

IAI Data and Information System (IAI-DIS): A Concept Proposal by Brazil/INPE

Revised June, 1996 in conjunction with the IAI Directorate

1. Introduction

Global Change investigators and programs must rely on an information base that is by its own nature spread over large areas, across countries and institutions.

Research methods in recent years have evolved to take advantage of timely, reliable satellite data and of a worldwide communications system that allows information to be relayed and shared among cooperating organisms with increasing efficiency, as well as to employ massive computer power to process, digest and correlate environmental data to forecast weather and climate phenomena in regional and global scales.

Integration is a key factor in this context of surveying, extracting and combining scientific information from different sources to support and enhance Global Change research work. Nowadays, scientific institutions and even individual researchers from the First World have their computers plugged into what seems to be a hyperspace where no data lie more than a few clicks away. In developed countries, scientists are getting used to have most of the data relevant to their domain at the reach of their desktop computer or workstation.

Unfortunately, in developing countries the situation is still way behind this reality. Institutions with efficient electronic mail access are still an exception, and the ability to inquire remote databases from the personal desktop is a privilege that few individual researchers can take for granted. On the other hand, there are several instances among the American countries where local organizations keep important data holdings which are not easily accessible (and sometimes not even known) to the outside, due to reasons that can be internal or external to the institutions, or both. Currently available technology can overcome these limitations at a moderate cost, but sometimes existing personnel or infrastructure constraints make it difficult or impractical to implement effective solutions in the short run.

The nature and scope of the IAI¹ can aid in establishing a regional network of data and information services that would help correct this situation by allowing global change researchers of both developed and developing countries to access and utilize each others' data for the benefit of scientific work.

2. Basic Requirements

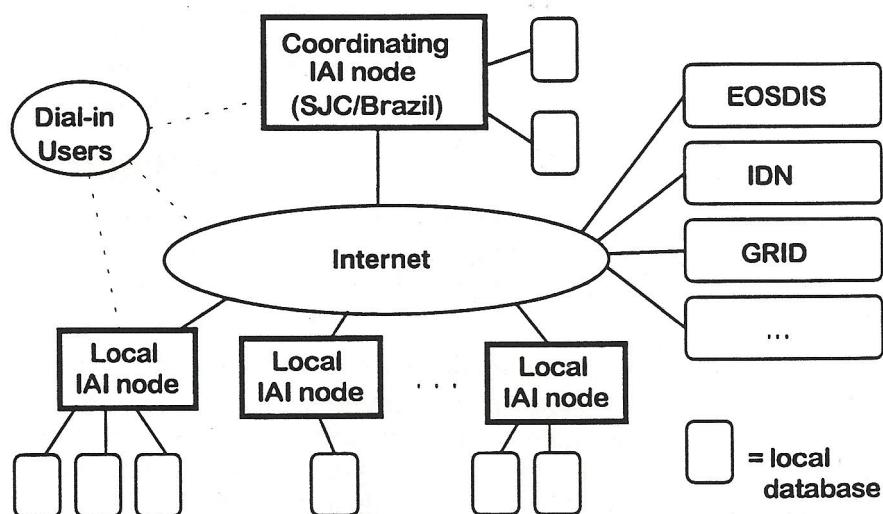
The list below summarizes the main requirements envisioned for an effective IAI Data and Information System (IAI-DIS):

- Distributed archive architecture to avoid the need to move or replicate data holdings;

¹ The Inter-American Institute for Global Change Research (IAI) was established with the objectives, among others, of "improving the scientific and technical capabilities and research infrastructure of the States of the region by identifying and promoting the development of facilities for the implementation of data management [...]" and "fostering standardization, collection, analysis and exchange of scientific data relevant to global change".

- Nodes at each member State holding relevant metadata in a format interoperable among all States and appropriate server software, capable of handling remote inquiries and of bridging requests for data or products to the respective data holders;
- International digital communications capability at 64Kbps as a minimum (employing direct satellite links if necessary) at each node, using the Internet as backbone;
- User-friendly client software at each node to support network and dial-up remote inquiries, searches, browses and request placement where applicable;
- Coordinating node at the IAI headquarters holding IAI-specific information (projects in course, contacts database, bulletin board etc) for use by IAI members and to support IAI administration;
- Client software at each node to allow access to other relevant international data systems such as NASA's EOSDIS, the CEOS IDN and GRID from UNEP.

The figure below suggests the architecture of a fully implemented IAI-DIS.



IAI-DIS Architecture Concept

Each IAI-DIS node should act in its country as a focal point of national support to Global Change studies. To achieve this role, it should seek cooperation with local holders of data relevant to Global Change, to generate and incorporate suitable metadata to make their holdings known — and available, if possible — to the scientific community, through IAI-DIS. The interconnection of all nodes through the Internet, employing adequate software, would allow that data from all nodes be searchable and visible from any node.

The Coordinating Node, in addition to its routine duties as part of the system, would oversee the operation of the whole system and be responsible for assisting and supervising each Local Node in the aspects of data management, node administration and compliance with the IAI Data Policy Statement, including the organization of training courses and seminars as necessary.

3. Criteria for the establishment of IAI-DIS Local Nodes

Taking into account that the effectiveness of the IAI-DIS as a whole relies on the contribution of each node, a set of eligibility criteria must be matched by organizations willing to host a Local Node.

It should be noted that dial-up and network access to the equipment is absolutely critical to the success of the system, as is the maintenance and creation of metadata.

Criteria for the location of the Local Nodes are as follows:

- Internet connection (64Kbps minimum, higher preferable)
- Dial-up facility available
- Existing relationships with data holders, or willingness and capability to develop them in short term
- Capability to support operating costs:
 - personnel (2 to 4 full-time people)
 - maintenance/replacements
 - communications
 - metadata preparation
 - local overhead

4. Operating Policy

The IAI-DIS would operate in accordance with the Data Policy Statements of the IAI. Automatic mechanisms and administrative procedures to ensure and verify this compliance (user registration and authentication, statistics reports, audit trails, etc.) should be an integral part of IAI-DIS.

Access to data generated by IAI-sponsored projects will be free and open, but we suggest that the Local Nodes also hold information on other databases of interest, whether they are freely available or not (e.g. met data, census data, etc.). All metadata should be searchable, visible and available to anyone accessing the system either through the Internet or via remote login, but actual data delivery through IAI-DIS would be subject to an authorization and request submission scheme managed by each node with regard to their respective holdings, in consonance with the Institute policy.

The above mentioned mechanisms and procedures would be defined and detailed in interaction with IAI specialists during the systems design phase of the IAI-DIS implementation.

5. Cost and Time Estimates

Preliminary studies indicate that a full-featured Coordinating Node at the IAI Directorate in Brazil, based on RISC workstations, can be implemented and phased into operation with a budget of about USD150K², including development, equipment, software and training costs. INPE, the IAI Directorate host organization, is willing to contribute a substantial part of that budget.

A conservative estimate for the implementation of the Coordinating Node is about one year's time. However, initial operation in an experimental (but useful) basis can be expected six

² Personnel, travel, communications, administrative costs USD80K; equipment USD20K (taking advantage of existing GEF equipment); commercial software USD20K; training USD30K.

months after project start, working upon a considerable amount of environmental data held by INPE.

Subsequent installation of local nodes at member countries would take advantage of the gained experience and development carried for the Coordinating Node, and hence it has a lower cost estimate of around USD50K³ per node, if installed upon existing IAI/GEF equipment. The IAI may wish to consider supporting some or all of the initial installation costs for the hardware as an incentive.

-.-

³ Infrastructure, installation and commercial software USD30K (taking advantage of existing GEF equipment); training USD20K.

IAI: EC-IV/4.WD/Julio 31, 1996
Español

**Sistema de Datos e Información del IAI (IAI-DIS):
Una propuesta conceptual del INPE/Brasil**

Revisado en junio de 1996 en combinación con el Directorio del IAI

1. Introducción

Los investigadores y programas interesados en el cambio mundial tienen que depender de una base de información que por naturaleza propia abarque grandes zonas, a través de países e instituciones.

Los métodos de investigación empleados en los últimos años han evolucionado para poder aprovechar los datos oportunos y fiables retransmitidos por satélite y un sistema mundial de comunicaciones que permite retransmitir y compartir información entre organismos cooperantes con una eficiencia cada vez mayor, así como emplear la enorme capacidad de las computadoras para procesar, abreviar y correlacionar datos sobre el medio ambiente para predecir fenómenos atmosféricos y climáticos a escala regional y mundial.

La integración es un factor clave en este contexto de selección, extracción y combinación de información científica de fuentes diferentes para respaldar y mejorar el trabajo de investigación sobre el cambio mundial. En la actualidad, instituciones científicas e incluso investigadores individuales del Primer Mundo tienen conectadas sus computadoras en lo que parece ser un hiperespacio dentro del cual los datos se pueden obtener oprimiendo unas pocas teclas. En los países desarrollados, los científicos se están acostumbrando a tener la mayoría de los datos de su dominio al alcance de sus computadoras o estaciones de trabajo.

Lamentablemente, en los países en desarrollo la situación todavía está lejos de esta realidad. Las instituciones con acceso eficiente al correo electrónico siguen siendo la excepción, y la aptitud para acceder a bases de datos remotas desde una computadora personal es un privilegio que pocos investigadores pueden dar por un hecho. Por otro lado, en los países americanos han varios ejemplos de organizaciones locales que mantienen importantes fuentes de datos a las cuales el mundo exterior no tiene fácil acceso (y algunas veces ni siquiera se conocen) por razones ya sea internas o ajenas a las instituciones, o ambas cosas. La tecnología actual puede superar estas limitaciones a un costo moderado, pero algunas veces las restricciones en personal o infraestructura hacen difícil o impracticable la implementación de soluciones eficaces a corto plazo.

La naturaleza y alcance del IAI¹ puede ayudar a establecer una red de datos y servicios de información regionales que ayudarían a corregir esta situación al permitir a investigadores del cambio mundial en los países desarrollados y en desarrollo tener acceso a, y utilizar, los datos respectivos para beneficio del trabajo científico.

2. Requisitos básicos

En la lista siguiente se resumen los requisitos principales previstos para un Sistema de Datos e Información eficaz para el IAI (IAI-DIS):

- Arquitectura distribuida de archivos para no tener que transferir o duplicar bancos de datos;
- Nodos en cada Estado miembro que mantenga metadatos pertinentes en un formato utilizable entre todos los Estados y software idóneo para servidor, capaz de atender solicitudes de lugares distantes y de transferir datos o productos a los bancos de datos respectivos;
- Capacidad para transmitir comunicaciones digitales internacionales a 64 Kbps como mínimo (empleando enlaces directos por satélite en caso necesario) en cada nodo, usando la red Internet como sistema primario;
- Software de fácil uso por los clientes en cada nodo para atender solicitudes de lugares remotos, búsquedas, hojeadas ("browsing") y colocación de peticiones por la red y el sistema telefónico, cuando corresponda;
- Nodo coordinador en las oficinas centrales del IAI con información específica del IAI (proyectos en marcha, base de datos de contactos, pizarra de anuncios, etc.) para el uso de los miembros del IAI y respaldar la administración de éste;
- Software para clientes en cada nodo para permitir el acceso a otros sistemas de datos internacionales pertinentes, como el EOSDIS de la NASA, el IDN de CEOS y el GRID del PNUMA.

La figura siguiente sugiere la arquitectura de un DIS totalmente implementado para el IAI.

- 1 Usuarios que llaman por teléfono
- 2 Nodo coordinador del IAI (SJC/Brasil)

¹ El Instituto Interamericano para la Investigación del Cambio Global (IAI) fue establecido con los objetivos, entre otros, de "mejorar la capacidad científica y técnica y la infraestructura de investigación de los Estados de la región mediante la identificación y promoción del desarrollo de instalaciones para implementar el manejo de datos [...]" y "promover la estandarización, recopilación, análisis e intercambio de datos científicos pertinentes al cambio mundial."

- 3 Internet
- 4 Nodo local del IAI
- 5 Nodo local del IAI
- 6 EOS-DIS
- 7 IDN
- 8 GRID
- 9 Nodo local del IAI
- 10 Base de datos local

Concepto de arquitectura para el DIS del IAI

Cada nodo del IAI-DIS debe actuar en su país como un punto central de apoyo nacional a los estudios sobre el cambio mundial. Para desempeñar esta función, debe solicitar la cooperación de retenedores locales de datos pertinentes relacionados con el cambio mundial, a fin de generar e incorporar metadatos idóneos para dar a conocer sus fuentes de datos -y ofrecerlas, de ser posible- a la comunidad científica por intermedio del sistema IAI-DIS. La interconexión de todos los nodos a través de Internet, empleando software adecuado, permitiría buscar, y observar desde cualquier nodo, datos procedentes de todos los nodos. El Nodo Coordinador, además de sus funciones rutinarias como parte del sistema, supervisará la operación de todo el sistema y sería responsable de ayudar y supervisar cada nodo local en los aspectos del manejo de datos, administración de nodos y cumplimiento con la Declaración de Política sobre Datos del IAI, incluida la organización de los cursos de capacitación y seminarios que sean necesarios.

3. Criterios para el establecimiento de nodos locales del IAI-DIS

Teniendo en cuenta que la eficacia del IAI-DIS en pleno depende de la contribución de cada nodo, es preciso que las organizaciones dispuestas a mantener un nodo local cumplan con un conjunto de criterios de elegibilidad.

Cabe señalar que el acceso al equipo por medio del sistema telefónico y de la red es absolutamente crítico para el éxito del sistema, al igual que el mantenimiento y creación de metadatos.

Los criterios para ubicar los nodos locales son los siguientes:

- Conexión a la red Internet (velocidad mínima de 64 Kbps, preferiblemente más alta)
- Facilidad para llamar por la red telefónica
- Relaciones existentes con retenedores de datos, o disposición y capacidad para establecerlas a corto plazo
- Capacidad para mantener los costos de operación:

- personal (2 a 4 empleados a tiempo completo)
- mantenimiento/repuestos
- comunicaciones
- preparación de metadatos
- gastos generales locales

4. Política operativa

El sistema IAI-DIS funcionaría en conformidad con las Declaraciones del IAI en cuanto a Política sobre Datos. Los mecanismos y procedimientos administrativos automáticos para asegurar y verificar este cumplimiento (inscripción y autenticación de usuarios, informes sobre estadísticas, comprobantes de auditoría, etc), deberían formar parte del IAI-DIS.

El acceso a datos resultantes de proyectos patrocinados por el IAI será gratuito y abierto, pero sugerimos que los nodos locales también mantengan información sobre otras bases de datos de interés, ya sea que estén libremente disponibles o no (por ej., metadatos, datos de censos, etc.). Cualquiera que tenga acceso al sistema, ya sea por Internet u otro sistema remoto, deberá poder buscar, ver y obtener todos los metadatos, pero la entrega de datos por intermedio del IAI-DIS estaría sujeta a un procedimiento de autorización y solicitud administrado por cada nodo con respecto a sus fuentes de datos respectivas, en consonancia con la política del Instituto.

Los mecanismos y procedimientos arriba mencionados se definirían y detallarían con la colaboración de especialistas del IAI durante la fase de diseño de sistemas de la ejecución del IAI-DIS.

5. Estimaciones de costos y tiempo

Los estudios preliminares indican que un Nodo Coordinador completo en el Directorio del IAI en el Brasil, basado en estaciones de trabajo RISC, se puede instalar y comenzar a prestar servicio gradualmente con un presupuesto de alrededor de US\$150.000², que incluye costos de desarrollo, equipo, software y capacitación. El INPE, organización anfitriona del Directorio del IAI, está dispuesto a contribuir con una parte considerable de ese presupuesto.

Una estimación moderada del tiempo que tomaría implementar el Nodo Coordinador es de alrededor de un año. Sin embargo, la operación inicial con carácter experimental (pero útil) se puede esperar que tome unos seis meses después de comenzar el proyecto, trabajando con una cantidad considerable de datos sobre el medio ambiente mantenidos por el INPE.

² Costos de personal, viaje, comunicaciones y administrativos: US\$80.000; equipo: US\$20.000 (aprovechando el equipo existente del FMAM); software comercial: US\$20.000; capacitación: US\$30.000.

La instalación subsiguiente de nodos locales en países miembros se valdría de la experiencia adquirida y el desarrollo llevado a cabo para el Nodo Coordinador, y por tanto su costo estimado es menor, o sea, alrededor de US\$50.000³ por nodo, si se instala en el equipo existente del FMAM del IAI. El IAI tal vez desee considerar pagar una parte o todos los costos iniciales de instalación de los equipos de computación como incentivo.

³ El costo de infraestructura, instalación y software comercial asciende a US\$30.000 (aprovechando el equipo existente del FMAM); el de capacitación, a US\$20.000.