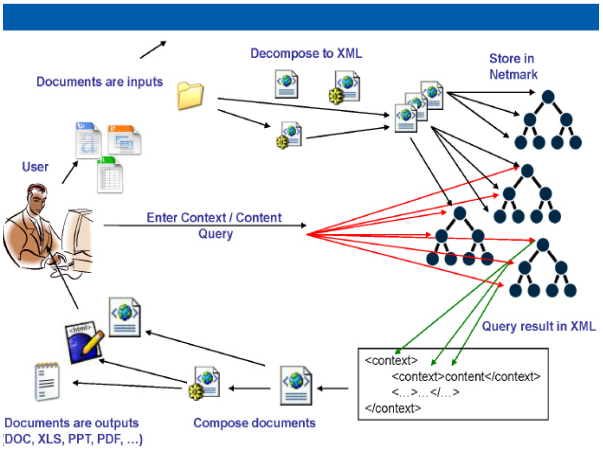
National Aeronautics and Space Administration

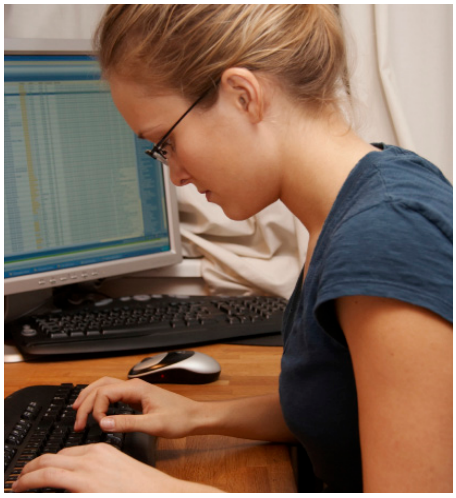
technology opportunity

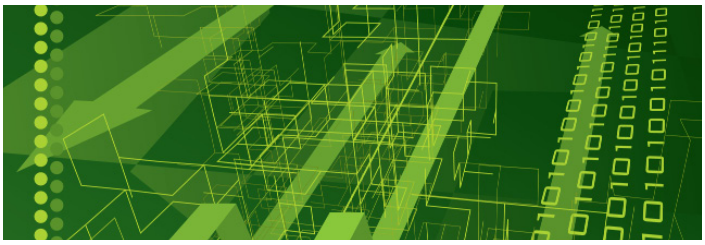
NETMARK

An Advanced XML Database Integration Technique for Managing Unstructured Documents



NETMARK's schema-less integration technique converts information from many different data types into a universal data type for unprecedented information assimilation and retrieval across the enterprise.





NASA's Ames Research Center offers for license its NETMARK software, a unique innovation designed to seamlessly integrate structured, semi-structured, and unstructured data and documents across enterprise organizations. Originally developed to integrate the vast quantities of complex, heterogeneous documents existing within NASA, this schema-less integration technique and framework offers a highly scalable, open enterprise database architecture that eliminates or reduces the need for database design and administration, and converts information from a wide range of data types into a single, universal data type for storage, retrieval, and content- and context-sensitive query and search. A production-ready, enterprise-level application, NETMARK rapidly assimilates and retrieves gigabytes of disparate information and can be easily integrated with existing applications as well as accommodate new data formats—fitting into the legacy data network while growing with evolving technologies and business practices.

www.nasa.gov

Benefits

- **Economical:** Eliminates the need to design, develop, and maintain expensive, highly structured relational databases, lowering both software and administrative costs
- **Flexible:** Combines information from heterogeneous structured, semi-structured, and unstructured data sources, and enables easy and unstructured data queries
- **Adaptable:** Enables query-based composition of Microsoft® Office documents, Macromedia® Flash® applications, HTML Web pages, and others, and runs queries on Flash, Perl, C, C++, Java, Visual Basic, and other environments that support http and https protocols
- **Secure:** Limits query results to the information that users and groups have permission to access
- **Custom:** Includes configurable databanks for tailored query workflows in diverse applications
- **User-friendly:** Offers toolbars for Microsoft Internet Explorer and HTML, and is built on W3C international standards, including HTTP, WebDAV, XML, XSL, and XSLT

Applications

- Enterprise knowledge management applications
- Document and content management systems

NETMARK supports a user-friendly portal for simple data inputs and outputs, which can include standard business documents such as Microsoft Office files.

Technology Details

How It Works

NETMARK takes advantage of an object-relational model and the eXtensible Markup Language (XML) standard, along with an open, extensible database framework to dynamically generate arbitrary schema stored within relational databases and an object-relational database management system. NETMARK maps XML-encoded information into a true data model by employing a customizable data type definition structure, defined by an SGML parser to model the hierarchical structure of XML data regardless of any particular XML document schema representation.

By achieving a true XML data model, NEMARK can help enterprise organizations make better use of the information they need to make business decisions by converting Web pages, text documents, PDF files, spreadsheets, presentations, and other document types into a single, universal data type, then storing it in an object-relational database. Users can query this database with searches that are based on content or contextual associations. Query results then can be composed into different data types, including presentations, spreadsheets, and text documents, enabling rapid reuse of information and broadening the scope of data from which users can gain knowledge and make decisions.

Why It Is Better

Most traditional document management systems do not provide an easy and efficient mechanism to store, manage, and query relevant information from heterogeneous and complex data types. To do so, database management systems need a standard for common data and exchange. The industry standard, XML, places structure within documents. The traditional mapping model is limited because the hierarchy is different for each set of XML documents. In contrast, NETMARK's SGML parser models the documents themselves, and its structure is the same for all XML documents, providing independence of any particular XML document schemas.

IP Protection

NASA has secured a patent for its NETMARK software: U.S. Patent No. 6,968,338.

Licensing and Partnering Opportunities

This technology is part of NASA's Innovative Partnerships Program, which seeks to transfer technology into and out of NASA to benefit the space program and U.S. industry. NASA invites companies to inquire about the licensing possibilities for the NETMARK technology (ARC-14662-1, ARC-15098-1, ARC-15089-1, and ARC-15370-1) for commercial applications.

For More Information

If you would like more information about this technology, please contact:

Andrew Vo
Technology Partnerships Division
NASA Ames Research Center
(650) 604-0004, andrew.vo@nasa.gov

NP-2009-11-02-ARC
11.05.09