

Xpiori XMS Powers Gene Expression Analysis Site

Meets the data demands of bioinformatics, providing essential speed & flexibility

Company

University of Denver, Center
For Computational Biology
(CCB)

Industry

Bioinformatics

Web Site

www.cudenver.edu/ccb

Background

The Human Genome Project generated vast amounts of data at un-precedented speeds. This drove a new revolution in molecular biology, spawning the field of computational biology.

Challenge

The Center for Computational Pharmacology (CCP), allied with the CCB, needed a flexible database backend to support a Web site housing gene expression data for a nationwide six-university consortium.

Solution

Xpiori XMS handles the volume and complexity of CCP's data types with high speed and minimal maintenance. Plus, it offers inherent flexibility to handle the center's evolving data needs.

Why XMS?

This high-performance native XML database supports the flexibility and scalability inherent in XML, a widely used data format in bioinformatics. Based on a patented technology, Xpiori XMS can handle large amounts of unstructured and highly dynamic data.

Molecular biology has increasingly become an information science, giving birth to the field of computational biology. Researchers are continuously creating high volumes of diverse data and must have the ability to store and analyze it quickly and easily.

Xpiori XMS, a self-constructing, native XML database, delivered the speed and flexibility to serve as the backend for a critical gene expression analysis Web site. With XMS managing their data, researchers can focus on finding answers to today's medical mysteries.

The Setting

Right now, molecular biologists are diligently working, laying the foundation for the future of health care. With the inception of the Human Genome Project, this field has experienced a revolution unlike any that it has seen before. Mapping the human genome created unprecedented amounts of data. Data that once took years to generate, now takes just days or minutes. Mathematics and computer science have taken on a new importance in analyzing and processing data. As a result, a new field of biology has been born – computational biology.

In response to this dramatic industry development, the Center for Computational Biology was established at the University of Colorado at Denver in 2001. Primarily, the CCB aims to foster collaborations among participating CU campuses by coordinating events and programs, and by creating key industry alliances. The core of CCB's scope is bioinformatics, the science of storing, extracting, organizing, analyzing, interpreting and using information about organisms. Ultimately, by combining efforts and technology, bioscience will drive important new discoveries in medical research.

"All the interesting problems that remain in medical research can't be solved by someone sitting alone in a lab like Jonas Salk did," explained Harvey Greenberg, director of the Center for Computational Biology. "To do serious work in biomedical research, you're going to have to involve computer science and mathematics."

The Opportunity

Like most organizations of its type, the Center for Computational Pharmacology (CCP) at the CU-Health Sciences Center, one of the institutions in alliance with the CCB, is actively involved in a data-intensive project. Under a grant from the Integrated Neuroscience Initiative on Alcoholism (INIA), the CCP and five other workgroups nationwide are investigating how alcohol affects the brain.

For this specific project, the CCP wanted to create a Web site to enable the researchers in the consortium to perform gene expression and other types of analysis. They hoped to build a site, on a flexible database backend, where they could directly upload and then analyze many different types of data. In the past, the center had tried to accomplish this goal using a relational database model. But biological data is complex to model due to the wide range of data types and their interrelationships. So, the relational database model was abandoned and an object-oriented database was tried. When Ron Taylor, Director of Gene Expression Analysis at the CCP, attempted to input one complete protein database and a variety of other data types, their initial object database system all but broke down, with the uploading and indexing process dragging out for days.

Knowing its data requirements would only be expanding in the future, the CCP needed a more powerful database solution, one that easily managed large, diverse data sets – while not consuming much staff programming and maintenance time. Also, since various XML markup languages are becoming important in bioinformatics, the CCP wanted an information storage technology that worked easily and directly with such markup languages as the MicroArray Gene Expression Markup Language (MAGE-ML), and XML query languages like XPath and XQuery. Additionally, they desired a database with the flexibility to handle both the consortium's current and future data needs as they evolved.

The Solution

With the CCP's objectives in mind, Greenberg connected the center with Xpiori LLC, maker of the Xpiori XML Information Management System (XMS), a high-performance native XML database designed to support the flexibility and scalability inherent in XML. Based on a patented technology, Xpiori XMS can handle large amounts of unstructured and highly dynamic data.

*The inherent flexibility of
Xpiori XMS could
handle the center's
continuously evolving
data needs*

As a self-constructing database, Xpiori XMS was the right backend for the volume and complexity of data types in the gene expression analysis Web site. Additionally, the native XML database provided critical support of the public XML markup and query languages that CCP and the other consortium workgroups want to use. Plus, the inherent flexibility of Xpiori XMS could handle the center's continuously evolving data needs.

As an added benefit, Xpiori has created a sequence search plug-in module to complement Xpiori XMS. In bioinformatics, sequence alignment tools are critical for enabling researchers to compare biological sequences, such as sections of DNA. Similar to a BLAST (Basic Local Alignment Tool), the Xpiori custom plug-in is based on Xpiori's patented digital pattern processing technology and allows CCP to do targeted sequence-based searches within the XML database. With it, researchers can



The Result

From the first tests, Xpiori XMS showed high speed and simplicity. When Taylor loaded into Xpiori XMS the same protein database that previously took days to input, the task was completed in just 45 minutes. That means that data is available on the Web site for researchers to access sooner, so their work isn't limited by the capabilities of the database.

Most important, Xpiori XMS has proven to be so simple to use and maintain that researchers can focus their time and energy where it matters most. "What's most valuable to us is the time of the people here in the lab," Taylor explained. "We're a fairly small lab with limited manpower, so flexibility and low maintenance are important to us. We want to concentrate on research rather than building and maintaining databases, and Xpiori XMS allows us to do that."

With a flexible backend for its Web site, the CCP feels confident that its system can handle future needs as well, such as storing data on protein levels, metabolic and signaling pathways in the brain and brain imaging data. The CCP is even looking toward Xpiori XMS as the possible future backend for the large knowledge base being built for the INIA consortium grant.

Xpiori XMS has proven it is a viable database solution not only for CCP, but for other CU institutions. This success for CCP is likewise a win for the Center for Computational Biology as it aims to connect University of Colorado researchers with the technologies to accelerate their progress – ultimately driving better health and quality of life.

***This success aims to
connect researchers with
the technologies to
accelerate their progress -
ultimately driving better
health and quality of life***

###



For more information about Xpriori, including our product brochures, white papers, and company information, please visit us at www.xpiori.com.

About Xpriori

Xpriori is a technology innovator and provider of a self-constructing XML database and related products that are powered by the company's patented Digital Pattern Processing technology.

Xpriori's products allow organizations to rapidly respond to changing business requirements by eliminating costly database development time and enabling the creation of adaptable applications – business applications that easily adapt to heterogeneous and ever-changing data sets.

Xpriori is privately held and headquartered in Colorado Springs, Colorado. For more information, visit us on the web at

www.xpiori.com.

Xpriori XMS

Xpriori® XMS, XML Information Management System, is a self-constructing native XML information management system built on Xpriori's patented Digital Pattern Processing technology (DPP) for the purpose of persisting, managing and developing solutions with XML.

Xpriori Content Manager, XCM, is only one of many applications built to leverage the power of DPP and XML using Xpriori® XML Information Management System.

With Xpriori® XMS, changes in business requirements or upgrades in functionality can be accomplished without traditional application and information design, significantly increasing an application's extensibility and flexibility while dramatically reducing its time to market and total cost of ownership (TCO). Using a pattern-centric approach, Xpriori customers realize a competitive advantage through their ability to deploy high-performance, customizable applications that fully leverage valuable XML information. Please contact Xpriori to find out how we can help you to build XML-centric applications.

Xpriori, LLC

2864 South Circle Drive, Suite 401

Colorado Springs, Colorado 80906

Phone 719-425-9840

www.xpiori.com

