Subsurface Sensing Technologies and Applications Vol. 4, No. 4, October 2003 (©2003)

A Database System to Advance Subsurface Sensing and Imaging

Huanmei Wu, Becky Norum, Betty Salzberg and David Kaeli*

Center for Subsurface Sensing and Imaging Systems, Northeastern University, Boston, MA 02115

Received September 26, 2002; revised May 17, 2003

The CenSSIS Image Database System is a scientific database that enables effective collaboration, scientific data sharing and accelerates fundamental research. We describe a state-of-the-art system that uses the Oracle RDBMS and J2EE technologies to provide remote, Internet-based data access. The system incorporates efficient submission and retrieval of images and metadata, indexing of metadata for efficient searching, and complex relational query capabilities.

Key Words. Hierarchical collections, web image database, metadata, XML tagging.

1. Introduction

The Center for Subsurface Sensing and Imaging Systems (CenSSIS) seeks to revolutionize our ability to detect and image biomedical and environmental-civil objects or conditions that are underground, underwater, or embedded within cells or inside the human body. Our unified, multidisciplinary approach combines expertise in wave physics, sensor engineering, image processing, and inverse scattering with rigorous performance testing to create new sensing system prototypes that are transitioned to our industry partners for further development.

A major barrier facing CenSSIS researchers is the storing, indexing, and sharing of subsurface image and sensor data. The geographical separation between and the diverse disciplines of CenSSIS members make collaboration a particular challenge. In addition, scientific disciplines such as biology and the earth sciences have recently been generating data at enormous rates,

*To whom all correspondence should be addresed. Phone: 617-373-5413; fax: 617-373-8970; e-mail: kaeli@ece.neu.edu

1566-0184/03/1000-0393/0 © 2003 Plenum Publishing Corporation

making it difficult for scientists to track and organize these vast repositories. The development of a centralized database system to store, organize and retrieve subsurface imaging data is key to addressing these challenges.

A centralized image database system has several benefits. First, it facilitates data collection for individual members by providing a framework for experimental annotations and variables. Also, it provides a valuable resource for the educational initiatives of CenSSIS by providing real data for students to use in the classroom. Thirdly, it minimizes the required effort of individual CenSSIS members to manage data sets, freeing their time for analysis and research. Fourth, it forces a consensus on data and imaging standards within the CenSSIS community. These standards will then facilitate the development of CenSSIS toolboxes and other data management tools.

This paper reports on the implementation of the CenSSIS Image Database System (CenSSIS-DB). We begin by discussing related work in Section 2. The data model design is explored in Section 3. Section 4 addresses the implementation of the system. Section 5 discusses potential applications. Future research topics are introduced in Section 6. We end by summarizing the project in Section 7.

2. Related Work

There are a variety of existing scientific databases that provide support for research efforts; many are available over the world wide web. While planning CenSSIS-DB, we investigated several existing scientific databases in order to ascertain their functionalities and limitations. Next we discuss some of these related efforts.

Butler *et al.* [1] briefly review current research work in Bioinformatics. They discuss how the rising rate of data acquisition in the biological sciences has increased the need for more complex data management and computational tools; findings which mirror our own. They describe several endeavours, primarily in the field of genomics, as typical Bioinformatics projects.

Well-known and publicly available databases include the National Center for Biotechnology Information (NCBI) [11] databases (e.g., GenBank, PubMed, MMDB) and the National Cancer Institute (NCI) Cooperative Breast Cancer Tissue Resource (CBCTR) [10]. NCBI is a national resource for molecular biology information, including genetic sequences, protein sequences and structures, and genetic disorders. The CBCTR provides clinical data for specimens distributed to clinical researchers, but images of specimens and/or microarrays are not available.

Other scientific databases include BioSig [12,13], the European Computerized Human Brain Database (ECHBD) [4,5], the Cell Centered Database (CCDB) [2] at the National Center for Microscopy and Imaging

Research, and the Alexandria Digital Earth ProtoType (ADEPT) [8,15]. These databases are all sophisticated online scientific image databases similar to our own. The BioSig system is an imaging bioinformatics system for studying phenomics and intracellular signaling. The ECHBD is customized for the viewing and analysis of brain specimens. The CCDB is an object-relational database system that stores high-resolution 3D light and electronic microscope reconstructions. ADEPT is a distributed digital library (DL) of spatial map sets covering most of the world and including images from satellite, space shuttle, aerial, and other sources. We studied each of these systems, focusing upon their data models and user interfaces, to determine best practices as we developed our own design. For example, while the CCDB data model is very similar to our design, the CCDB places no restrictions upon data access or submission, which we have identified as a key component of our architectural requirements.

The Distributed Metadata Server (DIMES) [17] is an earth science data system that is especially interesting because of its flexibility. DIMES was designed with extensibility in mind and accepts metadata submissions in any valid XML format, thus placing no restrictions on metadata entries. While a metadata design provides for great flexibility, this approach also decreases query performance. Since data is stored and queried in XML format, rather than a relational database management system optimized for complex query formation, system performance is far from optimal.

TaxaServer [9] is a Web-based taxonomy browser. It uses an explicit taxonomy as a key element in a system that integrates distributed heterogeneous biological databases. The system features a Web interface that enables a user to navigate a taxonomy tree and to post queries to distributed biological collections databases based on a selected node in the taxonomy tree. It is similar to our hierarchical view of the collections as discussed in subsection 3.4.3, except that TaxaServer is the taxonomy tree of databases while our hierarchical view is for data collections.

SkyServer [6,16] and SIMBAD [14] are examples of the online astronomical database. The SkyServer provides the online access to the public Sloan Digital Sky Survey (SDSS) data. The SIMBAD astronomical database provides basic data, cross-identifications and bibliography for astronomical objects outside the solar system. They both characterize huge amount of astronomical objects.

3. Data Model

Our key considerations in developing a data model and choosing a relational database system were flexibility, extensibility, and reliability. The broad research base of the CenSSIS community requires that a number of different types of image data are generated, each with unique metadata characteristics. Although we have incorporated several types in our data model, it is likely that additional image types will be identified in the future; therefore, designing for extensibility and flexibility in our model is imperative.

We have identified a set of common characteristics to be included with all data sets—these are the metadata. These include fields such as the generator of the image, the date the image was generated, and the instrument used to obtain the data, and the image category. Category refers to the image type and is particularly important, because it also tells us what additional metadata is required for a particular category. For example, we want to be able to retrieve hyperspectral data sets based not only on the core metadata set, but on the wavelength range the data set is generated within. We store the starting and ending wavelengths for all hyperspectral data sets in order to enable more extensive indexing and searching within a category as well as provide additional information to users of the dataset.

Figure 1 presents a partial data model of the system as an entity–relationship (ER) model. Each box in the diagram corresponds to an entity in the database (i.e., a table). An entity has attributes (table fields). Entities can be related to one another using relations. For example, the INSTRUMENT entity has a one-to-many relationship with the DATA entity. This means that an INSTRUMENT is related to many DATA entities, but every DATA entity corresponds to only one INSTRUMENT.

Two relationships are critical in our model, and are a key to its understanding. The first is the relationship between the DATA entity and its

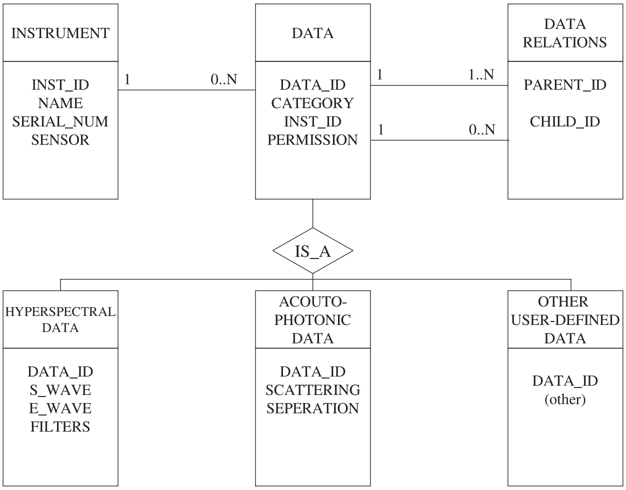


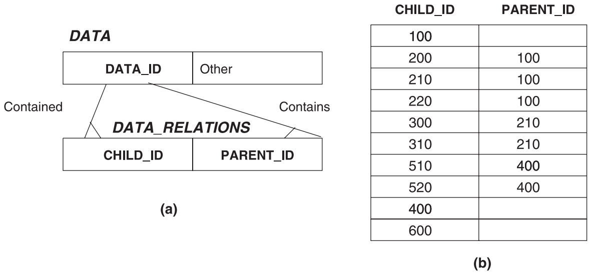
Figure 1. Partial CenSSIS-DB Data Model.

subtypes, represented by an *IS–A* relationship. A HYPERSPECTRAL entity is a subtype of the DATA entity, for example. The relationship between a row in HYPERSPECTRAL and DATA entities is one-to-one. The entities share the same primary key, the primary key of the supertype (DATA: Id). This design allows us to extend the DATA entity attributes by creating subtypes with minimal redundancy. This design also makes our model flexible and extensible, since we can create new subtypes quickly without negatively impacting the model.

The second interesting relationship is between the DATA and DATA_RELATIONS entities. This is a bill-of-materials (BOM) data structure, used to represent a data hierarchy. We want to be able to associate data sets with one another in a hierarchical fashion. To do so, we created a DATA_RELATIONS entity containing the attributes of parent and child. This allows us to generate an unlimited number of relationships between data entities (see Fig. 2) and thus enables clients to organize data sets into collections (Fig. 2).

The concept of a collection is fundamental to many of the data sets generated within CenSSIS. For example, in quadrature tomographic microscopy, multiple images of a sample are taken and reconstructed into a final processed image. It is important for us not only to store the initial raw images, but any and all reconstructions because additional reconstructions may be executed on the raw data.

In addition, we want to know all images and/or data sets related to a particular sample. The metadata model described above allows clients to create a hierarchy of images and/or data sets and organize them into experiments. The client creates a root collection (refer to nodes 100 and 400 in Fig. 3). Within this collection, the client can store images (for example,



CHILD_ID	PARENT_ID
100	
200	100
210	100
220	100
300	210
310	210
510	400
520	400
400	
600	

Figure 2. Representation of the Bill of Materials data structure in the CenSSIS-DB data model. (a) The data_id field in the DATA entity is the primary key. In order to establish relations between data, the DATA_RELATIONS entity contains two fields, both derived from the data_id field. (b) Observe that element 100 contains elements 200, 210 and 220 and element 300 is contained in element 210.

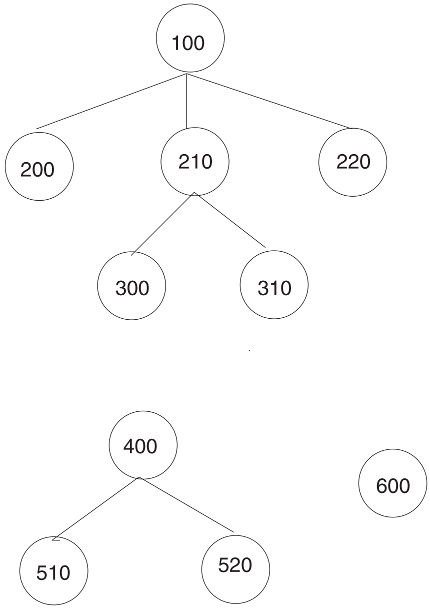


Figure 3. A hierarchical representation of the data relationships presented in Figure 2. The root nodes are 100, 400 and 600. Node 100 has three children: 200, 210 and 220. Node 210 has two children: 300 and 310. Node 400 has two children: 510 and 520. Node 600 has no children.

nodes 200 and 220 in Fig. 3), or additional collections (refer to node 210 in Fig. 3). This hierarchy is potentially limitless, allowing each client the flexibility of naming and organizing his or her own data sets, while adhering to the general metadata schema described above.

4. Implementation

Next, we describe the implementation of CenSSIS-DB. First, we describe the software architecture. Second, we describe the search and submission capabilities of the system. Third, we discuss the decision to store binary files separately from metadata. Finally, we the explain how CenSSIS-DB handles network security issues.

4.1. Architecture

CenSSIS-DB uses a standard client–server architecture (Fig. 4). The system is divided into layers comprising database access, application logic,

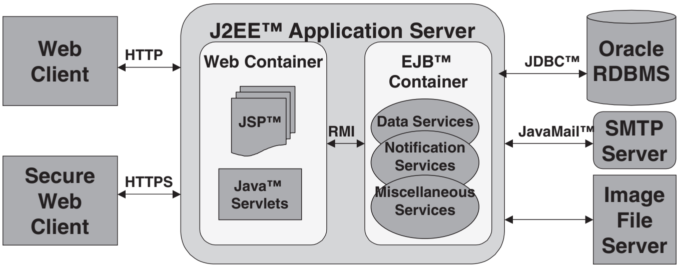


Figure 4. CenSSIS-DB System Architecture. The client can interact with the system via a secure (HTTPS) or insecure (HTTP) connection. Web pages are generated by the J2EE application server using Java Servlets and Java Server Pages. These services interact with the Enterprise JavaBeans in order to retrieve and submit data. The J2EE application server interacts with the database using JDBC, the SMTP server using JavaMail and accesses the file server.

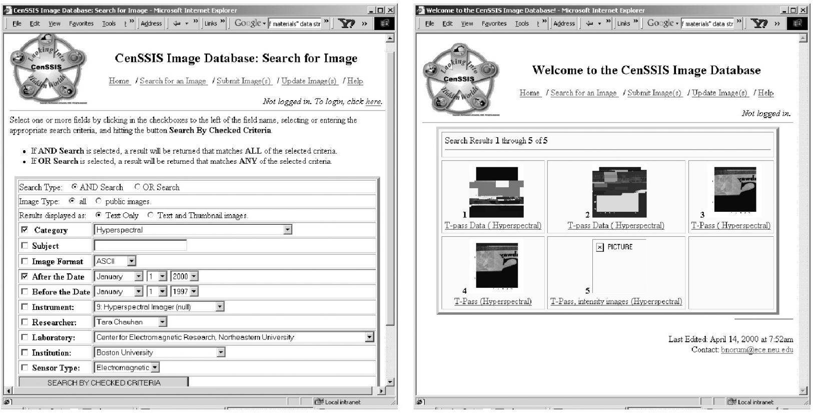
and presentation. The components are:

1. A user interface written in HTML;
2. Java source code, Java Servlets, Enterprise Java Beans (EJB), JDBC, and Java Server Pages (JSP);
3. Metadata stored in a relational database system (Oracle); and
4. Image and data files stored on a separate file server and referenced by pointers in the relational database system.

Clients access CenSSIS-DB through our web site (http://censsis-db1.ece.neu.edu). Clients can retrieve and submit information by clicking a hyperlink or submit button on the appropriate web page (i.e., the user interface). The request is received by the web server, which forwards the request for processing to the appropriate Java class. The Java class processes the request, generating the appropriate database query, and queries the relational database system using JDBC (the Java API for database systems). The database server generates a set of results and returns them to the requesting Java class. The class then processes the results for HTML display. The web server sends this dynamically generated web page back to the client's browser.

4.2. Searching

CenSSIS-DB permits clients to retrieve image/data sets based upon metadata. Clients can indicate whether results should be presented one at time or in a list format. Also, clients can select whether to have thumbnails of the data sets returned with the results for viewing. In order to ensure system flexibility, we have identified a few key types of queries that clients can use to retrieve data sets: by ID, by complex queries, or by subtree (the hierarchical view).


CenSSIS Image Database: Search for Image - Microsoft Internet Explorer
File Edit View Favorites Tools Address Links Google f materials" data str
Looking Into CenSSIS Hidden Worlds
Local intranet
Welcome to the CenSSIS Image Database! - Microsoft Internet Explorer
PICTURE

CenSSIS Image Database: Search for Image

Home / Search for an Image / Submit Image(s) / Update Image(s) / Help

Not logged in. To login, click here.

Select one or more fields by clicking in the checkboxes to the left of the field name, selecting or entering the appropriate search criteria, and hitting the button **Search By Checked Criteria**.

- If **AND Search** is selected, a result will be returned that matches **ALL** of the selected criteria.
- If **OR Search** is selected, a result will be returned that matches **ANY** of the selected criteria.

Search Type: ● AND Search ○ OR Search	
Image Type: ● all ○ public images.	
Results displayed as: ● Text Only ○ Text and Thumbnail images.	
☑ **Category**	Hyperspectral
□ **Subject**	
□ **Image Format**	ASCII
☑ **After the Date**	January 1 2000
□ **Before the Date**	January 1 1997
□ **Instrument:**	9: Hyperspectral Imager (null)
□ **Researcher:**	Tara Chauhan
□ **Laboratory:**	Center for Electromagnetic Research, Northeastern University
□ **Institution:**	Boston University
□ **Sensor Type:**	Electromagnetic

SEARCH BY CHECKED CRITERIA

Welcome to the CenSSIS Image Database

Home / Search for an Image / Submit Image(s) / Update Image(s) / Help

Not logged in.

Search Results 1 through 5 of 5

1 T-pass Data (Hyperspectral)
2 T-pass Data (Hyperspectral)
3 T-Pass (Hyperspectral)
4 T-Pass (Hyperspectral)
5 T-Pass, intensity images (Hyperspectral)

Last Edited: April 14, 2000 at 7:52am
Contact: bnorum@ece.neu.edu

Figure 5. An example of a complex query and its results.

4.2.1. ID Search

The simplest type of search is based upon the image id. Every image in the database is assigned a unique identification number upon submittal. This id can be used for later retrieval.

4.2.2. Complex Queries

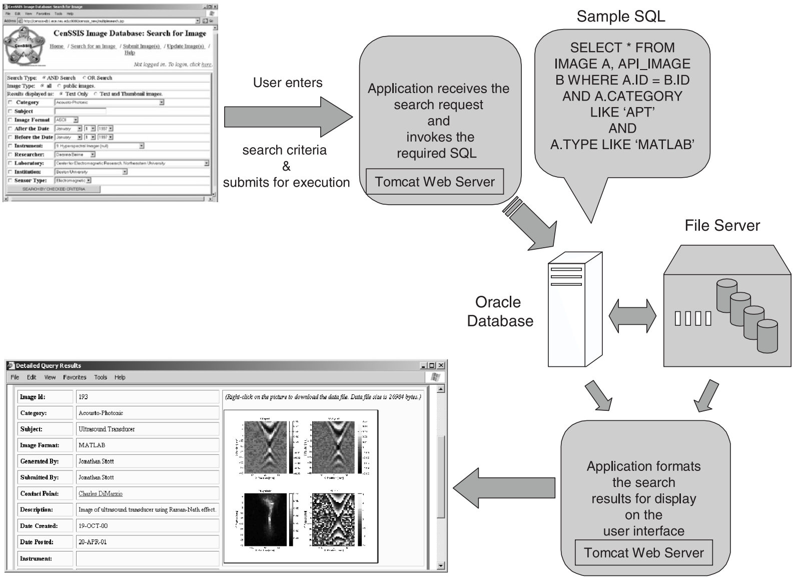
Clients can select multiple entries from a list of metadata, and enter criteria to search upon. Searches can be executed as AND or OR queries to return all of the selected criteria or any of the selected criteria, respectively. Figure 5 shows the execution of a complex query and its results. In addition, a textual search is available to search upon keyword and description fields. Figure 6 shows a more detailed flow of processing a complex query in our 3-tiered architecture.

4.2.3. Hierarchical View

Clients can select a data set as a root element and be given a tree presentation of all of its child nodes. This presentation can be expanded and reduced upon request. This is a way to present data sets in a way that is convenient for the client and easily navigable.

4.3. Data Submission

Data submission is a critical challenge for the CenSSIS-DB. The nature of the user interface (world wide web) places limitations on the design of submission forms. At present, the client can choose to create a new data



CenSSIS Image Database: Search for Image

Home / Search for an Image / Submit Image(s) / Update Image(s) / Help

Not logged in. To login, click here.

Search Type: ● AND Search ○ OR Search
Image Type: ● all ○ public images.
Results displayed as: ● Text Only ○ Text and Thumbnail images.

□ Category	Acousto-Photonic
□ Subject	
□ Image Format	ASCII
□ After the Date	January 1 1997
□ Before the Date	January 1 1997
□ Instrument:	9. Hyperspectral Imager (null)
□ Researcher:	Deanna Beirne
□ Laboratory:	Center for Electromagnetic Research, Northeastern University
□ Institution:	Boston University
□ Sensor Type:	Electromagnetic

SEARCH BY CHECKED CRITERIA

Detailed Query Results

Image Id:	193
Category:	Acousto-Photonic
Subject:	Ultrasound Transducer
Image Format:	MATLAB
Generated By:	Jonathan Stott
Submitted By:	Jonathan Stott
Contact Point:	Charles DiMarzio
Description:	Image of ultrasound transducer using Raman-Nath effect.
Date Created:	19-OCT-00
Date Posted:	20-APR-01
Instrument:	

(Right-click on the picture to download the data file. Data file size is 26984 bytes.)

User enters
search criteria
&
submits for execution
Application receives the search request and invokes the required SQL
Tomcat Web Server
Sample SQL
SELECT * FROM IMAGE A, API_IMAGE B WHERE A.ID = B.ID AND A.CATEGORY LIKE 'APT' AND A.TYPE LIKE 'MATLAB'
Oracle Database
File Server
Application formats the search results for display on the user interface
Tomcat Web Server
File Edit View Favorites Tools Help

Figure 6. The processing of a complex query, showing the flow through our 3-tiered architecture.

collection or add data to an already existing collection. First, the client selects the category of data to be submitted. The browser redirects the client to a form generated to address the category-specific metadata requirements. The client fills out the web form with the appropriate metadata, attaches relevant file(s) and submits the request.

The metadata goes through a quality check based upon expected values. If there is a problem, the browser redisplays the client's form and asks the client to correct the errors. If there are no errors, the metadata and a pointer to the corresponding files are stored in the database and the attached file(s) are stored on the file server.

CenSSIS-DB also allows users to upload multiple data files. If a data set is comprised of multiple data files, the client is prompted to group files together into a single file, and then upload the aggregate file along with the data set metadata.

Upon submission, a data set is available for retrieval immediately. However, it is associated with a conditional tag to alert clients that the data set has not yet been approved and may have problems. A system administrator approves the data submission. Once approved, the conditional tag is removed and the data set is available unconditionally.

The approval process was implemented in order to meet two conflicting system requirements. First, clients need to be able to share very large image data sets quickly. These files are often too large for standard email servers; CenSSIS-DB provides an online repository to facilitate sharing. However, in order to ensure system integrity, we need to have control over and be able to remove inappropriate submissions. Therefore, we designed the above approval process that provides the maximum flexibility and security.

4.4. File Server

The binary data files could be stored in the database itself or in a separate file system. There were several compelling reasons to store them in a separate file system, with links to the data stored with the descriptive metadata. First, storage of binary data is not standardized across relational database systems. Second, in many cases, we are talking about storing relatively large amounts of binary data (e.g., 100s of megabytes). A file server is more reliable. Third, by keeping the image data in binary on a file server, it remains easily accessible to other tools that need to manipulate the data. Fourth, if the metadata is isolated from the binary data, the size of the file containing the metadata is smaller and searches will be more efficient than they would be if the metadata were interleaved with large binary images.

4.5. Security

The security of CenSSIS-DB is of special concern because it is world wide web accessible. The search and retrieval areas of the system are publically accessible. Some CenSSIS clients, however, need to restrict access to their data sets. Not only do we need to restrict access of particular data sets; but also we want to be able to provide restricted access for the submission of data in order to minimize the need to curate data. For this reason, a client must select an access permission level when submitting a data set.

We have identified an initial set of groups to be part of CenSSIS-DB:

- Public—anyone client with a web browser,
- CenSSIS—registered CenSSIS users,
- Client—a registered client, and
- Group—a predefined group of users.

CenSSIS members must register for a username and password in order to access restricted areas. Registered users can create or join groups. Registered users can also submit new data, update data, and manage instrument profiles. This functionality allows CenSSIS members to create online communities where they can share privileged information, for example, within a particular research group.

5. Example Applications

This section discusses one application which will make extensive use of the CenSSIS-DB. Biologists and engineers at Northeastern University are investigating embryo viability and plan to use the CenSSIS-DB to facilitate their work [7]. The embryo viability research focuses on the genetic control of early mammalian development and why embryos are not rejected by the maternal immune system. A long-term goal of this work is to understand the genetic mechanisms of early development in order to contribute to design of new methods for contraception and the treatment of infertility.

Images of oocytes and embryos will be taken at varying time intervals and features will be compared. We expect that more than 50 GB of image data will be produced every year. In addition, some of these images will be sent to other research organizations for processing and statistical analysis.

Currently, embryo images are stored on CD-ROMs and metadata is recorded by hand in laboratory notebooks. However, the rapid increase in data acquisition rates requires investigation of alternative methods. A centralized image database system such as CenSSIS-DB will provide an improved data storage and retrieval system. CenSSIS-DB provides efficient storage, searching, and exchange of large quantities of both raw and processed image and sensor data. The CenSSIS-DB hierarchical data schema provides a good method of modeling scientific experiments and facilitates searching/browsing and storage.

We have developed a custom data type for this group. The new data type includes information about the developmental stage of the embryo, microscope staining information, and drug treatment data. In addition, we added a flexible data field that can be used for storing miscellaneous data specific to an experiment. These new fields can be queried as well as the initial set of metadata, thus providing a superior design for efficient organization and searching. An example oocyte record is shown in Figure 7.

In addition, we are currently developing a tool that this group and others can use to tag areas of an image. These tags will allow researchers to identify interesting regions of an image and flag them for future reference. In addition, the tags will be stored, indexed and queried as part of the metadata schema, enhancing the query capabilities of the system.

6. Future Research

Future enhancements proposed for CenSSIS-DB include content-based indexing and retrieval (CBIR), multidimensional database indexing and content-based image tagging and searching. We will develop new tools to ensure seamless image format interchange and develop an advanced graphical interface to allow researchers to annotate and query parts of any image.

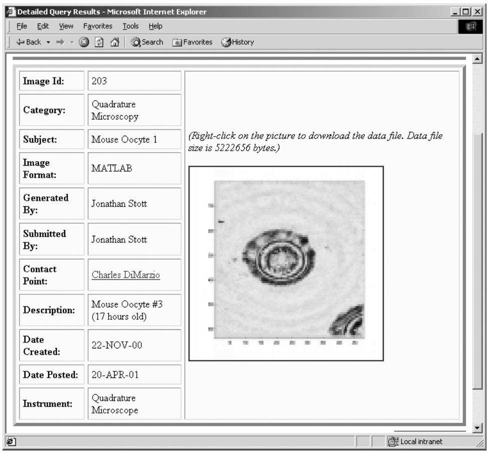

Detailed Query Results - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back
Search
Favorites
History
Local intranet

Image Id:	203
Category:	Quadrature Microscopy
Subject:	Mouse Oocyte 1
Image Format:	MATLAB
Generated By:	Jonathan Stott
Submitted By:	Jonathan Stott
Contact Point:	Charles DiMarzio
Description:	Mouse Oocyte #3 (17 hours old)
Date Created:	22-NOV-00
Date Posted:	20-APR-01
Instrument:	Quadrature Microscope

(Right-click on the picture to download the data file. Data file size is 5222656 bytes.)

Figure 7. Example of an oocyte image and its associated metadata.

CBIR is the ability to query images based upon primitive characteristics (shape, color) and logical characteristics (object identity) [3]. It has great import for clinicians and researchers, aiding the clinicians in diagnostic assessment, therapeutic approach and prognostic evaluation and providing researchers with enhanced means of gathering, assessing and collating relevant clinical data.

Other future work is to implement a web-based visualization tool. Many of our data sets involve 2-dimensional raw images that are reconstructed into a more meaningful 3-dimensional image. We plan to provide the client with the capability of performing and viewing that reconstruction online.

We will also continue to collaborate with other CenSSIS members to broaden the scope of our data collection. In addition, we plan to add the functionality of mass updates to the system in order to permit batch submissions of images and data. The distributed architecture of our application will make it relatively simple to connect from other networked client applications.

7. Conclusion

The CenSSIS-DB system provides for simple uploading and downloading of images and their associated metadata. It provides comprehensive searching capabilities through a web browser. The system adopts a three-tier client/server model according to J2EE specifications.

Our major contribution is providing a common model for scientific data sharing. We worked closely with biological researchers and CenSSIS engineers to develop a flexible and extensible system to meet current and future research storage requirements. The system is easily modifiable for the inclusion of new data types. Our user interface is friendly and easy to use.

In conclusion, we have successfully built a web-based database system for the storage, organization, and retrieval of subsurface sensing and imaging data. We have a flexible and easily extensible framework in place to handle the complex and diverse metadata of different categories of images. We have strong security measures in place in order to protect privileged information.

Acknowledgments

This work was supported by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the NSF (Award Number EEC-9986821).

References

1. Butler, G., Bornberg-Bauer, E., Grahne, G., Kurfess, F., Lam, C., Paquet, J., Shinghal, I.R.R., Tao, L., and Tsang, A., 2000, The BioIT Projects: internet, database and software technology applied to bioinformatics: Proc. of SSGRR'2000, Int. Conf. on Advances in Infrastructure for Electronic Business, Science and Education on the Internet.
2. Cell Centered Database (CCDB) at the National Center for Microscopy and Imaging Research (NCMIR), http://pamina2.sdsc.edu/CCDB/.
3. Eakens, J. and Graham, M., ■. Content-based image retrieval: A report to the JISC Technology Applications Programme: Northumbria Image Data Research Institute, http://www.unn.ac.uk/iidr/report.html. TQ1
4. European Computerized Human Brain Database (ECHBD), http://fornix.neuro.ki.se/ECHBD/Database/.
5. Fredriksson, J., Roland, P., and Svensson, P., 1999, Rationale and design of the European computerized human brain database system: Proc. of the 11th Int. Conf. on Scientific and Statistical Database Management, p. 148–157.
6. Gray, J., Szalay, A.S., Thakar, A., Stoughton, C., and van den Berg, J., 2002, Online scientific data curation, publication, and archiving: SIPE Astronomy Telescopes and Instruments.
7. Hardy, K., Warner, A., Winston, R., and Becker, D., 1996, Expression of intercellular junctions during preimplantation development of the human embryo: Mol. Human Reproduction, v. 2, no. 8, p. 621–632.
8. Janee, G. and Frew, J., 2002, The ADEPT Digital Library Architecture: JCDL, p. 342–350.
9. Leow, R. and Taylor, K., 2000, Efficient web access to distributed biological collections using a taxonomy browser: Proc. of the 12th Int. Conf. on Scientific and Statistical Database Management, p. 25–38.
10. Marshall, H., Meyer, J., Helms, C., and Donis-Keller, H., 1997, St. Louis Breast Cancer Tissue Registry and the Cooperative Breast Cancer Tissue Registry (CBCTR): a combined

resource for marker validation studies: Proc. of the 47th Annual Meeting of the Amer. Soc. of Human Genetics.

11. National Center for Biotechnology Information (NCBI), http://www.ncbi.nlm.nih.gov.
12. Parvin, B., Cong, G., Fontenay, G., Taylor, J.R., Henshaland, R., and Barcellos-Hoff, M.H., 2000, BioSig: a bioinformatic system for studying the mechanism of intra-cell signaling: Proc. of BIBE 2000, p. 281–288.
13. Parvin, B., Yang, Q., Fontenay, G., and Barcellos-Hoff, M.H., 2002, BioSig, an imaging bioinformatic system for studying phenomic: IEEE Computer, v. 35, no. 7, p. 65–71.
14. SIMBAD Astronomical Database, http://simbad.u-strasbg.fr/.
15. Smith, T.R., Janee, G., Frew, J., and Coleman, A., 2001, The Alexandria digital earth prototype: JCDL, p. 118–119.
16. Szalay, A., Gray, J., Thakar, A., Kuntz, P., Malik, T., Raddick, J., Stoughton, C., and Vandenberg, J., 2002, The SDSS SkyServer—Public Access to the Sloan Digital Sky Server Data: ACM SIGMOD.
17. Yang, R., Deng, X., Kafatos, M., Wang, C., and Wang, X., 2001, An XML-based Distributed Metadata Server (DIMES) supporting Earth science metadata: Proc. of the 13th Int. Conf. on Scientific and Statistical Database Management, p. 251–256.

Author Queries:

1. Please provide year of publication.
2. Please check if figures 1–7 are cited correctly.