

Los Estándares Estadísticos Internacionales y la Armonización de las Estadísticas Nacionales

Alberto Ortega Venzor
Director General de Vinculación Estratégica, INEGI, México

Abstract

This paper was prepared for the Seminar: "Development of the National Statistical Systems: Recent experiences, lessons and evaluation" during the Fifth Meeting of the Statistical Conference of the Americas of the Economic Commission for Latin America and the Caribbean, SCA-ECLAC. Santa Fe de Bogotá, Colombia, August 11 2009

The paper highlights the importance of the use of the International Statistical Standards for the harmonization of national statistics, as well as the problematic of its organization and efficient systematization. The terms harmonization and standardization are described, plus some tools to document metadata. The concept of the degree of concordance, to measure the existent variations, between international statistical standards and the national standards applied, is introduced. A methodological proposal is presented to help the statistical harmonization, based on the concordance measurement with respect to international statistical standards.

1. Introducción.

En México, en abril del 2006 el INEGI obtuvo de manera constitucional la autonomía técnica y de gestión. A partir de ese año se inició el debate legislativo de la nueva Ley bajo la cual operaría el Instituto y que normaría el Sistema Nacional de Información Estadística y Geográfica (SNIEG). Después de casi 2 años de trabajo legislativo, en el mes de abril del 2008, fue publicada la *Ley del Sistema Nacional de Información Estadística y Geográfica*¹ (*Ley del SNIEG*), como el instrumento jurídico para organizar el "Sistema" de acuerdo a los principios constitucionales de accesibilidad a la información, transparencia, objetividad e independencia. Los datos integrados al "Sistema" serán considerados oficiales y de uso obligatorio para la Federación, los Estados y los Municipios.

En esta Ley, se establece que la finalidad del *Sistema Nacional de Información Estadística y Geográfica (SNIEG)* es la de suministrar a la sociedad y al Estado, información de calidad, pertinente, veraz y oportuna, a efecto de coadyuvar al desarrollo nacional (Art. 3) y se confiere al INEGI las facultades suficientes para normarlo y coordinarlo.

El INEGI, conforme a los principios constitucionales, realizará las acciones tendientes a lograr: (a) la adecuación conceptual de la Información de Interés Nacional a las necesidades que el desarrollo económico y social del país impongan; (b) que la información sea comparable en el tiempo y en el espacio, y; (c) la adecuación de los procedimientos estadísticos y geográficos a estándares internacionales, para facilitar su comparación (Art. 54). Bajo este nuevo marco jurídico, el tema de los estándares y recomendaciones internacionales toma especial importancia en el ámbito de la estadística nacional.

¹ Ley del Sistema Nacional de Información Estadística y Geográfica. Diario Oficial de la Federación. 16 de abril del 2008. México.

2. La comparabilidad internacional de la información.

La necesidad de una armonización más profunda de las estadísticas sociales en el nivel internacional, es la consecuencia de los procesos de globalización y de la institucionalización progresiva de la política económica y social en una economía abierta dirigida por el mercado (Józef Olenski, 2003).

En el mundo actual, la información estadística oficial en áreas económica, social, ambiental y geográfica, juega un papel cada vez más importante para el entendimiento de fenómenos locales y globales, fluye con mayor frecuencia fuera de las fronteras y se integra para llevar a cabo procesos de análisis a diferentes niveles (tendencias, estimación de efectos, etc.). Sin embargo, la integración de información de diferentes fuentes puede ser un proceso complicado y dar como resultado interpretaciones erróneas cuando atributos aparentemente iguales son manejados bajo diferentes definiciones, códigos de clasificación o métodos de medición (Köhl, *et. al.* 2000). Para evitar esta situación, organismos internacionales, como la ONU, sugieren el uso de estándares y prácticas comunes para fomentar la coherencia² de los sistemas estadísticos a través de la aplicación de los *Principios Fundamentales de las Estadísticas Oficiales* (ONU, 1994). En la actualidad varios países contemplan en sus códigos de práctica la adopción de **referencias y estándares estadísticos internacionales (EEI)**, como parte de sus modelos de calidad de datos.

Además del fenómeno de *globalización*³, la emergencia de temas que afectan al mundo entero, como el *Cambio Climático*, el *Desarrollo Sostenible y la Pobreza*; el desarrollo de programas internacionales como *Los Objetivos del Milenio y La Agenda 21*; así como la construcción de Bases de Datos Mundiales, han tenido como consecuencia que la convergencia en el lenguaje estadístico sea una de las tendencias actuales en la evolución de los Sistemas Estadísticos Nacionales (Lehohla 2007).

En términos prácticos, la idea del uso de **EEI** es simple: ¿Podemos hablar el mismo lenguaje, de tal forma que se puedan hacer comparaciones válidas a nivel internacional de la información estadística que se genera en el plano nacional?

3. Los Estándares Estadísticos Internacionales.

3.1 ¿Qué son los Estándares Estadísticos Internacionales?

Integrar información nacional para fines comparativos a escalas regional e internacional, implica “homogeneizar” conceptos y métodos que permitan hacerla comparable. Los organismos internacionales proponen seguir un modelo común conformado por los denominados estándares y recomendaciones internacionales. En este sentido, entender ¿Qué es un estándar estadístico internacional?, ¿Quiénes y cómo los desarrollan y promueven?, ¿Cuándo son modificados?, ¿Quiénes y cómo los utilizan?, etc. es fundamental para que puedan ser debidamente incorporados dentro de un marco de calidad de datos acorde a las necesidades y situaciones particulares de cada país.

² Principio No. 9. La utilización por los organismos de estadística de cada país de conceptos, clasificaciones y métodos internacionales fomenta la coherencia y eficiencia de los sistemas estadísticos a nivel oficial (ONU, 1994). Dentro de los marcos de Calidad de Datos, el uso de estándares internacionales normalmente está relacionado con el término de “coherencia”.

³ Globalización: Tendencia de los mercados y de las empresas a extenderse, alcanzando una dimensión mundial que sobrepasa las fronteras nacionales. Diccionario de la Real Academia Española.

Por **Estándar Estadístico** se entiende, al “conjunto de guías para coleccionar información por fuentes administrativas o muestreos sobre un tópico particular y cuyos componentes son: (1) definiciones; (2) unidades estadísticas; (3) clasificaciones; (4) procesos de codificación; (5) cuestionarios, y; (6) resultados” (SDMX 2008). En tanto que se define como **Estándar Estadístico Internacional (EEI)**, al “conjunto comprensivo de guías estadísticas internacionales y recomendaciones que han sido desarrolladas por organismos internacionales con la participación de agencias estadísticas nacionales” (SDMX 2008).

De la primera definición hay dos aspectos relevantes:

(1) El que se refiere a los **tipos de estándar**:

- definiciones,
- unidades estadísticas,
- clasificaciones,
- procesos de codificación,
- cuestionarios y
- resultados

(2) El que está relacionado a los **metadatos**⁴, o elementos de información asociados a los estándares, y que son útiles para su conocimiento y aplicación; por ejemplo:

- ¿Quién(es) lo desarrolló(aron)
- ¿Quiénes colaboraron?,
- ¿Cuándo se desarrolló?,
- ¿Con qué objetivo?,
- ¿Existen otras versiones en uso?,
- ¿Quiénes los usan y cómo?, etc.

3.2 ¿Cómo están organizados los Estándares Estadísticos Internacionales (EEI)?

Los **EEI** son difundidos principalmente por los Organismos Internacionales (OI) a través de sus publicaciones y sitios en Internet, de acuerdo a sus propias estrategias y mecanismos de difusión. La gran cantidad de actores involucrados y de temas de aplicación, así como la propia dinámica de desarrollo por tema a lo largo del tiempo, han propiciado que los **EEI** se encuentren altamente dispersos en una gran variedad de fuentes de información (sitios en Internet), de medios de difusión (documentos, glosarios en línea, bases de datos, etc.) y con diferentes modelos de metadatos. Frecuentemente, los **EEI** vienen inmersos dentro de grandes documentos (Manuales, Marcos de Referencia, etc.) lo que hace que su seguimiento sea verdaderamente complicado. Para hacer evidente la problemática en su organización actual, se pueden mencionar los siguientes aspectos:

- (a) *Los diferentes grados de desarrollo de EEI por tema estadístico.* El tema de las Cuentas Nacionales ha sido abordado mucho tiempo antes que las Estadísticas Ambientales y cuenta con un marco de referencia o marco analítico⁵ bien desarrollado llamado *Sistema de Cuentas*

⁴ Metadato: son los datos que definen y describen otros datos y procesos (SDMX 2008).

⁵ Marco analítico: Un sistema conceptual estandarizado de definiciones y clasificaciones relacionadas a los datos (SDMX 2008).

Nacionales (SNA, por sus siglas en inglés) que compila los estándares y recomendaciones que le dan forma y que se difunde desde una sola fuente de información. Esto facilita su consulta y evolución en el tiempo. Por otro lado, el tema de las Estadísticas Ambientales, que ha sido abordado hace relativamente poco tiempo, 60 años de acuerdo con Trewin (2007), presenta aún varios marcos de referencia dispersos a través de diferentes actores (Tabla 1). Identificar si estos marcos son similares, complementarios y/o en qué medida comparten las mismas definiciones, clasificaciones y metodologías para estadísticas ambientales, no es una tarea sencilla, habría que llevar a cabo un análisis comparativo detallado.

Tabla 1. Marcos de Referencia Generales relacionados a Estadísticas Ambientales.		
Organismo Internacional	Año	Marco de Referencia
United Nations. Statistics Division	1984	The Framework for the Development of Environment Statistics (FDES)
United Nations. Statistics Division	1991	Statistics of the Natural Environment.
OECD. Organisation for Economic Co-operation and Development.	2003	Environmental indicators: Development, Measurement and Use.
United Nations. UNEP	2005	Towards a UNEP Environment Watch System.

- (b) *La diversidad de actores involucrados.* Dado que el desarrollo de estándares es llevado a cabo bajo la cooperación internacional por grupos de expertos de diferentes países, organismos e instituciones, es comprensible que una vez consensado un estándar o grupo de estándares, cada organismo participante lo publique o difunda de acuerdo a sus propios mecanismos y medios de difusión (incrementando la redundancia de información). Esto propicia que, si se siguen mecanismos diferentes de actualización que no estén debidamente coordinados entre los organismos, se eleve el riesgo de presentar versiones distintas acerca de un mismo estándar o tema de interés⁶. Un ejemplo pueden ser los diversos sitios referentes a los Indicadores de Desarrollo del Milenio cuya información, en ocasiones, no ha sido del todo homogénea aún a pesar de estar desarrollados bajo la misma organización internacional; la ONU (Tabla 2).

Tabla 2. Sitios en Internet de Naciones Unidas que difunden información de los indicadores de los Objetivos de Desarrollo del Milenio (Consulta realizada el 1° de abril del 2009).		
Sitio	OI responsable	Objetivos
United Nations site for the MDG Indicators. http://mdgs.un.org/unsd/mdg/Default.aspx	Naciones Unidas. División de Estadística. Departamento de Asuntos Económicos y Sociales.	Difusión de los indicadores (con metadatos), datos de cada país y reportes anuales.
United Nations Development Programme. Millennium Development Goals (MDGs). http://www.undp.org/mdg/ .	Naciones Unidas. Programa de las Naciones Unidas para el Desarrollo.	Difusión de los reportes anuales e indicadores por país (sin metadatos).

Un aspecto adicional que resulta del elevado número de actores y medios de difusión involucrados, es el relacionado con la información y los datos que se proporcionan de cada

⁶ Es claro que el uso de tecnologías de información ha facilitado la difusión de información y ha producido enormes beneficios a la sociedad pero también ha provocado una sobrecarga y redundancia de información al usuario final, fenómeno conocido como *information overload* (Heylighen 2002); y que no siempre se traduce en conocimiento (Hall 2007).

estándar (metadatos). Es frecuente encontrar información similar organizada bajo diferentes modelos de metadatos. (Tabla 3).

Tabla 3. Metadatos proporcionados por los Glosarios de Términos de OCDE, FAO y la División de Estadística de las Naciones Unidas, respecto de la definición de "Marco de Muestreo" (Sampling Frame). Consulta realizada el 1° de abril del 2009.		
OI responsable	Definición y medio de difusión	Metadatos
OECD Organisation for Economic Co-operation and Development.	Marco de Muestreo: "Una lista de todos los miembros de una población, usada como base para un muestreo. Sin tal marco, o sin su equivalente, los métodos de muestreo con propiedades de insesgamiento no estarían disponibles. El marco define a la población bajo estudio". Glossary of statistical terms http://stats.oecd.org/glossary/	<ul style="list-style-type: none"> - Título (con acrónimos) - Definición - Fuente de Publicación (título, autor, año; sin liga a la fuente) - Referencias cruzadas - Tema Estadístico - Fecha de creación. - Fecha de última actualización
United Nations Statistics Division	Marco de Muestreo: "Una colección de unidades (personas, hogares, instituciones, eventos, etc.) de las que una muestra puede ser extraída (Ver: Población)" UnData / Glossary http://data.un.org/Glossary.aspx	<ul style="list-style-type: none"> - Título - Definición - Fuente de Publicación (con liga a la fuente) - Términos relacionados
FAO	Marco de Muestreo: "Puede ser definido como una lista de unidades de muestreo que: (a) incluyen al (100%) de la población de interés sin duplicidad; (b) provee un medio claro para identificar a cada unidad de muestreo, y; (c) arregla estas características de tal forma que un muestreo probabilístico pueden ser llevado a cabo de manera eficiente. FAOSTAT / Glossary http://faostat.fao.org/site/379/default.aspx	<ul style="list-style-type: none"> - Nombre. - Definición. - Fuentes (autor, año y título, sin liga a la fuente) - Definiciones y conceptos relacionados

(c) *El elevado número de iniciativas internacionales que solicitan información nacional.* A la fecha existen muchas y variadas iniciativas internacionales (Convenios, Programas, Bases de Datos Mundiales, etc.) que requieren de información sobre fenómenos de estudio similares y que han dado como resultado soluciones diversas a un mismo problema.

Por ejemplo, en el año 2002 se hizo patente que existían diferentes definiciones acerca del término "BOSQUE" promovidas por dos diferentes organismos: FAO (Food and Agriculture Organization of United Nations) y UNFCCC (United Nations Framework Convention on Climate Change) para solicitar estadísticas nacionales. Mientras que la definición de FAO menciona que se deben considerar como *bosques* aquellas extensiones de tierra iguales o mayores a 0.5 hectáreas cubiertas por árboles con ciertas características, para el UNFCCC, el criterio establecía extensiones mínimas entre 0.05 y 1 hectárea, además de diferencias en cuanto al

tamaño de los árboles y su cobertura de dosel (Tabla 4). A este respecto, existe un estudio para los países de la Comunidad Europea que ha documentado el efecto que resultaría de aplicar distintas definiciones de “BOSQUE” en el cálculo del indicador de “área o superficie boscosa”, lo que consecuentemente llevaría a interpretaciones diferentes o erróneas (Köhl, *et. al.* 2000).

Tabla 4. Definiciones de “Bosque” de acuerdo a FAO y al Marco Convención de Cambio Climático. Tomado de Puustjärvi y Simula (2002).	
OI responsable	Definición y medio de difusión
FAO	<p>Bosque:</p> <p>“Incluye bosques naturales y plantaciones. Se usa para referir tierra con árboles con altura mínima de 5 metros y cobertura de dosel de más del 10% que ocupan un área mayor a 0.5 hectáreas.”</p>
UNFCCC	<p>Bosque:</p> <p>“Es un área mínima de tierra de entre 0.05 a 1 hectáreas con árboles con el potencial de alcanzar en madurez una altura mínima en sitio de entre 2 a 5 metros y cobertura de dosel de más del 10 al 30% ”</p>

Es evidente que, aunque ha habido un enorme progreso en el desarrollo de **EEI** en diversos temas, existen todavía retos que enfrentar con respecto a su eficiente organización (desarrollo, actualización y coordinación internacional), sistematización (modelos de metadatos) y difusión y, sobre todo, hay una necesidad de desarrollar un mecanismo que elimine la redundancia de información, facilite el acceso a las fuentes originales y considere un modelo común de metadatos.

4. La estandarización y la armonización.

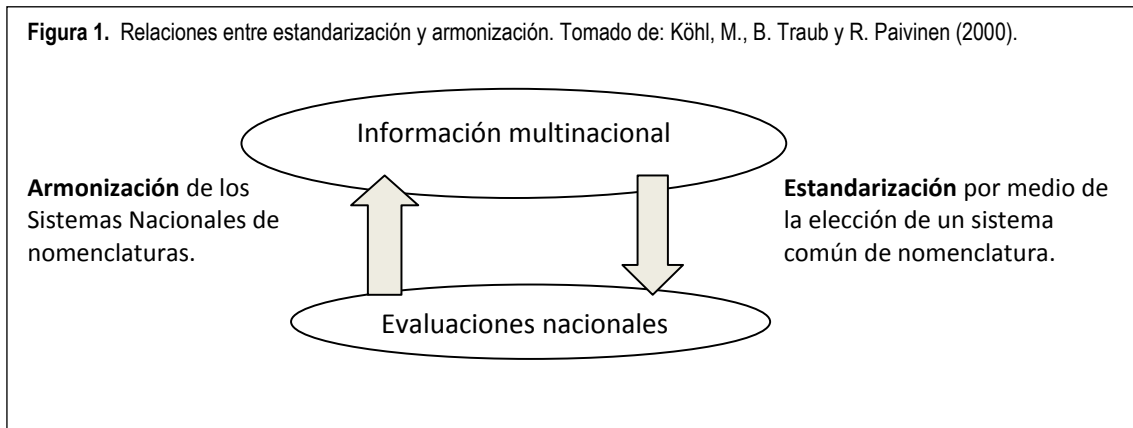
Dada la gran relevancia que han llegado a tener los temas de estandarización y armonización en las agendas de los Organismos Internacionales y en las Oficinas Nacionales de Estadística (ONE), es importante, en principio, establecer las diferencias entre ambos conceptos.

Por **estandarización** se entiende, la introducción de un conjunto común de estándares (definiciones, clasificaciones, métodos, indicadores, etc.) que pueden ser aplicados en los procesos de generación de información estadística nacional.

En tanto que la **armonización** es considerada como un proceso que busca un acuerdo común acerca de cómo los datos deben ser manejados, de tal forma que la información que se produce pueda ser fácilmente interpretada y comparable a niveles supranacionales (Köhl, *et. al.* 2000) (Figura 1).

La meta principal de ambos procesos es introducir, sistemáticamente, un conjunto de variables armonizadas en cada instrumento de captación (registros, encuestas, métodos de monitoreo, etc.) que facilite las comparaciones a través del tiempo y a lo largo de diferentes áreas geográficas y otros dominios de información (industrias, actividades, etc.). Los datos producidos por diferentes fuentes, y en diferentes tiempos, se podrían agrupar sin problemas para enriquecer las bases de datos usadas para análisis más detallados y para la toma de decisiones. Los beneficios se podrían hacer evidentes en varios aspectos: reducción de redundancias, eficiencia y calidad mejoradas y, por supuesto, aumento en la comparabilidad de la información estadística (Isfan 2007).

Figura 1. Relaciones entre estandarización y armonización. Tomado de: Köhl, M., B. Traub y R. Paivinen (2000).



Sin embargo, los procesos de armonización son frecuentemente lentos y complejos, ya que la búsqueda de estos acuerdos (que pueden ser a escala nacional, regional o internacional) debe partir de un diagnóstico previo muy detallado que incluya los siguientes aspectos:

- (a) Tener una idea clara de lo que se desea armonizar, qué temas, que subtemas, qué variables y seleccionar los estándares que serán tomados como base;
- (b) Recopilar todos los elementos conceptuales y metodológicos inmersos en las prácticas nacionales para la generación de las estadísticas de interés;
- (c) Evaluar las diferencias o variaciones entre los elementos conceptuales y metodológicos usados en las prácticas estadísticas nacionales (normas o estándares locales) y los estándares estadísticos seleccionados como referencia, lo que equivaldría a medir el **grado de concordancia con estándares internacionales**;
- (d) Conocer los instrumentos legales ligados a las prácticas estadísticas correspondientes: normas y leyes locales.

Es importante contemplar que los resultados de la armonización puedan ser llevados a la práctica sin impedimentos o contradicción con las normas y leyes locales.

5. Los Sistemas de metadatos estadísticos.

Para facilitar los procesos de armonización de las estadísticas oficiales, se han propuesto una gran cantidad de iniciativas regionales e internacionales (Tabla 5) que compilan información sobre los metadatos de las prácticas nacionales, es decir, sobre los aspectos relacionados a la recolección, el proceso, el almacenamiento, la difusión y el uso correcto de los datos estadísticos. La mayoría de éstas se han acompañado del desarrollo de sistemas de información denominados *Sistemas de Metadatos Estadísticos* o *Meta-sistemas* y constituyen una fuente de información fundamental para el uso correcto de los datos sobre los que estos informan (UNECE 1995).

Los metadatos han llegado a ser tan importantes que existen áreas especializadas en la administración de metadatos, las cuales están abocadas al contenido, estructura y diseño requeridos para manejar el vocabulario y terminología que describen a los datos estadísticos, su diseño y procesos asociados (Bargmeyer 2000).

Tabla 5. Algunas Iniciativas Internacionales relacionadas a Metadatos Estadísticos.	
Metasistema o Sistema	Breve descripción.
METIS	Grupo de Trabajo sobre Metadatos Estadísticos (METIS) de la UNECE (United Nations Economic Commission for Europe). Mantiene una base de conocimiento dinámica para facilitar la armonización de estructuras y modelos de datos para metadatos estadísticos en el contexto del procesamiento y la difusión de la información estadística.
SDMX	SDMX : Standard Data and Metadata Exchange. Desarrollo de estándares técnicos para el intercambio de datos y metadatos, con énfasis en datos agregados. Promovido por: BIS - ECB - EUROSTAT - IMF - OECD - UN - World Bank (siglas en Inglés).
Dublin Core	Dublin Core : Es una organización que desarrolla estándares de metadatos para datos de un amplio marco de tipos de datos y modelos de negocios. Integrado por Bibliotecas Nacionales (Singapur, Nueva Zelanda, Finlandia y Corea).
Otras iniciativas concluidas	GESMES : servicio usado por el Banco Central Europeo para el intercambio de datos y metadatos con sus socios en el Sistema Europeo de Bancos Centrales. Ahora incorporada al proyecto SDMX . AMRADS : (Accompanying Measure to Research and Development in Official Statistics) Medidas empleadas para la investigación y desarrollo de estadísticas oficiales (2000 - 2003). MetaNet : red creada para sintetizar y armonizar los desarrollos en metadatos estadísticos (2000-2003). COSMOS : Cluster of Systems of Metadata for Official Statistics. Grupo de Sistemas de Metadatos para estadísticas oficiales (junio, 2003). METAWARE : Statistical Metadata Support for Data Warehouses. Metadatos estadísticos para almacenes de datos. (febrero, 2003).

Sin embargo, en muchos de estos Meta-sistemas, las prácticas estadísticas quedan documentadas con modelos de metadatos que no necesariamente detectan las posibles discrepancias con estándares o recomendaciones internacionales, ya que no se solicita directamente que se hagan explícitas estas variaciones. Este aspecto, que en este artículo se le denomina “**concordancia**”, es la pieza fundamental que se considera que hace falta a los sistemas de metadatos para facilitar los diagnósticos, que son la base de los procesos de armonización. Posiblemente este concepto de “**concordancia**” no se ha incorporado a los Meta-sistemas dado que es un conocimiento que es específico a cada contexto, es difícil de formalizar, de codificar y, por lo tanto, de registrar. Revisando algunos Meta-sistemas se resalta este problema.

5.1 Ejemplos de iniciativas de metadatos para estadísticas económicas.

En el campo de las estadísticas económicas, el Fondo Monetario Internacional mantiene los sistemas denominados *Special Data Dissemination Standard* (**SDDS**)⁷ y el *General Data Dissemination System*⁸ (**GDDS**). El primero fue iniciado en 1996 cubriendo datos macroeconómicos clave, en tanto que el sistema **GDDS** inició en 1997 y es menos prescriptivo que el primero. Los temas y las dimensiones de los metadatos que cubren son diferentes y lo mismo aplica para otros metasistemas y sistemas mixtos

⁷ Special Data Dissemination Standard (SDDS) Site: <http://dsbb.imf.org/Applications/web/sddshome/>

⁸ General Data Dissemination System (GDDS) Site: <http://dsbb.imf.org/Applications/web/gdds/gddshome/>

(aquellos que registran tanto los datos de las variables estadísticas como sus metadatos) como **Laborsta**⁹, que es una base de datos operada por el Consejo de Estadística de la Organización Internacional del Trabajo (Tabla 6).

Tabla 6. Metadatos de algunos metasistemas.		
Metasistema o Sistema	Temáticas consideradas	Modelo de metadatos (Dimensiones)
SDDS	Sector Real, Sector Fiscal, Sector Financiero, Sector Externo, Información Socio-demográfica (Población)	<ol style="list-style-type: none"> Características de los Datos: <ul style="list-style-type: none"> - Cobertura - Periodicidad - Disponibilidad Acceso: <ul style="list-style-type: none"> - Liberación simultánea - Información anticipada de liberación Integridad: <ul style="list-style-type: none"> - Términos y condiciones, etc. Calidad <ul style="list-style-type: none"> - Fuentes metodológicas - Validación cruzada - Observaciones Metodología <ul style="list-style-type: none"> - Marco analítico: conceptos, definiciones y clasificaciones. - Alcance de los datos - Convenciones contables - Naturaleza del dato básico - Prácticas de compilación - Otros aspectos
GDDS	Sector Real, Sector Fiscal, Sector Financiero, Sector Externo, Información Socio-demográfica (Población, Salud, Educación, Pobreza)	<ol style="list-style-type: none"> Características de los Datos: <ul style="list-style-type: none"> - Cobertura - Periodicidad - Disponibilidad Calidad: <ul style="list-style-type: none"> - Documentación de la metodología Planes de Mejoramiento Formatos de Difusión.
LABORSTA	Variables del tema de Empleo Desempleo, así como algunas variables económicas (índice de precios al consumidor).	<ol style="list-style-type: none"> Definiciones Clasificaciones Fuentes y Métodos

Realizando una consulta a los metadatos para la variable: *Índice mensual de la actividad industrial* (Production Index)¹⁰ en el sistema **SDDS**, en el caso de México se puede corroborar que los datos que se usan para el cálculo de este índice son clasificados de acuerdo al *Sistema de Clasificación Industrial de América del Norte* (SCIAN); es evidente entonces que para esta variable se aplica el estándar SCIAN y es comparable (está armonizada) con Estados Unidos y Canadá; al menos hasta el

⁹ LABORSTA: <http://laborsta.ilo.org/>

¹⁰ Subscribing Countries' Production index metadata: Summary Methodology. Production index
<http://dsbb.imf.org/Applications/web/sddscategorycountrylist/?strcat=IND00&strcatname=Production%20index>

nivel de industria (hasta el dígito 5 en los códigos de esta clasificación)¹¹. Si se realiza la misma consulta pero para otro país; por ejemplo, Argentina, se encuentra que la clasificación utilizada es la *International Standard Industrial Classification of all Economic Activities* (ISIC Rev.3), que recomienda las Naciones Unidas. Para el caso de México, surgen entonces varias preguntas: ¿Sería este indicador comparable con los países que, como Argentina, siguen la clasificación recomendada por Naciones Unidas?, ¿A qué nivel?, ¿Cómo se puede hacer comparable la información?, ¿Existe alguna tabla de equivalencias? Esta información no queda explicada en el sistema **SDDS**, de tal forma que para responder a éstas preguntas habría que consultar más fuentes de información.

En el caso de la base de datos **LABORSTA**, ocurre algo similar. Las posibles concordancias o discrepancias con algún estándar o recomendación internacional no quedan visibles a un nivel tal que sean evidentes de manera sencilla para cualquier usuario potencial, tenga o no conocimientos especializados en el tema. Por ejemplo, haciendo una consulta de los metadatos para México, respecto de las definiciones que se siguen para el Censo de Población y Vivienda¹², se describe a detalle la definición de **Empleo** en uso (Tabla 7).

Tabla 7. Metadatos en LABORSTA para México en el caso de las definiciones utilizadas en los Censos de Población. (Consulta realizada el 1° de abril del 2009)	
Definición de empleo que aparece en LABORSTA ¹²	
<p>Definición Nacional de Empleo:</p> <p>Empleo: Se consideran empleadas u ocupadas las personas de 12 años y más, que respondieron afirmativamente a las preguntas de Condición de actividad: “La semana pasada, esta persona: ¿trabajó?; ¿tenía trabajo, pero no trabajó?”. Incluye también dentro de la población con empleo a las personas que declararon a las preguntas anteriores “no”, esto es, no reconocieron trabajar, pero en la pregunta de Verificación de actividad respondieron realizar alguna actividad económica en la semana de referencia: “¿La semana pasada... ayudó en un negocio familiar?; ¿vendió algún producto?; ¿hizo algún producto para vender?; ¿ayudó a trabajar en el campo o en la cría de animales?; ¿a cambio de un pago realizó otro tipo de actividad?”. Se consideran ocupadas las personas de 12 años y más, que realizaron alguna actividad económica, al menos una hora en la semana de referencia.</p> <p>Quedan incluidas las siguientes categorías:</p> <ul style="list-style-type: none"> i) personas que trabajan sin remuneración en un negocio familiar; ii) personas en la producción de bienes primarios para el autoconsumo; iii) personas ocupadas pero temporalmente ausentes de su trabajo; iv) trabajadores estacionales u ocasionales que trabajaron en la semana de referencia; v) personal de las fuerzas armadas. 	

Esta definición, ¿está apegada a la definición internacional de empleo recomendada por la OIT? (descrita en “*La Resolución sobre estadísticas de la población económicamente activa, del empleo, del desempleo y del subempleo*”, adoptada por la decimotercera Conferencia Internacional de Estadísticos del Trabajo en Ginebra, 1982). Para detectar las divergencias sería necesario realizar un análisis comparativo adicional.

5.2 Ejemplos de iniciativas de metadatos regionales y nacionales.

En América Latina, La Comisión Económica para América Latina y el Caribe (CEPAL), difunde estadísticas oficiales de la región a través de diversos medios, uno de ellos es el sistema

¹¹ La estructura del SCIAN se puede consultar en: <http://www.naics.com/info.htm>

¹² LABORSTA. Metadatos para México. <http://laborsta.ilo.org/applv8/data/SSM5/S/MX.html>

CEPALSTAT¹³ compuesto por bases de datos de las diferentes Divisiones de la Comisión y por publicaciones estadísticas. Para las variables e indicadores que integran la base de datos, el modelo de metadatos incluye los siguientes elementos: (1) definición; (2) unidad de medida; (3) metodología de cálculo; (4) tipo de dato; (5) comentarios, (6) fuentes, y; (7) dimensiones y desagregaciones. La información se presenta a nivel general, sin detalle de las particularidades nacionales de los países de los que se integra la información.

Por otro lado, La Conferencia Estadística de las Américas de la CEPAL (CEA-CEPAL), mantiene una base de datos de Proyectos Multilaterales (**BADEPRO**¹⁴) en la que se hace mención a 21 proyectos de armonización registrados desde finales de los años 90 y para diferentes temas estadísticos. Desafortunadamente, la información detallada acerca de cómo se llevaron a cabo los procesos y los resultados obtenidos, no están disponibles en la misma Base de Datos referida.

Finalmente, a nivel nacional, también existen modelos de metadatos que han ido evolucionando a lo largo del tiempo para proporcionar información detallada de las variables que son objeto de medición. Sin embargo, aún existen modelos de metadatos que no permiten detectar la aplicación de estándares internacionales. Un ejemplo lo constituye el *Sistema Nacional de Indicadores Ambientales (SNIA)*¹⁵ de la SEMARNAT, en el que para algunos indicadores no es factible identificar la aplicación de estándares y/o recomendaciones internacionales (Figura 2).

Figura 2. Metadatos proporcionados para el indicador: “*Extensión de bosques y selvas*”. Tomado de la fuente: SNIA. Sistema Nacional de Indicadores Ambientales. SEMARNAT. 2009. Liga: <http://www.semarnat.gob.mx/informacionambiental/snia/Pages/index-snia.aspx>. Se observa que no hay referencias documentales a la definición de los términos “bosque” y “selva”, objetos de medición del indicador (Consulta realizada el 1° de abril del 2009).

Metadatos	
Nombre:	Extensión de bosques y selvas.
Definición breve:	Superficie cubierta por bosques y selvas.
Unidad de medida:	km ²
Objetivos y metas:	No definidos.
Definiciones y conceptos:	No aplica.
Método de medición:	La superficie de bosques y selvas se estimó a partir de la Cartografía de uso del suelo y vegetación serie III 2002 del INEGI a partir de la interpretación de imágenes.
Periodicidad:	No definida.
Limitaciones del indicador:	No aplica.
Fuentes de datos:	Con base en: Instituto Nacional de Estadística, Geografía e Informática, Conjunto de datos vectoriales de uso de suelo y vegetación, 2002-2005. Escala 1:250,000, Serie III. 1a edición, INEGI, México, 2005.
Referencia:	No aplica.

¹³ CEPALSTAT. Sistema de Bases de Datos y Publicaciones Estadísticas de la CEPAL:

<http://www.eclac.org/estadisticas/default3.asp>

¹⁴ BADEPRO. Base de Datos de Proyectos Multilaterales. <http://www.eclac.cl/deype/ceacepal/proyectos.htm>

¹⁵ SEMARNAT. Sistema Nacional de Indicadores Ambientales (SNIA).

<http://www.semarnat.gob.mx/informacionambiental/snia/Pages/snia.aspx>

5.3 Las iniciativas de metadatos y el problema de la documentación del uso de estándares.

Por el momento, la evidencia muestra que muchas prácticas estadísticas están documentadas aunque con modelos de metadatos distintos. Sin embargo, la medición de convergencias y/o divergencias respecto de los **EEI, o grado de concordancia**, que permita identificar de forma más sencilla en dónde es factible enfocar los esfuerzos de armonización, o no han sido identificadas o no han sido difundidas de manera precisa.

Considerando la importancia del papel de los metadatos para facilitar el diagnóstico de los estándares y recomendaciones internacionales que siguen los países para la producción de estadísticas oficiales, es necesario trabajar en un modelo de metadatos que incorpore la información acerca de cuáles **EEI** se aplican y el grado de concordancia con los estándares o normas nacionales, así como la factibilidad de realizar ajustes en función de los instrumentos jurídicos locales o nacionales. Es claro que se trata de un problema complejo que requiere de múltiples aproximaciones. En países como Australia, por ejemplo, su sistema de metadatos ha sufrido muchos cambios importantes desde 1905 hasta terminar con una estrategia de largo plazo que incluye documentar a nivel de variables las variaciones que se tienen respecto de estándares internacionales, usando como base en el estándar **ISO/IEC 11179**¹⁶ para documentar variables (Australian Bureau of Statistics, 2008).

6. Hacia la medición de la concordancia de las estadísticas nacionales con respecto a los estándares estadísticos internacionales.

Una de las obligaciones actuales que tiene el INEGI consiste en dar cumplimiento a la *Ley del SNIEG*, respecto de lo establecido para el uso de estándares y recomendaciones internacionales¹⁷, con el objetivo de mejorar la integración regional e internacional de la información estadística oficial del país.

El papel del INEGI como coordinador del SNIEG, le da la oportunidad de proponer métodos de trabajo que eliminen la duplicidad de esfuerzos y de recursos, por lo que es importante considerar el desarrollo de un modelo único de metadatos para todo el Sistema, con una visión de largo plazo; que pueda ser alimentado por todos los generadores de información (Unidades de Estado¹⁸); y, que haga visibles las variaciones que existen entre las prácticas metodológicas nacionales y los estándares internacionales. A nivel operativo, sería conveniente el diseño de una estrategia de trabajo que contemple los siguientes puntos:

- (1) Tener una identificación de los temas, subtemas y variables que conformarán la Información de Interés Nacional;

Para los temas, subtemas y variables propuestos:

- (2) Identificar los **EEI** existentes (tipos de estándares y versiones) que pudieran estar relacionados y que debieran ser tomados en cuenta;

¹⁶ ISO/IEC 11179 Metadata Registry (MDR) standard. Normas para el registro de metadatos: cómo clasificar los datos y cómo almacenar y administrar las descripciones de los datos.

¹⁷ Los estándares y recomendaciones internacionales son referidos en: Artículo 30, Fracciones IV y VI; Artículo 33, Fracción IV; Artículo 36, Fracciones III y IV; Artículo 54, Fracción III; Artículo 55, Fracciones I y II; Artículo 57 y Artículo 89 de la Ley del SNIEG. Diario Oficial de la Federación. 16 de Abril del 2008.

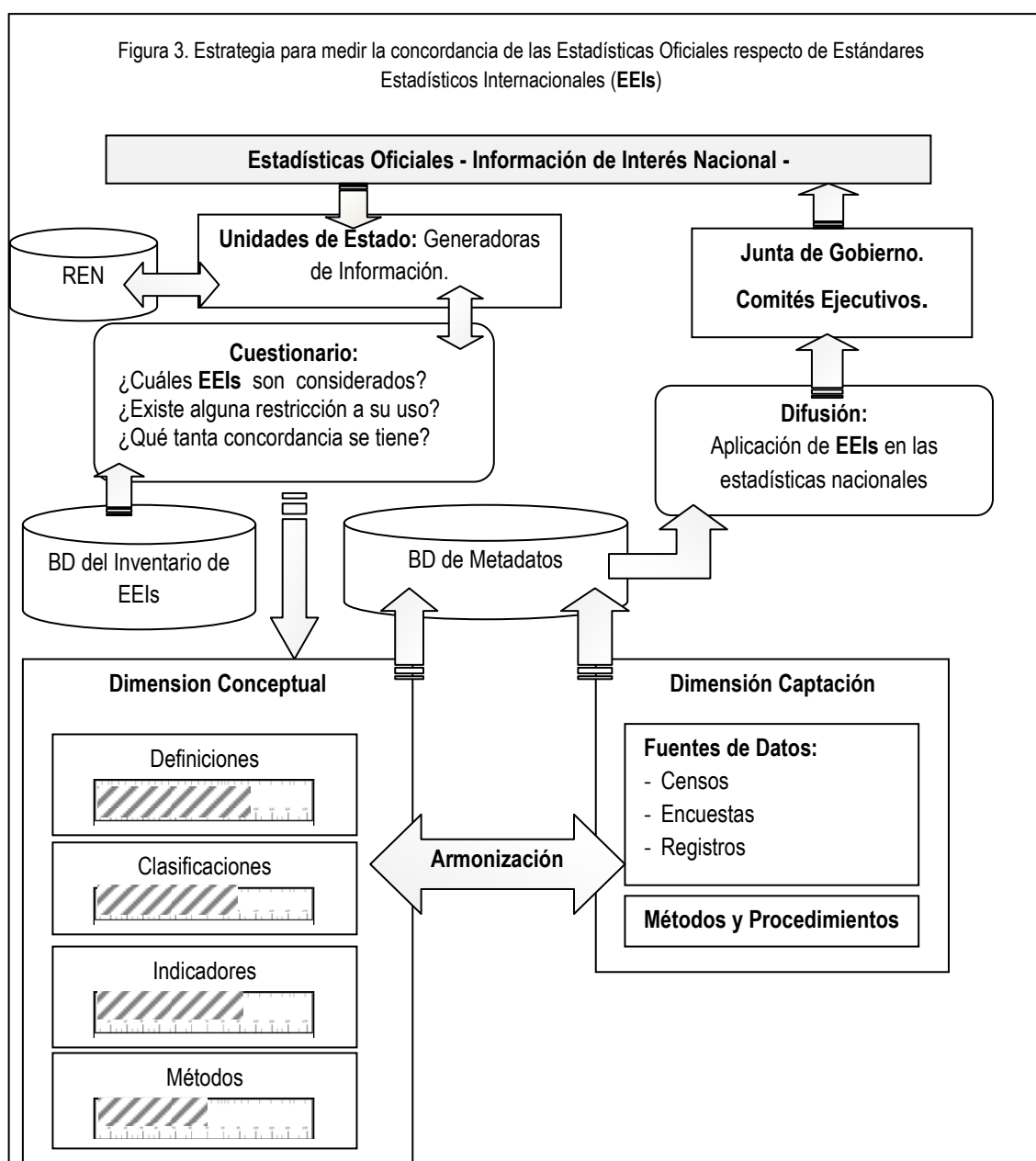
¹⁸ Las Unidades del Estado son las áreas administrativas con atribuciones para desarrollar actividades estadísticas y geográficas para obtener Información de Interés Nacional. Artículo 2 de la Ley del SNIEG. Diario Oficial de la Federación. 16 de Abril del 2008.

- (3) Llevar a cabo un diagnóstico de los estándares internacionales aplicados y las posibles variaciones respecto de los mismos (**medición del grado de concordancia**), por medio de la aplicación de cuestionarios a los generadores de la información (Unidades de Estado);
- (4) Documentar las variaciones respecto de los **EEI** y sistematizar y almacenar la información en un solo modelo de Metadatos.
- (5) Difundir la información sobre la concordancia con los **EEI** a los Comités Ejecutivos y Técnicos Especializados, para que de acuerdo a sus criterios, pueda o no ser empleada en la generación de normas técnicas.

El primer paso está previsto en la *Ley del SNIEG*, que otorga facultades al Consejo Consultivo Nacional¹⁹ de proponer los temas, la información y los indicadores que puedan considerarse de Interés Nacional para su aprobación por la Junta de Gobierno del INEGI. Para el segundo paso: la identificación de estándares internacionales existentes para los temas de interés nacional, se propone el uso del **Inventario de Estándares Estadísticos Internacionales del INEGI** que facilite su búsqueda, identificación y conocimiento. Respecto de los puntos (3) y (4) se requiere, en principio, adoptar un modelo de metadatos adecuado a las necesidades de integración nacional, trabajar posteriormente en el fortalecimiento de las capacidades institucionales de las diferentes Unidades de Estado para consolidar el uso y la administración del modelo propuesto y aplicar una metodología para diagnosticar las variaciones existentes entre los diversos elementos de las prácticas estadísticas nacionales respecto de los estándares y recomendaciones internacionales, a través de cuestionarios especializados. Esto significaría un gran avance hacia el fortalecimiento de la infraestructura estadística nacional.

La estrategia está inmersa dentro del marco de la nueva *Ley del SNIEG* tomado en consideración a los nuevos actores (Unidades de Estado, Comités Ejecutivos, etc.) y las herramientas ya desarrolladas (El Inventario de Estándares) (Figura 3).

¹⁹ Art. 2 de la Ley del SNIEG. Diario Oficial de la Federación. 16 de Abril del 2008.



Una vez determinada la Información de Interés Nacional, se avanzará hacia la generación de los diagnósticos del grado de concordancia con los estándares internacionales, para tener una visión clara de sus posibilidades de comparación internacional.

En principio, se pueden ubicar las Unidades de Estado que participan en la generación de información a través del Registro Nacional de Información Geográfica y el Registro Estadístico Nacional (REN) para, posteriormente, preparar un cuestionario de diagnóstico de la aplicación de estándares y sus grados de concordancia a nivel de la dimensión conceptual (Definiciones, Clasificaciones, Métodos de cálculo e Indicadores) por cada instrumento de registro (dimensión de captación o fuentes de datos). Los estándares y tipos de estándares pueden ser identificados fácilmente a través del **Inventario de Estándares Estadísticos Internacionales del INEGI**. Los resultados podrían ser almacenados en un Meta-sistema único para todo el Sistema, que, con una adecuada difusión hacia los Comités Ejecutivos

temáticos, sería de gran ayuda para decidir los pasos y acciones necesarias a tomar hacia una armonización de los diferentes instrumentos de captación.

7. El Inventario de Estándares Estadísticos Internacionales del INEGI.

Para construir un conjunto efectivo de estándares y clasificaciones alrededor de un tema, un buen punto de partida sería contar con un marco de referencia global que cubra toda la información de interés relacionada al tema, que proporcione un mapa gráfico del terreno, que delimite los conceptos más importantes y que muestre en una forma lógica cómo están interrelacionados.... Para ser efectivos, los estándares deben estar no solamente bien desarrollados, sino también estar actualizados y rápidamente accesibles a usuarios potenciales. (Linacre & McColl, 2007).

Con el objetivo de organizar los **EEI** en una sola base de datos, que compile los diferentes tipos de estándares bajo un modelo único de metadatos (Tabla 8), que facilite el acceso a las fuentes originales de información y que elimine redundancias e inconsistencias en su difusión, el INEGI ha puesto en marcha el desarrollo de un **Inventario de Estándares Estadísticos y Geográficos Internacionales (Inventario)**²⁰. El **Inventario** es un sistema de información que concentra los estándares y recomendaciones internacionales de diversos temas, subtemas y variables estadísticas, que permite revisar brevemente el contenido de un estándar así como los atributos que le son relevantes (metadatos). La información de cada estándar se presenta a manera de “Ficha Técnica” que facilita la comprensión del mismo y que mantiene el acceso a la(s) fuente(s) de información original(es) y relacionadas (Figura 4).

Los estándares fueron clasificados, con base en la definición de estándares de la SDMX en las siguientes cinco categorías:

- Conceptos y definiciones;
- Clasificaciones;
- Metodologías y procedimientos;
- Fuentes de datos;
- Indicadores.

Adicionalmente, el Inventario también es una puerta de entrada única a las diferentes formas de difusión de los estándares:

- Marcos de referencia generales;
- Marcos estadísticos o analíticos
- Glosarios de términos
- Sistemas de Información o de Bases de Datos.

²⁰ Inventario de Estándares Estadísticos y Geográficos Internacionales: <http://mapserver.inegi.org.mx/estandares/>

Tabla 8. Metadatos proporcionados por el Inventario	
Elementos del metadato	Descripción
Nombre del Estándar	Nombre del Estándar
Tema estadístico	De acuerdo a la clasificación de temas del INEGI
Organismos Internacional que lo publica	OI responsable de publicarlo
Fuente de información original	La fuente puede ser documental (manual, guía, marco de referencia, etc.) un glosario de términos, un sistema de información o una bases de datos. En algunos casos éstos son Marcos de Referencia que proporcionan una guía conceptual o teórica para la generación de información y liga en INTERNET (URL).
Contribuidores	Otros OI que contribuyeron al desarrollo del estándar
Tipo de estándar	<ul style="list-style-type: none"> • Conceptos y definiciones • Clasificaciones • Fuentes de datos • Metodologías y procedimientos: • Indicadores
Versión	Se refiere al año de la publicación del documento que contiene el estándar.
Versión previa	Se refiere al año de la publicación de la versión previa inmediata anterior que contiene el estándar.
Objetivo	Describe el propósito más general que persigue el estándar.
Descripción	Representa una breve explicación de lo planteado por el estándar.
Fecha de elaboración	Fecha en la que se elaboró la ficha técnica.

Figura 4. Ejemplo de Ficha Técnica en el Inventario

DEFINICIÓN DE ÁREAS PROTEGIDAS (UNEP-1994)
DEFINING PROTECTED AREAS (UNEP-1994)

TEMA: Estadísticas Ambientales - Ambiente Natural - Biodiversidad

Temas relacionados

Temas Geográficos - Datos Fundamentales - Recursos Naturales y Medio Ambiente

ORGANISMO INTERNACIONAL: United Nations Environment Programme (UNEP)

FUENTE:

1 UNEP, "Guidelines for Protected Area Management Categories ". 1994.

http://www.unep-wcmc.org/protected_areas/categories/eng/index.html

CONTRIBUIDORES: The World Conservation Union (IUCN), World Comission on Protected Areas (WCPA)

TIPO DE ESTÁNDAR: Conceptos y Definiciones

VERSIÓN PREVIA DEL ESTÁNDAR: 1978

ÚLTIMA ACTUALIZACIÓN DEL ESTÁNDAR: 1994

OBJETIVO: Proporcionar la definición de áreas protegidas para el mantenimiento de la biodiversidad.

DESCRIPCIÓN:

Un área protegida se define como una superficie de tierra y/o mar especialmente dedicada a la protección y el mantenimiento de la diversidad biológica, así como de los recursos naturales y culturales asociados y manejados a través de medios jurídicos u otros medios eficaces.

Esta definición incluye el "universo" de áreas protegidas pero en la práctica, el propósito preciso para el cuál el área protegida es administrado, difiere enormemente. Los principales propósitos considerados son:

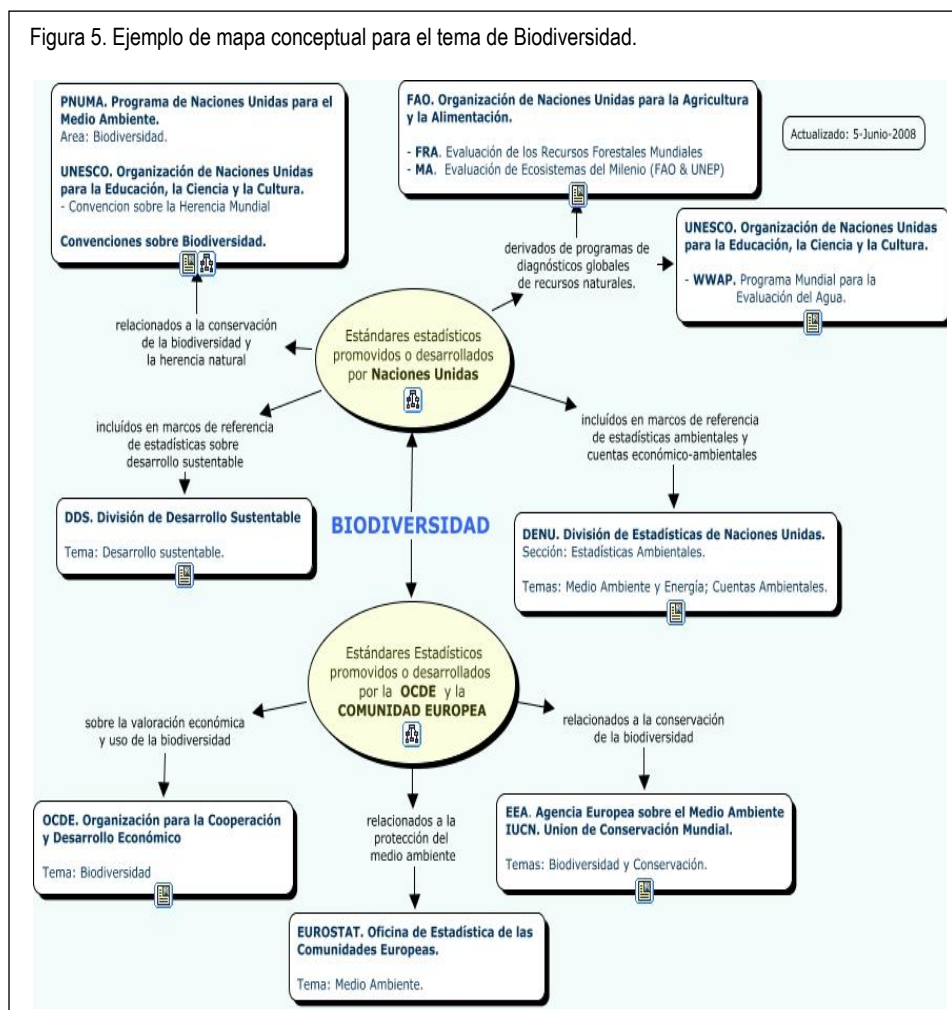
- 1) Investigación científica.
- 2) Protección de la vida salvaje.
- 3) Preservación de especies y diversidad genética.
- 4) Mantenimiento de los servicios ambientales.
- 5) Protección de características naturales y culturales específicas.
- 6) Turismo y recreación.
- 7) Educación.
- 8) Uso sustentable de recursos provenientes de ecosistemas naturales.
- 9) Mantenimiento de atributos culturales y tradicionales.

FECHA DE ELABORACIÓN DE LA FICHA TÉCNICA: 30-05-07

Por último y buscando un mejor entendimiento de la organización de los estándares, se desarrollan "mapas o diagramas conceptuales"²¹ que puedan explicar de una manera más sencilla cómo se relacionan los organismos internacionales y otros actores con los **EEI**. Esto es útil en áreas en donde hay múltiples marcos de referencia (Figura 5).

²¹Los mapas o diagramas conceptuales están elaborados en IHCM-tools, una herramienta de uso libre cuando no se persiguen fines comerciales: <http://cmap.ihmc.us/conceptmap.html>

Figura 5. Ejemplo de mapa conceptual para el tema de Biodiversidad.



Aunque el Inventario es una iniciativa desarrollada bajo una perspectiva de uso nacional, su uso puede extenderse más allá de este ámbito. En este sentido, el Inventario ha sido utilizado por el proyecto *Integración y Armonización de Instrumentos para las Estadísticas Sociales*²², de la Comisión de Estadística de las Américas de la CEPAL (CEA-CEPAL) y dirigido por el Departamento Administrativo Nacional de Estadística de Colombia (DANE), como un medio de acceso rápido a los diferentes **EEI** en los temas de Mercado Laboral, Salud y Educación.

8. Conclusiones

A través del uso de los estándares recomendados por los principales organismos internacionales en la generación de información nacional, se podrá fomentar la coherencia y eficiencia de los sistemas de información estadística. De esta forma es posible integrar adecuadamente información con atributos comunes de tal manera que se haga más comparable a diversas escalas; estatal, nacional, regional e internacional.

²² Integración y Armonización de Instrumentos para las Estadísticas Sociales (INAES). <http://www.dane.gov.co/inaes/>

La documentación adecuada de los diferentes atributos de la información (metadatos) es un procedimiento que puede ayudar en este esfuerzo. En este sentido es que se propone la aplicación de una metodología que integre los estándares estadísticos internacionales en un Inventario: una sola base de datos sistematizada por cinco tipos de estándares y por temas, subtemas y variables específicas. Se emplee estos estándares para medir su concordancia con las prácticas estadísticas nacionales, usadas para producir, integrar, conservar y difundir información; y se documenten las posibles variaciones respecto de los estándares internacionales dentro de un modelo de metadatos único a nivel nacional.

El INEGI, con sus facultades para producir información, coordinar y normar el Sistema Nacional de Información, está en posibilidad de coadyuvar al desarrollo de la información estadística y geográfica en México mediante la adecuación de los procedimientos estadísticos y geográficos a estándares internacionales, así como a las mejores prácticas en la materia para facilitar su comparación.

Referencias.

Australian Bureau of Statistics, 2008. METIS Case Studies. Publicado en línea para el Statistical Metadata (METIS) System - Common Metadata Framework. United Nations Economic Commission for Europe (UNECE).

<http://www1.unece.org/stat/platform/display/metis/Australian+Bureau+of+Statistics>

Bargemeyer, B.E. and D. W. Gillman. 2000. Metadata Standards and Metadata Registries: An Overview. Trabajo presentado en la *International Conference on Establishment Surveys II*, del U.S. Bureau of Labor Statistics. Buffalo, Nueva York. <http://www.bls.gov/osmr/abstract/st/st000010.htm>

Hall, Jon. 2007. A Global Project on Measuring the Progress of Societies, OECD World Forum on Statistics, Knowledge and Policy.

Heylighen, Francis. 2002a. Complexity and Information Overload in Society: why increasing efficiency leads to decreasing control. <http://pespmc1.vub.ac.be/Papers/Info-Overload.pdf>

Isfan, T.M. 2007. Variables - Harmonization and Normalization. Trabajo presentado en la Sesión número 56 del International Statistical Institute (ISI). Lisboa, 22-29 de Agosto del 2007. Referencia: CPM157.

Köhl, M., B. Traub y R. Paivinen. 2000. Harmonisation and Standardisation in Multi-National Environmental Statistics - Mission Impossible? *Environmental Monitoring and Assessment* 63: 361–380.

Lehohla, P. 2007. Seminar on "Evolution of National Statistical Systems". 38th Session of the United Nations Statistics Commission. United Nations.

Linacre, S. y B. McColl, 2007. International Standards for Social Statistics: Addressing the Challenges. Trabajo presentado en la Sesión número 56 del International Statistical Institute (ISI). Lisboa, 22-29 de Agosto del 2007. Referencia: IPM18: Updating international statistical standards: does the process work?

Olenski, J. 2003. SSDIS. Global Standard for harmonization of Social Statistics, with special reference to transition and globalization processes. Expert Group Meeting on Setting the Scope of Social Statistics. United Nations Statistics Division in collaboration with the Siena Group on Social Statistics. New York. SSDIS ESA/STAT/AC.88/10 - 7 Abril del 2003.

ONU. 1994. Principios Fundamentales de las Estadísticas Oficiales. División de Estadística de las Naciones Unidas. <http://unstats.un.org/unsd/methods/statorg/FP-Spanish.htm>

Puustjärvi, E. and M. Simula. 2002. Development of Common Framework for Forest-Related Definitions. Comparative framework and Options for harmonization of definitions. En: FAO, 2002. Second Expert Meeting on Harmonizing Forest-Related Definitions for use by various stakeholders. Helsinki, 30 September 2002. <http://www.fao.org/docrep/005/y4171e/Y4171E06.htm#TopOfPage>

SDMX. 2008. SDMX Content-Oriented Guidelines. Annex 4. Metadata Common Vocabulary. http://www.sdmx.org/index.php?page_id=11

Trewin, D. 2007. "The evolution of National Statistical Systems: Trends and Implications" http://unstats.un.org/unsd/statcom/statcom_seminar/australian.pdf

UNECE, 1995. Guidelines for the Modelling of Statistical Data and Metadata. United Nations Statistical Commission / Economic Commission for Europe. Geneva, 1995 <http://www.unece.org/stats/publications/metadatamodeling.pdf>