

COMISIÓN DE LA VERDAD Y RECONCILIACIÓN
PERÚ

ANEXO 6

BASE DE DATOS PARA LA INVESTIGACIÓN CUALITATIVA (BDI)

INTRODUCCIÓN

La idea básica de la Base de Datos para la Investigación Cualitativa (BDI) de la CVR fue unificar el procesamiento de la información recolectada de un conjunto de áreas sobre la base de un mismo programa informático y de un mismo libro de códigos.

Por su parte, la Base de Datos de Sistemas de Información está construida a partir de resúmenes de testimonios de víctimas o familiares de víctimas en los que básicamente se dan datos en torno a crímenes o violaciones a los derechos humanos específicos. Sin embargo, existe una vasta información recopilada por diversas áreas en la que se puede encontrar ese tipo de datos pero también otro tipo de material. Se trata, en buena parte, de material de investigaciones y de diálogos entre investigadores e informantes calificados sobre diversos temas de interés para la CVR. Allí se recoge la interpretación de la población y de diversos actores¹ sobre el proceso de violencia, sus antecedentes y sus consecuencias en diversas zonas del país, así como sus propuestas para el futuro y para una posible reconciliación. Este material incluye entrevistas en profundidad, grupos focales, talleres, libretas de campo y reseñas bibliográficas.

En la segunda mitad de 2002 se efectuaron diversas reuniones a las que asistieron responsables de diversas áreas de la CVR que habían recopilado este tipo de información o que requerían analizarla.² Se hizo evidente la necesidad de intercambiar información entre los distintos grupos de trabajo, ya que lo que debía analizar cada área iba más allá de la propia información recopilada y podía ser encontrada en los trabajos de otros grupos de trabajo. Del mismo modo, era evidente que el intercambio de los informes elaborados no era suficiente por dos razones: en primer lugar, muchos de éstos iban a estar disponibles en una época tardía, cuando los demás ya no tuvieran tiempo de incorporarlos a su propio trabajo; y, en segundo lugar, si bien los resultados de dichos informes interesaban específicamente al área que los elaboró, existía un gran volumen de esa información podía ser de utilidad a otras áreas.

La solución fue elaborar un sistema unificado de procesamiento de fuentes primarias al cual cada área tuviera acceso. Antes de esas reuniones, varios grupos habían recopilado información pero todavía no la habían procesado; otros, lo habían hecho de modos diversos. Ante el gran volumen de información que se debía procesar, la única alternativa fue la de utilizar un *software* especializado en el procesamiento de información de tipo cualitativo. Por suerte, varias de las áreas que habían procesado información por separado habían recurrido a un mismo programa informático, el Atlas ti de Scientific Software. Debido a que el programa permitía la fusión de esos diversos proyectos y un trabajo relativamente rápido de conversión de códigos, se decidió utilizarlo para almacenar las diversas informaciones en un archivo o base de datos posteriormente llamado Base de Datos para la Investigación Cualitativa (BDI).³

¹ Entre los cuales están miembros o ex miembros de grupos subversivos o de las fuerzas del orden, ronderos, representantes de instituciones, de organizaciones sociales o partidos políticos, de la Iglesia Católica u otros grupos religiosos, etc.

² Estudios en Profundidad, Secuelas y Reparación, Reconciliación, Salud Mental, Proceso Nacional, Historias Regionales, Género, Equipo Jurídico, Sistemas de Información, Unidad de Documentación, Audiencias Públicas y Comunicación.

³ Para una mayor información sobre el Atlas ti de Scientific Software, puede ingresar a la siguiente página web: <www.atlasti.de>.

Del mismo modo, en esas reuniones se decidió crear una unidad que se dedicaría a la tarea de elaborar la base de datos y coordinar con diversas áreas el trabajo.⁴ En un principio, aquéllos que procesaban la información eran los propios investigadores que la recolectaron o dirigieron los estudios. Sin embargo, existía un amplio material que no había sido trabajado en Atlas ti; en consecuencia, se optó por adiestrar grupos de voluntarios para que procesaran la información que faltaba. Paralelamente, se contrató un grupo reducido de profesionales que se dedicaría a la tarea de controlar este trabajo. Los voluntarios eran en su inmensa mayoría estudiantes de ciencias sociales y algunos profesionales jóvenes.⁵

En esas reuniones se decidió, además, que se debía procesar la información de las audiencias públicas y algunas de las entrevistas hechas en penales. Este trabajo también estuvo a cargo de voluntarios. Cabe señalar que las entrevistas o testimonios de penales también se encuentran en la Base de Datos de Sistemas de Información pero no son las únicas que se encuentran en ambas bases de datos. El equipo de Salud Mental de la CVR también trabajó con testimonios que se encuentran en Sistemas de Información, pero que fueron procesados simultáneamente en Atlas ti y fueron incluidos en la BDI. En ambos casos la información recibe un tratamiento diferente en cada una de las bases de datos.

El primer paso del trabajo de construcción de la BDI fue elaborar un libro de códigos común compartido por las diversas áreas. En Atlas ti los archivos de bases de datos son denominados unidades hermenéuticas (*hermeneutic units*). Se decidió que los trabajos de los diversos equipos serían procesados en dos formatos. Cada grupo haría un trabajo con el libro de códigos compartido para la unidad hermenéutica que unificaba todos los trabajos y también podría elaborar unidades hermenéuticas propias a sus intereses particulares utilizando códigos específicos. Estos códigos no serían compartidos por todos, pero entrarían en una de las casillas del libro de códigos común. Por ejemplo, si en uno de los estudios en profundidad se tenía la necesidad de poner como códigos nombres propios de personas (Hildebrando Pérez (líder del PCP-SL), Epifanio Curitumay (autoridad local), comandante *Camión* (mando militar), etc.), cuando pasaran al libro de códigos unificado, las citas codificadas con esos nombres se ubicarían en categorías más amplias (líderes del PCP-SL, autoridades locales, mandos militares, etc.).

La elaboración del libro de códigos común significó una tarea que involucró a diferentes áreas. A partir de una plantilla inicial, se pidió a diversos investigadores sugerencias sobre cambios de códigos, elaboración de nuevos códigos y precisión en la definición de los conceptos. Hubo reuniones semanales en las que se presentaban nuevos cambios a la plantilla inicial. Estos cambios se elaboraban sobre la base de las sugerencias enviadas durante el transcurso de la semana por el personal de diversas áreas. En un determinado momento, cuando se detuvieron las sugerencias de cambios, se decidió cerrar la discusión grupal y dejar a la unidad encargada de elaborar la base de datos la tarea de hacer nuevos cambios. Los cambios que se realizaron en los meses siguientes no fueron numerosos.

CARACTERÍSTICAS DE LA BASE DE DATOS

Documentos primarios

La base de datos es, estrictamente, una herramienta de análisis cualitativo, aunque puede tener también otras funciones. Es una forma de almacenar textos, en su gran mayoría transcripciones completas de entrevistas, grupos focales y talleres. Estos textos aparecen en su integridad en el *software*. Cada uno de ellos tiene un número y un identificador o nombre y un acceso independiente. En el software los textos son llamados documentos primarios. En la base de datos, existe un espacio en el cual se colocan comentarios a cada uno de los textos. En ese espacio se pueden encontrar datos de los entrevistados y un resumen de los puntos significativos del documento.

Citas

El primer trabajo que se hace con ellos es la selección de pasajes significativos (citas). Éstas mayormente se asocian a códigos aunque puede haber citas asociadas a memos o citas independientes. Una misma cita puede asociarse a uno o

⁴ Posteriormente, se decidió que esa unidad formara parte del área Núcleo de Informe Final (NIF).

⁵ El número total de voluntarios que han colaborado en este trabajo supera las cincuenta personas.

varios códigos o memos. Cada cita también tiene un número específico. Si bien no existe la obligación de citar todos los pasajes que contiene un texto, en su gran mayoría éstos están citados por lo que el número de citas es considerable como se verá más adelante.

Códigos

Señalamos anteriormente que la edificación del libro de códigos de la base de datos fue un trabajo interáreas. Finalmente, el número de códigos que resultó del proceso fue de 328. En nuestra base de datos, los códigos se diseñaron como etiquetas de tipo descriptivo que abarcan diferentes temas generales tratados por las diversas áreas de la CVR. En principio, no se intentó ingresar en ese espacio conceptos muy elaborados o abstractos, tampoco hipótesis de trabajo. Éstos se han colocado en los memos, de los que trataremos más adelante. A continuación, haremos una lista de las grandes categorías en las que están agrupados los códigos:

- Actores (autoridades estatales, actores económicos, partidos políticos, comunidades campesinas o indígenas, instituciones estatales, fuerzas del orden, grupos paramilitares, grupos religiosos, grupos subversivos, líderes o dirigentes, medios de comunicación, grupos étnicos o de género, ONG, organizaciones de derechos humanos, organizaciones campesinas o indígenas, organizaciones de base, sector educación, rondas/CAD, sindicatos/gremios, narcotraficantes)
- Condición de víctimas (muertos, requisitoriados, arrepentidos, refugiados, etc.)
- Contextos y procesos sociales (conflictos, relaciones, vacíos de poder, militarización, relaciones de género, desplazamiento poblacional, etc.)
- Crímenes y violaciones a los derechos humanos (ejecuciones extrajudiciales, desaparición forzada, tortura, etc.)
- Ámbito temporal (período anterior a la violencia, período posterior a la violencia y situación actual)
- Ámbito espacial (nivel familiar, nivel comunal, nivel regional y nivel nacional)
- Ámbito social (economía, política, religión, etc.)
- Secuelas (código transversal cruzado generalmente con otros como economía, política, etc.)
- Expectativas y reconciliación (acceso a infraestructura y servicios, justicia, reparación, perdón, olvido, etc.)
- Justificaciones y motivaciones de responsables (autocrítica, justificación de actos, negación de responsabilidades, etc.)
- Estrategias y *modus operandi* (amenazas, hostigamiento, cupos, etc.)
- Características, estructuras y organización interna (adocctrinamiento, reclutamiento, preparación militar, organización, etc.)
- Presencia de los actores (bases militares, campamentos, columnas, etc.)
- Percepciones y opiniones (generalmente asociadas a otros códigos como actores, procesos, etc.)
- Respuestas de la población (apoyo, rechazo, denuncia, etc.)
- Salud física (heridos lisiados, enfermedades, salud física en general)
- Salud mental (miedo, enfermedades psicosomáticas, problemas graves de salud mental, etc.)

Un principio para la formación de los códigos fue que el número debía ser el mínimo posible para abarcar la mayor cantidad de temas. Esto en la práctica quiere decir que en lugar de utilizar códigos sobre temas muy específicos se hagan códigos más elementales, pero que puedan ser combinables. Así pues, en lugar de poner en una cita un código como «Incursiones de FFAA en comunidades nativas» se ponen varios códigos que expresen la misma idea: «FFAA», «Incursiones» y «Comunidades nativas». Cada uno de estos puede ser combinable con muchos otros en otras citas. Si para cada tema específico se pusiera un código, la base de datos constaría de cinco mil o diez mil códigos, con lo que sería prácticamente inmanejable y poco útil.

Memos

Éstos constituyen un elemento importante de la base de datos. Memo quiere decir en principio anotación. Allí sí se pueden colocar ideas, comentarios, hipótesis de trabajo, conceptos abstractos o reflexiones metodológicas. Los memos también se asocian a citas. Éstas deberían reflejar o discutir lo que sugiere el memo. En el caso de los memos, los

investigadores o codificadores sí tienen libertad para resaltar lo que consideran relevante sin ceñirse a la «camisa de fuerza» (o número delimitado) de los códigos.

Familias

Otro elemento importante de la base de datos son las familias. Dentro del Atlas ti una familia es la agrupación de elementos que comparten una característica en común. En nuestro caso hemos elaborado familia de textos (documentos primarios) y de códigos.

Familias de textos

Las familias no son excluyentes entre sí. Cada uno de los textos ingresados pertenece a varias familias según las características que tenga. La creación de familias permite la realización de búsquedas específicas dentro de un grupo de textos determinado. Incluso, se pueden combinar las familias en las búsquedas como veremos más adelante. Los criterios de agrupación han sido los siguientes:

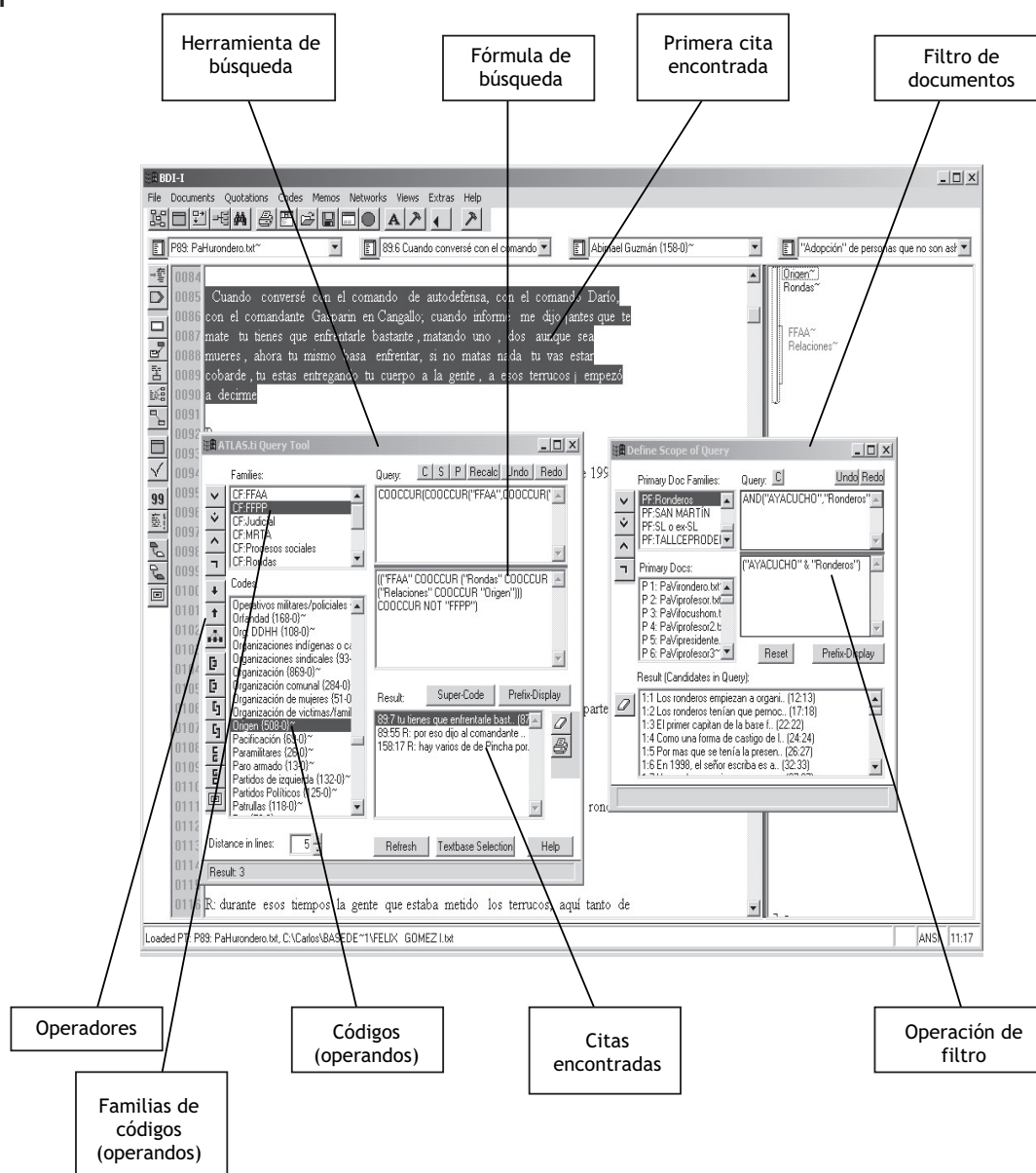
- Origen dentro de la CVR. Si se trata de un grupo de entrevistas de un estudio particular o del trabajo de un área específica. Allí tenemos, por ejemplo, la familia «Estudio Arrepentimiento-Aries» o la familia «Audiencias Públicas» o «Entrevistas en penales».
- Áreas geográficas que se mencionan. El criterio no es el origen geográfico de los entrevistados ni el lugar en donde se encuentran, sino los lugares de los que hablan. Se consideran los departamentos. Allí tenemos, por ejemplo, la familia «Huanuco» o «Puno».
- Género de los entrevistados o participantes. Cada texto se coloca, ya sea en la familia «Hombres» o en la familia «Mujeres» dependiendo del género de los entrevistados, no de las personas a las que se refieren éstos. Por ejemplo, si una mujer habla mayormente de su hijo desaparecido, el texto va a pertenecer siempre a la familia «Mujeres».
- Otras características de los entrevistados o participantes. Enumeraremos las diferentes familias dentro de este rubro: «Autoridades» (incluye ex autoridades), «Desplazados», «Estudiantes-ex estudiantes», «FFOO-ex FFOO» (miembros o ex miembros de las fuerzas del orden), «MRTA-ex MRTA» (miembros o ex miembros del MRTA), «Profesores» (incluye a ex profesores), «Ronderos» (miembros o ex miembros de los comités de autodefensa o rondas campesinas) y «SL-ex SL» (miembros o ex miembros del PCP-SL).

Familias de códigos

Los códigos también comparten una característica común: pueden ser agrupados en familias. Esto se decide de acuerdo con los intereses de búsqueda de información. Simplemente se crean familias de códigos para simplificar las búsquedas. Por ejemplo, si queremos hacer una búsqueda de citas correspondientes a diversos códigos relacionados a salud mental, lo que hacemos es crear una familia que contenga esos códigos y buscar en esa familia y ya no en cada uno de esos códigos. En la base de datos existen algunas familias de códigos básicamente ligadas a agrupaciones de actores o de temas específicos. Las familias que existen son las siguientes:

- Fuerzas Armadas (Fuerzas Armadas, Marina, Ejército, etc.)
- Fuerzas Policiales (Fuerzas Policiales, Guardia Civil, DINCOTE, etc.)
- Judicial (Acción legal, denuncia, Poder Judicial, etc.)
- MRTA (MRTA, UDP, Pueblo en Marcha, etc.)
- Procesos sociales (militarización, organización comunal, escuelas, etc.)
- Rondas o comités de autodefensa (rondas, rondas urbanas, etc.)
- Salud mental (alcoholismo, problemas graves de salud mental, efectos en autoestima e identidad, etc.)
- PCP-SL (Sendero Luminoso, Feliciano, Abimael Guzmán, etc.)
- Violaciones de los derechos humanos (ejecuciones extrajudiciales, desaparición forzada, tortura, etc.)

Gráfico 1



Herramientas de búsqueda

El motivo principal para la creación de la Base de Datos para la Investigación Cualitativa fue servir como instrumento importante en el análisis de la información recolectada por la CVR. Ésta fue finalmente usada por los diversos equipos o consultores dedicados a elaborar documentos que formarían parte del Informe Final. Los textos nos remiten no sólo a los procesos y consecuencias vividos por la población y diversos actores, sino también a su propia interpretación de los mismos así como a su visión del futuro y sus expectativas. La base de datos representa un instrumento que nos permite llegar de manera ordenada y sistemática a lo que todos tenían que decir sobre las grandes y pequeñas preguntas que se planteó la CVR en un principio y profundizó a lo largo de su proceso de trabajo.

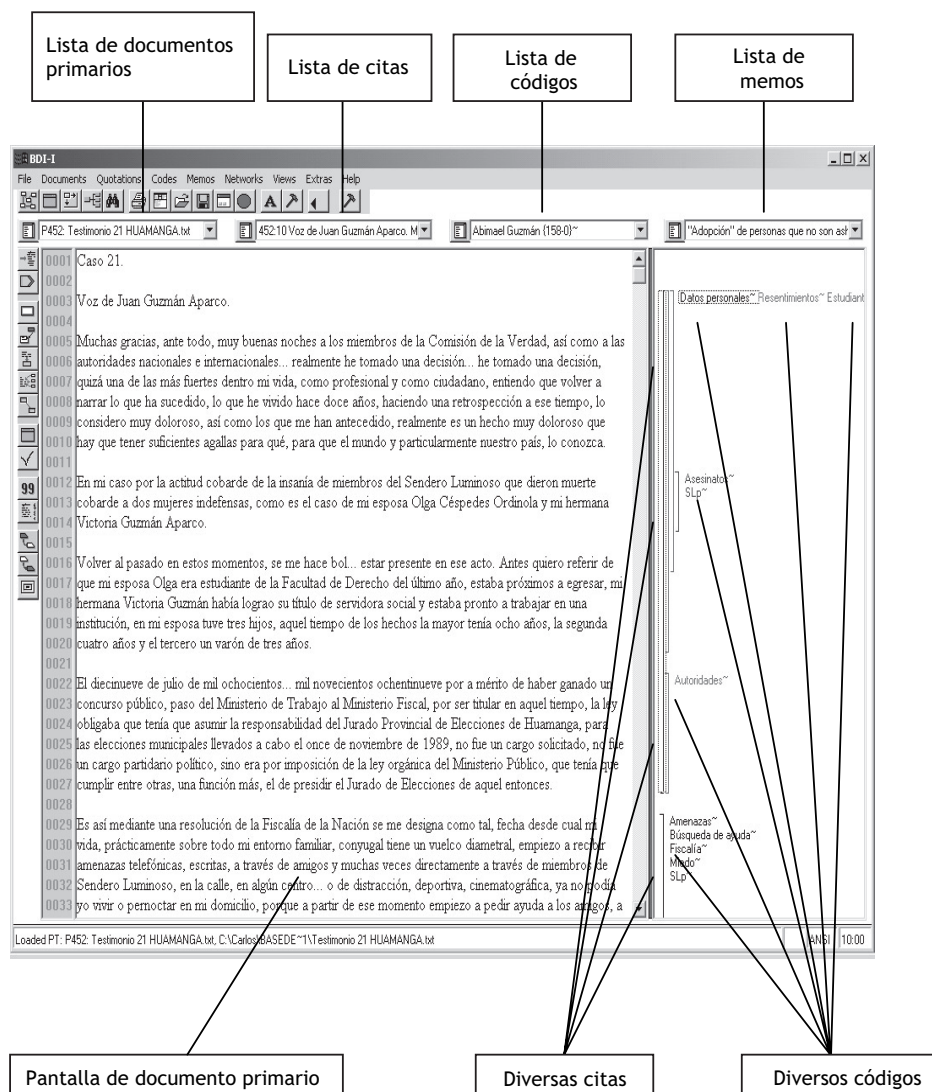
Búsqueda por códigos y familias de códigos

Dentro del Atlas ti existe una herramienta llamada Query tool que permite el planteamiento de fórmulas complejas de búsqueda por códigos y familias de códigos y que a su vez permite realizar filtros en familias de documentos. Por ejemplo, si nos interesa indagar exclusivamente lo que piensan los ronderos de Ayacucho sobre las relaciones entre FFAA y población para la formación de Comités de Autodefensa, pero donde no haya habido intervención de las FFPP, podemos elaborar una fórmula en ese sentido. Utilizamos las familias de códigos siguientes: «FFAA», «FFPP» y

«Rondas» y los códigos siguientes «Origen» y «Relaciones». Pedimos al programa buscar las citas codificadas con todos estos elementos (co-ocurrencia): «FFAA» (varios códigos), «Rondas» (varios códigos), «Origen» y «Relaciones». Luego le pedimos al programa no incluir las citas que tengan alguno de los códigos de la familia «FFPP». La fórmula está concluida en ese momento pero nos va a arrojar citas sobre lo que dijeron todos los actores de todos los rincones del Perú donde se recolectó información. Dentro del Query tool existe un espacio en el cual podemos filtrar los documentos que nos interesan más precisamente. De ese modo, ingresamos en ese espacio y allí elaboramos otra fórmula: le pedimos al programa que sólo nos proporcione citas de la fórmula previa, pero solamente de los documentos que pertenezcan al mismo tiempo a la familia de textos «Ayacucho» y a la familia de textos «Ronderos». De ese modo, el programa nos va a dar exactamente el tipo de citas que nos interesan. Si queremos, podemos guardar en un archivo o imprimir un reporte de la búsqueda, el cual incluye el conjunto de las citas completas con sus identificadores respectivos de modo que podemos ubicar los textos de donde provienen y su lugar específico dentro de esos textos. Si queremos mayor información sobre esos textos, podemos ingresar a las reseñas por texto que se ubican en otro espacio del programa. Si queremos explorar más el sentido de la cita o su contexto, podemos navegar a través del documento del cual proviene.

Lo anterior no es más que una muestra de las diversas posibilidades de búsqueda. En ese ejemplo sólo hemos utilizado tres tipos de operadores (co-ocurrencia de varios códigos, negación de la familia «FFPP» e intersección de las dos familias de textos) de un total de quince operadores que permiten otros tipos de búsquedas. Al mismo tiempo se utilizaron solamente cinco operandos («FFAA», «FFPP», «Rondas», «Origen» y «Relaciones») cuando podría utilizarse un número mucho mayor de ellos, diez o veinte o más. De otro lado, solamente se trabajó con dos familias de textos cuando hubiéramos podido utilizar más.

Gráfico 2



Búsqueda textual

Por diversas razones un investigador puede estar interesado en la ocurrencia de una palabra o un grupo de palabras en el conjunto de los textos. El programa tiene un espacio en donde se puede realizar ese tipo de búsquedas. Esto es útil por ejemplo cuando se quiere encontrar el nombre de personas o un grupo de personas incluyendo sus seudónimos o cuando se quieren encontrar menciones de fechas específicas. Este tipo de búsquedas son, sobre todo, comunes en investigaciones específicas sobre crímenes o violaciones a los derechos humanos o sus posibles perpetradores o víctimas. De otro lado, es posible que el investigador quiera encontrar ciertas palabras referentes a ciertos temas que le interesa investigar. Por ejemplo, diferentes insultos referidos a la condición étnica o a la condición de mujer, por mencionar algunos temas. Si se desea acotar la búsqueda, el programa también permite filtrar las familias de documentos primarios que nos interesa explorar más específicamente.

VOLUMEN DE LA INFORMACIÓN

Al momento de la redacción de este documento todavía se seguía ingresando información a la base de datos aunque ésta ya se encontraba en la parte final de elaboración. Por este motivo no nos atrevemos a dar cifras definitivas sobre el número exacto de codificaciones realizadas y el número de memos, aunque podemos hacer un cálculo bastante aproximado: en el caso de las codificaciones se calcula ingresar alrededor de 145,000 y en el caso de memos el número calculado es de 1,300. De otro lado, la cantidad de documentos que han sido ingresados es de 1984. Los documentos varían en su número de páginas, pero se calcula que en total están almacenadas alrededor de 35,000.