, the answer will be a word ("Muslim" or "Methodist").*
>
> *Social research is based on dening variables, looking for associations among them, and trying to understand whether—and how—variation in one thing causes variation in another. Some common variables that you'll .nd in social research are age, sex, ethnicity, race, education, income, marital status, and occupation" (Bernard, 2006: 28)*

Las discusiones colaterales generadas por las dimensiones de las variables, por el tipo de valores que asumen (nominales, ordinales o por intervalos) no son pocas y de hecho llenan los manuales de metodología, pero no serán objeto de controversia aquí.

Lo que me interesa destacar, sin embargo, es que el proceso de recolección de datos es mucho más amplio y complejo que lo que induce a pensar, por supuesto, su forma tabular final. En realidad, las decisiones metodológicas más trascendentes consisten en segmentar el campo perceptivo en unidades

tratables y autónomas, con propiedades mensurables en un tiempo que necesariamente tiene que ser finito. Además de esto, enlazar a estas unidades de modo teóricamente significativo no resulta, tampoco, un esfuerzo menor, y esto trasciende a cualquier evaluación de mensurabilidad. En términos más realistas, la mensurabilidad última de nuestras variables aparece como un aspecto secundario y no central de cualquier proceso de recolección de datos en su fase inicial, ya que las primeras preguntas que nos hacemos tienen mucho que ver, en cambio, con los rasgos de nuestras unidades de análisis y su integración dentro de sistemas más amplios.

Recapitulando, para el ejemplo ya visto, tendríamos cuatro tipos de objetos, que serían la aldea, los sub-grupos que la conforman, las unidades domésticas e individuos. Al vincular la entidad central "aldea" con cada uno de los tres tipos de objetos, podemos transformar en propiedades mensurables de ella a la cantidad de individuos, de grupos sociales y de unidades domésticas que le pertenecen.

Los distintos niveles de agregación informativa

Por otro lado, la posibilidad de deducir las propiedades que vinculan a dos tipos cualquiera de objetos en estos diagramas arroja una certeza nada trivial: si el dato es colectado en su nivel inferior, en este caso el nivel 4 (individuos), va a poder ser transferida al resto de la jerarquía de objetos que integran el esquema, pero si no es así esto no será posible.

En efecto, si disponemos de la cantidad de individuos por unidad doméstica, y sabemos que unidad doméstica integra que aldea (suponiendo que haya más de una aldea), dada una aldea determinada deduciremos que cantidad de individuos la componen mediante una simple suma.

La sensibilidad del promedio es, desde ya, el problema de base de esta imposibilidad de extrapolación "hacia abajo" de los resultados, pero nos alerta sobre una regla de oro de la recolección de información, ya que, como afirma Bernard:

> *"Remember this rule: No matter what you are studying, always collect data on the lowest level unit of analysis possible.*
> *Collect data about individuals, for example, rather than about households.*
> *If you are interested in issues of production and consumption (things that make sense at the household level), you can always package your data about individuals into data about households during analysis. But if you want to examine the association between female income and child spacing and you collect income data on households in the first place, then you are locked out.*
> *You can always* ***aggregate*** *data collected on individuals, but you can never* ***disaggregate*** *data collected on groups" (Bernard 2006: 51)*

En efecto, al promediar los datos del nivel inferior perdemos la posibilidad de conocer con precisión la información de menor rango.

Es conveniente, entonces, no sólo colectar la información con el mayor grado de detalle, sin promediar o totalizar valores, sino que también conviene desagregar los datos de la manera formalmente más compleja para luego agregarlos en una etapa analítica posterior[4]. Como afirma Bernard, si uno obtiene información sobre individuos y no sobre hogares, por ejemplo, podría inferirla sobre estos últimos a partir de datos de menor nivel, lo que no será posible inversamente.

La agregación, el promedio y la totalización, como operaciones lógicas y aritméticas, pueden desarrollarse de los niveles de menor jerarquía hacia los superiores, pero nunca al revés

---

[4] Ya he realizado, en otro lugar y a propósito del análisis de redes sociales, una observación análoga:

> "Hay, entonces, una relación procesual asimétrica entre la recolección de datos agregada y desagregada. Lo desagregado se puede agregar, llevar a un nivel de expresión más sintético y luego recorrer el camino inverso, pero lo agregado no se puede desagregar si no contamos con los datos desagregados desde el comienzo.
> Una pregunta crucial en cualquier diseño de investigación de redes es, entonces, aquella dirigida a establecer el nivel de agregación correcto o la escala adecuada en la cual quiero obtener mis datos primarios." (Miceli 2008)

El contexto que aquí planteo es distinto, pero estimo que el espíritu de la apreciación es sin dudas el mismo.

Sin embargo, es un hecho innegable que no todas las propiedades de un modelo muestran esta posibilidad de agregación o desagregación en el pasaje de un nivel a otro de la estructura. En nuestro ejemplo, si bien puede inferirse la cantidad de personas o la cantidad de ganado de una aldea si conocemos la cantidad que forma cada unidad doméstica y el número de unidades domésticas que forman cada grupo, la propiedad "ubicación geográfica" no es extrapolable de un nivel a otro. En efecto, conocer la ubicación de los grupos sociales dentro de la aldea no permite que sepamos, por ningún procedimiento de totalización, la ubicación exacta de la aldea en un contexto más amplio. La *composicionalidad*, permite, entonces, que para algunas propiedades apelemos a la deducción para inferir sus valores en los niveles superiores desde los inferiores, pero que no recorramos el camino inverso. Otras propiedades, en cambio, no son afectadas por estas reglas de composicionalidad, y muestran una existencia solamente anclada en cada nivel del modelo.

Figura 1: Diagrama relacional de objetos en modelo artificial de aldea

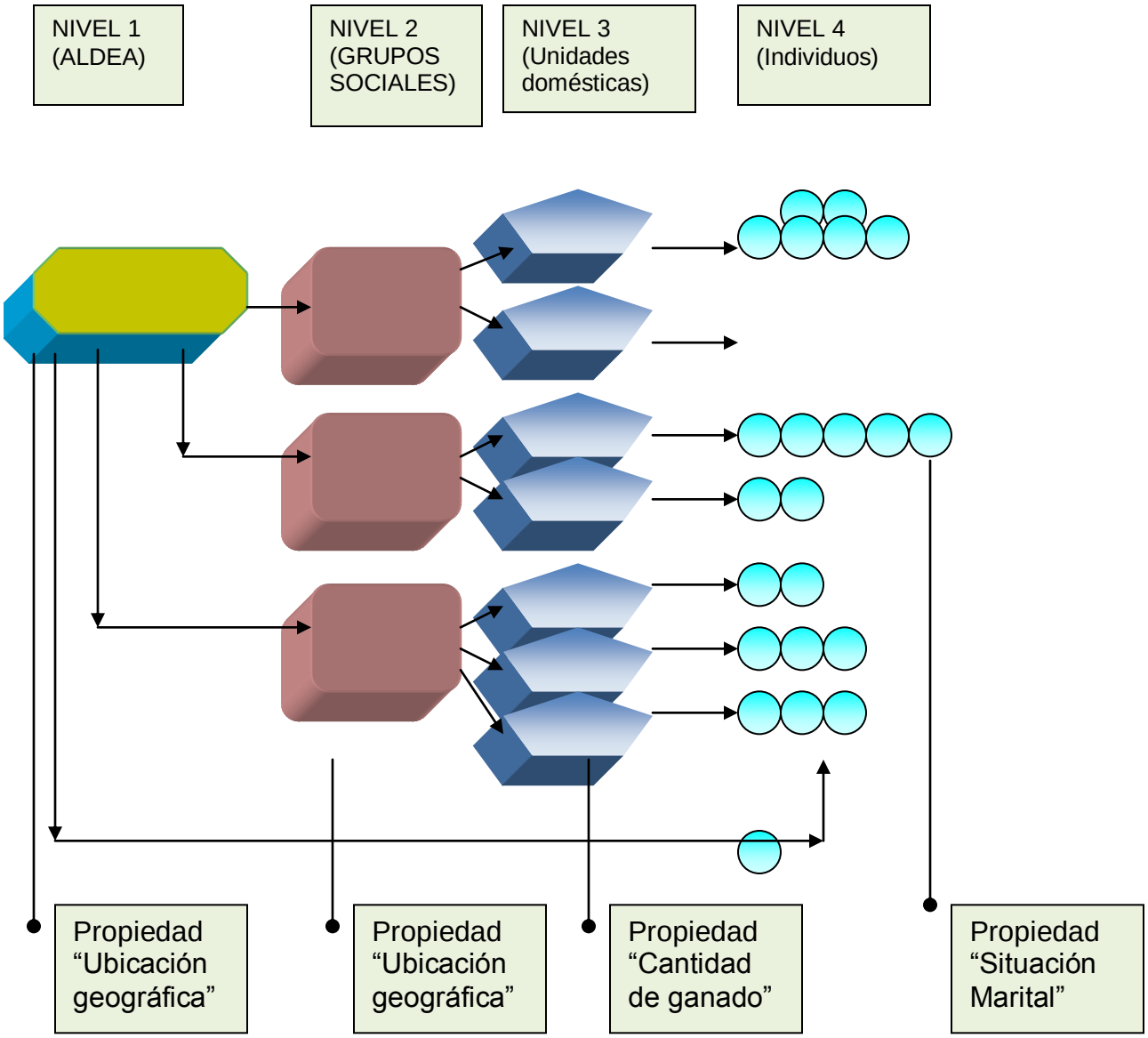

Tabla 1: Objetos, propiedades e interacciones y conceptos de datos y metadatos aplicados a cada uno de ellos

	Metadatos	**Datos**	**Ejemplos de Metadatos**	**Ejemplos de datos**
Objetos	Definición de lo que es un objeto del modelo.	Información que se ajusta a la definición del objeto	Definición de Individuos, Grupos Sociales, Unidades Domésticas.	La aldea A tiene X individuos. La aldea A tiene X Grupos sociales. La aldea A tiene X Unidades domésticas.
Propiedades	Definición de lo que es una propiedad de cada objeto del modelo	Información que se ajusta a la definición de la propiedad de los objetos	Definición, para individuos, de Estado marital. Definición, para Grupos Sociales, de Valor de sus posesiones. Definición, para Unidades Domésticas, de Individuos disponibles para trabajar.	De los X individuos de la aldea A, hay X casados y X solteros. El Grupo 1 de la Aldea A tiene X vacas. El Grupo 2 de la Aldea A tiene X cabras. La unidad doméstica 1 de la aldea A tiene X individuos disponibles para trabajar.

Interacciones	Definición de lo que es una interacción de los objetos del modelo.	Información que se ajusta a la definición de la interacción entre objetos	Definición, para nivel interindividual, de la interacción Casamientos. Definición, para Grupos Sociales, de la interacción Comercio. Definición, para Unidades Domésticas, de la interacción Prestaciones de trabajo.	En la aldea A se produjeron X casamientos, y ahora esa aldea tiene X individuos casados y X solteros. El Grupo 1 y el Grupo 2 de la Aldea intercambiaron ganado. El Grupo 1 quedó con X vacas, y el Grupo 2 quedó con X cabras. La Unidad Doméstica 1 le brindó a la Unidad Doméstica 2 X individuos para tareas de pastoreo durante X días.

La modelización: generando una representación formal de los datos colectados a partir de la sensibilidad escalar

Sin embargo, la modelización de los datos no está implicada, *strictu sensu*, en esta recolección inicial. Aunque toda distinción en el universo perceptivo es significativa en esta primera etapa, la modelización opera subsumiendo esta información primaria en circuitos más complejos de vinculación causal. La construcción de un modelo, en este sentido, no es más que una puesta en escena bien formal de aquello que Gregory Bateson designó como "la diferencia que hace la diferencia".

En efecto, según Bateson:

> *"1. Toda información digital tiene que ver con la diferencia. En las relaciones de mapa y territorio (de cualquier clase y en el más amplio sentido) aquello que pasa del territorio al mapa es siempre y necesariamente noticia de la diferencia. Si el territorio es homogéneo, no habrá marcas en el mapa. Una sucinta definición de información podría ser "una diferencia que hace una diferencia a cierta distancia"*
>
> *2. El concepto de diferencia entra dos veces en la comprensión del proceso de percepción: primero, debe haber una diferencia debe ser convertida en un evento dentro del sistema de percepción, es decir, la diferencia debe sobrepasar un umbral, debe ser diferente del valor de un umbral (G. Bateson y M.C. Bateson 1994: 126). "*

Antes que nada, la idea de "cierta distancia" es aquí crucial y remite a los instrumentos de medición o procesos perceptivos que se están utilizando para obtener valores de las propiedades de los objetos. En el campo, y también en los modelos, sucede que no todas las diferencias son significativas.

Pensemos un poco más en profundidad esta idea. Cuando observamos una escena, como cuando la describimos, no todas las inflexiones que ingresan en nuestro aparato perceptivo son dignas de nuestra atención. Lo interesante es que esto no sólo nos sucede a nosotros, sino también a aquellas personas que entrevistamos en nuestro proceso de recolección de datos. No sólo los procesos

psicológicos atencionales (lo que nos interesa, lo que nos atrae sensorialmente en un momento dado) son responsables de estas asimetrías de énfasis, sino que el mismo lenguaje y las instituciones humanas cumplen, de algún modo, la función de cosificar, de eliminar diferencias de modo sesgado en cada acto de habla. Las consecuencias de esta eliminación de diferencias permanecen parcialmente ocultas hasta que la recolección de datos y la confección de un modelo las interroga de modo crucial.

El sociólogo francés Bernard Lahire expresa esto con notable lucidez, y sus conclusiones tienen fuertes implicancias para la tarea modélica:

> *"Desde ese punto de vista, las prácticas y los saberes se hacen más visibles y declarables en la medida que son claramente sostenidos por instituciones. Cuanto más ligados están la práctica y el saber a tiempos y lugares específicos, relativamente autónomos, son más visibles y designables como tales. Por ejemplo, si en el continuo de actividades se recortan sin dificultad las "prácticas laborales", es porque existen fábricas, oficinas, empresas, etc., es decir, lugares y tiempos que autonomizan y con ello hacen visibles estas prácticas. De igual manera, si hoy es fácil hablar de "prácticas de esparcimiento" (por más que en la historia no haya sido siempre así, y el fenómeno siga siendo variable según el grupo social de pertenencia), es porque el mundo social separó tiempos y lugares específicamente consagrados a tal efecto (horas cotidianas por fuera de la jornada de trabajo o días no laborables que permiten frecuentar clubes o gimnasios, parques de entretenimiento, centros o asociaciones culturales, clubes vacacionales, etc.)." (Lahire 2006: 139)*

El lenguaje, entonces, desarrolla un doble proceso de rotulamiento y homogeneización que nos presenta el mundo de las prácticas de modo continuo cuando, en lo real, hay cortes y rupturas que nuestro modelo tal vez deba tener en cuenta, y este requisito es necesario cumplirlo independientemente de que queramos construir un proceso de simulación o desarrollar una investigación a la manera clásica. Si interrogamos a alguien sobre actividades laborales en un momento específico del día, probablemente, como señala Lahire, contestará

“estaba trabajando” (Lahire 2006: 139), pero después de una breve secuencia de preguntas más específicas, evocará actividades no incluidas en esa categoría “oficial” como haber tenido un sueño, o haber tenido conversaciones informales con compañeros de trabajo.

El establecer cortes en el continuum perceptivo no es, por ende, una característica exclusiva de la recolección de datos o de la modelización, sino, más bien, de la experiencia lingüística más básica. En términos batesonianos, los seres humanos no podemos dejar de producir mapas de nuestra experiencia cotidiana, y la relación con el universo referencial nunca es absolutamente atómica o indescomponible.

La actividad modelizadora, sin embargo, debe interrogar los mapas de los actores o informantes para trazar sus propias versiones de la realidad a la que se accede. A veces, sin embargo, puede no ser suficiente “creer” en el mapa del actor en base a una interrogación fina de sus acciones, y tal vez haya que observarlo actuar sin mediación alguna. Lógicamente, si trabajamos sobre un universo empíricamente muy grande, será imposible corroborar esto observacionalmente y para todos los casos, pero tendremos dos salidas posibles ante este dilema: extraer una muestra de nuestro universo, o acotar el mismo universo de modo radical.

En el cuadro siguiente, y siguiendo los señalamientos de Lahire, vemos como diferentes descripciones de lo que realizó una persona en el mismo segmento temporal podrían resultar en mapeos bastante divergentes de los mismos hechos de acuerdo a la técnica que se use para relevarlos. En el primer caso, la interrogación no detallada de una acción conduce a una primera visión homogénea de las acciones que esa persona llevó a cabo. En el segundo, la descomposición de esa macro-acción en micro-acciones genera una descripción, sustentada en el actor, mucho más precisa que la primera. En el tercer caso, la observación directa conduce a una reformulación aún más diferenciada de esas mismas acciones:

Figura 2: Descripción progresivamente más "granular" de la actividad de un agente

Primer mapa de la actividad agente:

Trabajar (1 hora)

Segundo mapa de la actividad del agente:

Trabajar (35 m.) – Charlar (15 m.) – Planear actividades (10 m.)

---------- ---------- ----------

Observación de la actividad del agente:

Trabajar (25 m.) – Charlar (10 m.) – Ir al baño (10 m.) – Hablar con el jefe (15 m.)

---------- ---------- ---------- ----------

Como vemos en este caso, el "instrumento" de medición que usamos, nuestro "microscopio" sociológico o antropológico, hace variar el resultado de las mediciones de acuerdo a su precisión relativa. Esto sucede todos los días, sin que nos demos cuenta, pero nuestro universo contingente y variable es reducido de distintas maneras cuando lo describimos. En efecto, decimos que "vamos a trabajar todos los días", y hasta podemos decir que "un día es igual a otro", y esto seguramente será, en términos de verdad ontológica, estrictamente falso.

Lo que sucede en esos casos no es nada más ni nada menos que la noción de "diferencia" batesoniana que usamos es muy gruesa, muy poco sensible a las diferencias sutiles que afectan a la experiencia. También estamos manejando, implícitamente, una noción de *umbral* a partir del cual las modificaciones de algún orden se tornan significativas. En este ejemplo, sólo dos días son sustancialmente distintos si no están dedicados al trabajo. Si utilizamos reglas descriptivas cada vez más sensibles, por ejemplo, distinguiendo jornadas en las cuales el trabajo ha resultado más sencillo o más complejo, podríamos introducir distinciones progresivamente más sutiles y obligarnos a reevaluar, de este modo, nuestros supuestos categoriales más amplios.

¿Nos hace caer esto en conclusiones epistemológicamente relativistas, en una suerte de constructivismo radical que sostendría que ninguna descripción tiene validez si no es en referencia a la perspectiva subjetiva de quien evalúa? Entiendo que no. Lo que estamos estableciendo es la relatividad de cada percepción y medición respecto de los instrumentos, categorías y condiciones en que cada una de ellas se realiza, ya que la simple nominación lingüística común puede ocultar, tanto en el sentido común como en la ciencia, el despliegue de metodologías de observación, registro y formalización muy diferenciadas.

Tras la referencia a las mismas palabras, e incluso a los mismos procedimientos, la existencia de distintos mapas puede pasar desapercibida si nuestro propio examen de esos mapas no es, a su vez, lo suficientemente atento y detallado. Esto abre, imprevistamente, un nuevo flanco de análisis en el cual la idea de recursividad no pasa desapercibida; en efecto, la similitud formal de dos mapas sólo puede analizarse de cara a un tercero que los compare.

Esta idea, planteada de un modo tan "local" a las ciencias sociales, tal vez parezca algo obvia, pero sus correlatos formales son poderosísimos y en alguna medida aleccionadores.

En un área aparentemente tan distante de estas reflexiones como la medición de la longitud de la costa de los países, la importancia diferencial de los patrones usados se pone de manifiesto, inquietantemente, en el concepto de "dimensión fractal". En efecto, y como sostiene el antropólogo Carlos Reynoso:

> *"Cuando Benoit Mandelbrot se lanzó a buscar soluciones para una pregunta clásica formulada por el meteorólogo inglés Lewis Frey Richardson (¿Cuánto mide la costa de Gran Bretaña?) encontró que la medida depende de las convulsiones de la línea de la costa y de la sensibilidad de la regla usada para la medición. Descubrió también que la variación de la longitud que resulta de cambiar el tamaño de la regla linealmente a ½, ¼, 1/8, 1/16....no es lineal. Para obtener una línea recta, hay que trazar un gráfico que relaciona el logaritmo de la longitud de la costa contra la longitud de la regla (...) Si se aumenta la resolución al doble, la longitud no aumentará en la misma proporción; puede que se*

incremente poco o mucho, dependiendo de lo accidentado que sea el objeto a ser medido. Richardson ya sabía que cada país asigna un valor diferente a la longitud de su frontera común: España afirma que su frontera con Portugal mide 987 kilómetros, mientras Portugal alega que alcanza 1214(....) Todos tienen razón. Una curva natural no posee una longitud "objetiva"; pero tampoco es subjetiva, sino estrictamente relativa a la sensibilidad de la medida. Como en la paradoja de Aquiles y la tortuga, cuando la regla tiende a cero, la longitud tiende a infinito." (Reynoso 2006: 333)

La moraleja es clara: como en este caso la longitud de nuestra regla determina el largo de las costas medidas, el nivel de minuciosidad de nuestras observaciones, comenzando por el valor referencial concreto de los usos lingüísticos vulgares, determina los resultados no sólo del registro de la conducta de las personas sino, llegado el caso, de su misma percepción u ordenamiento cognitivo del mundo.

Esto introduce la indispensable sutileza en el debate sobre la fidelidad de las descripciones que los estudiosos dan del universo empírico. La veracidad de las afirmaciones descriptivas siempre se produce de cara a los instrumentos, procesos de medición y teorías puestos en juego, nunca en relación a la subjetividad a secas.

Aunque Bateson aplica este concepto a las facetas elementales del proceso perceptivo, podemos extender este principio, asumiendo los efectos colaterales de esta extensión, a la misma modelización de un sistema, ya que no nos interesan, en nuestro modelo, todos los objetos, propiedades e interacciones, sino aquellos que produzcan, en su conjunción, resultados formalmente significativos.

Conclusiones

En este repaso he querido integrar algunas reflexiones respecto de la confluencia entre la modelización tradicional y la modelización más detallada necesaria para el desarrollo de una simulación. Es mi propósito destacar la continuidad formal y ontológica en estos dos tipos de procedimientos, y la necesidad de incluirlos en una misma heurística global. El mandato batesoniano de distinguir entre mapa y territorio implica, sin duda alguna, una permanente vigilancia epistemológica respecto de la importancia de registrar diferencias que "hagan la diferencia" a efectos analíticos. Lejos de oponer lo cualitativo a lo cuantitativo, o lo sincrónico a lo diacrónico, entiendo que la comprensión profunda de la necesidad ideal de analiticidad (división en objetos, propiedades e interacciones), mensurabilidad (adjudicación de valores a las propiedades y a las interacciones) y dinamismo (subsunción temporal de la simulación) en nuestros procedimientos de recolección y mapeo nos plantea desafíos metodológicos de amplio alcance.

Sin embargo, entiendo que no debemos generalizar estos requisitos y tener una visión rígida de ellos. Quizás las variables con las que tratamos son poco mensurables, o nos cuesta incluir estos procesos en el tiempo de un modo completo y exhaustivo, o tal vez la delimitación de nuestro sistema sea muy precaria y la apertura hacia el contexto sea demasiado importante como para aislar los componentes de influencias externas. Si lo que analizamos son estructuras mentales o esquemas simbólicos, tal vez nuestra investigación tenga que ser encarada de una manera absolutamente distinta, pero creo que las consideraciones precedentes ofrecen un buen marco para tener una visión más integrada de la actividad modelizadora y más allá de sus propósitos programáticos o convencionales.

Viéndolo de manera más global y prescindiendo de los contextos de aplicación, resulta notable que ninguno de los principales requisitos de la modelización computacional tenga implicancias ajenas a las de una investigación etnográfica clásica:

1- En principio, la ***complejidad*** en el proceso de recolección de datos etnográfico no está dada por una cualidad intrínseca de la información colectada, sino por los modos relativamente más intrincados por los cuales se obtiene si se los compara con una investigación tradicional.

2- La necesidad de capturar la información con el mayor ***nivel de desagregación*** es un requisito de ambas perspectivas. Aunque en la experiencia etnográfica el nivel de detalle y la confiabilidad de la información dependan de la estadía prolongada, lo recomendable es siempre actuar con la máxima "granularidad" en la etapa de recolección de datos. Lejos de aspirar a una "confusión categorial" como preámbulo de una apertura de la mirada, como sostiene Rockwell, considero que esta apertura es más factible cuanto más clara sea la operacionalización de los preconceptos que portamos de antemano.

3- Entiendo que la distinción entre ***datos* y *metadatos***, propia de ambientes informáticos, es aplicable a investigaciones tradicionales, distinguiendo entre los objetos y las interacciones de un sistema (metadatos), que son los conceptos formales que el investigador construye o lleva a su investigación, y los valores que estos componentes adquieren en la experiencia investigativa (datos).

4- Hemos visto que la ***sensibilidad escalar*** adecuada, en estos términos y considerando las observaciones de Bernard Lahire, también es un producto de la habilidad y profundidad con la que el etnógrafo interroga a los informantes, y no sólo un producto factual que se "impone" al investigador. Los mismos hechos, observados con aproximaciones diferentes, pueden generar informaciones de diferente especificidad.

Bibliografía

Agar, M., (inédito) “Nos hemos encontrado con el Otro y somos todos no lineales: La etnografía como un sistema dinámico no lineal”

Ashby, R., 1960 [1956], *Introducción a la cibernética*, Buenos Aires: Nueva Visión.

Bateson, G. y Bateson, M. C. 1994, *El temor de los ángeles* Barcelona: Gedisa.

Cuadra, F. De, 2001, “Límites éticos de la Inteligencia Artificial”, en *Razón y Fé.*

Epstein, J. M. 2008, “Why Model?”. Journal of Artificial Societies and Social Simulation 11(4)12.

“Introducción a Los Metadatos”, en http://sistemasavanzadosderecuperaciondeinformacion.iespana.es/, visitado el 2 de Junio de 2010.

Grupo Antropocaos, 2007, “Exploraciones en Antropología y Complejidad”, Buenos Aires: Editorial SB.

Lahire, B., 2006, *El espíritu sociológico*, Buenos Aires: Ediciones Manantial.

McCarthy, J. Y Hayes, P.J., 1985, “Some Philosophical Problems from the Standpoint of Artificial Intelligence” en Readings in Artificial Intelligence, San Francisco: Morgan Kaufmann.

Miceli, J., 2008, “Los problemas de validez en el Análisis de Redes Sociales: algunas reflexiones integradoras”, Revista hispana para el análisis de redes

sociales, Vol.14,#1, Junio 2008. Dirección Web: http://revista-redes.rediris.es/html-vol14/vol14_1.htm ISSN: 1579-0185.

Popper, Karl, 1998 [1994] “La responsabilidad de vivir”, Barcelona: Altaya.

Radcliffe Brown, A.R. 1986 [1956], *Estructura y función en la sociedad primitiva.* Buenos Aires: Planeta-Agostini.

Reynoso, C., 1998, *Corrientes en Antropología Contemporánea.* Buenos Aires:Biblos.

Reynoso, C., 2006, *Complejidad y Caos: una exploración antropológica*. Buenos Aires: Editorial SB.

Rockwell, E., 2009, *La experiencia etnográfica. Historia y cultura en los procesos educativos*. Buenos Aires, Paidós.

Russell, B., 2006 [1940], *Research Methods in Anthropology*. Oxford: Altamira Press.