

Fundamentos de análisis de datos cualitativos para el desarrollo de software para ciencias sociales.

Pablo Gustavo Rodriguez.

Cita:

Pablo Gustavo Rodriguez (Julio, 1997). *Fundamentos de análisis de datos cualitativos para el desarrollo de software para ciencias sociales. V Congreso Argentino de Antropología Social. Facultad de Ciencias Naturales y Museo (UNLP) y Colegio de Graduados en Antropología, La Plata.*

Dirección estable: <https://www.aacademica.org/pablo.gustavo.rodriguez/16>

ARK: <https://n2t.net/ark:/13683/pymh/mbG>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

V Congreso de Antropología Social

La Plata - Argentina

Ponencias publicadas por el Equipo NAYa

<http://www.naya.org.ar/>

info@naya.org.ar

Julio-Agosto 1997

INICIO	NAYa en INTERNET	NAYa en CDROM	EQUIPO NAYa
------------------------	----------------------------------	-------------------------------	-----------------------------

FUNDAMENTOS DE ANÁLISIS DE DATOS CUALITATIVOS PARA EL DESARROLLO DE SOFTWARE PARA CIENCIAS SOCIALES.

Lic. P. Gustavo Rodriguez

Investigación cualitativa e investigación cuantitativa

El objetivo de este trabajo es señalar la necesidad de un mayor aprovechamiento de las computadoras en los estudios cualitativos. Al hacerlo nos motiva el convencimiento de que el desarrollo de las ciencias reclama mayor rigor, precisión y exhaustividad en los análisis sociales. Como una cantidad importante de ellos tienen como materia prima y como resultado final material discursivo (textos) la abundancia de datos hace que su procesamiento resulte una tarea que insume un tiempo desproporcionadamente grande en relación a la duración habitual de los proyectos de investigación, que ronda los dos años, si consideramos la mayoría de las becas o el programa nacional de incentivos.

Esto produce que el investigador se sienta abrumado por la gran cantidad de información, o que restrinja en extremo la amplitud de su investigación por razones prácticas antes que teóricas. La alternativa de emprender proyectos de gran envergadura a largo plazo como para dar tiempo al análisis detallado de gran cantidad de información puede hacer que los resultados de ese estudio resulten obsoletos ó inútiles a su conclusión.

La solución de que trabajen varias personas simultáneamente, multiplica la capacidad de trabajo de un investigador individual, pero también la disparidad de criterios para la toma de las decisiones técnicas de codificación, categorización, clasificación, etc. La computadora multiplica igualmente estas capacidades pero manteniendo la homogeneidad de criterios adoptada por una persona. Al automatizar las tareas más mecánicas y rutinarias el investigador ahorra tiempo para poder invertirlo en la toma de las decisiones más importantes, en la elaboración de interpretaciones de los datos y en la construcción de teoría, tareas en las que no puede ser sustituido por la máquina.

En la prosecución de estos objetivos los paquetes comerciales de software tipo "office" (procesador de texto, base de datos y planillas de cálculo) pueden ser de alguna utilidad, pero para un real aprovechamiento de las capacidades de la máquina es conveniente disponer de programas desarrollados especialmente para este tipo de análisis en ciencias sociales.

Los paquetes estadísticos como el SPSS también tienen una aplicación limitada, puesto que son aprovechables en la medida que nuestros datos cualitativos puedan ser cuantificados, o nuestro interés sea contabilizar la ocurrencia de determinadas categorías o establecer la ubicación relativa de alguna medida de tendencia central

en una escala ordinal. El tratamiento estadístico de variables cualitativas obliga a transformarlas en variables por conteo por lo que la utilización exclusiva de estos paquetes fuerza o sesga la investigación hacia aquellos tipos de análisis que pueden ser realizados con esos programas, cuando son los programas los que deben estar hechos a la medida de las necesidades de las investigaciones y no las investigaciones diseñadas de acuerdo a las posibilidades del software.

Pero para disponer de programas de análisis que respondan a los intereses de nuestras investigaciones debemos disponer de técnicas de análisis con rutinas precisas y detalladas, en las que ningún procedimiento quede librado a la intuición. Y es un hecho que existe una desproporción muy grande entre la literatura sobre técnicas de recolección de datos, que es abundante y la de análisis de datos, que es exigua. Por este motivo un segundo objetivo de este trabajo es repasar los aspectos característicos de la investigación cualitativa que el software especializado debe considerar.ii Si bien es cierto que en todas las ciencias (sean "blandas" o "duras") hay cierta cuota de "arte", de creatividad y de libertad individual, que es positiva y necesaria, y que no debe ser cercenada por procedimientos demasiado rígidos y estereotipados, ello es sólo conveniente en ciertos puntos cruciales del análisis en los que la computadora no puede hacer el trabajo humano. Pero simultáneamente los aspectos más rutinarios y mecánicos de la investigación deben ser definidos con el mayor rigor a fin de alcanzar un nivel de objetividad adecuado.

Hace falta creatividad e imaginación, y mucha, para inventar esas técnicas de análisis en las que resulte transparente el proceso por el que se arriba partiendo de determinados datos empíricos, y dados tales y cuales supuestos teóricos, a determinadas conclusiones.

Cuando esta parte del trabajo puede ser confiada a la computadora, el investigador queda libre para aplicarse a la más importante tarea de tomar las decisiones que justifican la elección de cierto corpus, la realización de sus mediciones, y la elección de las variables, y de las teorías desde las que recorta la realidad e interpreta los resultados de su investigación.

La máquina no puede tomar ninguna decisión por nosotros, a menos que le suministremos instrucciones y criterios precisos para hacerlo, por lo cual la responsabilidad última por cualquier falla en el diseño de la investigación, la elección de los criterios o la interpretación de los datos será exclusivamente nuestra.

En última instancia, la seriedad de un estudio no radica en el uso de la computadora por puro snobismo tecnológico, ni en el despliegue de cuadros estadísticos y coeficientes, sino en esas decisiones previas que justifican (o no) que los datos representan lo que pretendemos que representen.

"Un análisis estadístico, por muy elaborado que sea, raramente o nunca llegará a compensar las fallas de un proyecto mal concebido, o de un instrumento de recolección de datos deficiente. [...] la estadística puede ciertamente constituir un auxiliar valioso de un acertado discurrir juicioso, pero nunca, en cambio, un sustituto del mismo. Desde el punto de vista del sociólogo no es más que un instrumento" (Blalock, 1986: 20).

Siempre existe el riesgo de usar mecánicamente las pruebas estadísticas, de olvidar que se basan en una modelización de la realidad cuyos supuestos no están justificados en muchos casos, ya sea porque los eventos estudiados no son aleatorios o no son replicables, o bien porque no conocemos los parámetros poblacionales, por citar sólo algunos de los supuestos en los que se basan las pruebas estadística más comunes. La replicabilidad de los experimentos depende de que puedan obtenerse a los mismos resultados siguiendo los mismos procedimientos en "condiciones semejantes". Pero estrictamente hablando estas son condiciones ideales, que no siempre podemos asumir como válidas.

"Sin duda las condiciones cambian en la realidad, pero ha de ser posible imaginar, por lo menos que no lo hacen [...]. Prescindimos del hecho de que en el proceso del lanzamiento de la moneda real podría acabar por gastarse de modo irregular, o que en ocasiones se pudiera mantener de canto. Aprendemos a concebir un juego de naipes perfectamente barajado, en el que ninguno de ellos tienda a pegarse a otro, pese a que semejante juego no se encontrará nunca en la vida real" (Blalock, 1986: 129).

No es nuestro objetivo oponer los métodos cuantitativos a los cualitativos como alternativas excluyentes. No pensamos que ninguno de estos enfoques sea tan omnipotente como para dar por sí sólo una imagen completa de la realidad social. Pensamos que son enfoques complementarios que ayudan a dar cuenta de diferentes aspectos de la realidad. Los investigadores individualmente pueden mostrar preferencia por uno u otro de estos enfoques o intentar combinarlos equilibradamente, pero el mayor desarrollo de las técnicas cuantitativas no justifica una restricción de los problemas que pueden ser abordados o resueltos con este enfoque, así como el menor desarrollo de las técnicas cualitativas no justifica la proliferación de ensayos ilustrados, con frecuencia brillantes pero apoyados más en la intuición, que en una investigación empírica y metodológicamente sólida, encaminada a desarrollar técnicas de análisis más precisas.

Recientemente parece estar recibiendo crecientes adhesiones una posición que justifica la especificidad de los métodos cualitativos en base a una supuesta naturaleza radicalmente distinta entre el objeto de estudio de las ciencias sociales y el de las naturales con los siguientes argumentos: "Los fenómenos sociales son distintos de los naturales y no pueden ser comprendidos en términos de sus relaciones causales mediante la subsunción de los hechos sociales a leyes universales porque las acciones sociales están basadas e imbuidas de significados sociales: intenciones, actitudes y creencias" (Fielding y Fielding, 1986; citado en Vasilachis de Gialdino 1992:57).

Esta posición tiene una larga tradición ya que se remonta a la separación entre ciencias nomotéticas y ciencias ideográficas, a la que subyace otra división entre "ciencias" (a secas) a las que corresponde explicar por medio de leyes causales y "humanidades" (no ciencias), a las cuales, imposibilitadas de explicar, sólo les queda comprender a su objeto. Actualmente suele vincularse a las ciencias nomotéticas o "duras" con las técnicas cuantitativas y a las ideográficas o "blandas" con los estudios de tipo cualitativo y los defensores de estas posiciones citan a P. Ricoeur, M. Ponty y Heidegger como sus precursores. Uno de los puntos en discusión es la capacidad de las ciencias sociales para generar explicaciones causales. Huberman y Miles lo plantearon en estos términos: "¿Pueden los estudios cualitativos establecer relaciones causales? Frecuentemente se ataca esta posibilidad tanto desde la derecha ("Sólo los experimentos controlados cuantitativos pueden hacerlo") como desde la izquierda ("La causalidad es un concepto fuera de lugar en el comportamiento humano - las personas no son bolas de billar")" (Huberman y Miles, 1995: 434).

Vasilachis de Gialdino (1992) basándose en esta tradición, sostiene en sus tesis que aquellas ciencias duras de la naturaleza desarrolladas en el positivismo abordan su objeto desde una metodología hipotético - deductiva y con técnicas cuantitativas, mientras que para el objeto de las ciencias sociales son más apropiadas las técnicas cualitativas y la metodología inductiva correspondiente al paradigma interpretativo. A pesar de ello su última tesis admite que la coexistencia de los paradigmas en sociología (esto podría hacerse extensivo a todas las ciencias sociales) permite la combinación de las "metodologías cuantitativas y cualitativas" (Vasilachis 1992: 10) en lo que denomina "la estrategia de la triangulación". Esta estrategia permitiría controlar los respectivos sesgos metodológicos de cada enfoque, haciendo más riguroso el análisisiii .

Por nuestra parte consideramos ante una diversidad de los objetos de estudio y una diversidad de estrategias de abordaje de esos objetos, una unidad de los mecanismos de pensamiento que los construyen a través de dichas estrategias. Por lo tanto, la elección de enfoques predominantemente cualitativos o cuantitativos, o la combinación de ambos en diversa medida, no depende exclusivamente de la naturaleza del objeto de estudio, sino también de las preferencias, objetivos y opciones paradigmáticas del investigador.

Nada impide que describamos el comportamiento de los sujetos según modelos de caja negra como el del comportamiento de las bolas de billar y buscar regularidades y leyes. De hecho existen enfoques semejantes. La imposibilidad de predecir el comportamiento individual no es una limitación exclusiva de las ciencias sociales. Ningún físico puede predecir el punto exacto en el que caerá una piedra determinada arrojada al aire al ser atraída por la fuerza de gravedad, ni la fracción de segundo exacta en la que tocará el suelo, ni el punto y la hora exactos en el que comenzará un tornado. Todo lo más, se puede predecir con un cierto grado de probabilidad que el valor observado caerá dentro de un cierto intervalo de valores teóricos.

"[...] Nos serviremos del término probabilidad para referirnos no a acontecimientos particulares (llover hoy, Jones asesino), sino a un gran número de acontecimientos o a lo que ocurre a la larga." (Blalock, 1986: 129. Subrayado agregado por nosotros).

Este grado de exactitud en la precisión "a la larga" no es inalcanzable para las ciencias sociales. Pero es casi seguro que no satisfaga nuestros objetivos. Cuando un partido político encarga una encuesta de opinión quiere saber qué probabilidades tiene de ganar esta elección, no qué porcentaje de elecciones puede llegar a ganar si se presenta un número suficientemente grande de veces. Por eso la elección de una aproximación cuali o cuantitativa depende fundamentalmente de los problemas que nos proponemos resolver, que a su turno dependen de nuestras preferencias teóricas y paradigmáticas más que de la naturaleza del objeto de estudio.

Los estadígrafos señalan que la necesidad de utilizar técnicas estadísticas se incrementa en la medida en que lo hace la cantidad de variables o condiciones de experimentación no controladas. Blalock (1986: 18) recuerda que "Algunas de las ciencias naturales se han desarrollado sin duda por el espacio de siglos sin el empleo de la estadística inductiva. [Esto se debe a que] en la medida en que imperan condiciones de laboratorio escrupulosamente controladas, la necesidad práctica de las técnicas estadísticas es menor".

Plantear una diferencia radical entre los objetos y los métodos de las ciencias naturales y las sociales deja la puerta abierta a la descalificación de éstas últimas, porque sí es cierto que la estrategia cuantitativa, el método hipotético - deductivo y el paradigma positivista hegemónicos en las primeras, son hegemónicos también en los niveles "gerenciales" del campo científico (Bourdieu, 1994) desde donde se distribuyen los recursos y se legitima el hacer científico. En consecuencia, si al planteamiento de esta división radical le sumamos la definición unilateral y excluyente de científicidad del enfoque positivista obtenemos el desconocimiento de los estudios basados en técnicas cualitativas, por no ajustarse al modelo normativo de investigación establecido.

"La posición que asumimos aquí es que los estudios cualitativos son particularmente apropiados para hallar relaciones causales [...]. En efecto, nosotros nos metemos en la caja negra; podemos saber no sólo que ha ocurrido algo en particular, sino también cómo y por qué ha ocurrido" (Huberman y Miles, 1995) Veamos entonces las características generales de los estudios cualitativos por contraste con un diseño de investigación positivista y cuantitativo típico.

Es verdad que estos estudios siguen generalmente una estrategia hipotético

deductiva. Esto es, partiendo de una teoría o modelo teórico (preferentemente matemático, Bunge 1981) se deducen sus consecuencias observacionales (valores esperados) siguiendo las "reglas de correspondencia" (Klimovsky, 1971). Seguidamente se diseña un experimento consistente en realizar un muestreo o una observación controlada. Dicho diseño incluye el uso de técnicas de estadística inductiva y la formulación de hipótesis nula e hipótesis alternativa en relación al intervalo de confianza dentro o fuera del cual se espera que caigan los valores observados, de acuerdo a un determinado nivel de significación convencionalmente determinado como aceptable. Se culmina con la recolección de datos, los cuales corroborarán provisoriamente o refutarán definitivamente la hipótesis nula. Hay que destacar que este modelo normativo de estudio se sigue incluso en los casos en los que la variable de estudio se mide en alguna escala nominal u ordinal. En estos casos lo cualitativo se transforma en cuantitativo por la contabilización de frecuencias.

De modo que en este tipo de estudios la estadística como recurso y un "marco teórico" o modelo como punto de partida resultan indispensables, y su objetivo es poner a prueba la teoría existente (verificarla o refutarla, según los matices). Pero cuando no contamos con esa teoría o cuando una estadística de casos no parece resultar adecuada a nuestros propósitos necesitamos recurrir a otra estrategia metodológica que nos permita generar teoría en el transcurso de la investigación, partiendo a los sumo (ya que hacerlo desde un vacío teórico es imposible) de algunos "conceptos ordenadores" provisorios. Cuanto más diversas y contradictorias resulten estas perspectivas de referencia más productiva puede llegar a ser la investigación. Estos conceptos guiarán la recolección de datos y la formulación de problemas de investigación en una etapa inicial, pero se irán modificando, enriqueciendo, complejizando, reduciendo a medida que avanza la investigación. Y con su modificación vendrá también la reformulación de los problemas y de las preguntas y la búsqueda de nuevos datos pertinentes para contestar estas preguntas, en un ciclo inductivo - deductivo, un ir y venir de los datos a la teoría y viceversa, a lo largo del cual irá cobrando forma algún tipo de formulación teórica original. El detalle de este proceso fue formulado hace tres décadas por dos sociólogos norteamericanos en una obra ya clásica cuyas ideas centrales resumiremos a continuación.

La "Grounded theory methodology"

(GTM) En 1967 B. Glaser y A. Strauss publicaron "The discovery of grounded theory: strategies for qualitative research", donde exponían lo que consideran "(...) una metodología general para desarrollar teoría basada en datos recolectados y analizados sistemáticamente" en proyectos de investigación empírica (Strauss y Corbin, 1992: 273) y que desde entonces ha servido de orientación a muchos investigadores sociales en sus estudios cualitativos.

A fin de exponer sintéticamente los principios de esta metodología es necesario precisar algunos de los conceptos utilizados por los mencionados autores. Por teoría "fundada" ("grounded") ó también teoría "sustantiva" entienden una que es "[...] generada y desarrollada a través del interjuego con los datos recolectados durante proyectos de investigación" (idem p. 275). El sentido de esta expresión se entendía por la oposición a las que los mismos autores denominaban las "Grandes Teorías", haciendo referencia con esta expresión a las teorías dominantes funcionalistas y estructuralistas de autores como Merton, Parsons y Blau, a las que consideraban "altamente especulativas", abstractas y formales.

Glaser y Strauss usaban esta expresión en un sentido muy próximo al de "Teorías de alcance intermedio" que Merton definía como : "[...] teorías intermedias entre las estrechas hipótesis de trabajo que se producen abundantemente durante las diarias rutinas de la investigación y las amplias especulaciones que abarcan un sistema

conceptual dominante, del cual se espera que se derive un número muy grande de uniformidades de conducta social empíricamente observadas". (Merton 1964: 15-16) "Teorías especiales aplicables a campos limitados de datos - teorías, por ej. de dinámica de clases, de presiones de grupos antagónicos, o de la corriente de poder y el ejercicio de la influencia interpersonal - y no buscar inmediatamente la estructura conceptual 'integrada' suficientemente para sacar de ella todas esas y otras teorías" (Idem p. 19).

Merton también llamaba "teorías especiales" a las que nos ocupan, y señalaba su creencia de que el desarrollo de sistemas conceptuales eficaces en sociología se efectuaría sobre la base del trabajo sobre teorías especiales. Y concluía: "La teoría sociológica debe avanzar sobre estos planos interconectados: a través de teorías especiales suficientes para campos limitados de datos sociales, y a través de la evolución de un sistema conceptual más general, adecuado para unificar grupos de teorías especiales" (Idem p. 19- 20).

Para Strauss y Corbin (1992) las teorías basadas en los datos se caracterizan además por su "densidad conceptual". Con esta expresión, que no se confunde con la "descripción densa" de Geertz, desean indicar que tales teorías abundan en relaciones entre conceptos y grupos de conceptos, expresadas en proposiciones. De este modo identifican el proceso de producción teórica con la generación de conceptos inductivamente a partir de los datos de investigaciones empíricas y el establecimiento de múltiples relaciones entre dichos conceptos y entre los grupos de conceptos constituidos a partir de las relaciones previamente identificadas. La GTM sería la metodología desarrollada para pensar y conceptualizar los datos. Sus rasgos definitorios, "[...] aquellos que consideramos tan centrales que su abandono podría significar un gran alejamiento" de lo que podría llamarse GTM con propiedad, son: 1) la construcción de la teoría sobre los datos a través del interjuego entre datos y teoría, 2) el constante establecimiento de comparaciones, 3) la formulación de preguntas teóricas, 4) la codificación teórica y 5) el desarrollo de teoría" (Strauss y Corbin, 1992: 283).

El procedimiento puede resumirse entonces en un ciclo que va de la codificación de los datos a su análisis, elaboración de categorías y establecimiento de relaciones entre estas (hipótesis) inductivamente, a la búsqueda de nuevos datos que corroboren, refuten o permitan reformular las hipótesis planteadas a través de una secuencia de agregado de casos denominada muestreo teórico. Si los conceptos elaborados o las relaciones establecidas resultan insuficientes o incorrectas se hace necesario efectuar modificaciones o reformulaciones, que nos conducirán a un nuevo muestreo teórico y así recursivamente hasta alcanzar un punto en que los nuevos casos no aportan más información ni obligan a un replanteo conceptual. A este punto denominan Glaser y Strauss saturación teórica.

Para finalizar con las precisiones terminológicas señalamos que el término codificación ha sido propuesto por Lazarsfeld y, de acuerdo a Merton, se opone a derivación formal en el mismo sentido en que inducción se opone a deducción : "Mientras la derivación formal enfoca nuestra atención sobre las implicaciones de una teoría, la codificación intenta sistematizar las generalizaciones empíricas de que se dispone en esferas aparentemente diferentes de la conducta" (Merton, 1964: 110).

Glaser y Strauss hablan de codificación teórica para indicar que a fin de generar teoría no basta con codificar los datos en categorías descriptivas y por ello se requiere de la elaboración de categorías teóricas. "Los códigos teóricos conceptualizarían cómo los códigos descriptivos pueden relacionarse entre sí como hipótesis a ser integradas en la teoría" (Glaser, 1978; citado en Glaser y Corbin, 1992). Si bien no ignoran la carga teórica que se halla presente aún en el lenguaje descriptivo más elemental, intentan tomar distancia de esta manera de aquellos investigadores que persiguen fines meramente empáticos de identificación con su

objeto de estudio y de comprensión, limitándose a describir y "dar la voz" a los sujetos y absteniéndose de elaborar sus propias interpretaciones de los procesos observados.

En el capítulo V de la obra citada exponen la metodología que actualmente recogen muchos de los sistemas informáticos de ADC y que podría resumirse en lo siguiente: Elaborar categorías con definiciones provisorias.

Codificar los datos en las categorías disponibles o en categorías elaboradas especialmente.

Al hacerlo, comparar los datos ingresados con los codificados previamente en la misma categoría para controlar la homogeneidad del conjunto. Elaborar memos (fichas, notas) con las ideas que surjan de la lectura de los datos codificados: posibles contradicciones, variación en el énfasis del propio pensamiento, casos dudosos que requieren posterior revisión, heterogeneidad de los datos que indiquen la necesidad de una subdivisión de categorías o la conveniencia de una redefinición, relaciones entre categorías, etc.

Discutir las ideas de los memos, el sistema de categorías, la definición de cada una de ellas y las codificaciones efectuadas con algún colega. Ellos pueden marcar lagunas, descubrir incongruencias, sugerir casos contrarios, casos límites, relacionados, etc. iv de la misma investigación o de otra.

Integrar las categorías y sus propiedades, pasando de la comparación de casos (datos) entre sí a comparar datos con categorías resultantes de anteriores comparaciones entre datos. Al relacionar categorías entre sí estamos elaborando hipótesis.

Continuar el relevamiento de datos mediante el muestreo teórico. De este modo las preguntas que surgen del análisis orientan la recolección de más datos en la búsqueda de casos nuevos que permitan llenar los vacíos o precisar la teoría.

Delimitar y reducir la teoría. Reducción del número de conceptos explicativos, y generalización a otros campos de aplicación. Tras varios ciclos recursivos el sistema de categorías se va haciendo más y más sólido y estable a medida que los nuevos casos ya no introducen modificaciones importantes en el sistema. Entonces "[...] el analista puede descubrir las propiedades subyacentes en el conjunto original de categorías o sus propiedades, y puede luego formular la teoría con un conjunto más pequeño de conceptos de alto nivel". A este proceso llaman reducción.

Escribir la teoría. Con lo cual el resultado de un estudio sobre datos cualitativos (generalmente textos, bajo la forma de respuestas a cuestionarios, entrevistas, o notas de campo) es un texto.

En todo este planteo subyace la concepción de los sistemas conceptuales como sistemas clasificatorios susceptibles de ser representados por conjuntos en sentido matemático. Es esta propiedad la que permite a los sistemas informáticos efectuar operaciones booleanas con los datos codificados. Por este motivo haremos un breve repaso a los principios clasificatorios generales.

Los sistemas de clasificación

De acuerdo a la Taxonomía, que "[...] es el estudio teórico de la clasificación, incluyendo sus basas, principios, procedimientos [y] reglas" (Crisci, 1983: 4), la clasificación es el agrupamiento de objetos en clases sobre la base de atributos que poseen en común y/o sus relaciones (idem p. 3). Desde un punto de vista, por clasificación entendemos tanto un procedimiento (el mencionado agrupamiento) así como el resultado obtenido a partir de ese procedimiento, al que llamaremos

sistema clasificatorio (Lahitte et al 1987:15).

Generalmente se utilizan indistintamente los términos clasificación y tipología, pero para algunos autores no designan lo mismo: "Cuando se distinguen los elementos de un conjunto en función de un criterio único se habla de clasificación. [...] Cuando se utilizan varios criterios de clasificación para distribuir individuos y la combinación de criterios permite definir clases no jerarquizadas, se habla más bien de tipología que de clasificación. Pero importa notar que las tipologías sólo son una forma particular de clasificación" (Boudon y Bourricaud, 1990: 648).

Por su parte Crisci (1983) diferencia clasificación de determinación señalando que "este último término se define como la ubicación de un objeto no identificado en la clase o grupo al que corresponde conforme a una clasificación construida previamente". Las clases serían categorías que a modo de cajones nos sirven como receptáculos donde "guardar" los items clasificados y poseen una definición consistente en el conjunto de atributos que un objeto debe presentar a fin de poder ser incluido en dicha clase.

Una propiedad importante de los sistemas clasificatorios de especial interés para la producción de teoría son las relaciones de inclusión entre categorías tal como señalaba Musso (1980): "Un conjunto de categorías constituye un sistema (o escala) de clasificación cuando, de acuerdo con sus definiciones, se ha logrado agruparlas de tal modo que todas ellas sean subdivisiones de un mismo criterio ordenadas por sus relaciones de inclusión y exclusión. De este modo, si se afirma que un objeto pertenece a una categoría, quien conozca el sistema (las relaciones de inclusión y exclusión existentes entre sus categorías) puede deducir inmediatamente un conjunto de propiedades que el objeto poseerá o no según que estas estén incluidas o no en la categoría en la cual se lo clasificó. En cambio no podrá deducir nada respecto de otras propiedades que el objeto podrá poseer o no porque son contingentes respecto de esa categoría." (Musso R. 1980:94) El mismo autor distingue entre criterio de clasificación y categoría indicando que "Llamamos 'criterio' a la categoría genérica que designa la propiedad o las propiedades sobre cuya base se clasifica a los objetos, y 'categorías' a las distintas clases en que se subdivide a los valores posibles de dicha propiedad o propiedades" (idem p. 95).

El conjunto de los objetos incluidos en la misma clase puede constituir un grupo monotético o politético. En el primer caso, los individuos (casos, objetos, items en general) que lo integran comparten un conjunto de atributos que son condición necesaria y suficiente para convertirse en integrante del grupo. "Grupo politético es un grupo donde sus miembros comparten un gran número de atributos, pero ninguno de esos atributos es compartido por todos los miembros con la excepción del atributo de pertenecer a ese grupo [...] cada miembro posee algunos atributos de ese conjunto. por lo tanto ningún atributo es condición necesaria o suficiente para convertir a un objeto en integrante del grupo" (Crisci, op. cit p. 4).

De acuerdo a Williams y Dale (1965; citado en Crisci op. cit p. 4) los requisitos lógicos que debe cumplir un agrupamiento de objetos en clases para que sea considerado una clasificación se resumen en tres axiomas:

"a) En cada clase de más de un objeto (no unitaria) debe existir, para cada objeto de la clase, otro distinto que comparta con él como mínimo un atributo.

b) El ser miembro de una clase no es en sí mismo un atributo.

c) Cada objeto de cualquier clase debe diferir, al menos en un atributo de cada objeto de cualquier otra clase".

De acuerdo al primer axioma no podemos clasificar un objeto acerca del cual no tenemos información. El segundo axioma nos impide hacer agrupamientos totalmente arbitrarios. El tercer axioma establece que los objetos idénticos no

pueden ser asignados a clases diferentes.

Por otra parte las categorías de una buena clasificación deben ser exhaustivas, inambiguas y mutuamente excluyentes. Por exhaustivas entendemos que no debe quedar ningún ítem (caso, objeto) del universo en estudio que no pueda ser incluido en alguna categoría. El sistema, como un todo, debe prever todos los rasgos posibles de los ítems reales. De este modo los casos no previstos nos obligan a crear nuevas clases en las que puedan ser incluidos a fin de satisfacer el requisito de exhaustividad.

Un conjunto de clases es mutuamente excluyente cuando cada individuo del universo en estudio puede ser asignado a una y solo una clase del sistema. Ello se logra si las categorías pertenecientes a un mismo nivel de abstracción responden todas al mismo criterio. Por ej. no puede clasificarse a un grupo de personas en rubios, pelirrojos y obesos, puesto que la última clase no responde al mismo criterio que las otras dos (color de pelo) y una persona obesa y rubia podría incluirse en las dos categorías (Blalock, 1986:).

Desde el punto de vista pragmático, la mejor clasificación es la más estable, robusta y predictiva.

"Entendiéndose por estable que no se modifica en forma drástica por la incorporación de nueva información; por robusta que no es alterada radicalmente por la incorporación de nuevas entidades, y por predictiva que una propiedad conocida para la mayoría de las entidades de un grupo está garantizada con una gran probabilidad de que existe en aquellas entidades de ese grupo todavía no examinadas con miras a buscar esa propiedad" (Crisci, 1983:6).

Un sistema clasificatorio elaborado, en la medida en que cumple con los requisitos mencionados y consta de categorías teóricas así como descriptivas, precisamente definidas y saturadas puede aproximarse a un objeto-modelo en el sentido de Bunge (1981). Generar categorías, aún los más simples descriptores al igual que identificar categorías nativas es una actividad teórica y una contribución a la teoría social. Del mismo modo la determinación de cuáles pasajes de texto de una entrevista o documento son relevantes a cada categoría conlleva decisiones teóricas.

Al mismo tiempo, determinar la pertenencia de un individuo a una categoría es efectuar una medición en una escala nominal u ordinal. Y en esta operación el investigador es el instrumento de medición, por lo que para garantizar la confiabilidad y precisión del instrumento debe poseer una sólida formación teórica y criterios definidos explícitamente. Las categorías y sus relaciones pueden representarse simbólicamente. Las diversas operaciones que pueden efectuarse con ellas han sido exploradas por la lógica y la matemática y las computadoras poseen la capacidad como par que podamos reproducirlas cuantas veces resulte necesario. Para asistir al especialista en ciencias sociales en la realización de esta tarea han sido diseñados varios sistemas que explotan estas posibilidades. La exposición detallada de los distintos tipos de sistemas existentes que explotan estas posibilidades escapa al marco de este trabajo. Aquí es suficiente recordar que casi todos ellos reproducen la técnica de fichado textual (Técnica del fichado Los estudios cualitativos trabajan sobre textos. Pero el discurso en bruto no es utilizable. Antes debe ser procesado, segmentado y ordenado a fin de que pueda ser recuperado como dato significativo para el análisis. A este fin ha servido la técnica del fichado que describimos a continuación.

A medida que se van desgrabando y leyendo detenidamente las entrevistas u otros documentos se van buscando los principales términos nativos y transcribiendo en fichas de cartulina los pasajes correspondientes o bien se marcan en el margen con algún código de colores o de otro tipo los fragmentos esclarecedores sobre el significado de los términos en cuestión. Otro tanto se hace con aquellos pasajes que

podrían considerarse ejemplos relevantes a la aplicación de los conceptos ordenadores de los cuales parte la investigación. Estas serían las fichas temáticas de transcripción de entrevista.

Otro tipo de fichas son aquellas en las que transcribimos fragmentos de nuestra libreta de campo sobre las circunstancias que rodearon la entrevista, o cualquier información contextual importante, descripción de los entrevistados, etc. A estas se denominan fichas descriptivas o contextuales y pueden incluir también impresiones personales, sentimientos y percepciones.

En las fichas de registro de texto se anotan las direcciones en las que se pueden hallar los fragmentos que nos interesan. Una dirección de texto se expresaría por el código del entrevistado, el número de entrevista, número de página y de párrafo o línea según la unidad textual que utilizemos. Estas fichas reemplazan a las fichas temáticas de transcripción cuando no disponemos del tiempo o del personal suficiente como para transcribir todos los fragmentos pertinentes,. Ellas permiten efectuar el fichado en menos tiempo, pero complican la consulta de los fragmentos porque obligan a recurrir al original cada vez.

Las fichas de resumen sirven para volcar las ideas principales expresadas en un documento muy extenso o al que no podremos acceder con facilidad cuantas veces sea necesario.

Finalmente están las fichas de reflexión también conocidas como memorandums o simplemente memos, en las que registramos nuestras ideas, intuiciones e interpretaciones surgidas en el campo, durante la transcripción y lectura de las entrevistas, en estado de insomnio o en cualquier otra circunstancia de nuestra vida.

El formato básico para todas ellas consiste en consignar en un ángulo, generalmente el superior izquierdo, el tema y subtema (criterios y categorías) bajo los cuales se clasifica la ficha, en el ángulo opuesto, por ej. el superior derecho, los datos que permiten ubicar el fragmento en el contexto de la entrevista o documento del cual fue extraído, y la fecha en que se confeccionó la ficha. Y en el cuerpo principal de la ficha el contenido de la transcripción o reflexión.

Para analizar la información que hemos registrado es necesario organizarla. Si nuestra intención es producir teoría, lo mejor es hacerlo en temas y subtemas, categorías o conceptos, considerados como clases en un sistema clasificatorio, como ya hemos visto (Hammersley y Atkinson 1994:184). Para toda variable relevante al análisis generamos una categoría y para todo valor posible de la variable una subcategoría, ya sea que se trate de variables nominales, ordinales o de intervalos, en cuyo caso conviene trabajar con datos agrupados en intervalos discretos.

La asignación de un segmento de texto a una categoría, es decir, su determinación como miembro de una clase dada es denominada por algunos autores codificación o indexación.

Cada ficha constituye un registro y cada tema o subtema (variable o categoría) es el equivalente a un campo. La base de datos se ordena por lo general alfabéticamente por temas. Dado que un mismo pasaje puede ser relevante a varias categorías de análisis, puede ser necesario confeccionar con él más de una ficha.

Algunas alternativas a la transcripción utilizadas ocasionalmente son el fotocopiado y posterior pegado de los segmentos textuales, lo cual puede resultar muy costoso para grandes cantidades de información, y la anotación en el margen.

Un método alternativo a la clasificación manual (y un tanto más ingenioso) es descrito por Hammersley y Atkinson (1984: 187-188) y consiste en la perforación de las fichas de cartulina: "Con esta técnica sólo se requiere de una copia de la

información adicional al registro original. Cada segmento de información se añade a una ficha agujereada. Los agujeros en el margen de la ficha corresponden a las diferentes categorías, y se elabora un índice que indique (sic) a qué categoría corresponde cada agujero numerado. Cuando un segmento de información es relevante a una categoría se hace un corte en el agujero; cuando no lo es, se deja intacto el agujero, o viceversa.

"Con este sistema todos los datos se pueden guardar en fichas sin ningún orden en particular. Cuando se necesita consultar el material relevante a una categoría en particular se introduce una larga aguja a través del agujero adecuado y se levantan o se dejan las fichas pertinentes (dependiendo de si los agujeros que representan a la categoría que nos interesa han sido cortados o dejados intactos). Además, al igual que en la clasificación manual, todas las informaciones relevantes pueden examinarse simultáneamente pero sin tener que hacer copias múltiples y ahorrando un espacio considerable. Con este sistema se pueden identificar subtipos de información relevante a dos o más categorías".

Desde luego, esto es exactamente lo que hacían las viejas computadoras que trabajaban con fichas perforadas, con la sola diferencia de que esas computadoras se limitaban a contar las fichas recogidas que satisfacían las condiciones indicadas. El proceso que siguen las actuales computadoras personales sería mucho más complejo de explicar y está fuera de mis capacidades hacerlo, pero la lógica en la que se basan los programas que realizan codificación y recuperación de segmentos de texto es en lo fundamental la misma. No obstante, aún sin recurrir a los programas especializados, el software "integrado" de los paquetes tipo "office" consistente en planilla de cálculo, base de datos y procesador de texto también puede ser de gran utilidad y superando la capacidad del sistema de fichas de cartulina.

Conclusiones

Hasta aquí no hemos hecho más que recordar o explicitar procedimientos por todos conocidos y utilizados. Lo que deseamos destacar es que para sacar el mayor provecho de las capacidades de las computadoras personales en nuestra labor de investigación necesitamos programas que realicen por nosotros la parte más mecánica de todo este proceso, que es también la que consume más tiempo: hacer copias de los segmentos de texto, confeccionar las fichas, llevar una lista de los códigos, un control sobre la solidez de nuestras clasificaciones, la derivación de consecuencias lógicas de nuestras hipótesis, la intersección de categorías, etc.

Simultáneamente, los programas desarrollados para ADC deben ser lo suficientemente flexibles como para permitir ese ir y venir entre datos y teoría aceptando las modificaciones del sistema clasificatorio y la incorporación de nuevos datos en cualquier momento del análisis. Todo ello sin riesgos de perder información.

Con algo menos que esto tendremos programas que limitarán la investigación antes que facilitarla. Un conocimiento adecuado y un uso apropiado de las posibilidades que nos brindan las nuevas tecnologías puede potenciar la riqueza y profundidad de los estudios cualitativos.

Bibliografía citada

Blalock, Hubert M. (1986): Estadística social. Ed. Fondo de Cultura Económica. México.

Boudon R. y F. Bourricaud (1993): Diccionario crítico de sociología. Ed. Edicial, Bs. As.

Bourdieu, Pierre (1994): "El campo científico". En: Redes. Revista de estudios sociales de la ciencia. Vol. 1, N° 2. Bs. As.

Bunge , Mario (1981): Teoría y realidad. Ed Ariel, Barcelona.

Crisci, Jorge Víctor (1983): Introducción a la teoría y práctica de la taxonomía numérica. Monografía n° 26, serie Biología. O.E.A. Washington D.C.

Forni , Floreal y M. Antonia Gallart y otro (1992): Métodos cualitativos II. La práctica de la investigación. Ed. CEAL, Colecc. Los fundamentos de la ciencias del hombre, nro. 57, Buenos Aires.

Glaser, Barney y Anselm Strauss (1967): The discovery of grounded theory: strategies for qualitative research. Aldine Publishing Company, N.Y.

Hammersley, Martyn y Paul Atkinson (1994): Etnografía. Métodos de investigación. Ed. Paidós, Biblioteca Paidós básica n° 69, 1º ed., Barcelona.

Hempel, Karl G. (1977): Filosofía de la ciencia natural. Ed. Alianza, Madrid.

Huberman A. Michael y Matthew B.Miles (1995): "Data management and analysis methods". En: Methods of collecting and analyzing empirical materials, Cap. 27, p.428- 444.

Klimovsky, Gregorio (1971): "Estructura y validez de las teorías científicas". En: Ziziemsky, E. (comp.) Métodos de investigación en psicología y psicopatología. Ed. Nueva Visión.

Lahitte,H.; J. Hurrel y A. Malpartida (1987): Relaciones. De la ecología de las ideas a la idea de ecología. Mako Editora. La Plata.

Merton, Robert K. (1964): Teoría y estructura sociales. Ed. Fondo de cultura económica, México D.F.

Musso, R. (1980): Problemas y mitos metodológicas de la Psicología y la psicoterapia. Ed. Psiqué, Bs. As.

Ontoria, A. et al. (1995): Mapas conceptuales. Una técnica para aprender. Ed. Narcea S.A., Madrid.

Rodriguez, P. Gustavo (1997): "El recurso informático en el procesamiento de datos cualitativos en ciencias sociales".

Strauss, Anselm y Juliet Corbin (1992): "Grounded theory methodology". En: Strategies of inquiry, Cap. 17, pp 273-285.

Vasilachis de Gialdino, Irene (1992): Métodos cualitativos I. Los problemas teórico - epistemológicos. Ed. C.E.A.L., Bs. As.

Vessuri, Hebe M.C. (1973) "Técnicas de recolección de datos en la antropología social". En: Relaciones Tomo 18, art. n° 75.

Wilson, J. (1969): Thinking with concepts, Cambridge.

NOTAS

i Escuela Sup. de Trabajo Social, U.N.L.P.

ii En otro trabajo hemos reseñado las capacidades que diversos programas especializados y no especializados ofrecen para procesar datos cualitativos

(básicamente texto) en la investigación social (Rodríguez. P.G. 1997).

iii "El conocimiento es más científico cuantas más perspectivas contempla" (V. de Gialdino, 1995 com. pers.) iv En un trabajo ya clásico en el que exponía entre otras técnicas, las de análisis conceptual, Hebe Vessuri definía cinco posibles tipos de casos que pueden resultar útiles identificar al trabajar con la GTM. Ellos eran: Casos paradigmáticos: aquellos que se reconocen fácilmente como instancias del concepto en cuestión. "Algo de lo que podemos decir: 'Bueno, si eso no es un ejemplo de tal y tal cosa entonces no hay nada que los sea'".

Casos contrarios: Aquellos de los cuales se puede afirmar que no son instancias del concepto considerado.

Casos relacionados: Otros conceptos asociados, similares, complementarios, englobantes, que, juntos con el estudiado, forman un paradigma o red o constelación de conceptos,.

Casos límites: Aquellos de los que no estamos seguros dónde clasificarlos.

Casos inventados: Diseñados por el investigador para poner a prueba la clasificación, a fin de verificar que ésta cumpla con las condiciones de una buena clasificación científica.

Lic. P. Gustavo Rodriguez

 [VOLVER](#) /  [ARRIBA](#) /  [IMPRIMIR](#)

INICIO	NAyA en INTERNET	NAyA en CDROM	EQUIPO NAYA
------------------------	----------------------------------	-------------------------------	-----------------------------

Noticias de Antropología y Arqueología:
Congresos en CDROM - Versión 2.2 - ON-LINE
Equipo NAYa - 2002
<http://www.naya.org.ar/congresos/>
info@naya.org.ar

