

# INFRAESTRUCTURA COLOMBIANA DE DATOS ESPACIALES – ICDE



Traducción libre del documento "POLICY FRAMEWORK FOR INFORMATION SHARING", Biodiversity Conservation Information System (BCIS), 1999.

*Autores: Dora Inés Rey Martínez  
Martha Ivette Chaparro Domínguez.  
Iván Alberto Lizarazo Salcedo*

*Agosto de 1999*

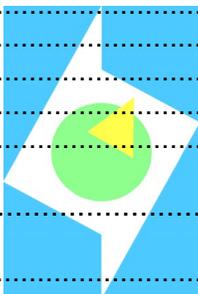


# TABLA DE CONTENIDO

POLITICAS DE INFORMACIÓN GEOGRAFICA.....	4
1 INTRODUCCIÓN.....	4
1.1 IMPORTANCIA DE UNA POLÍTICA DE INFORMACIÓN.....	4
1.2 IMPLICACIONES NACIONALES E INTERNACIONALES Y ACTIVIDADES RELACIONADAS.....	5
1.3 POTENCIAL COMERCIAL.....	5
1.4 RELACIÓN ENTRE DATOS E INFORMACIÓN.....	6
2 MIEMBROS.....	6
3 PRINCIPIOS Y REGLAS DE JUEGO PARA COMPARTIR INFORMACIÓN.....	6
3.1 META.....	6
3.2 ESTRATEGIAS.....	7
3.3 PROCESO DE GESTIÓN DE LOS DATOS.....	8
4 ESTRUCTURA PROPUESTA.....	8
4.1 INTRODUCCIÓN.....	8
4.2 COORDINACIÓN.....	9
4.2.1 Comité de dirección de la Infraestructura.....	9
4.2.2 Grupo de gestión de los datos.....	9
4.3 INFRAESTRUCTURA DE GESTIÓN DE LOS DATOS.....	9
4.3.1 Los productores de datos.....	10
4.3.2 Los propietarios de los datos.....	10
4.3.3 Custodios de los datos.....	10
4.3.4 Agencias de información.....	12
4.3.5 Infraestructura física.....	12
4.3.5.1 Redes electrónicas.....	12
4.3.5.2 Tecnología.....	12
4.4 CONJUNTOS DE DATOS BÁSICOS.....	13
4.5 HERRAMIENTAS DE ANÁLISIS Y DIFUSIÓN.....	14
4.6 ESTÁNDARES, LINEAMIENTOS Y PROCEDIMIENTOS.....	14
4.6.1 Aseguramiento de la calidad.....	15
4.6.2 Estándares.....	15
4.6.3 Documentación de los datos.....	15



4.6.4	Intercambio de datos .....	15
5	RETOS DE IMPLANTACIÓN.....	17
5.1	LIBERANDO EL FLUJO DE DATOS .....	17
5.2	RECURSOS PARA GESTIÓN DE DATOS.....	17
5.3	SOLUCIÓN SOBRE PROPIEDAD Y CUSTODIA.....	18
6	PROCEDIMIENTOS GENERALES PARA LA GESTIÓN DE DATOS .....	18
6.1	GESTIÓN DE DATOS "DE EXTREMO A EXTREMO".....	18
6.2	METAS ESPECÍFICAS DE LA GESTIÓN INTEGRAL DE DATOS.....	19
6.3	EL PROCESO DE GESTIÓN DE DATOS "DE EXTREMO A EXTREMO".....	19
6.3.1	Documentación de la gestión de los datos.....	20
6.3.2	Lineamientos .....	20
6.3.2.1	Fase preparatoria .....	21
6.3.2.1.1	Plan de gestión de los datos .....	21
6.3.2.2	Recolección de datos.....	22
6.3.2.3	Procesamiento de datos .....	22
6.3.2.4	Remisión de datos .....	22
6.3.2.5	Gestión de datos.....	23
6.3.2.6	Intercambio de datos .....	23
7	CUSTODIA .....	23
7.1	PRINCIPIOS DE CUSTODIA.....	25
7.2	FUNCIONES DE UN CUSTODIO.....	26
7.2.1	Responsabilidades y derechos .....	26
7.2.2	Custodios y Propietarios .....	27
8	PROPIEDAD DE LOS DATOS .....	27
9	ACCESO A LOS DATOS .....	27
10	CALIDAD .....	27
11	RESPONSABILIDAD .....	27
12	COMERCIALIZACION .....	27
13	PRIVACIDAD .....	27



# POLITICAS DE INFORMACIÓN GEOGRAFICA

## 1 INTRODUCCIÓN

### 1.1 *Importancia de una política de información*

La efectiva utilización y protección del ambiente depende de un claro entendimiento del mismo, asegurando que el desarrollo sostenible ocurra y que los conflictos de uso sean resueltos desde una posición de conocimiento. El desarrollo de políticas económicas, sociales y ambientales depende de la existencia de una "infraestructura de información" o de una "base de conocimiento".

Parte de esta "base de conocimiento" comprende los datos e información existentes y futuros. Para que ellos puedan usarse deben estar organizados y ser de fácil acceso, consistentes y bien documentados de manera que una multitud de usuarios puedan entender fácilmente los datos que ellos están accedendo. El concepto detrás de una política de datos es hacer que ellos sean accesibles fácilmente y que se puedan utilizar por todos los niveles de usuarios potenciales en diferentes aplicaciones.

La implantación de una política de datos ocurrirá únicamente si ella es parte integral de los planes y acciones de una organización y tiene apoyo institucional adecuado. Para que la gestión de los datos tenga éxito ella debe ser de alto perfil y debe estar soportada en un mandato institucional. El éxito también depende de que existan los recursos necesarios para implantar la infraestructura indicada (construida a partir de los desarrollos existentes) y requiere de la reglamentación de ... El éxito también requiere un cambio mental en las actitudes frente a la gestión de los datos.

Esto puede ser ayudado mediante el suministro de capacidad interna para desarrollar y soportar la administración de los datos. (Cuadro 1)

#### **Construcción de la Capacidad para la Gestión de información**

Una política de datos ayuda a desarrollar capacidad en la siguiente forma:

- Identifica fortalezas y debilidades existentes en la gestión de información
- Identifica, documenta y facilita el intercambio de las mejores prácticas para la gestión de la información
- Maximiza la coordinación y minimiza la duplicación y el traslapo en la producción de datos
- Identifica oportunidades para la sinergia y la producción de valor agregado
- Proporciona un mecanismo para resolver conflictos



Los miembros de una Infraestructura de Datos Espaciales, sólo darán prioridad a la gestión de los datos si ellos observan beneficios directos tangibles que superen los costos. Estos beneficios de invertir en datos con mejor gestión incluye la obtención más rápida y económica mediante:

- Utilizar datos estandarizados de fuentes externas.
- Integrar datos internos y externos
- Generar productos de información más atractivos y fáciles de comprender
- Proporcionar datos a otros ya sea mediante convenios de cooperación o para algún propósito de bienestar público.

Pero sobre todo, la mejora en la gestión de la información aumentará el perfil de la entidad como una organización competitiva con mejores oportunidades de negocio. Únicamente mediante el establecimiento de una infraestructura de información y de datos se utilizarán al máximo los recursos de las entidades participantes que son muy limitados en comparación con los retos que ellas afrontan.

## **1.2 Implicaciones nacionales e internacionales y actividades relacionadas**

El desarrollo de una política de datos no puede ser realizado de manera aislada a otras actividades relacionadas. La política debe incorporarse y desarrollarse a partir de las estructuras existentes o en proceso de ser implementadas en otros proyectos.

Otras iniciativas proporcionan una fuente de datos valiosos y la interacción con esos programas internacionales también proveen acceso a tecnología y conceptos que pueden ayudar a administrar a los datos.

Los protocolos, estándares y guías internacionales para la gestión de los datos tienen un papel que jugar en una política de información.

## **1.3 Potencial comercial**

Aunque una política de datos se enfoca hacia el mejoramiento del acceso y utilización de los datos para soportar, entre otras cosas, el desarrollo sostenible hay otros beneficios menos obvios que deben verse como oportunidades comerciales:

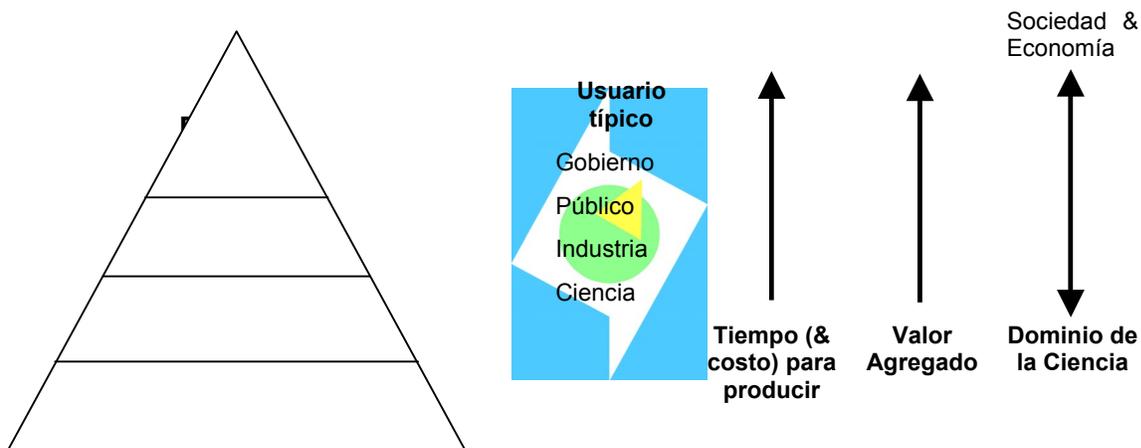
- Venta de productos y servicios
- Programas de entrenamiento y educación (nacionales e internacionales)
- Desarrollo de capacidades a través de la aplicación de propiedad intelectual y destrezas.



## 1.4 Relación entre datos e información

Hay una compleja relación entre datos e información y la diferencia entre ellas es a menudo difusa. La generación de información empieza con datos primarios que son obtenidos de diversa manera incluyendo diversas técnicas de observación con instrumentos. Estos datos son transformados en datos administrados mediante procesos de aseguramiento de la calidad y la incorporación en conjuntos de datos estructurados. Los datos son entonces integrados, analizados, interpretados y transmitidos a los usuarios como información. Mientras los datos son almacenados y administrados de forma estándar, útiles por muchos años sino siglos, la información está orientada al usuario y tiene comparativamente una vida útil más corta.

El siguiente diagrama demuestra algunos atributos de la relación entre información y datos y su relevancia para una variedad de propósitos



## 2 MIEMBROS

## 3 PRINCIPIOS Y REGLAS DE JUEGO PARA COMPARTIR INFORMACIÓN

### 3.1 Meta

La meta de una política de datos es asegurar que el proceso de obtención, gestión, distribución de datos e información no sólo que se conduzca eficientemente, de manera coordinada y de acuerdo a los objetivos de los miembros, sino que también mejore la capacidad de atender temas prioritarios a diferentes niveles. La política es alcanzada a través del establecimiento de infraestructuras y procedimientos basados en estándares y lineamientos desarrollados a través de un proceso de consulta y que toma en cuenta los objetivos más amplios posibles.

La infraestructura comprende las entidades que capturan y administran datos y sus redes, incluyendo los custodios de los datos y aquellas agencias que procesan datos y elaboran servicios y productos de valor agregado. Los estándares y procedimientos cubren la obtención de los datos, su procesamiento incluyendo el aseguramiento de la calidad, la gestión y difusión, y lo más importante, la documentación asociada con cada una de éstas actividades. Ellos incluyen acuerdos sobre acceso a los datos y licenciamiento que especifican derechos, responsabilidades y obligaciones de los diferentes actores en el proceso de transformar datos crudos en información de valor agregado.

La política está dirigida a proporcionar los datos y la información necesaria para obtener una perspectiva balanceada de los requerimientos a menudo conflictivos de los proveedores y usuarios de los datos.

La meta en política es obtener "una capacidad integrada y tecnológicamente avanzada de adquisición, procesamiento, gestión y difusión de datos e información que apoye las necesidades de la organización, de los usuarios y de los tomadores de decisiones para desarrollar e implementar políticas de desarrollo sostenible a todos los niveles".

El objetivo político se alcanza a través de la formulación de la infraestructura, la definición de lineamientos, estándares y procedimientos para soportar la gestión y procesamiento de los datos y la generación y distribución de datos y productos de información. Este proceso debe ser implementado por un grupo distribuido de custodios de los datos y de otras entidades relacionadas. El proceso de gestión de los datos debe integrarse con otras redes e iniciativas similares.

### **3.2 Estrategias**

Una política de datos debe apoyarse en las siguientes estrategias para cumplir su objetivo:

1. Crear una red de datos distribuída, fácilmente accesible, par administrar y diseminar los datos que se recolecten como parte de los proyectos y actividades que apoyen los objetivos de la red.
2. Desarrollar conjuntos de datos fundamentales que sirvan de base para actividades de conservación de la biodiversidad.
3. Proporcionar las mejores prácticas en términos de mecanismos y procedimientos de aseguramiento de la calidad para crear conjuntos de datos bien documentados y validados que satisfagan las necesidades prioritarias de información.
4. Almacenar donde sea más apropiado todos los datos recolectados para asegurar la disponibilidad de ellos para uso múltiple y para salvaguardarlos para las generaciones futuras.



5. Mejorar la relación costo beneficio de las actividades de producción e datos a través de la coordinación de actividades.
6. Proporcionar oportuna y operacionalmente, datos y productos de información relevantes para los proyectos de desarrollo económico sostenible.
7. Hacer que los datos de los miembros estén disponibles a un rango amplio de usuarios, con la mayor calidad y oportunidad posibles.
8. Proporcionar servicios y productos de información de interés público.

### **3.3 Proceso de gestión de los datos**

La gestión de los datos es un proceso continuo que empieza con el diseño del programa de adquisición de los datos y finaliza con el acceso del usuario a un conjunto de datos comprensible de alta calidad y completamente documentado. La gestión de los datos también tiene una función de valor agregado que ayuda a maximizar el retorno de la inversión mediante la posibilidad de usar varias veces los conjuntos de datos a través de técnicas efectivas de integración y de administración.

La gestión de los datos es una actividad especializada que se desarrolla en un marco de programas, lineamientos, estándares y políticas intra e interinstitucionales.



## **4 ESTRUCTURA PROPUESTA**

### **4.1 Introducción**

Los principios de gestión de los datos cubren los siguientes elementos principales:

- Un mecanismo de coordinación
- La Infraestructura física de los miembros y organizaciones asociados que participan en la red.
- Los conjuntos de datos fundamentales o básicos
- Las herramientas de análisis y difusión.
- Los estándares, lineamientos y procedimientos dentro de los cuales operan las entidades involucradas.

Estos principios se basan en las siguientes premisas:

- La infraestructura de información es un activo.
- Los datos básicos deben ser fácilmente accesibles.



- Existen beneficios, tanto para los miembros como para la sociedad, que resultan del uso múltiple de los datos por un amplio número de usuarios y de asegurar su disponibilidad para las generaciones futuras.

## **4.2 Coordinación**

La adquisición, gestión y difusión de los datos requiere coordinación para asegurar una operación efectiva y eficiente. Adicionalmente, el desarrollo y mejoramiento continuo de la gestión de los datos y de los procesos relacionados requiere un mecanismo de coordinación que tenga en cuenta e incorpore las perspectivas y necesidades de todos los participantes. El mecanismo de coordinación debe establecerse con base en las estructuras existentes, tener un mandato apropiado y unas funciones específicas.

### *4.2.1 Comité de dirección de la Infraestructura*

Este Comité, constituido de acuerdo con los principios de la infraestructura, será responsable de los lineamientos políticos generales.

### *4.2.2 Grupo de gestión de los datos*

Este grupo que reporta al Comité de dirección de la infraestructura, proporciona un mecanismo informal para mejorar la gestión de los datos y que sirve de conexión entre la dirección y la infraestructura de información. Este grupo está abierto a personas de las entidades involucradas que estén interesadas en el desarrollo de la infraestructura de información. El grupo de gestión de los datos realizará la coordinación y asesoría de los aspectos técnicos. Un número de grupos técnicos, asesores informales, puede usarse para aumentar la capacidad del grupo de gestión de los datos en áreas específicas como: calidad, metadatos, formatos de intercambio y tecnologías de comunicaciones.

El grupo de gestión de los datos trabajará en la elaboración de lineamientos y procedimientos en las prioridades establecidas, los cuales deben ser aprobados por el comité directivo.

## **4.3 Infraestructura de gestión de los datos**

La infraestructura comprende los siguientes componentes:

- Las entidades que producen los datos
- Los custodios de los datos
- Otras entidades relacionadas con el manejo de la información
- La infraestructura física



#### 4.3.1 *Los productores de datos*

Son aquellas entidades productoras e individuos tomadores de los datos para un gran número de propósitos. El aseguramiento del uso múltiple de los datos y la simplificación a su acceso por otros usuarios comienza desde el proceso de adquisición. Es esencial que las actividades que realizan los productores miembros del Comité para la recopilación de los datos estén en concordancia con los principios generales y procedimientos acordados. Mediante la promoción de la "mejor práctica" se espera asegurar un alto nivel de estandarización, y específicamente de documentación, que reduzca los esfuerzos futuros y los costos, minimice la redundancia y maximice las oportunidades.

Las entidades participantes tienen la responsabilidad de asegurar que los datos obtenidos durante sus proyectos de campo estén administrados y documentados de acuerdo a los estándares establecidos. Igualmente, en concordancia con estos principios, los datos deben ser remitidos a los custodios apropiados dentro de los plazos acordados.

#### 4.3.2 *Los propietarios de los datos*

Los temas relativos a la propiedad de los datos son complejos, especialmente cuando los propietarios no son los custodios. Los propietarios tienen el derecho absoluto de decidir quién tiene acceso a sus datos y bajo qué circunstancias, excepto cuando se haya decidido con su consentimiento otra cosa. Otros tienen la idea que todos los datos ambientales son un componente fundamental de los "acuerdos globales" y que deben estar disponibles sin costo alguno a todas las personas. Una política debe reconocer esta tensión y animar la provisión de datos dentro del dominio público, reconociendo que los derechos de propiedad intelectual y otros temas reclamados legítimamente por los propietarios deben ser respetadas.

#### 4.3.3 *Custodios de los datos*

Es esencial que la función de gestión de los datos se desarrolle alrededor de las capacidades y experiencias existentes. La cantidad y diversidad de tipos de datos que se adquieren en el dominio de interés de la infraestructura son muy grandes para que cualquier miembro los maneje de manera aislada. La única solución viable es mediante un sistema distribuido de "custodios". Los custodios se definen como la entidad o persona responsable por el desarrollo y/o la gestión de un conjunto de datos y que tiene el derecho de determinar las condiciones en las cuales esos datos pueden ser entregados al público y usados.

Los custodios son entidades líderes en determinadas disciplinas, que se hacen cargo de la gestión de los datos que estén dentro de su competencia. Mediante la implementación del concepto de "custodia", un número de entidades serán autorizadas para administrar conjuntos de datos específicos en representación de



la infraestructura. La "autorización" a las entidades requiere un proceso de revisión y consulta de alto nivel a través del comité de dirección.

Los custodios pueden ser de cualquier clase, inclusive individuos. No obstante, es indispensable que su papel sea claramente reconocido y aprobado y que ellos sean capaces de garantizar que los recursos necesarios para realizar su tarea estarán disponibles. La gestión de los datos es una tarea de largo plazo que requiere que los recursos necesarios no fluctuen significativamente debido a cambios políticos u organizacionales.

Los custodios requieren aprobación del comité de dirección y deben operar bajo los lineamientos y estándares desarrollados por el grupo de gestión de datos el cual debe desarrollar un papel de coordinación, asistencia técnica y monitoreo de las actividades de custodia de una manera consistente con los principios de gestión de los datos.

El grupo de gestión de los datos debe apoyar a los custodios mediante el monitoreo de los recursos destinados a su gestión, asegurando que los procedimientos y lineamientos sean seguidos, vigilando y reportando la efectividad de los custodios.

La infraestructura básica para desarrollar un sistema de custodia distribuido ya existe y se reconoce un número de entidades como expertas en determinados campos. Estas entidades ya han creado infraestructuras importantes para el procesamiento y gestión de sus propios datos. Esta capacidad debe expandirse para poder soportar las necesidades de la infraestructura.

Los custodios deben cumplir con las siguientes obligaciones:

- Desarrollar un conjunto de datos (con socios cuando sea apropiado)
- Mantener un conjunto de datos (haciendo las actualizaciones, asegurando su calidad, reestructurándolo si fuere necesario)
- Mantener la seguridad del conjunto de datos (protegiendo la integridad física de pérdida o daño, incluyendo el acceso no autorizado)
- Proporcionar la documentación adecuada del conjunto de datos (Metadatos)
- Proporcionar un acceso adecuado al conjunto de datos (Legitimación de usuarios)
- Suministrar asesoría para el uso apropiado de un conjunto de datos (usos sugeridos, impropios, inadecuados)
- Coordinar de desarrollos posteriores de un conjunto de datos (con socios apropiados y con la asistencia de usuarios claves)



#### 4.3.4 *Agencias de información*

Muchos usuarios no requieren acceder a conjuntos de datos individuales, es decir a datos sin procesar. Los tomadores de decisiones dentro de las entidades gubernamentales, el público en general y en muchos casos las instituciones educativas, requieren datos procesados y productos de información. Estos productos se derivan de la integración y análisis de una variedad de diferentes datos, a menudo que cubren varias disciplinas. Este es el componente "fábrica de información" que forma parte del ciclo de información.

Un gran número de entidades realiza esta tarea de valor agregado que se alcanza mediante el procesamiento y unión de los datos usando técnicas de fusión de datos para producir información con múltiples propósitos, incluyendo el desarrollo de políticas, la definición de prioridades, el soporte de decisiones y la concientización. Otras entidades desarrollan procesos de fusión de datos muy importantes sobre un conjunto muy amplio de datos multidisciplinarios con el objeto de soportar los requerimientos de información de los políticos. Dentro de la infraestructura de información hay que considerar el rango completo de proveedores de información existentes.

#### 4.3.5 *Infraestructura física*



##### 4.3.5.1 *Redes electrónicas*

El desarrollo de una infraestructura física es necesario para enlazar los diferentes instrumentos distribuidos, tales como los custodios de los datos y las entidades de información. El mecanismo más flexible es usar INTERNET y las potencialidades de la World Wide Web teniendo en cuenta que la mayoría de las entidades líderes cuentan con dichas tecnologías, aunque se necesitan recursos para conectarlas de manera que se forme un sistema distribuido "sin costuras", también la mayoría de los clientes están conectados a INTERNET. Sin embargo, debe tenerse el cuidado de asegurar que la tecnología sea usada en detrimento de quienes no tienen acceso a ella. Por lo tanto debe encontrarse otros mecanismos para distribuir los datos y la información tales como las publicaciones en papel y los CD-ROMs.

Es esencial que se destinen recursos para crear este sistema distribuido y que se desarrollen los estándares y protocolos que aseguren facilidad de uso e interoperabilidad.

##### 4.3.5.2 *Tecnología*

Para implementar las estrategias indicadas anteriormente, lo más conveniente es trabajar sobre la infraestructura física existente y usar la experiencia técnica existente en la gestión y difusión de datos e información. Los avances importantes en informática, en particular los conceptos de sistemas distribuidos, hacen posible

crear una red electrónica efectiva que interconecte a los proveedores de información y a los administradores de los datos.

Es responsabilidad del grupo de gestión de los datos que la capacidad de INTERNET será utilizada efectivamente y que otros medios se utilizarán para difundir la información. Este grupo debe también definir los lineamientos para facilitar la interconexión electrónica entre entidades y bases de datos.

Se requieren recursos para establecer el sistema distribuido de datos e información. Sin embargo, como ya existe alguna capacidad, los recursos deben concentrarse en crear las conexiones, llenar los vacíos e introducir prácticas estandarizadas más que intentar el desarrollo de una infraestructura de gran escala. Ya existe considerable experiencia dentro de los miembros que se puede enfocar hacia una efectiva transferencia y asesoría.

El desarrollo de esta infraestructura de gestión de datos también proporcionará la experiencia y los incentivos necesarios para el desarrollo de redes temática y regionales que faciliten la construcción de una "Red de redes" que sea interoperable.

#### **4.4 Conjuntos de datos básicos**

Aunque la gestión de todos los datos recolectados en las áreas de interés de la infraestructura es importante, es necesario concentrarse en la obtención de conjuntos de datos específicos, que se consideran fundamentales o que son el "núcleo", que sirven de cimiento al desarrollo de conjuntos de datos especializados, productos y servicios de información. La composición de un conjunto de datos básico debe ser determinada mediante un proceso de consulta. Los ejemplos de los tipos de datos que probablemente deben incluirse son los siguientes:

- Límites nacionales y político administrativos
- Clima
- Topografía básica, incluyendo cuencas
- Hidrología superficial (Ríos, Lagos, etc.)
- Subsuelo (Geología, suelos, etc.)
- Infraestructura humana (Asentamientos, transporte y corredores de comunicación)
- Cobertura y uso de la tierra
- Regiones de biodiversidad (Biomás, ecosistemas, tipos de vegetación, etc.)



Es importante que estos y otros conjuntos de datos básicos sean producidos de las mejores fuentes disponibles y que se facilite el acceso a todos los usuarios potenciales utilizando todos los mecanismos apropiados, incluyendo INTERNET. Sin embargo, es importante que se tenga el cuidado de asegurar que la resolución temporal y espacial de éstos datos se ajuste a los requerimientos de los usuarios, sin comprometer la integridad y los estándares de los datos.

#### **4.5 Herramientas de análisis y difusión**

Como en el caso del conjunto de datos básicos, es necesario concentrarse en el desarrollo de herramientas específicas, que deben incluir para el análisis básico de los datos, lo mismo que las requeridas para el desarrollo de conjuntos de datos especializados y productos y servicios de información. Las herramientas clave deben ser determinadas mediante un proceso de consulta, pero probablemente incluirán las siguientes:

- Procesadores de palabra y hojas de cálculo
- Sistemas manejadores de bases de datos
- Sistemas de manejo de datos espaciales (Incluyendo SIG's)
- Sistemas de manejo de gráficos, video y audio (multimedia)
- Sistemas estadísticos, de modelamiento, Sistemas expertos y Sistemas de soporte a las decisiones
- Sistemas de empaquetamiento y difusión de información (para www, cd-rom, etc.)

#### **4.6 Estándares, lineamientos y procedimientos**

La comparación y el análisis eficiente de datos "crudos" y el desarrollo de productos integrados de información requiere la atención a los siguientes aspectos:

- Aseguramiento de la calidad
- Estándares
- Documentación de los datos
- Intercambio de datos



#### 4.6.1 Aseguramiento de la calidad

La calidad de los datos es una medida de su " aptitud de uso" y por lo tanto es relativa al propósito de uso. Los datos pueden ser de alta calidad para un propósito y de baja calidad para otro. Aunque los datos deberían ser de la más alta calidad alcanzable, lo más importante es documentar su calidad de manera que un usuario tenga suficiente información para decidir si los datos se ajustan o no al uso propuesto, más que intentar alcanzar una calidad teórica ideal.

#### 4.6.2 Estándares

El objetivo de los estándares es "reducir los costos de operación de usar los datos". Nuevamente, el objetivo es facilitar la integración de los datos mediante su armonización más que imponer restricciones burocráticas y costos innecesarios a los custodios de los datos.

Dentro de la infraestructura, hay adoptar o adaptar estándares existentes. El grupo de gestión de los datos debe examinar el desarrollo de otros estándares apropiados, en colaboración con otros. Únicamente como un último recurso se desarrollará estándares dentro de la infraestructura y ellos probablemente tendrán solamente un uso interno.

#### 4.6.3 Documentación de los datos

Los metadatos se están reconociendo cada vez más como un componente importante de una gestión responsable de los datos ya que contribuyen a los siguientes objetivos:

- Comercialización de información y de datos mantenidos por un custodio o a través de la red
- Soporte para la gestión interna de los conjuntos de datos
- Provisión de descripciones detalladas de elementos de datos para soportar aplicaciones de usuarios

#### 4.6.4 Intercambio de datos

Las políticas expuestas recomiendan acceso rápido e irrestricto a los conjuntos de datos básicos por parte de todos los usuarios legítimos, en el interés más amplio del público. El intercambio de datos debe ser administrado mediante acuerdos de intercambio de datos. El propósito de estos acuerdos es facilitar el flujo de datos antes que inhibirlo. Los temas relacionados con propiedad intelectual y recuperación de costos es necesario considerarlos también.

Ya sea que los datos sean cobrados o no, la política de precios que debe usarse es un tema complejo. En general, el usuario secundario de los datos no debería



pagar doble vez los costos de recolección de los datos (doble contabilidad). Los costos de mantenimiento representan un costo adicional que debería ser recuperado más que de los usuarios del propio sistema de ventas de la entidad.

Uno de los objetivos de las políticas de gestión de datos es que se obtenga el máximo beneficio de los gastos destinados a recolección y análisis de los datos, mediante el uso múltiple de dichos datos. La habilidad para disminuir la inversión inicial a través de una gestión de datos efectiva es uno de los focos del marco de referencia de gestión de información de la infraestructura. Aunque esto puede ocasionar un costo adicional los beneficios son considerables, ya que se evita la duplicación y se reduce el costo de oportunidad.

Uno de los principales obstáculos para el intercambio de datos son las dificultades entre los custodios y los usuarios - entender los temas y negociar acuerdos equitativos de acceso a los datos. Otro importante obstáculo es la diversidad de los formatos usados encargadas de producir y administrar datos. Los principales retos son:

- La medición de entidades ambientales en el mismo sitio o de las mismas entidades pero de diferente manera y para diferentes propósitos (a menudo muy pobremente documentados).
- Diferentes rangos o valores de dominio para los atributos de esas entidades.
- Estructuras de bases de datos incompatibles.
- Diferentes formatos de intercambio de datos.

La mayor responsabilidad del grupo de gestión de datos es soportar el desarrollo de estándares para el intercambio de datos de tal manera que los recursos valiosos no sean dilapidados en convertir datos de un sistema a otro. No obstante, la comunidad internacional está intentando resolver este problema y la infraestructura debe continuar el monitoreo de las mismas y la cooperación con las actividades relevantes.

El uso creciente de INTERNET y el concepto de bases de datos distribuídas, disminuye algunos aspectos del problema del formato, pero no los soluciona completamente. El grupo de gestión de los datos requiere concentrarse en desarrollar una solución para las diversas dimensiones del tema de intercambio de datos.



## 5 RETOS DE IMPLANTACIÓN

### 5.1 Liberando el flujo de datos

Una de las barreras en la implementación de una política de datos es la resistencia de los recolectores de datos a someter sus datos a un custodio. Esto ocurre por varias razones, pero las principales son:

- La creencia que los datos pertenecen al recolector y que él es el único que los entiende.
- La existencia de pocos incentivos para hacer los datos disponibles a otros usuarios.

Un mecanismo para solucionar esto proviene a través de las fuentes de financiación. Una política de datos debe sugerir que los lineamientos de los programas de financiación incluyan una sección sobre las actividades de gestión de datos en donde se describan los procesos que se seguirán y se indique a cuáles custodios se entregarán finalmente los datos. Este plan de gestión de datos también debe identificar los recursos necesarios para crear un conjunto de datos de calidad adecuada (de acuerdo con los estándares apropiados incluyendo los metadatos) para ser almacenados en custodia. La aprobación de las propuestas debe considerar el cumplimiento de este procedimiento. Asegurando la provisión de elementos de gestión de datos de largo plazo, en este caso la custodia de los datos, dentro del presupuesto del proyecto, los administradores son capaces de obtener recursos para desarrollar esa tarea y también aceptar las responsabilidades para suministrar los datos.

En los casos en que los datos tengan usos operacionales posteriores, el plan de gestión de los datos debe identificar el mecanismo por medio del cual los datos estarán disponibles para las entidades operativas apropiadas en tiempo casi real. El conjunto de datos final debe ser entregado a la agencia que hace la custodia dentro de un tiempo preestablecido, por ejm 2 años. Esto cubre el tema de "derecho a publicar", proporcionando tiempo suficiente para análisis y publicación al mismo tiempo que se satisfacen necesidades más amplias.

Los custodios designados deben desarrollar el monitoreo de este tipo de datos para asegurar que todos los datos recolectados sean almacenados de manera apropiada. El grupo de gestión de los datos debe coordinar esta tarea y apoyar a los custodios en el proceso de monitoreo.

### 5.2 Recursos para gestión de datos

La obtención de recursos para la gestión de los datos en sus diferentes niveles presenta retos a todas las organizaciones. La consecución del personal y los recursos financieros necesarios para implementar una política adecuada de datos y participar efectivamente en una red colaborativa puede ser muy difícil. Aún



cuando muchos de los directivos aceptan en principio, que la mayor eficiencia y efectividad se obtiene a través de una mejor organización y colaboración, esos beneficios son a mediano y largo plazo, mientras que los costos deben cancelarse por anticipado. La financiación a través de fuentes externas, tales como donaciones o cofinanciación, puede ser crucial para acelerar el proceso. Donde estos recursos no estén disponibles el avance será desalentadoramente lento.

### **5.3 Solución sobre propiedad y custodia**

Las organizaciones involucradas con la gestión de los datos a menudo tienen una trayectoria larga de trabajo con otros en una variedad de acuerdos bilaterales y multilaterales. Pocas veces estos acuerdos tienen una intención estratégica y de largo plazo. La mayoría de ellos se iniciaron y desarrollaron de manera fortuita, de acuerdo a los imperativos del tiempo. Muchos de esos acuerdos son frágiles y aceptados con muchas tensiones, incluso conflictos. Aunque pueden existir acuerdos formales por escrito, pocas veces consideran los temas más relevantes y a menudo el lenguaje usado confunde o inhibe la resolución de dichos temas.

Nos enfrentamos a una herencia espantosa de conjunto de datos que han sido aportados por una diversidad de entidades pero donde los temas de propiedad y custodia de los mismos nunca han sido considerados, mucho menos, resueltos. A menos que sean manejados de manera inteligente, esos conjuntos de datos pueden ser una fuente de fricción entre las partes.

Allí hay dos niveles en los que estos problemas pueden ser resueltos. En el nivel operacional el entendimiento mutuo de los principios claves de propiedad y custodia de los datos puede ayudar. También, estos aspectos pueden ser motivo de acuerdos en el nivel de gestión de los datos y posteriormente aprobados por el Comité de Dirección.

## **6 PROCEDIMIENTOS GENERALES PARA LA GESTIÓN DE DATOS**

La política de datos de la infraestructura considera la gestión de los datos como un proceso integral ("de un extremo a otro") y proporciona lineamientos y estándares que cubren un rango completo de actividades, desde la planeación, la recolección de los datos, su procesamiento, el aseguramiento de la calidad, la documentación, la administración y el intercambio, lo mismo que la generación y difusión de productos.

### **6.1 Gestión de datos "de extremo a extremo"**

Esta gestión es una filosofía para mejorar la adquisición, procesamiento, almacenamiento e intercambio de datos e información. Sus objetivos son:

- Acelerar la recolección de datos prioritarios.



- Reducir pérdidas y demoras en la producción de datos.
- Minimizar operaciones redundantes.
- Promover el intercambio de datos y software entre usuarios y custodios.
- Mantener metadatos mínimos sobre las fuentes, el procesamiento y el control de calidad de los datos

La gestión integral de los datos aplica a todos los datos, independientemente de cómo o cuándo se produjeron y asegura que los datos y la información fluyan rápida y fácilmente de los productores a los usuarios.

## **6.2 Metas específicas de la gestión integral de datos**

Las metas de esta gestión son las siguientes:

- Definir procesos que el productor pueda llevar a la práctica y que le permitan minimizar pérdidas de datos, mejorar la calidad de los datos y acelerar el flujo de datos del productor al usuario.
- Implantar métodos de aseguramiento de la calidad de los datos y de los metadatos y aumentar la disponibilidad a los usuarios.
- Máxima coordinación e intercambio de datos, software y responsabilidades entre centros de información lo mismo que conjunto de datos agregados desarrollados fácilmente y con una relación costo beneficio alta a partir de datos primarios.
- Productos de datos de alta calidad entregados a los usuarios a tiempo y dentro del presupuesto definido.

## **6.3 El proceso de gestión de datos "de extremo a extremo"**

Los componentes claves de la mejor gestión de los datos comprenden:

- Un sistema efectivo de monitoreo de los datos que proporcione información sobre dónde están los datos y en que estado de desarrollo se encuentra un conjunto de datos o un producto.
- La documentación que cubra todos los procesos desde la planeación hasta el procesamiento y la gestión.

Con una adecuada documentación debes ser posible para otros usuarios localizar los datos y evaluar su utilidad para sus propósitos específicos.

Hay que desarrollar un método para asegurar el seguimiento de los datos, desde la recolección hasta el almacenamiento. Hay que definir un mecanismo de



monitoreo del flujo de los datos. La creación de tales metodologías es parte del componente de infraestructura de la política de datos.

### 6.3.1 Documentación de la gestión de los datos

Los datos son recolectados para un amplio rango de propósitos. Cada proyecto difiere en escala, justificación científica y objetivos de corto y largo plazo. Los proyectos varían desde grandes programas multidisciplinarios que involucran una gran cantidad de entidades y que cubren procesos regionales hasta científicos o técnicos que individualmente recolectan datos en áreas pequeñas.

Debido a la diversidad de proyectos no es muy práctico proporcionar especificaciones detalladas para elaborar planes de gestión de datos que se ajusten a todas las actividades. En lugar de ello es mas apropiado proporcionar lineamientos que ayuden a los planeadores de proyectos a desarrollar su plan específico de gestión de datos ajustado a sus necesidades particulares pero apto para utilizar de manera apropiada con otros propósitos.

Los objetivos básicos de la gestión de los datos son:

- Maximizar la consistencia de los datos recolectados en el proyecto asegurando que ellos están documentados de acuerdo a un estándar y que existe suficiente información para documentar la calidad y las limitaciones de los datos y por tanto su utilidad para una tarea particular.
- Facilitar el intercambio de datos e información dentro del proyecto para maximizar tanto el uso de los datos como la probabilidad de detectar y corregir problemas.
- Asegurar la vejez de los datos y el potencial de reutilización múltiple de los datos.

### 6.3.2 Lineamientos

Se requiere desarrollar componentes específicos para la gestión de los datos dentro de cada una de las fases de los procesos de los datos:

- Fase preparatoria
- Recolección de datos
- Procesamiento de datos
- Remisión de datos
- Gestión de datos



- Intercambio de datos

### 6.3.2.1 Fase preparatoria

La gestión de los datos empieza antes de iniciar el trabajo de recolección de los datos. Hay que desarrollar un plan de gestión de los datos como parte del plan general del proyecto. Este plan debe considerar aspectos como la determinación de la disponibilidad de datos relevantes existentes o de la existencia de actividades de recolección de datos con el objeto de reducir la duplicación de esfuerzos y gastos innecesarios. En particular el plan de gestión de los datos debe documentar cómo el proyecto contribuirá en el desarrollo de la infraestructura de información.

Después de que el proyecto haya sido terminado, hay que describir los procesos tal y como ellos sucedieron. Otro propósito del plan de gestión de los datos es proporcionar unos metadatos adecuados que documenten de manera completa la historia de los datos. Esto es esencial si se quiere que usuarios potenciales puedan obtener completo beneficio de los conjuntos de datos y las entidades deben asegurar que su inversión en este aspecto sea protegida.

#### 6.3.2.1.1 Plan de gestión de los datos

Este plan debe describir de manera detallada cuáles son los productos que tendrá el proyecto. En un proyecto grande y multidisciplinario que involucre a varias entidades es conveniente conformar un grupo asesor que asegure que todas las actividades de gestión de los datos sean consideradas dentro del proyecto.

El plan de gestión de los datos debe considerar los siguientes elementos:

- Elementos seleccionados para ser adquiridos y razones para ello
- Flujo de datos e información
- Documentación de los datos
- Calidad de los datos
- Elementos técnicos (bases de datos) y otras aplicaciones desarrolladas como parte del proyecto
- Difusión de información
- Almacenamiento final
- Indicadores de desempeño

Una vez que los datos han sido recolectados y procesados exitosamente, el plan proporciona la base para la descripción de los metadatos. Como es probable que los conjuntos de datos finales difieran en algo de lo que se proyectó inicialmente, se requiere modificar los elementos para reflejar lo que realmente sucedió. Las soluciones existentes, los conjuntos de datos y otras facilidades potenciales también deben examinarse para asegurar un uso costo efectivo de los recursos financieros.

#### 6.3.2.2 *Recolección de datos*

Durante el proceso de recolección de datos hay que registrar los siguientes detalles:

- Estrategias y métodos de muestreo
- Nivel de precisión y calibración de los instrumentos
- Exactitud de posición y método de determinación
- Condiciones ambientales generales que pueden afectar el trabajo de recolección de datos
- Cualquier anomalía o evento que pueda afectar la exactitud de las observaciones
- Cualquier entrega de datos internos o de "baja resolución" a las agencias competentes en tiempo cercano al real para propósitos operativos (monitoreo, respuesta a emergencias)

#### 6.3.2.3 *Procesamiento de datos*

Las descripciones adicionales que se requieren en esta fase son las siguientes:

- Cualquier procedimiento de aseguramiento de la calidad que se aplique a los datos (incluyendo indicaciones de integridad de la recolección de los datos dentro de un dominio especificado)
- Cualquier algoritmo usado para convertir los datos (ejm. Correcciones instrumentales, conversión de unidades de campo a unidades estándares)
- Artículos científicos y reportes relacionados con los datos y otros metadatos relevantes

#### 6.3.2.4 *Remisión de datos*

Esta remisión puede ocurrir de dos maneras:

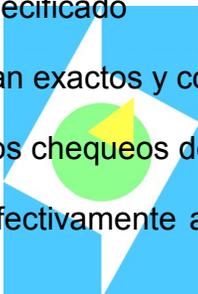


- Remisión del conjunto de datos completo con sus respectivos metadatos al sitio de almacenamiento apropiado, una vez que ha sido procesado de acuerdo al estándar definido. Al recibir los datos la entidad competente debe notificar a la entidad que financió los datos que ellos han sido entregados y que el proyecto está completo. La entidad que custodia puede determinar que durante un período razonable los datos no serán liberados o entregados al público (mientras la investigación se termina y se hacen las publicaciones).
- Retención de los datos en la entidad productora por un período de tiempo corto (no mayor de 2 años) y luego remitirlo a la entidad que hace la custodia para una difusión amplia.

### 6.3.2.5 Gestión de datos

Luego de recibir un conjunto de datos la entidad que hace la custodia debe ser requerida para asegurar:

- Que los datos estén en un estándar adecuado y que estén completos dentro del dominio especificado
- Que los metadatos sean exactos y comprensibles
- Que se hayan hecho los chequeos de aseguramiento de la calidad
- Que los datos sean efectivamente administrados y posean copias de respaldo
- Que los datos sean entregados o estén disponibles para el uso de otras entidades
- Que los metadatos sean entregados o estén disponibles en los directorios o clearinghouses apropiados.



### 6.3.2.6 Intercambio de datos

## 7 CUSTODIA

Los miembros de la infraestructura reconocen que, aparte de sus recursos humanos, los datos y la información están entre sus mayores activos. En gran medida, la oportunidad de aumentar el valor de sus recursos de información mediante la armonización y la integración de ellos con otros fue una de las motivaciones más importantes para unirse a la infraestructura. Ellos también son conscientes que los costos de mantener esos activos y de explotar las oportunidades que ellos generan son muy importantes. Dados esos costos, los miembros buscan obtener la mayor utilidad de sus bases de datos. De manera

creciente, ellos están usando datos para propósitos múltiples y están reconociendo la necesidad de desarrollar documentación y asegurar el cumplimiento de los estándares que están emergiendo. Ellos también están empezando a entender los beneficios de compartir datos y colaborar con otros a través de las redes de información.

Desde luego, las entidades pueden dedicarse a recolectar los datos de manera autónoma y para satisfacer sus propias necesidades, pero esto conduce a una costosa duplicación de datos que no pueden ser integrados fácilmente con otros para producir productos de valor agregado. El uso de los datos que han producido con un propósito en otras aplicaciones es muy rentable. **Sin embargo, para aprovechar al máximo este principio, la información debe ser consistente, cumplir algunos estándares, su existencia debe ser conocida y su acceso debe ser fácil. Esta es la esencia de la planeación de la infraestructura de información.** La infraestructura proporcionará el marco institucional y técnico para asegurar el cubrimiento, el contenido y la consistencia necesaria para cumplir los requerimientos de los usuarios. También asegurará que todos los esfuerzos de recolección y de mantenimiento sean desarrollados de acuerdo con los más amplios intereses de la conservación de la biodiversidad.

El desarrollo de información relevante para las políticas ambientales requiere a menudo acceso a una gran variedad de fuentes de datos, de numerosas organizaciones y múltiples disciplinas. Si se quiere la producción de información sea eficiente y costo-efectiva, esas fuentes deben ser fácilmente accesibles y deben existir las personas y las herramientas requeridas para convertir los datos en información para el soporte de decisiones a todos los niveles.

Las fuentes preferidas deben ser aquellas organizaciones que estén en la mejor posición para asegurar la calidad y la accesibilidad de sus conjuntos de datos y para asesorar sobre su utilización apropiada. Ellos se refieren aquí como los custodios.

Los custodios pueden ser definidos como "la persona o entidad designada como responsable para el desarrollo y/o la administración de un conjunto de datos y que tiene el derecho a determinar las condiciones en las cuales los datos pueden ser usados o entregados al usuario".

Si una custodia no se ha asignado o no se está manejando apropiadamente, entonces los usuarios pueden enfrentar una variedad de datos que están incompletos y son incompatibles, con documentación inadecuada y con procedimientos de acceso inconsistentes o definidos precariamente. Como resultado los usuarios no podrán integrar datos para obtener información relevante para sus decisiones dentro de las limitaciones de tiempo usuales.

La custodia es el medio por el cual la responsabilidad para la gestión de un conjunto de datos es asignada y aceptada por la entidad más apropiada. Su fines principales son:



- Minimizar la duplicidad de esfuerzos y el traslapeo entre datos
- Asegurar que los datos están disponibles para su uso (es decir, que pueden localizarse y accederse de manera fácil)
- Garantizar que los datos tienen la calidad asegurada (es decir, que sean ellos validos, mantenidos, documentados y seguros)

La custodia provee un mecanismo para asegurar que importantes conjuntos de datos están desarrollados, mantenidos y sean accesibles para usuarios autorizados. Esto asegura el aprovechamiento y disponibilidad de un conjunto de datos dentro de una jurisdicción específica (es decir un sector, disciplina o tema), para asegurar que los productos de información usados por los gobiernos y otros tomadores de decisiones sean exactos, completos, identificables, accesibles y auditables. En resumen, la custodia es el corazón de una infraestructura de información eficiente, responsive y capaz de servir a los intereses de organizaciones individuales o redes. " Esta es simplemente la única forma que los datos puedan ser administrados en una manera ordenada" (ANZLIC 1998)

### **7.1 Principios de custodia**

Los siguientes principios de custodia pueden servir de guía para los desarrollos operacionales:

- Los datos deben ser fácilmente accesibles a aquellos que tengan la necesidad de tomar decisiones con mejor soporte de información.
- Los datos deben ser manejados por la organización que esté en la mejor condición de hacerlo.
- Los datos deben ser manejados de conformidad con los derechos de propiedad intelectual y con los acuerdos hechos con los productores.
- Los datos deben ser manejados de una manera costo - efectiva por personal que entienda los datos: qué intentan representar?, cuáles son sus características?, cómo fueron producidos?, qué procedimientos de aseguramiento de la calidad se aplicaron?, cuáles son sus limitaciones?.
- La custodia no debería ser duplicada o fragmentada en diferentes lugares.

Los principios mencionados no deben entenderse como reglas rígidas, La clave es manejar los datos de una manera tal que ellos puedan convertirse fácilmente en una variedad de productos de información, para una variedad de usuarios y asegurar que los datos sean flexibles para responder a las demandas de la toma de decisiones. Idealmente cada conjunto de datos debería tener un custodio reconocido, aunque por razones de prioridad los custodios de conjuntos de datos fundamentales o esenciales, deberían determinarse primero.



Los conjuntos de datos fundamentales son aquellos vitales para la operación de una organización de manera que ellos justifiquen el esfuerzo y los costos debidos a su colección, almacenamiento y aseguramiento de la calidad. Los conjuntos de datos fundamentales son la base para el desarrollo de múltiples productos de información, para múltiples usuarios y de esta manera tienen una naturaleza de largo plazo. Por el contrario los conjuntos de datos no fundamentales se producen mediante procesos no documentados con el fin de obtener resultados rápidos, transitorios y de calidad incierta. Desafortunadamente muchas organizaciones no son conscientes de la distinción entre los conjuntos de datos.

## **7.2 Funciones de un Custodio**

### **7.2.1 Responsabilidades y derechos**

La custodia de un conjunto de datos conlleva ciertas responsabilidades y derechos:

- Responsabilidades
  - Crear un conjunto de datos (con socios cuando apropiado)
  - Respetar la propiedad intelectual y otros derechos de los propietarios y proveedores de datos
  - Mantener un conjunto de datos (es decir, actualizar, asegurar su calidad, reestructurarlos).
  - Asegurar un conjunto de datos (es decir, garantizar la integridad física y proteger los datos de pérdida daño y acceso no autorizado)
  - Proporcionar la documentación adecuada del conjunto de datos (Metadatos)
  - Proporcionar un acceso adecuado al conjunto de datos (Legitimación de usuarios de acuerdo a los protocolos establecidos)
  - Suministrar asesoría para el uso apropiado de un conjunto de datos (usos sugeridos, impropios, inadecuados)
  - Coordinar de desarrollos posteriores de un conjunto de datos (con socios apropiados y con la asistencia de usuarios claves)
- Derechos
  - Regular el acceso a un conjunto de datos (dependiendo de la categoría del usuario)



- Salvaguardar la propiedad intelectual (es decir, reconocimiento de créditos, regulación de copias o de distribución futura)
- Recuperar costos (es decir, valor comercial, inversión original, costos marginales de distribución)

### 7.2.2 Custodios y Propietarios

Los derechos de propiedad sobre los datos son un tema muy complejo. La siguiente es una guía pragmática que intenta clarificar los elementos más importantes.

Según Eldred (1995) uno de los principios fundamentales sobre los cuales se basa la Ley de Copyright es que ella no protege ideas, cifras o información. Únicamente protege una expresión particular de esas ideas, cifras o información.

Los principios de propiedad intelectual tradicionales no funcionan muy bien dentro de los dramáticos desarrollos de la tecnología de la información. Como resultado, a menudo no es claro cuánta protección le da la ley a las bases de datos, por ejemplo a los Sistemas de Información Ambiental, la WIPO está examinando este tema actualmente. Bajo la ley Británica actual por ejemplo un examen muy simple "sweat of the brow"

## 8 PROPIEDAD DE LOS DATOS

## 9 ACCESO A LOS DATOS

## 10 CALIDAD

## 11 RESPONSABILIDAD

## 12 COMERCIALIZACION

## 13 PRIVACIDAD

