# REVOLUTIONARY ARCHITECTURES FOR FUTURE EARTH OBSERVATIONS: EARTH SCIENCE INFORMATION WEB AND KNOWLEDGE CREATION

**Daniel S. DeVito, Janice K. Smith, Dennis J. Andrucyk, William J. Campbell**
National Aeronautics and Space Administration, Goddard Space Flight Center
Greenbelt, MD 20771, USA

## ABSTRACT

Significant advances in Earth science and applications depend on comprehensive information gathering of the global environment and the rapid translation of that information into useful, widely available knowledge. The goal of creating an adaptive web of small intelligent sensors and information systems will extend our ability to explore, comprehend, predict, and respond to our dynamic global environment. This extension, facilitates the human perception and cognition of our environment, while providing sophisticated interpretation of the data and thus improve our ability to model and forecast global events. The Earth Science Information Web ultimately provides a highly robust capability of an “information anywhere/information anytime” concept.

## 1 INTRODUCTION

With ever increasing demands on our Earth’s limited, natural resources, the need for knowledge may appear boundless. Applications such as precision agriculture, fishing, natural disaster prediction, weather forecasting and nowcasting continue to help Governments, industries, academia, and the public perform strategic and tactical planning in order to protect lives, property and investments. Long term science goals of understanding Earth and its environment as a system and the concomitant ability to use this model are necessary for predicting future behavior of both local, rapidly changing phenomena and long-term, gradual shifts in the global ecosystem. Significant advances in Earth science and applications depend on comprehensive information gathering of the global environment and the rapid translation of that information into useful knowledge. The concept of an inter-operable intelligent sensorweb combined with advanced information translation and data reduction techniques provide the Earth science user community as well as Government and industrial communities the ability to efficiently and effectively monitor, explore, respond and predict changes to the Earth environment. The advances in sensors, flight data systems, and communications as noted in The Intelligent Sensorweb paper [Reference 1], presents a partial story to the improvement of our Earth observational processes. The space asset technologies afford an increased volume of data, significant to our understanding of the Earth, but far more data than ever envisioned. Thus, to complete the observational story, it is necessary to acquire the observational parameters and convert them into useful knowledge (Figure 1).

A sensor web provides the capability of continuously observing the Earth environment by acquiring thousands of observational parameters, yielding petabytes of raw data per day. By increasing the intelligence and data reduction ability of the sensor web assets, the informational parameters may be retrieved from the observed data and made available to the rest of the information web. The resulting terabytes of information parameters are ultimately saved as part of the historical record of the Earth environment. The subsequent use of data fusion, historical modeling, and forecast modeling techniques, the information parameters are further reduced to gigabytes of impact parameters. These impact parameters, representing changes to previously acquired information, are mined and visualized by a wide variety of users seeking knowledge relevant to their interest. Ultimately, a finished product (varying from kilobytes to megabytes) is generated seamlessly for the user from the petabytes of available data.

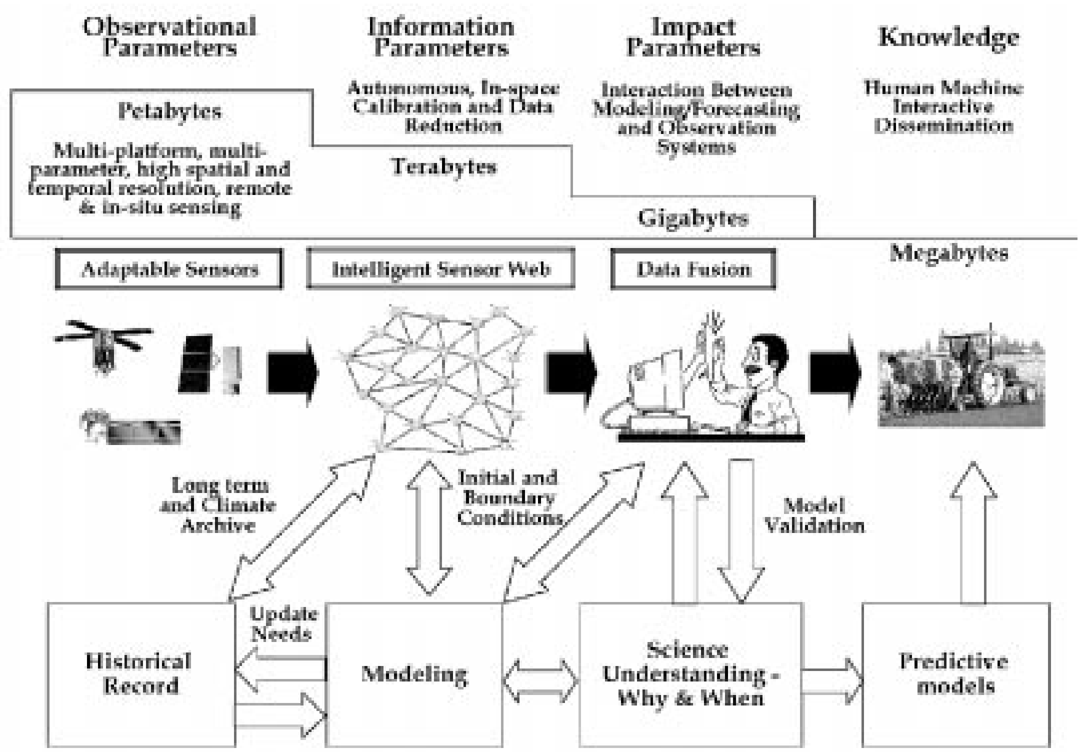


Figure 1. Managing End-to-End Information Flow

## 2 ARCHITECTURAL FEATURES

Major challenges to achieving an Earth Science Information Web include: enabling interoperability with heterogeneous sensors and information systems; performing efficient transformation of collected data into knowledge; providing advanced environments for knowledge dissemination; infusing knowledge feedback into the sensorweb and intelligently archiving key data sets and information.

### 2.1 Onboard Information Systems

Onboard information systems, with sufficient capacity, allow value-added providers and end users to upload data requests and processing instructions, in order to receive tailored products. Automated planning and scheduling systems coordinate observations within/between sensors to accommodate requests. Data from observations are stored on terabit-sized, solid-state recorders with associated metadata recorded within freely accessible databases. These temporary digital libraries are available for search and order of raw, observational data, or for information processed to higher levels by reconfigurable processing farms. Requested information may be directly delivered via RF/optical link, or indirectly delivered through a crosslink to a space-based Internet. Temporary digital library holdings are maintained for a pre-determined period of time. Unique event data or critical information parameters may be identified for longer retention given a tiered temporary storage concept.

## 2.2 Heterogeneous Digital Libraries

Digital Libraries focused on a variety of interests receive terabytes of information from the intelligent sensor web per day. Prospecting, mining and subsequent fusion of dissimilar datasets increases along with the emphasis on inter-disciplinary science and applications. Utilization of dissimilar data sets has long been recognized as a manually intensive undertaking due to the heterogeneous nature of the digital libraries which house the data. The development of catalog interoperability, data format standards, and metadata format standards between digital libraries of similar datasets and technology are successful. However, the effectiveness of these techniques decreases sharply as the heterogeneity of digital libraries increases. Therefore, intelligent, advertising agents for digital libraries may act as multi-lingual educators for potential customers. The advertising agent not only describes the type of data holdings within a digital library, but instructs a potential customer on how to natively interact with a digital library.

## 2.3 Metadata Warehousing

The highly distributed nature of the Earth Science Information Web poses significant challenges to efficient data searching. Current concepts for data mining aid in retrieving information from libraries only if the user knows where to look and what specific datasets are desired. The more difficult challenge of searching for data/information across many digital libraries within the information web requires advanced capabilities such as data prospecting and metadata warehousing. Metadata warehouses collect only metadata and browse images from temporary libraries within the sensor web and ground based digital libraries. They have no direct role in maintaining the distributed archive, but provide information about the location and attributes of different types of data collections within the archive (e.g. data relevant to land utilization). Metadata warehouses enable the user to perform complex, high speed, searches across hundreds of libraries. This capability enables efficient, exhaustive data prospecting across the entire information web without detailed knowledge about the underlying architecture. If a user is interested in obtaining specific data as a result of a search, the user is directed by the metadata warehouse to the appropriate digital library.

## 2.4 Value-added Providers

Value-added providers provide a critical role in the Earth Science Information Web. Many users lack the computational capability, expertise, or desire to reduce data/information into knowledge. Value added providers develop tools and services for millions of users of the information web. For example, a farmer may subscribe to a service, which makes recommendations on what crop (or strain of crop) to plant by calculating the seasonal forecast specific to the farm's location and monitoring crop futures. Once planted, the value-added provider continually monitors the farms local weather, soil moisture, growth, and diseases to make detailed recommendations for maximizing profit.

## 2.5 Earth Science Research Assistant

While metadata warehousing and value-added providers aid in the utilization of the Earth Science Information Web, a personalized, integrated suite of tools for the end user are key to broad cultural acceptance. The integrated tool suite acts as an Earth science research assistant by helping the end user; plan and coordinate activities; search for and order data/information; receive data/information directly from sensor web assets; process information; maintain a personal digital library; analyze and visualize results; and share knowledge with other users. The centerpiece of the Earth science research assistant is the intelligent, research agent, which provides a single interface for the owner to the sensor and information webs.

Given broad objectives and guidelines by the owner, the agent actively carries out tasks on the Earth Science Information Web. Coordinating with the various heterogeneous digital library and metadata warehouse, advertising agents, the research agent performs prospecting and mining activities. The research agent also carries the owner's credentials and is capable of requesting authorized observations and services by the sensor web and can negotiate for services with value-added providers.

The intelligent research agent learns and adapts to changes in owner habits and to the sensor and information webs. The research agent can continually perform passive prospecting and mining for new sources of information and services of potential interest to the owner. If the agent is unable to perform a requested function the agent will request additional guidance and/or clarification from the owner.

### 2.6 Advanced Visualization

How information is presented and shared is important to the end user obtaining knowledge relevant to their interest. Advanced visualization technologies, such as the immersive environments under consideration for the Digital Earth Initiative, allow seamless access to information and knowledge. Immersive environments provide users with the ability to manipulate large amounts of information based on body movements and eye motion. As stated in the United States Vice President's address to the California Science Center, "(Users of Digital Earth will be able to) zoom in, using higher and higher levels of resolution, to see continents, then regions, countries, cities, and finally individual houses, trees, and other natural and man-made objects." Distributed immersive environments facilitate the sharing of knowledge and experience between users by providing the capability for people to work interactively and remotely with information and each other.

## 3 ACHIEVING THE VISION

The vision of an Earth Science Information Web may be realized by employing the same model used to establish the World Wide Web. Governments need to establish the foundation by investing in the basic research of key standards, tools, and technologies, and making them available to industry. Industry will tend to adopt the foundation where viable markets for Earth science information products and services are foreseen. As individual users discover how the products and services aid in every day life, the Earth Science Information Web will take root and flourish.

1. S.P. Neeck, R.L. Taylor, J.B. Garvin, M.G. Ryschkewitsch, W. Wiscombe, *Revolutionary Architectures for Future Earth Observations: The Intelligent Sensor Web*, 2nd IAA Symposium on Small Satellites for Earth Observation, (April 1999)