

Comisión de Estadística
Trigésimo octavo periodo de sesiones
27 de febrero al 2 de marzo de 2007

Traducción no oficial

Punto 3(j) de la agenda provisional
Puntos para debate y adopción de decisiones:
Problemas de gestión en las oficinas nacionales
de estadística: Acceso a los Microdatos

Principios y Directrices para la Gestión de la Confidencialidad Estadística y el Acceso a los Microdatos

Nota: Este documento se basa en el informe proporcionado por el Grupo de Trabajo de la UNECE sobre el Acceso a los Microdatos. Se ha ampliado para incluir los puntos de vista de otras regiones.

Contenido

I.	INTRODUCCIÓN.....	3
II.	¿POR QUÉ LAS OFICINAS NACIONALES DE ESTADÍSTICA DEBERÍAN APOYAR A LA COMUNIDAD DE INVESTIGADORES?.....	4
III.	PRINCIPIOS FUNDAMENTALES.....	6
IV.	LA LEGISLACIÓN DE APOYO Y LA RESPONSABILIDAD DE LAS ONEs COMO GUARDIANES DE LA CONFIDENCIALIDAD.....	7
V.	MÉTODOS PARA APOYAR A LA COMUNIDAD DE INVESTIGADORES.....	8
VI.	MANEJO DE LAS TENSIONES ENTRE LAS OFICINAS NACIONALES DE ESTADÍSTICA Y LOS INVESTIGADORES.....	15
VII.	CUESTIONES ADMINISTRATIVAS VINCULADAS CON EL ACCESO A MICRODATOS.....	19
VIII.	ALGUNAS CUESTIONES ESPECIALES.....	20

I. INTRODUCCIÓN

1. Históricamente, la protección de la confidencialidad ha sido principalmente un tema nacional. Sin embargo, ahora también se está convirtiendo en un tema internacional, por las siguientes razones: se está incrementando la difusión de datos a través de Internet; una gran parte de la colaboración entre los miembros de la comunidad de investigadores es a nivel internacional; y las comparaciones internacionales son de mayor interés. En consecuencia, existe una creciente demanda de microdatos de los países. El acceso a tales datos es no sólo de interés para los investigadores académicos, sino también para las agencias internacionales, las cuales desean utilizar los microdatos para propósitos de investigación, particularmente para las comparaciones internacionales. Tales estudios son generalmente de interés para los países participantes, pero hasta ahora el acceso a los microdatos de los países está frecuentemente limitado por el temor de que la protección a la confidencialidad no pueda ser garantizada.

2. Una pregunta clave que este informe intenta abordar es si es posible llegar a acuerdos a nivel internacional sobre algunos principios comunes para la difusión de los microdatos. Esta pregunta ha sido considerada por la Conferencia de Estadísticos Europeos (CES), en sus sesiones desde 2003, cuando se constituyó un Grupo de Trabajo para producir un anteproyecto sobre un conjunto de principios y directrices vinculadas que fueran lo suficientemente generales para ser adaptados fácilmente por los países que participan en la Conferencia de Estadísticos Europeos. El informe original está disponible en el sitio Web de la CES.

3. Gran parte de este informe se basa en el trabajo del Grupo de Trabajo mencionado, el cual ha sido revisado para tomar en cuenta las circunstancias de las Oficinas Nacionales de Estadística (ONEs) en otras regiones¹. El informe reconoce que el apoyo a la investigación es una función importante de las Oficinas Nacionales de Estadística (ONEs), y que, generalmente, las ONEs podrían hacer mucho más para satisfacer las necesidades de los usuarios de acceso a los microdatos. Asimismo revisa las maneras en las que actualmente se ponen los microdatos a disposición de los usuarios nacionales e internacionales, y esboza procedimientos para salvaguardar la confidencialidad de los informantes individuales en los microdatos. Su énfasis primordial está en la administración de riesgos, y tiene tres objetivos principales:

- (a) Facilitar un mejor acceso a los microdatos por parte de la comunidad de investigadores cuando los proyectos de investigación lo ameriten;
- (b) Permitir a los países mejorar sus procedimientos para proporcionar el acceso a los microdatos; y
- (c) Fomentar una mayor uniformidad en las prácticas de acceso a los microdatos entre los países.

4. Aunque se espera que estos principios propuestos conduzcan a una mayor uniformidad en los enfoques de trabajo, algunos países podrían no tener los sistemas o los recursos para mantener las salvaguardas necesarias de la confidencialidad. Las medidas precisas para el acceso a los microdatos varían de un país a otro, dependiendo de factores tales como el contexto histórico, el nivel de desarrollo estadístico, la legislación, las actitudes públicas y la capacidad para apoyar a la comunidad de investigadores. Los estudios de caso sobre diversas opciones descritas en este informe se enumeran en el anexo. Varios países en vía de desarrollo están considerando actualmente una revisión de sus legislaciones estadísticas y se espera que ante tal oportunidad, se llegue a precisar o cambiar las legislaciones que rigen la confidencialidad o el acceso a los microdatos, los estudios de caso relacionados pueden ser útiles en la determinación de los cambios apropiados.

¹ Los cambios incluyen la abreviación de algunas secciones, la revisión de algunos párrafos y una sección sobre otras formas de acceso a los datos, los cuales son específicos para algunos países en vía de desarrollo fuera de Europa.

Uso de términos clave

5. Oficinas Nacionales de Estadística (ONEs): No obstante que en este informe la referencia es hacia las ONEs, en muchos países, particularmente aquéllos con sistemas descentralizados, existen varios productores de estadística. La referencia a las ONEs debe ser leída como aplicable a todos los productores de la estadística oficial.

6. La comunidad de investigadores: Incluye aquellos que trabajan en instituciones académicas, así como a los investigadores que trabajan en organizaciones no gubernamentales y en las agencias internacionales. Además, algunos de los investigadores que requieren el acceso a los microdatos podrían trabajar al interior de las agencias e instituciones financiadas por el gobierno. Todos estos investigadores están considerados como parte de la “comunidad de investigadores”, aún cuando los temas pertinentes pudieran tener cierta variación entre los diversos elementos de la comunidad de investigadores.

7. El término microdatos tal como es utilizado en este informe, se refiere a los datos acerca de un individuo, hogar, negocio u otra entidad. Pueden ser datos recopilados directamente por la ONE u obtenidos por otras fuentes, por ejemplo los registros administrativos.

8. Anonimizados: El término implica, que no sólo los nombres y direcciones han sido eliminados de los registros individuales del conjunto de microdatos, sino que se han tomado otras medidas (por ejemplo, la compresión de detalles geográficos, edad, lugar de nacimiento y ocupación) para garantizar que la identificación de individuos sea altamente improbable.

II. ¿POR QUÉ LAS OFICINAS NACIONALES DE ESTADÍSTICA DEBEN APOYAR A LA COMUNIDAD DE INVESTIGADORES?

9. En casi todos los países, las estadísticas oficiales son recopiladas no sólo para los gobiernos, sino también para el uso de la comunidad. La comunidad de investigadores desempeña un papel importante en la estimulación del análisis de las políticas y su debate. Esto es importante particularmente para las democracias donde las estadísticas oficiales se pueden utilizar para evaluar la eficacia de las políticas y de los programas del gobierno; ellas proporcionan un espejo para la sociedad, y así, es de interés público que los datos estén siendo analizados y presentados desde diversas perspectivas. Más aún, si de esta manera los datos de las encuestas son más utilizados, esto puede proporcionar un nivel adicional de protección contra las reducciones del presupuesto para los programas estadísticos.

10. Además, proporcionar al investigador el acceso a los microdatos puede ser una manera de extraer un valor adicional al alto costo del levantamiento de las estadísticas oficiales, ampliando la gama de productos estadísticos, y obteniendo conocimientos valiosos sobre la calidad de los datos y de cómo se pudieran mejorar los estudios estadísticos. En cambio, si los investigadores no tienen acceso a los datos estadísticos oficiales relevantes, existe el riesgo de que pudieran obtener sus propios datos, los cuales estarían a menudo referidos a muestras más pequeñas y con una calidad inferior respecto a la de las encuestas oficiales. Estos levantamientos generan costos adicionales para ambos, el recolector de datos y el informante, aumentando la carga de requerimientos de información impuesta a la comunidad.

La perspectiva de la Oficina Nacional de Estadística

11. Mientras que las ONEs están reconociendo cada vez más el valor agregado que la comunidad de investigadores le puede proporcionar al levantamiento y procesamiento de los datos de la ONE, y

aprecian la importancia de apoyar tales esfuerzos, existe un número de asuntos con los que la ONE debe lidiar con respecto al acceso a los microdatos. Además de los asuntos de confianza, lo cuál es fundamental para la eficacia de la ONE, existen las cuestiones de la autoridad legal, la calidad y los costos.

12. Las ONEs deben mantener la confianza de los informantes para que continúen cooperando en sus recopilaciones de datos. La protección de la confidencialidad es el elemento clave de esa confianza. Otra preocupación de las ONEs es si tienen suficiente autoridad, ya sea a través de un mandato legal u otra forma de autorización, para brindar a los investigadores el acceso a los microdatos.

13. Algunas ONEs están preocupadas porque la calidad de sus microdatos pudiera no ser bastante buena para su difusión. Mientras que la calidad podría ser suficientemente precisa para sustentar las estadísticas agregadas, pudiera haber inconsistencias entre los resultados de una investigación basada en los microdatos y los datos agregados publicados. En algunos casos, se pueden haber hecho ajustes a las estadísticas agregadas en la etapa de edición de salida, sin la correspondiente enmienda a los microdatos. Otras pueden preocuparse de que algunos investigadores no sean suficientemente cuidadosos en preservar la validez de sus análisis, y que se aventuren más allá del nivel de desagregación que el diseño de la muestra puede soportar.

14. Las ONEs pueden estar también preocupadas por los costos, no únicamente para la creación y documentación de archivos de los microdatos, sino también para la creación de herramientas de acceso y de salvaguarda, y para la atención y autorización de las solicitudes hechas por la comunidad de investigadores; los nuevos usuarios de los archivos de datos necesitan ayuda para navegar por las complejas estructuras de archivos y las definiciones cambiantes. Aunque los costos son sufragados por las ONEs, generalmente éstas no están provistas de un presupuesto complementario para hacer este trabajo adicional. Además, los investigadores no tienen a menudo el financiamiento para contribuir sustancialmente con estos costos.

15. Conceder el acceso internacional presenta en ocasiones problemas más serios que el acceso al interior del país debido a las restricciones legales. Pero para las ONEs de muchos países en vía de desarrollo, los recursos y las capacidades para proporcionar productos no estandarizados o analíticos, no existen o, en el mejor de los casos, son limitados. Por ende, esta clase de productos podrían ser producidos por los investigadores que radican fuera del país, de lo contrario, hay un serio riesgo de que los datos de las encuestas, recolectados a menudo a un gran costo, sean subutilizados.

La perspectiva de la comunidad de investigadores

16. Desde la perspectiva de la comunidad de investigadores, el apoyo a la investigación basada en microdatos debe ser un componente importante de cualquier sistema estadístico oficial. Los beneficios incluyen:

- (a) Los Microdatos le permiten a los diseñadores de políticas, plantear y analizar preguntas complejas.
- (b) El acceso a los microdatos permite a los analistas calcular y analizar efectos marginales en vez de únicamente los de promedio.
- (c) Un acceso ampliamente disponible a los microdatos, hablando en términos generales, permiten replicar importantes investigaciones.
- (d) El acceso a los microdatos con fines de investigación, y la retroalimentación resultante, puede contribuir a la mejora de la calidad de los datos.
- (e) La gama de productos derivados de los trabajos estadísticos y por lo tanto, el valor total del dinero invertido en esas actividades se incrementa.

17. Las ONEs pueden desempeñar un papel muy útil en el abastecimiento de una fuente de datos aceptado y con autoridad, así como de alta calidad, para la comunidad de investigadores. En ausencia de una fuente así de confiable, los investigadores tendrían que utilizar diversos conjuntos de datos para analizar tópicos específicos. Los investigadores sin embargo sienten que las ONEs han sido en líneas generales demasiado conservadoras en cuanto al acceso que proporcionan a los microdatos. Destacan que ellos no están interesados en identificar a individuos y por lo tanto los costos del aseguramiento de los microdatos podrían ser innecesarios.

III. PRINCIPIOS FUNDAMENTALES

El sexto principio fundamental de las Naciones Unidas sobre Estadística Oficial es muy claro respecto a la confidencialidad estadística:

“Los datos individuales recogidos por las agencias estadísticas para la compilación estadística, ya sea que se refieran a las personas físicas o legales, deberán ser estrictamente confidenciales y utilizados exclusivamente para propósitos estadísticos.”

Cualquier principio para el acceso a los microdatos debe ser consistente con este Principio Fundamental.

19. Los siguientes principios deberán ser utilizados para manejar la confidencialidad de los microdatos.

- (a) Principio 1: Es procedente, siempre y cuando se proteja la confidencialidad, que los microdatos recolectados para elaborar la estadística oficial, sean utilizados en el análisis estadístico para apoyar a la investigación.
- (b) Principio 2: Los microdatos podrán ser proporcionados sólo para propósitos estadísticos.
- (c) Principio 3: El acceso a los microdatos debe ser consistente con las regulaciones legales y otras que fueran necesarias para garantizar que la confidencialidad de los microdatos proporcionados esté protegida.
- (d) Principio 4: Los procedimientos para el acceso del investigador a los microdatos, así como sus aplicaciones y los usuarios de microdatos deben ser transparentes, y estar disponibles al público.

20. El Principio 1 no constituye una obligación de proporcionar microdatos; y hacerlos disponibles para la investigación no está en contradicción con el sexto Principio Fundamental de la ONU siempre y cuando no sea posible identificar los datos que se refieren a un individuo o a una entidad. Puede haber otras preocupaciones por ejemplo, la calidad, que hagan inadecuado proporcionar el acceso a los microdatos. O puede haber ciertas personas o instituciones a quienes sería inadecuado proporcionar los microdatos. La ONE debe ser la entidad que decida si es pertinente o no proporcionar microdatos.

21. Por el Principio 2, si los microdatos no son utilizados con fines estadísticos o analíticos, entonces el acceso no debe ser otorgado. El objetivo del uso estadístico o analítico es el de derivar estadísticas que se refieran a un grupo ya sea de personas o entidades legales, y que pueda ser llevado a cabo por los investigadores y otras agencias estadísticas. Esto incluye la compilación de agregaciones estadísticas en varias formas y de distribuciones estadísticas; el ajuste de modelos estadísticos, o el análisis de diferencias estadísticas entre las subpoblaciones. Estas aplicaciones forman parte del meollo de las investigaciones que suministran información para la toma de decisiones en beneficio de la sociedad; y debe ser diferenciadas de las aplicaciones administrativas. En cambio, el objetivo del uso administrativo de datos es derivar la información sobre una persona particular o una entidad legal para tomar una decisión que pueda acarrear una ventaja o un daño al individuo, por ejemplo, peticiones a nivel individual de datos por orden judicial. Por el interés de la confianza pública en el sistema estadístico oficial, estas

peticiones, por ser contrarias a este principio, tendrían que ser rechazadas. De no ser así, la cooperación de los informantes disminuirá y la calidad de la estadística se vería afectada. Los comités de ética u otros arreglos similares podrían apoyar en situaciones en donde existiera incertidumbre acerca de proporcionar el acceso o no a los microdatos.

22. Con respecto al Principio 3, idealmente las medidas legales para proteger a la confidencialidad deben estar en vigor antes de que cualquier microdato sea proporcionado. Sin embargo, las disposiciones legales tienen que ser complementadas con medidas administrativas y técnicas para regular el acceso a los microdatos y para asegurar que los datos individuales no puedan ser identificados. La existencia y la visibilidad de tales arreglos (ya sea en la ley o en las regulaciones complementarias, ordenanzas, etc.) son necesarias para aumentar la confianza pública respecto a que los microdatos serán utilizados apropiadamente.

23. El Principio 4 es importante para aumentar la confianza pública en que los microdatos estén siendo utilizados adecuadamente y para demostrar que las decisiones acerca del acceso a los microdatos están siendo tomadas sobre una base objetiva. Estará a cargo de las ONEs la decisión acerca de cómo y cuáles microdatos podrían ser proporcionados. Pero sus decisiones deben ser transparentes. El sitio Web de las ONEs es una manera eficaz de asegurar su cumplimiento y también para proporcionar la información sobre cómo tener acceso a los reportes de la investigación basados en los microdatos proporcionados.

24. Para que las ONEs puedan derivar los mayores beneficios de esta asociación con la comunidad de investigadores, se deben tomar ciertas medidas con aquellos a quienes se les ha proporcionado el acceso a los microdatos para asegurar que los productos estadísticos resultantes de su trabajo se proporcionen a la ONE. Los usuarios de los microdatos también deben ser motivados para proporcionar una retroalimentación constructiva con respecto a la calidad de los datos como resultado de su análisis. Tal información puede ser invaluable al diseñar las futuras actividades estadísticas.

IV. LEGISLACIÓN DE APOYO Y LA RESPONSABILIDAD DE LA ONE COMO GUARDIÁN DE LA CONFIDENCIALIDAD.

25. La legislación que sustenta el acceso a los microdatos es muy importante según lo indicado en el Principio 3 (véase el capítulo III). Existen varias razones:

- (a) generar confianza pública en los procedimientos de acceso - que existen limitaciones legales que determinan lo que puede y no puede realizarse;
- (b) promover la comprensión mutua sobre los procedimientos entre las ONEs y los investigadores;
- (c) proporcionar una mayor consistencia en la manera como las propuestas de investigación son dictaminadas; y
- (d) establecer una base para enfrentar los incumplimientos.

La reputación de la ONE estaría en riesgo si el acceso a los microdatos, aún cuando estos datos se hayan anonimizado, no estuviera controlada por alguna autoridad.

26. Las disposiciones generales no forzosamente necesitan existir en la ley; y las detalladas pueden ser mejor recogidas en regulaciones, ordenanzas, etc., que tienen aún cierto impacto legal. Si no hay tales provisiones legales, es esencial que haya algunas otras formas de autorización. Los acuerdos especiales se deben determinar y convenir fuera de la ONE, (por ejemplo, por la Procuraduría General) como sea más apropiado. Los procedimientos de acceso deben incluir un compromiso o convenio firmado por los investigadores que tendrían acceso a los datos y por algún director ONE. El compromiso debe cubrir el tipo de aspectos mencionados en el párrafo 42, y debe estar disponible públicamente para evitar suspicacia sobre dichos arreglos.

27. Como mínimo, el acceso a los microdatos debe estar apoyado por una cierta forma de autoridad. Los arreglos legales pueden no ser factibles en algunos países, en cuyo caso algunas medidas administrativas tendrían que ponerse en su lugar. Los arreglos legales u otros deben también tener el visto bueno de las autoridades de privacidad de los países en donde existan, antes de que sean establecidos por la ley. Si no existen tales autoridades, puede haber ONGs que tengan un papel de “defensor” en materia de la privacidad. Sería razonable conseguir su ayuda para cualquier arreglo legal u otro, o al menos atender cualquier preocupación seria que pudieran tener.

28. Es importante que la legislación o autorización cubra los siguientes aspectos:

- (a) qué es lo que se puede y lo que no se puede hacer y para qué propósitos;
- (b) las condiciones del acceso; y
- (c) las consecuencias que acarrearía el incumplimiento a dichas condiciones.

29. Los acuerdos deben efectuarse, no sólo con los investigadores sino también con los directivos de las instituciones que representan. Cualquier incumplimiento se debe tomar seriamente. Aún cuando no fuera posible tomar acción legal, puede haber otras medidas que puedan ser tomadas. Tres posibilidades existen:

- (a) cancelar el acceso futuro al investigador y a su institución;
- (b) reportar el incumplimiento al directivo de la institución y solicitar que se tome una acción disciplinaria;
- (c) cancelar el acceso general a la comunidad de investigadores por un período específico, para asegurarse de que los individuos y las instituciones se sientan también responsables de proteger la confidencialidad.

30. Las ONEs deben considerar seriamente su gran responsabilidad como guardianes de la confidencialidad de los informantes individuales en sus conjuntos de datos, antes de emprender cualquier arreglo para conceder el acceso a los microdatos al público, los investigadores y las agencias internacionales. Un paso esencial en este proceso es la construcción de la confianza de los informantes, con la seguridad de que los datos confidenciales sean manejados completamente de acuerdo con el sexto de los Principios Fundamentales de las Naciones Unidas en materia de Estadísticas Oficiales. Esto es especialmente relevante para los países en los que la diferencia entre el uso estadístico y no estadístico de los microdatos no tiene una larga tradición, o no está asentada en ninguna legislación.

V. MÉTODOS PARA APOYAR A LA COMUNIDAD DE INVESTIGADORES

31. Existen varias formas en que las Oficinas Nacionales de Estadística pueden apoyar el trabajo de investigación.

- (i) Productos estadísticos para ser utilizados fuera de la ONE:

Medio de Distribución	Notas
Tablas estadísticas y Cubos de Datos	Esto puede incluir tanto tablas estándar como tablas especiales (o hasta análisis especiales), generados a petición del investigador. Algunas oficinas ahora divulgan matrices muy detalladas, conocidas como cubos de datos, que los investigadores pueden manipular para apoyar sus propias necesidades. Sin embargo, si éstos son muy detallados, el nivel del riesgo de la confidencialidad podría ser similar al de

	los microdatos.
Archivos de Microdatos Anonimizados (AMFs) - Archivos de Uso Público (PUFs)	Estos son archivos de microdatos que se difunden para uso público general fuera de la ONE. Ellos han sido anonimizados y a menudo se divulgan en un medio tal como el CD-ROM, a veces a través de un archivo de datos. El nivel de protección de la confidencialidad en los Archivos de Uso Público debe ser tal que la identificación no sea posible aún cuando se cruce con otros archivos de datos.
Archivos de Microdatos Anonimizados - los archivos bajo licencia	Los archivos bajo licencia son también anonimizados, pero distintos de los Archivos de Uso Público ya que en ellos su uso se ha restringido a investigadores autorizados. Incluso si estuvieran anunciados como generalmente disponibles para el público, no son proporcionados si antes no existe la firma de un compromiso o un contrato por parte del investigador.

- (ii) Una ventana de servicio a través de la cual los investigadores pueden someter su solicitud de datos

Servicio	Notas
Facilidades de Acceso Remoto (RAFs)	Son procedimientos que permiten a los investigadores generar productos estadísticos a partir de los archivos de microdatos, mediante redes de computadoras, sin que los investigadores realmente “vean” los microdatos. Debido a los controles adicionales que están disponibles a través de las RAF, y al hecho de que los microdatos no salen realmente de la ONE, se puede proporcionar de esta manera un acceso a microdatos más detallados.

- (iii) Procedimientos para permitir que los investigadores trabajen en las instalaciones de la Oficina Nacional de Estadística

Servicio	Notas
Laboratorios de Datos (DL)	Brinda acceso a microdatos más identificables, dentro de las instalaciones, típicamente con estricto rastreo para fines de auditoría y supervisión de la ONE. El acceso a datos más detallados crea una cierta inconveniencia al investigador, debido al requerimiento de trabajar en la ONE, o en un enclave de la ONE.

32. Un punto importante a enfatizar es que los archivos de microdatos no se anonimizan sólo quitando nombres y direcciones, sino también tomando otras medidas, (por ejemplo, comprimiendo el detalle geográfico). Esto podría ser llamado eliminación de la identificación espontánea. La posibilidad de identificación de los informantes individuales es mayor en poblaciones más pequeñas, donde pueden ocurrir identificaciones involuntarias con muy serias consecuencias. Atención particular también debe darse a las poblaciones con una distribución altamente asimétrica de alguna característica específica o con presencia de valores atípicos.

33. Aún cuando se hayan tomado medidas como la anonimización y otras para asegurar que la identificación de informantes individuales sea altamente improbable cuando los archivos se utilicen de manera aislada, éstos pueden contener datos que potencialmente permitan ligarse a otros archivos de microdatos del sector público o privado, e identificar a dichos informantes individuales. Hay estudios que demuestran que a través de cruces estadísticos, entre los archivos de microdatos de la ONE y los archivos existentes, se pueden identificar registros únicos. El número puede ser apreciable dependiendo de la cantidad de detalle disponible en el archivo de microdatos de la ONE. En términos relativos, el número de casos únicos será mayor para los países más pequeños. Ésta es una razón por la que se requiere un compromiso o un contrato preventivo antes de proporcionar acceso a los datos. Estos riesgos pueden ser reducidos si se utilizan técnicas tales como la perturbación o el intercambio de datos en los archivos de microdatos de la ONE.

Tablas estadísticas y Cubos de Datos

34. Las tablas estadísticas siguen siendo la manera más económica de satisfacer muchas de las necesidades de la investigación. Su importancia no debe ser subestimada. El advenimiento de los cubos de datos (tablas multidimensionales muy detalladas) ha aumentado la utilidad de las tablas estadísticas para propósitos de la investigación dado que permiten a los investigadores manipular los cubos de datos para satisfacer sus propias necesidades.

35. Todavía persisten problemas de confidencialidad en las tablas estadísticas y en los cubos de datos, y la mayor parte de la legislación estadística requiere que datos identificables no sean divulgados a través de las tablas estadísticas. Pero la “confidencialización” se hace previamente a la divulgación. Existen actualmente algunos sistemas de software para conservar la confidencialidad de las tablas estadísticas y se están desarrollando métodos más avanzados. Estos son llamados frecuentemente métodos para evitar la revelación de información.

Archivos de Microdatos Anonimizados – Archivos de Uso Público

36. Estos son vistos por los investigadores como un servicio muy valioso. Sin embargo, a la luz de las crecientes posibilidades de interrelacionar datos, la tendencia podría ser reducir la cantidad de datos disponibles en los archivos de uso público y depositar mayor confianza en los archivos de microdatos bajo licencia, en los RAFs y en los Laboratorios de Datos para el acceso de los investigadores. La eficacia de los archivos bajo licencia depende de que los investigadores honren los contratos o compromisos que han asumido de no intentar identificar las unidades del conjunto de datos. Estos compromisos son a menudo una pieza clave en el acceso a un AMFs bajo licencia (Ver la sección siguiente).

37. Aunque las ONEs proporcionan generalmente igualdad en el acceso a los usuarios de sus estadísticas, esto podría no ser apropiado para los microdatos. Una actitud diferente podría tomarse ante usuarios que no tienen buenas cartas credenciales de investigación o si tienen acceso a las bases de datos con las que sea fácil vincular los AMFs. Para fortalecer la credibilidad de las ONEs y obligar al cumplimiento entre

los usuarios, sería útil que se intercambiara información acerca de los incumplimientos de las comunidades de investigadores, y que se excluyera a las instituciones infractoras de cualquier acuerdo de acceso.

38. El acceso a los Archivos de Uso Público (PUFs) está deliberadamente establecido para ser amplio y se aprecia enormemente en aquellos países en donde ya existe y se usa generalizadamente. Los investigadores han acentuado la importancia de los PUFs, pero no deberían esperar que todos los países divulguen estos PUFs. Podría ser fácil para alguien que esté inclinado a identificar públicamente algunos informantes individuales a través del cruce estadístico con otras bases de datos. Previo a la difusión de los PUFs debe haber un examen minucioso de las condiciones bajo las cuales son divulgados para un mejor manejo de los riesgos de una violación a la confidencialidad. Por ejemplo, un convenio con fuerza legal podría ser uno de los requisitos del acceso. Debería ser posible implementar un procedimiento en el que se necesite firmar un acuerdo previo incluso donde el acceso a los PUFs sea a través de Internet.

39. Generalmente, el potencial para la identificación es mucho mayor para países con poblaciones más pequeñas y aquellos con registros de población. El riesgo de la identificación se puede reducir mediante el uso de técnicas tales como la perturbación de datos o el intercambio de datos. Estas técnicas, por ejemplo, son frecuentemente utilizadas en los Estados Unidos. La desventaja es que estas técnicas podrían reducir la utilidad de los microdatos así tratados.

40. Existe una extensa literatura disponible sobre los métodos para anonimizar los archivos de microdatos. Un buen resumen está disponible en Willenburg, L y de Waal, T (2001), *Elementos de Control de la Revelación Estadística*. El paquete de software, u-ARGUS, está diseñado para la protección de los microdatos contra la revelación de los informantes e incluye varias técnicas para anonimizar los archivos de microdatos. Es generalmente una buena práctica obtener el punto de vista de un experto acerca de si el archivo de microdatos está suficientemente anonimizado o no. Esta precaución no debe causar un retraso demasiado largo en cuanto a que socave la oportunidad de los datos que se divulgarán.

Archivos de Microdatos Anonimizados - archivos bajo licencia

41. Este es un procedimiento en donde usuarios específicos son autorizados o facultados para la utilización de los archivos anonimizados de microdatos después de suscribir un compromiso o un contrato relevante. Aunque estos archivos se han anonimizado y los individuos no pueden ser identificados en estos archivos de microdatos considerados aisladamente, sí sería posible realizarlo a través de un cruce (estadístico) con otros archivos, por esta razón es necesaria de una licencia. Habrá condiciones vinculadas con dicha licencia, mismas que podrán ser especificadas en el compromiso o en el contrato firmado por el investigador o su institución. Las condiciones pueden variar de un país a otro o incluso de un investigador a otro dependiendo de la propuesta de investigación y posiblemente de la institución de donde provenga.

42. Las condiciones pueden incluir algunas o todas de las siguientes:

- (a) que exista un acuerdo del investigador, al cual él o ella tendrá que sujetarse conforme a las condiciones del acceso;
- (b) que no se realizará ninguna tentativa para identificar nominalmente personas u organizaciones;
- (c) que la información será utilizada únicamente para propósitos estadísticos o de investigación;
- (d) que los microdatos no serán proporcionados a otras personas;
- (e) que los microdatos serán devueltos a la ONE cuando se termine el proyecto de investigación; y
- (f) que no se efectuará ninguna tentativa de vinculación estadística con otras bases de datos sin el permiso correspondiente.

43. Es una buena práctica que dicho compromiso tenga cierto sustento legal, por ejemplo, tener disposiciones para tales compromisos dentro de la legislación en la materia. Esto permitiría tomar acciones legales respecto a los incumplimientos de las condiciones del compromiso. Esto no excluye otras acciones que se pudieran tomar dados los incumplimientos, tales como no proporcionarle ningún servicio más al investigador, y posiblemente a la institución del investigador. Estos puntos se discuten en el capítulo VII.

44. Debería ser posible divulgar más datos por medio de archivos bajo licencia, que a través de archivos de uso público, si se pudiera tener además confianza en el compromiso para asegurar la protección de la confidencialidad de los datos. Esto es, en los casos donde algunos de los datos sean potencialmente identificables, al ligarlos con otros archivos.

Microdatos en donde la identificación pudiera ser posible.

45. Algunos países divulgan archivos de microdatos con fines estadísticos o de investigación que contienen datos que pudieran ser identificables aunque bajo estrictas licencias. Los acuerdos que norman estas licencias deben incluir las condiciones conforme a las cuales los datos puedan ser utilizados, y el procedimiento debe ser cubierto específicamente por la legislación. Es necesario un procedimiento estricto a fin de mantener la confianza de los informantes y la confianza del público en general. Las Facilidades de Acceso Remoto y los Laboratorios de Datos son otras maneras de lidiar con este tipo de situación.

Facilidades de Acceso Remoto (RAFs)

46. Estas facilidades están llegando a ser cada vez más importantes, pero la manera en como las RAFs son instrumentadas varía considerablemente de un país a otro. La característica clave es que los investigadores no tienen el acceso a los microdatos en sí mismos, pero algunas tareas utilizando esos microdatos pueden ser solicitadas de manera remota a través de Internet. A menudo hay un arreglo contractual entre la ONE y el investigador o la institución del investigador.

47. Existen dos tipos básicos de RAF.

- (a) Ejecución remota, donde un investigador presenta un programa y recibe posteriormente el resultado a través de Internet.
- (b) Facilidades remotas, mediante las cuales el investigador lleva a cabo el análisis y puede ver inmediatamente la respuesta en la pantalla.

48. A manera de ilustración, la Oficina de Estadística de Canadá proporciona a los investigadores archivos de microdatos ficticios y les permite presentar corridas contra el archivo real y completo por medio de la red de computadoras. La Oficina de Estadística de Canadá corre las peticiones fuera de línea y envía los resultados de regreso, vía las redes de computadoras, después de comprobar la confidencialidad. Aunque existen medidas similares en la Oficina Australiana de Estadística, hay algunas diferencias importantes. Los archivos de microdatos son confidencializados para prevenir la identificación espontánea antes de llegar a ser accesibles mediante las RAFs. Sin embargo, se permiten corridas de ensayo dentro de los archivos de la RAF y se permite la descarga de una pequeña cantidad de registros de unidades no identificadas para explorar valores atípicos o extremos y similares. El resultado se verifica antes de ser enviado al investigador. El sistema funciona actualmente mediante procesamiento de lotes, pero se está considerando una versión interactiva. Los procedimientos en la Oficina de Estadística de Dinamarca son diferentes. Consisten en un sistema en línea en donde los investigadores pueden correr análisis contra el archivo completo de microdatos. Esto se hace de tal forma que los

investigadores no pueden descargar los microdatos mismos. Para un manejo adicional de riesgos, se apoyan en los acuerdos hechos con las instituciones y en la sanción (particularmente la negación del acceso futuro) si hubiera incumplimiento de tales acuerdos.

49. Aunque solamente está disponible hasta ahora en muy pocos países, y aunque los modelos y los enfoques varían según lo ilustrado anteriormente, la experiencia hasta la fecha con las Facilidades de Acceso Remoto (RAFs) ha sido generalmente positiva. Desde la perspectiva del costo, las RAFs son preferibles a los Laboratorios de Datos (véase abajo) pues el acceso supervisado en una RAF es menos intensivo en trabajo que el que implica la supervisión del uso del Laboratorios de Datos.

50. Si estas facilidades no eliminan completamente el riesgo de la identificación, debe haber aún un acuerdo hecho con los investigadores para garantizar que ellos están completamente conscientes de sus obligaciones. Es una buena práctica proporcionar solamente el acceso a aquellos investigadores que han firmado un acuerdo que delimite las condiciones del acceso. La educación de los usuarios respecto al uso de las RAF es también importante, junto con el monitoreo constante y la comprobación del uso de dichas facilidades.

Laboratorios de Datos

51. Los Laboratorios de Datos (DLs) han estado funcionando por muchos años en algunas ONEs y han sido eficaces en controlar el riesgo de la identificación, a la vez que permiten el acceso del investigador particularmente a conjuntos de datos en los que proporcionar el archivo de microdatos confidencializados no es posible. Se requieren todavía condiciones de acceso para proporcionar un nivel adecuado de protección. La crítica principal hacia los DLs ha sido los inconvenientes que implica para el investigador, incluyendo algunas ocasiones en las que el investigador ha sido obligado a utilizar software desconocido para el análisis de datos. Son también costosos de manejar para las ONEs comparado con otras opciones.

52. Algunas ONEs han establecido nuevas instalaciones para los laboratorios de datos en ubicaciones que son más convenientes para los investigadores (conocidos a veces como Centros de Investigación de Datos), pero esto puede ser también una opción costosa a menos que se le proporcione a la ONE el financiamiento específico.

53. ¿Cuáles son las condiciones más importantes para el acceso a los microdatos a través de los Laboratorios de Datos? Éstas podrían ser las siguientes:

- (a) Documentación del beneficio público que proporcionará la investigación,
- (b) Descripción de cómo serán accesibles al público los resultados,
- (c) Evidencia de la buena fe de los investigadores,
- (d) Compromiso legalmente vinculante, y
- (e) Requerimientos para la supervisión por parte del personal de la ONE.

Contratación de un investigador como miembro temporal del personal de la ONE

54. Otra manera en que los investigadores pueden tener acceso a microdatos es mediante su contratación como miembros temporales del personal de la ONE y sujetándolos a los mismos lineamientos de confidencialidad que el personal de la ONE. Esto no debería realizarse a menos de que el investigador esté participando en el trabajo de la ONE, o que el arreglo fuera autorizado de alguna otra manera, de lo contrario, el arreglo se podría considerar como una farsa. Si este tipo de simulación ocurriera y llegara a hacerse pública, la confianza en la ONE disminuiría.

55. El involucramiento del investigador podría darse a iniciativa de la ONE, si el investigador es visto como alguien que puede aportar sus habilidades especiales al trabajo de la ONE y ampliar la utilidad del conjunto de datos. Por otra parte la propuesta puede provenir del investigador. Pero la ONE debe aceptar el mérito de la propuesta e incorporarla dentro de las actividades de su programa de trabajo. Es más fácil demostrar que los investigadores están contribuyendo a la ONE si un producto publicado por la ONE fuera el resultado del trabajo (aún cuando se etiquetara de manera diferente a los productos publicados normalmente). Por supuesto que habrá beneficios para los investigadores de tales arreglos y podría acordarse que publicaran los productos de su investigación de alguna otra manera, (quizás después de su revisión por la ONE).

Manejo de los datos de las empresas

56. Hay algunas cuestiones especiales asociadas a los datos de las empresas, incluyendo las empresas agrícolas. Las empresas, y particularmente las grandes empresas, son más fácilmente identificables que los datos de hogares o de personas, especialmente sobre una base espontánea, debido a que la distribución de sus características es mucho más asimétrica. También, en muchas encuestas de empresas, las grandes empresas son seleccionadas con certeza. En algunos países, son más accesibles las bases de datos de empresas, de tal modo que se facilita el cruce estadístico (con microdatos anonimizados de la ONE). Además, muchos investigadores académicos pudieran también fungir como consultores de las empresas y aún un acceso de buena fe a microdatos de empresas, podría ser incompatible con dicho rol de consultor. Además aún, los países pueden tener problemas de competitividad económica (y posiblemente incluso de seguridad), debido a que se compartan datos de empresas identificables con los investigadores de otros países.

57. Desde el punto de vista del acceso del investigador, las principales diferencias entre los datos de hogares o los de personas y los datos de las empresas, son que los medios de acceso que proporcionan una mayor protección son más relevantes para los datos de las empresas.

58. En términos de los medios de acceso:

- (a) Las tablas estadísticas siguen siendo relevantes, aunque el más alto nivel del riesgo de identificación implica que las tablas con mayor detalle no estarán generalmente disponibles en lo que se refiere a las empresas.
- (b) Los archivos de los microdatos que se han anonimizado podrían ser relevantes tan solo para las empresas más pequeñas. Para cierto tipo de investigación, las pequeñas empresas pudieran ser un grupo de interés particular para los investigadores. Incluso entonces existirá la necesidad de que haya una “distorsión” de ciertos datos, (por ejemplo, de los datos financieros) para evitar la vinculación con otras bases de datos, (como es el caso de los datos de los impuestos). Una alternativa sería presentar los datos por rangos, aunque de este modo se limitaría el uso de los archivos de microdatos anonimizados.
- (c) Por razones similares, las RAFs podrían ser relevantes solamente para los archivos de los microdatos de las empresas más pequeñas. Por lo menos, el uso de estos procedimientos habilita a las ONEs para controlar el riesgo de identificación, por lo que no sería necesaria la “distorsión” de los datos para proteger su confidencialidad. Pero, si las grandes empresas están incluidas, podría ser difícil confidencializar los resultados aún cuando los investigadores pudieran no tener directamente el acceso a los microdatos.
- (d) Los Laboratorios de Datos son probablemente los más pertinentes para el acceso a los archivos de los microdatos de las empresas.

59. Algunos estudios de investigación podrían llevarse a cabo con el consentimiento de las empresas implicadas.

VI. MANEJO DE LAS TENSIONES ENTRE LAS OFICINAS NACIONALES DE ESTADÍSTICA Y LOS INVESTIGADORES

60. La cultura y el sistema de valores de la comunidad de investigadores son muy distintos a los de las ONEs. Mientras que los investigadores miran a menudo algunos de los controles inherentes al acceso de los microdatos como una burocracia innecesaria, la violación de la confidencialidad tiene consecuencias graves para las ONEs. Si los informantes creen o perciben que una ONE no protegerá la confidencialidad de sus datos, será menos probable que cooperen o proporcionen datos exactos. Incluso un simple incidente, particularmente si recibe una fuerte atención de los medios masivos de comunicación, podría tener un impacto significativo en la cooperación del informante y por lo tanto, sobre la calidad de las estadísticas oficiales.

¿Cómo podrían resolverse las tensiones entre las perspectivas de las ONEs y del investigador?

61. Esto se realizará lo más eficazmente posible cuando las ONEs cambien la estrategia de evitar el riesgo, por una estrategia de gestión del riesgo.

62. Hay riesgos que definitivamente tienen que ser gestionados. La rápida expansión de las bases de datos, conteniendo datos sobre personas identificables, significa que es virtualmente imposible evitar totalmente la identificación a través de la vinculación de datos de un número significativo de personas aun cuando se hayan eliminado los nombres y las direcciones, particularmente si la estructura del hogar está contenida en los archivos. Muchas de estas bases de datos se encuentran en poder del sector privado donde generalmente los controles para su uso son menos rigurosos que los del sector público. Además, los avances de la tecnología han hecho más fácil la vinculación de datos, ya sea a través de un apareamiento preciso o mediante las técnicas de cruces estadísticos (las cuales podrían conducir a identificaciones exactas en casos únicos). Evitar el riesgo, significa en esencia no permitir que los microdatos identificables salgan de las instalaciones de la ONE, a menos de que se dieran otros pasos tales como la perturbación o el intercambio de los datos. Los riesgos variarán de acuerdo con el tamaño del país, entre otras cosas. En los países más pequeños, el riesgo será relativamente más alto porque hay casos relativamente únicos.

63. Sin embargo, el acceso a los microdatos proporcionado por las ONEs no parece haber sido un área de controversia pública. Implícitamente parece haber un nivel razonablemente alto de aceptación pública de las prácticas actuales aunque se carece de información acerca de los países en donde ha habido un extenso debate público. Pero las preocupaciones generales de la comunidad sobre la privacidad sugieren que existe un límite a lo que el público pudiera aceptar. Se podría detonar fácilmente un debate (a través de las fronteras nacionales), por un incidente desafortunado. El nivel de cooperación en la recolección de estadísticas podría ser consecuentemente afectado.

64. La transparencia es importante para evitar acusaciones respecto al secreto. Por lo tanto, es una buena práctica para las ONEs ser transparentes en destacar que uno de los usos valiosos de los datos de algunas recopilaciones estadísticas será proporcionar acceso a los investigadores a los microdatos confidencializados bajo condiciones controladas para propósitos específicos. Esto tiene que ser manejado cuidadosamente o los abogados de la privacidad podrían influir en la opinión pública. La ayuda de personas respetadas y autorizadas respecto a los beneficios de la investigación es muy importante.

¿Cómo gestionan las ONEs los riesgos de acceso a los microdatos?

65. Algunas sugerencias se resumen a continuación:

- (a) Convenir sobre un conjunto de principios que deberían seguirse para proveer el acceso a los microdatos (tales como los esbozados en el capítulo III).
- (b) Asegurarse de que existe una sólida base legal y ética (así como las herramientas técnicas y metodológicas), para la protección de la confidencialidad. Esta base legal y ética requiere de una valoración equilibrada entre el bien público de la protección de la confidencialidad por una parte, y los beneficios públicos de la investigación por la otra. Una decisión acerca de proporcionar o no el acceso debería depender de los méritos de la propuesta específica de investigación y de la credibilidad del investigador, lo anterior debería estar contemplado, en las disposiciones legales. El acceso no se debe considerar como automático.
- (c) Tener un proceso que brinde holgura a la ONE para equilibrar estos dos bienes públicos. Es una buena práctica implementar un comité interno para discutir estas materias y para hacer recomendaciones al jefe de la ONE. Dado que hay juicios involucrados los comités de ética o cuerpos similares podrían apoyar en situaciones donde se deba ponderar la decisión de proporcionar el acceso o no. Los argumentos sobre el bien público son más fuertes si los resultados de la investigación serán del dominio público.
- (d) Ser completamente transparente acerca de los usos específicos de los microdatos para evitar suspicacias de mal.
- (e) Estar preparado para proporcionar mayor acceso a través de las Facilidades de Acceso Remoto y de los Laboratorios de Datos en caso de que no fuera posible que los microdatos de divulgación pública fuesen no identificables, sin una distorsión considerable en los mismos. Explorar otras posibilidades para utilizar los progresos tecnológicos con el objeto de mejorar el acceso a los microdatos de una manera tal que se proporcione una adecuada protección a la confidencialidad.
- (f) Pasar algo de la carga de la responsabilidad a la comunidad de investigadores. Asegurar que los investigadores entiendan las razones por las que las ONEs protegen tanto la confidencialidad y que están enterados de las consecuencias para ellos y para su institución en caso de incumplimientos. Aplicar el castigo apropiado si se presentan tales incumplimientos.

66. Muchos investigadores no ven la razón de estos controles. Aunque no hay incidentes conocidos de investigadores que usen su acceso a los microdatos para identificar deliberadamente a individuos, ha habido incidentes en los cuales los microdatos proporcionados a ellos en una base exclusiva, se han puesto a disposición de otros investigadores sin permiso, o de casos donde los microdatos se han cruzado estadísticamente con otros datos para producir conjuntos de datos más completos, sin permiso. Los investigadores en cuestión pueden sentir que no han hecho nada malo, pues no han intentado identificar a individuos. Sin embargo, incidentes de este tipo, si llegaran a ser públicos, podrían socavar la confianza pública y deben ser tratados seriamente. Las ONEs y los investigadores funcionan con diversas culturas y asumen diversas perspectivas del riesgo de los incidentes. Esto tiene que ser considerado en la determinación de los procedimientos para el acceso a microdatos.

¿Cómo pueden las ONEs compartir el riesgo con los investigadores?

67. El daño potencial de una divulgación no autorizada no puede ser subestimado, particularmente si está hecha deliberadamente. Tales situaciones deben ser tratadas muy seriamente. Las acciones pudieran incluir:

- a) Pedir a los investigadores que prueben su buena fe, que demuestren los beneficios públicos de su investigación y la necesidad de los microdatos para su realización.

- (b) Hacerles firmar a los investigadores un compromiso legalmente vinculante con penalizaciones similares a las que funcionan para el personal de la ONE si incumplieran alguna de las disposiciones de confidencialidad.
- (c) Explicar a los investigadores las razones de la cautela de las ONEs. Asegurarse de que ellos están completamente enterados de sus obligaciones a través de un entrenamiento apropiado. Dar un seguimiento eficaz con procedimientos del monitoreo y supervisión, (aunque esto pudiera no ser posible para muchas ONEs debido a la carencia de recursos y a la falta de transparencia de parte de los investigadores). Podría ser útil establecer un Código de Conducta en colaboración con la comunidad de investigadores.
- (d) En el momento en que ocurran las infracciones, cancelar todos los servicios actuales y futuros al investigador y posiblemente a su institución por un período del tiempo, (por ejemplo, hasta que la institución haya emprendido la acción disciplinaria apropiada contra el infractor). Hacer que se percaten de que los futuros accesos a los microdatos a cualquier investigador están en riesgo si existe una fuerte crítica pública. Llevar a cabo una demanda legal cuando sea pertinente.

Un factor que puede desalentar a las ONEs para proporcionar los microdatos es la demora de tiempo entre el uso incorrecto de los datos y el descubrimiento de la violación.

68. La realidad es que una combinación de medidas legales, administrativas y técnicas será necesaria para asegurar la confianza pública en los procedimientos. Además, la comunidad de investigadores debe aceptar que no tiene derecho automático al acceso. El acceso del investigador debe ser a discreción de la ONE, y existirán responsabilidades asociadas directamente con dicho acceso. En particular, los investigadores tendrán una responsabilidad compartida para mantener y cumplir las condiciones bajo las cuales se les ha proporcionado el acceso. Las limitaciones y las salvaguardas podrán ser más restrictivas que con otros conjuntos de datos a los cuales tengan acceso, pero esto será por una justa razón.

Otros asuntos

El Consentimiento

69. Se discute a veces que el consentimiento del informante debería buscarse antes de que la divulgación de los microdatos saliera de la ONE. Los defensores de este punto de vista argumentan que los informantes tienen el derecho de decidir cómo es que sus datos deberían utilizarse aunque no fueran identificables. Esto debería evitarse, ya que:

- (a) Existen cuestiones prácticas importantes relacionadas con la búsqueda y el manejo del consentimiento.
- (b) Los datos que han sido proporcionados son inidentificables y se están utilizando solamente para propósitos estadísticos, lo cual es consistente con el propósito de la recopilación de datos.
- (c) Es muy difícil proporcionar toda la información requerida por un informante para que tome una decisión realmente informada, - y así que muchos informantes dirán “no” sólo como precaución. Una muestra pronto llegaría a ser no representativa si esta incluyera sólo a aquellos que dan su consentimiento.

Sin embargo, hay una obligación, según lo indicado previamente, de ser transparente con los informantes acerca de los procedimientos de acceso. Por estos medios, podría argumentarse que se ha obtenido su consentimiento pasivo.

70. Si es permitido por la ley, el consentimiento informado sería apropiado en una situación donde la publicación de pequeños agregados permitiera a los usuarios inferir la situación de una simple unidad de

la muestra, (por ejemplo, de personas o empresas) que es parte de este agregado. Esta situación se aplica más probablemente a la estadística de las empresas.

Datos administrativos

71. Hay otra perspectiva sobre el tema del consentimiento. Los datos de una ONE pueden abarcar tanto los datos recogidos directamente por ella misma como los datos recogidos por autoridades administrativas y transferidos a la ONE. A menos que haya disposición específica en la legislación o un protocolo que dictara lo contrario, una ONE no debería divulgar datos provenientes de fuentes administrativas en forma de microdatos sin el consentimiento de la autoridad administrativa (la cuál podría sentirse incapaz de dar este consentimiento debido a las promesas hechas a sus informantes). Aún cuando los datos administrativos ya sean del dominio público, sería una cortesía notificar a las autoridades que los administran para darles la oportunidad de que comenten al respecto. De otra manera, se podrían presentar dificultades con el suministro de los datos administrativos. Las agencias administradoras también tienen que controlar su propia privacidad y los asuntos de confidencialidad.

Planeación de las contingencias

72. Es importante que las ONEs realicen cierta planeación de contingencias para el evento de que el acceso a los microdatos se convirtiera en un asunto de debate público. No deberán asumir que dicho debate no sucederá. ¿Cuáles son algunas de las defensas clave?

- (a) Las ONEs pueden señalar el cuidado que tienen en proveer la protección a la confidencialidad a través de dispositivos tales como anonimizar los microdatos, proporcionar una importante seguridad física y tener el cuidado de diseñar un proceso para la valoración del equilibrio entre los bienes públicos en conflicto de la protección de la confidencialidad y los beneficios públicos de la investigación.
- (b) Si ha ocurrido alguna infracción, la ONE debe ser abierta sobre la infracción y sobre las penalizaciones aplicables, cuando se le cuestione. Debería aclarar que el incumplimiento es responsabilidad del investigador pero que la ONE tomará las acciones apropiadas en relación con dicho incumplimiento.
- (c) Las ONEs deberán señalar los beneficios públicos generales de proporcionar el acceso a los microdatos, particularmente en la situación en la que se haya dado la infracción, y dar algunos buenos ejemplos.
- (d) Deberán estar preparadas algunas personas ampliamente conocidas y respetadas que estén dispuestas a apoyar públicamente los procedimientos de acceso. Las autoridades responsables de la privacidad podrían ser de particular importancia en este respecto.

VII. CUESTIONES ADMINISTRATIVAS VINCULADAS CON EL ACCESO A MICRODATOS

Administración de la toma de decisiones sobre confidencialidad

73. Hay siempre una cierta probabilidad para la identificación, aunque sea muy pequeña. El software ahora existente puede estimar la proporción de registros que son únicos y por lo tanto en cierto riesgo de identificación.

74. Es el director estadístico o sus delegados quienes necesitan tomar la decisión sobre el acceso a los archivos de microdatos, ya sea que se realice a través de un archivo de microdatos anonimizados (de uso

público-o bajo licencia), de facilidades de acceso remoto o mediante los Laboratorios de Datos. Con el fin de tomar esta decisión, el director estadístico necesita ser asesorado sobre si, por ejemplo:

- (a) el riesgo de la identificación es suficientemente pequeño;
- (b) los ajustes hechos a los datos no han dañado indebidamente al archivo de los microdatos para los propósitos de la investigación; y
- (c) las variables que han sido compactadas son las más apropiadas, tomando en consideración ambas necesidades, las de los investigadores y las del riesgo de identificación.

75. Se debe disponer de los procedimientos e instancias apropiados para proporcionar esta asesoría de manera consistente. Esto a menudo necesita ser apoyado por una capacidad de investigación y podría estar ubicada dentro de alguna área metodológica.

Administración de los metadatos

76. Si los usuarios van a hacer un uso eficaz de los microdatos, deben tener acceso a los metadatos apropiados. Esto incluiría:

- (a) una descripción de las encuestas, incorporando información adicional sobre calidad;
- (b) una lista de las variables y de las clasificaciones utilizadas, (algunas veces referidas como un “diccionario de datos”); y
- (c) definiciones de las variables.

Disponer de (a) ayudará a asegurar que los microdatos no sean utilizados si los datos realmente no encajan con los propósitos previstos. Es también crítico que las restricciones de los datos sean precisadas, en particular, el nivel de la agregación sustentada por el diseño de la muestra.

77. Como los microdatos se suministran electrónicamente, los metadatos se deben proporcionar de una manera que sea accesible. De ser posible, los metadatos se deberían divulgar en vinculación con los microdatos. La copia impresa en papel podría ser todavía un medio eficaz, aunque el sitio Web de la ONE está llegando a ser cada vez más útil para estos propósitos.

Administración de los incumplimientos del investigador

78. Se deben realizar esfuerzos para reducir la probabilidad de incumplimientos conforme a lo señalado en el capítulo anterior. Sin embargo, los incumplimientos pueden ocurrir y se deben determinar los procedimientos para afrontarlos.

79. Los incumplimientos deben tratarse seriamente. Si esto no se hace, se erosionará la confianza pública en los procedimientos. De igual forma, los incumplimientos ocurrirán más a menudo si no son tratados seriamente.

80. Existe un gran número de maneras de lidiar con los incumplimientos. Por ejemplo, si ha ocurrido una infracción legal, se debe considerar una demanda legal. Esto es costoso pero es indispensable en los casos más severos para demostrar la importancia que la ONE otorga a la confidencialidad, y para reducir la probabilidad de futuras infracciones.

81. También, el investigador deberá quedar impedido para un acceso futuro a los microdatos. Ésta debería ser la acción mínima que se adopte.

82. Se debería también considerar impedir el acceso a la institución del investigador, por lo menos hasta:
- (a) que la institución haya tomado las medidas apropiadas para hacer frente a la infracción cometida por el investigador; y
 - (b) que la ONE tenga la certeza de que se han realizado los arreglos pertinentes dentro de la institución para reducir al mínimo las probabilidades de incumplimiento.

83. La comunidad de investigadores apoyará en términos generales la puesta en práctica de acciones severas en contra del número relativamente pequeño de infractores que pudieran estropear el buen nombre de la comunidad de investigadores, pues tal es su interés al largo plazo.

84. Para incumplimientos de menor importancia, podría ser suficiente una advertencia.

VIII. ALGUNAS CUESTIONES ESPECIALES

El Acceso Internacional

85. Las comparaciones internacionales son importantes para entender la eficacia de las políticas y de los programas adoptados por los países. Los Gobiernos, en particular, encuentran esto útil para el trabajo de evaluación de sus políticas. Los beneficios de proporcionar el acceso a los microdatos anonimizados a los investigadores que trabajan en las comparaciones internacionales y en las agencias internacionales son claros, pero existen también los riesgos, (por ejemplo el proporcionar microdatos sin autorización). La principal dificultad es que hay una posibilidad más limitada para aplicar las penas merecidas por los incumplimientos cometidos por parte del personal de las agencias o de los investigadores internacionales que viven en otros países. A excepción de EUROSTAT, el personal de las agencias internacionales no está sujeto a ninguna legislación nacional o internacional, más que a las reglas aplicables al personal de su organización. Por otra parte, la probabilidad de su identificación es mucho más baja, (mientras el investigador no pase los microdatos de nuevo a una tercera persona en el país anfitrión). Otra dificultad es que muchos países no tienen la autoridad legal para proporcionar datos a las agencias o a los investigadores internacionales fuera de su país.

86. Estas directrices han sugerido encaminarse hacia una gestión del riesgo en vez de un enfoque de elusión del riesgo en cuanto al suministro de microdatos. Los riesgos son más bajos o percibidos como más bajos, si la agencia receptora tiene credibilidad y una reputación digna de confianza. Los riesgos también se justifican más fácilmente si el propósito para el cual se proporcionan los microdatos está dentro del interés del país que suministra los datos. Esto podría ser porque los datos formen parte de un estudio internacional llevado a cabo por una agencia internacional o por una colaboración internacional de investigación respetada, (por ejemplo, el Estudio de Ingresos Luxemburgo- Luxemburg Income Study).

87. La Globalización está aumentando la importancia de tales estudios internacionales. Debería ser legítimo para las ONEs que pudieran tener la posibilidad de apoyar estos tipos de estudios suministrando el acceso a los microdatos. Pero debería ser un suministro con carácter habilitante (esto es, que la ONE debería ser capaz para decidir si proporciona los datos o no) que también especificara las salvaguardas y condiciones para el acceso. Las ONEs estarían más preparadas para proporcionar microdatos anonimizados donde sean más bajos los riesgos y los beneficios más altos.

88. ¿Cuáles son las opciones para que los investigadores tengan acceso a conjuntos de datos de otros países? ¿Cómo pueden las agencias internacionales obtener el acceso a los microdatos con fines estadísticos y de investigación? Las opciones incluyen:

- (a) Los datos son recopilados directamente por la agencia internacional (o el investigador) o a través de intermediarios (por ejemplo, una organización especializada en encuestas), de manera tal que el destinatario de microdatos se de a conocer en el momento de la recopilación de datos;
- (b) Archivos de Uso Públicos en donde estén disponibles;
- (c) Archivos de Microdatos Anonimizados bajo licencia en los países en los que sea posible hacer esto;
- (d) Facilidades de Acceso Remoto con las salvaguardas pertinentes;
- (e) Colaboración con un investigador adscrito a la ONE o al país de la ONE, quien tenga acceso a los microdatos.

89. Desde el punto de vista del acceso a los microdatos, las encuestas del tipo (a) son preferibles para los investigadores internacionales. El estudio de la OECD-PISA es un buen ejemplo. Sin embargo, hablando en términos generales, estos datos no serán recopilados bajo la legislación estadística que funciona en varios países. La calidad de los microdatos, y particularmente las tasas de respuesta, podrían afectar a algunos estudios. Esto dependerá del tipo de estudio y de la reputación de la organización que lleve a cabo el estudio. También dependerá del compromiso de los recopiladores de datos en cada uno de los países. Éste es un factor que los investigadores internacionales necesitan considerar antes de decidirse a adoptar este enfoque. Puede haber necesidad de elegir entre el acceso y la calidad.

90. En algunos casos, puede ser posible satisfacer el requerimiento del estudio, proporcionando datos muy detallados para el análisis, pero no microdatos. Este enfoque se utiliza en el Programa de Comparación Internacional y en el Estudio de la Paridad del Poder de Compra de la OCDE/Eurostat

91. Los Archivos de Uso Público están solamente disponibles para algunos países. Los archivos de microdatos anonimizados bajo licencia pueden ser una opción si no están restringidos por la legislación de la ONE. Si las ONEs están capacitadas para proporcionar los microdatos de esta manera, los factores que debieran tomarse en consideración son:

- (a) la confianza en el investigador y en su institución;
- (b) si el estudio es de importancia para el país; y
- (c) si el suministro sería consistente con las promesas hechas a los informantes al momento de la recopilación de los datos.

92. Es probable que muchos países se sientan más cómodos brindando acceso a ciertas agencias internacionales o a proyectos de investigación específicos que hacerlo en términos generales a la comunidad internacional de investigadores. También, podría haber algunas condiciones que aplican a investigadores particulares. Por ejemplo, algunos países pueden sentirse cómodos solamente proporcionando acceso a los investigadores vía la ONE del país de origen del investigador. Pero en todos los casos de acceso a los microdatos, sería buena práctica hacerlo solamente conforme a condiciones específicas utilizando también un compromiso o memorando de entendimiento. Algunas de las condiciones podrían incluir:

- (a) limitar el acceso a divisiones particulares de las agencias internacionales y prohibir pasar los datos a otras;
- (b) limitar los propósitos para los cuales los microdatos podrían utilizarse sin permiso;
- (c) acordar que los microdatos deban ser devueltos al solicitarlo (por ejemplo, podrían contener errores);
- (d) permitir la posibilidad de comentar respecto al material publicado que utiliza los microdatos; y
- (e) aclarar puntualmente las consecuencias de no respetar las condiciones del acceso.

93. La manera más eficaz de ocuparse de los incumplimientos a las condiciones de acceso será suspender los accesos posteriores. El asunto deberá ser llevado a la atención del personal de más alto nivel de la institución. Para las agencias internacionales, esto podría realizarse a través de los funcionarios

diplomáticos para las infracciones más serias. Pero el punto clave es que los incumplimientos no se pueden permitir sin que den lugar a una reacción. De lo contrario, simplemente se repetirán.

94. Para muchos países, el uso de las facilidades del acceso remoto podrían ser la ruta favorita para proporcionar el acceso a los investigadores internacionales. Bajo tales arreglos, existen más controles, y la posición de las ONEs respecto al acceso internacional a los microdatos se defiende más fácilmente si es cuestionada. Sin embargo, la utilidad de estos arreglos para el acceso internacional todavía necesita ser mejorada. La experimentación es importante.

95. Otra posibilidad para los investigadores internacionales, incluyendo a las agencias internacionales, es trabajar a través de las redes de investigadores nacionales para llevar a cabo el análisis deseado utilizando microdatos. Ciertamente, estos investigadores nacionales podrían ubicarse en la ONE para los estudios internacionales de particular importancia.

96. Las ONEs necesitarán decidir si pueden o no proporcionar el acceso a los investigadores internacionales, tomando en consideración la gama de asuntos discutidos en esta sección. Deben tener presente que se está fomentando un enfoque de control del riesgo. Para algunos proyectos de investigación, los beneficios podrían justificar los riesgos implicados, siempre y cuando el arreglo sea legal. Los riesgos pueden ser más bajos para algunas instituciones que para otras. Las ONEs también necesitarán decidir la forma más apropiada para el acceso en colaboración con el investigador internacional. Para ayudar a garantizar una toma de decisión consistente, los países deberán desarrollar las directrices para el acceso a los investigadores internacionales y a las agencias internacionales que sean coherentes con su propia legislación; o pueden decidirse a cambiar su legislación para habilitar el acceso en los casos justificados.

97. El acceso de los investigadores internacionales a los microdatos es mucho más fácil de justificar si existe algún beneficio para el país. Esto se podría asegurar si los productos estadísticos son retroalimentados al país de origen, de manera que sus beneficios fueran tangibles. Es importante que los resultados útiles de la investigación sean hechos del dominio público, o retroalimentados hacia los países que proporcionaron los datos. Es una forma de aumentar el valor del servicio estadístico nacional.

Nuevos procedimientos para el acceso a datos en algunos países en vía de desarrollo

98. Las opciones para proporcionar el acceso a los microdatos están mucho más limitadas todavía en muchos países en vía de desarrollo. Aparte de los recursos limitados para la recopilación de datos, a menudo el análisis y la difusión de los resultados de las encuestas no son presupuestados adecuadamente. Es aún más improbable que se hayan previsto recursos para la divulgación de los microdatos a los posibles investigadores, nacionales e internacionales. Éstas, las restricciones tanto de recursos como técnicas, han conducido a muchas ONEs, a considerar incluso opciones que podrían disminuir su control sobre los microdatos que son suministrados.

99. Una manera de actuar para las agencias financieras, o sus instituciones representativas, organizaciones internacionales, y las Comisiones Regionales de las Naciones Unidas, sería ser depositarias de algunos de sus archivos de microdatos. Otro enfoque es que las ONEs lleguen a establecer acuerdos con una /institución de investigación privada y/o académica, en los que la ONE produciría archivos de microdatos anonimizados, como Archivos de Uso Público, en intercambio por servicios de la institución que incluyan la custodia y preservación de los censos. Algunos de los repositorios de datos ofrecen otros servicios complementarios, incluyendo un catálogo central de encuestas y de censos². Algunas de estas

² http://www.internationalsurveynetwork.org/home/index.php?option=com_frontpage&Itemid=1

medidas son el resultado de los acuerdos derivados del financiamiento de encuestas³ específicas. Sin embargo, debe hacerse notar, así como en el párrafo 85, que estas instituciones podrían no tener las salvaguardas legales necesarias para garantizar la seguridad de los conjuntos de datos.

100. Bajo algunos arreglos, la ONE conserva el derecho de autorizar o de negar la divulgación de los datos, mientras que bajo otros, la ONE renuncia a cualquier autoridad sobre los datos una vez que estos han sido cedidos a la agencia que proporciona el servicio de repositorio. En cualquier caso, la ONE tiene poca o ninguna autoridad para hacer cumplir los acuerdos, y no tiene ningún recurso en caso de incumplimiento, excepto suspender dicho arreglo. Sin embargo, es esencial que la ONE esté informada acerca de cualquier acceso que se conceda. Las ONEs que incorporan tales arreglos necesitarían por lo tanto, asegurarse de que la institución intermediaria proporcione la garantía necesaria y demuestre compromiso para tomar acciones en contra de cualquier incumplimiento, con respecto a los datos para los que se ha dado el acceso. En los casos en los que la recopilación de datos sea financiada en parte o enteramente por una agencia externa, los términos que rigen el acceso a los microdatos deben ser convenidos mutuamente, y la ONE deberá recibir garantías de que la confidencialidad del informante no está comprometida.

101. Un argumento que nos obliga y una motivación para este tipo de arreglos es asegurar que se está obteniendo el valor máximo de las actividades estadísticas, las cuáles se llevan a cabo con un costo significativo. Los beneficios potenciales de conceder el acceso a los microdatos podrían ser socavados si los servicios de repositorio de datos no se sujetan a los siguientes acuerdos para:

- (a) garantizar que sea preservada la confidencialidad;
- (b) requerir que los resultados del uso de los microdatos sean proporcionados a la ONE;
- (c) informar a las ONEs sobre los accesos a sus microdatos; y
- (d) impedir la transferencia de microdatos a otros con los cuales la institución no tenga ningún acuerdo.

102. El acceso inicial y continuado a los datos deben por lo tanto estar condicionado a la retroalimentación a las ONEs sobre los resultados de la investigación, así como al cumplimiento de los beneficiarios. Los investigadores y las agencias internacionales, podrían desarrollar también un juicio sobre la calidad en los datos, esto también debería ser retroalimentado de una manera constructiva para auxiliar en el diseño de futuros levantamientos. Entre algunas de las acciones que podrían emprenderse en caso de un incumplimiento del acuerdo están: no proporcionar datos al servicio de repositorio de datos en el futuro; suspender los acuerdos; y promover una queja ante la máxima autoridad de la institución anfitriona.

103. El resguardo de conjuntos de datos en otro sitio es visto generalmente como buena práctica, como una medida contra la contingencia de que los datos sean destruidos por catástrofe u otros medios. En los países desarrollados, esto se hará generalmente en otro sitio dentro del país. Pero muchos países en vía de desarrollo podrían no tener esta capacidad de resguardo y el uso del repositorio de datos en otro país se podría justificar por estas razones. Pero esto debería suceder solamente cuando las salvaguardas adecuadas se hayan puesto en pie.

104. En los casos en los que la legislación de la estadística esté siendo desarrollada o revisada, se deben incluir disposiciones sobre repositorios de datos y procedimientos de acceso.

La vinculación de datos

³ Por ejemplo, Encuestas sobre Salud y Demografía, Indicadores Múltiples de Encuestas de Conglomerados, Encuestas sobre Estándares de Vida, etc.

105. La vinculación de conjuntos de datos, ya sea por cruce exacto o estadístico, puede facilitar una gama mucho mayor de análisis, y agregar un valor considerable a los mismos. De manera creciente, los investigadores están buscando utilizar conjuntos de datos vinculados que incluyan ligas a los conjuntos de microdatos de la Oficina Nacional de Estadística o de otras agencias estadísticas, (incluyendo los censos de población en algunos países). La investigación sobre salud, en particular, es un área en la que los conjuntos de datos vinculados pueden ser particularmente valiosos.

106. Así como existen claros beneficios en la vinculación de datos, también existen riesgos, particularmente si el guardián del archivo ligado no cuenta con la protección de la confidencialidad que a menudo existe en la ONE. Estudios en muchos países demuestran gran preocupación pública sobre la vinculación de las bases de datos. Las percepciones también son importantes. Es de particular importancia que los cuatro principios resumidos en el capítulo III sean aplicados para los conjuntos de datos ligados.

107. Es conveniente que las ONEs se involucren en la vinculación de conjuntos de datos para fines estadísticos. La agencia estadística debe ser el guardián de estos conjuntos de datos ligados. Puede también haber situaciones en donde ésta sea el guardián idóneo de los conjuntos de datos ligados aún cuando provengan de fuera de las agencias estadísticas, esto debido a las salvaguardas y la confianza pública que ya existen.

108. En aquellos países en donde existan, las Comisiones de Privacidad o sus órganos equivalentes deben apoyar las medidas para la vinculación de los datos.