

## **Perspectiva CTS en el desarrollo del sistema Mercado de Datos para el estudio y análisis de las estructuras de proteínas**

**CTS Perspective on the development of Data Mart for study and analysis of structure of proteins.**

**Leiny Ruíz Cervera<sup>1</sup>, Ariel Labrada Delgado<sup>2</sup>, Ernesto Camilo González Casanova<sup>3</sup>, Adalennis Buchillón Soris<sup>4</sup>,**

<sup>1</sup> Centro de Geoinformática y Señales Digitales. Departamento de Geoinformática, facultad 6. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños km 2 1/2, La Lisa, La Habana, Cuba, CP. 19370. [lcervera@uci.cu](mailto:lcervera@uci.cu).

<sup>2</sup> Centro de Geoinformática y Señales Digitales. Departamento de Geoinformática, facultad 6. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños km 2 1/2, La Lisa, La Habana, Cuba, CP. 19370. [alabradad@uci.cu](mailto:alabradad@uci.cu).

<sup>3</sup> Centro DATEC. Departamento de Almacenes de datos, facultad 6. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños km 2 1/2, La Lisa, La Habana, Cuba, CP. 19370. [abuchillon@uci.cu](mailto:abuchillon@uci.cu)

<sup>4</sup> Empresa XETID. Proyecto Genesig. Carretera a San Antonio de los Baños km 2 1/2, La Lisa, La Habana, Cuba, CP. 19370. [eccasanovag@uci.cu](mailto:eccasanovag@uci.cu).

---

### **Resumen:**

El presente trabajo aborda las principales experiencias de la incorporación de los estudios en Ciencia, Tecnología y Sociedad en el desarrollo del sistema informático “Mercado de datos para el estudio y análisis de estructuras de proteínas”, software implementado con el objetivo de desarrollar un mercado de datos que forme parte de la Plataforma de Servicios Bioinformáticos de la UCI para facilitar el acceso y búsqueda de información de los datos biológicos presentes en el Banco de Datos de Proteínas (PDB). En la investigación se identifican aspectos importantes, basados en la experiencia, que muestran a los profesionales de la Universidad de las Ciencias Informáticas (UCI) la importancia de la relación entre Ciencia, Tecnología y Sociedad presente en el desarrollo de aplicaciones informáticas.

**Palabras claves:** CTS, Mercado de datos, PDB, Bioinformática, Proteínas.

### **Abstract:**

This paper discusses the main lessons of the incorporation of the studies in Science, Technology and Society in the development of computer system “Data Mart for the

study and analysis of protein structures”, software implemented with the aim of developing a data mart part of the Service Platform bioinformatic UCI to facilitate access and search for information from biological data present in the Protein Data Bank (PDB). The research identified important aspects, based on experience, that show professional at the University of Information Sciences (UCI) the importance of the relationship between science, technology and society present in the development of computer applications.

**Keywords:** *STS, Data Mart, PDB, Bioinformatics, Proteins.*

## **Introducción**

El extraordinario progreso de las tecnologías ha sido un paso de avance en el desarrollo del conocimiento del ser humano. Hoy, cuando se habla de avances tecnológicos, no se puede dejar atrás el desarrollo que se ha ido adquiriendo en el campo de la Informática, que integrada a ciencias básicas como la matemática, la física, la biología y las tecnologías de la información, permite el surgimiento de una nueva ciencia emergente denominada Bioinformática. Las metas fundamentales de la Bioinformática son la predicción de la estructura tridimensional de las proteínas a partir de su secuencia, la predicción de las funciones biológicas y biofísicas a partir de la secuencia o la estructura, así como simular el metabolismo y otros procesos biológicos basados en esas funciones [1]. Esta ciencia en sus comienzos se restringía prácticamente a la creación y mantenimiento de bases de datos para almacenar información biológica. El desarrollo de estos tipos de bases de datos implicaba el diseño de interfaces complejas que permitieran el acceso a la información, la adición de los nuevos datos, la revisión y la actualización de los ya existentes. Estas tareas siguen siendo un objetivo fundamental de la Bioinformática, pero la prioridad actual es el desarrollo de herramientas computacionales que permitan el análisis e interpretación de este gran volumen de información [2]. En los últimos años se han creado un inmenso número de aplicaciones con fines de este tipo, entre los que se puede mencionar RASMOL [3] que es una herramienta científica para la visualización de las proteínas y además el sitio oficial del Protein Data Bank (PDB) que surge debido a una pequeña pero creciente colección de datos de estructura de proteínas para la estandarización de los archivos. El desarrollo sin precedente de este nuevo campo, se ha convertido en una realidad científica, productiva y económica, no solo para Cuba sino también para otros países subdesarrollados e incluso desarrollados. Varias instituciones en el país comienzan a realizar investigaciones con los datos presentes en los archivos en formato PDB, entre los que se encuentra la UCI, entre sus proyectos se encuentra el desarrollo del portal de Servicios Bioinformáticos de la UCI, en el cual se trabaja con las informaciones, que se registran referentes a las moléculas, que se encuentran presentes en el PDB,

estos datos están almacenados en más de ochenta mil archivos de texto, uno por cada estructura molecular. Dichas informaciones se determinan por difracción de rayos X o espectroscopia de resonancia magnética nuclear (RMN) y son de vital importancia para la creación de nuevos fármacos y entender la interacción entre diferentes proteínas.

El portal web presenta varias funcionalidades que son utilizadas para el estudio de las estructuras moleculares, pero carece de funcionalidades que permitan realizar estudios y análisis estadísticos de los datos de proteínas presente en el PDB. En la actualidad los investigadores realizan este tipo de estudio utilizando diversas aplicaciones tales como: RASMOL [3] y el sitio oficial del PDB [4], dichas aplicaciones contienen todas sus funcionalidades predeterminadas. También los investigadores pueden desarrollar programas informáticos que les permitan obtener resultados adicionales, a los que no pueden acceder utilizando las aplicaciones. El uso de este portal permite el acceso a los investigadores de todo el país a los servicios que se brindan, mediante la red nacional, lo que elimina la dependencia del uso de la Internet y además posibilita adicionar nuevas funcionalidades según el interés de cada usuario.

Teniendo en cuenta lo antes mencionado, surgen las interrogantes:

- ✓ ¿Cuáles son los impactos del sistema: Mercado de datos para el estudio y análisis de estructuras de proteínas?
- ✓ ¿Cuáles son los problemas sociales y éticos que pueden desencadenarse con la utilización del sistema?
- ✓ ¿Es suficiente la tecnología para mejorar la gestión de los datos presentes en los archivos PDB?

Siendo el conocimiento la materia prima de la sociedad, la investigación se presenta como una invitación a comprender la naturaleza de la actividad científica y tecnológica generadoras de conocimiento y a forjarse una visión general de su impacto en la sociedad. En el presente trabajo, se da respuesta a cada una de las interrogantes planteadas, teniendo como objetivo principal promover la apropiación de los estudios en Ciencia, Tecnología y Sociedad a partir de una presentación de las experiencias más significativas en el desarrollo del sistema: "Mercado de datos para el estudio y análisis de estructuras de proteínas".

### **Valoración del impacto de las tecnologías en el desarrollo del sistema.**

La utilización del sistema incide positivamente en la calidad de las investigaciones de los datos biológicos de proteínas ya que los investigadores pueden realizar profundos estudios que contribuyen a la salud del hombre, pudiendo realizar consultas simples y complejas de los datos, analizar y visualizar los resultados estadísticos.

El uso de este tipo de sistema permite reducir notablemente el número de errores en los archivos de proteínas, ya que se pueden tener reglas de validación incorporadas que eviten o hagan más improbables los errores y las omisiones en el momento en que el investigador facilita los datos. Esta es una de las virtudes fundamentales ya que asegura la calidad de los datos registrados en la base de datos en mayor medida de lo que se consigue cuando se trabaja en papel.

Con la obtención del mercado de datos que contiene información de las estructuras de proteínas presentes en el PDB se facilita la búsqueda de informaciones estadísticas a los investigadores del país que podrán acceder al mercado mediante el portal de Servicios Bioinformáticos de la Universidad de Ciencias Informática.

Las principales ventajas que ofrece el uso de la tecnología “Mercado de datos para el estudio y análisis de las estructuras de proteínas” a las organizaciones que gestionan este tipo de información, además de una mayor calidad de los datos es contar con la reducción del tiempo y el costo para obtener la información, así como el fácil acceso a los datos a través de la red nacional sin tener que prescindir de la conexión a internet. El sistema garantizará:

- ✓ Respuesta rápida a problemas que en el modo tradicional tardan semanas y meses en conocerse y, en consecuencia, más tiempo en solucionarse.
- ✓ Facilitar la comunicación entre todos los centros que participen en la obtención de análisis estadísticos de los datos de estructuras de proteínas.
- ✓ Facilitar y agilizar el proceso en la obtención de los resultados.
- ✓ Proveer mayor seguridad a la información.
- ✓ Permitir un acceso inmediato y actualizado de los datos.

### **Problemas sociales y éticos que pueden desencadenarse con la utilización del sistema.**

Cuba, por ser un país bloqueado, posee aún más restricciones, por lo que existe la posibilidad de que algunas de las instituciones que realicen estudios con los datos de estructura de proteínas no puedan acceder al PDB, ya que se requiere de la conexión a Internet.

Contar con un sistema propio, permite que el sistema se integre de manera orgánica con los esfuerzos que hoy realiza el país en la informatización y lograr así la soberanía tecnológica que se quiere. El sistema posibilitará la adquisición de un software libre, sin restricciones y garantizando a su vez soporte o mantenimiento sin incurrir en grandes gastos por concepto de licencias o servicios.

La utilización del sistema logrará un impacto positivo en la aceleración del desarrollo de varios productos y, por ende, un aumento de la competitividad de las instituciones como empresas biotecnológicas. Al contribuir al desarrollo más certero, con una mayor validez en la investigación, permitirá a los especialistas estimar con mayor veracidad el impacto de las estructuras de proteínas, acelerando la obtención de fármacos en un menor tiempo.

### **Circunstancias que limitan la utilización de los sistemas de gestión.**

A pesar de las magníficas credenciales que hacen de los sistemas de Mercados, tecnologías altamente útiles para cualquier institución, existen diversas circunstancias que dificultan su más amplia difusión entre todas las organizaciones y especialistas, entre las que se destacan:

- ✓ Problemáticas técnicas: el ancho de banda disponible es insuficiente para acceder con rapidez al Portal de servicios Bioinformáticos.
- ✓ Falta de formación: la necesidad de conocimientos teóricos y prácticos en el uso de la nueva tecnología. Uno de los factores para el éxito de este tipo de software es la formación del personal en la utilización del sistema, permitiendo que su labor sea más fácil y efectiva. Es necesario lograr una alfabetización en el uso de las TIC, logrando actitudes favorables en la utilización de las nuevas herramientas.
- ✓ Barreras económicas: el rápido proceso de deterioro de las computadoras y la necesidad de ir mejorando las condiciones técnicas para lograr mejor tiempo de respuestas de los sistemas, aconseja la renovación de estos equipos.

### **Conclusiones**

En el presente trabajo se logró, a partir de la presentación de las experiencias más significativas de la incorporación de los estudios CTS en el desarrollo del sistema “Mercado de datos para el estudio y análisis de estructuras de proteínas” identificar elementos importantes que sirven para entender la importancia de la relación entre Ciencia, Tecnología y Sociedad.

Finalizada la investigación y cumplidos los objetivos planteados para la misma se arriba a las siguientes conclusiones:

- ✓ La Universidad de las Ciencias Informáticas cuenta con recursos necesarios para formar a los profesionales no solo en la implementación de aplicaciones sino también para enfrentar proyectos informáticos donde predominen enfoques CTS en las investigaciones y desarrollo de software.
- ✓ No sólo debe valorarse si un software cumple con el objetivo trazado, sino que también debe tenerse en cuenta sus impactos en la organización, la economía y la sociedad.

## Referencias

1. Febles Rodríguez, Juan Pedro y González Pérez, Abel. Aplicación de la minería de datos en la Bioinformática. [En línea] [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-943520020002000003#cargo](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-943520020002000003#cargo).
2. Facultad de biología. [En línea] [Citado el: 5 de marzo de 2013.] <http://fbio.uh.cu/sites/bioinfo/index.html>.
3. Home Page for RasMol and OpenRasMol. [En línea] [Citado el: 5 de marzo de 2013.] [http://translate.google.com/cu/translate?hl=es&sl=en&u=http://ras-mol.org/&ei=eF\\_AT8fnNdL3ggfbqo3HCQ&sa=X&oi=translate&ct=result&resnum=2&ved=0CGYQ7gEwAQ&prev=/search%3Fq%3Drasmol%26hl%3Des%26biw%3D1366%26bih](http://translate.google.com/cu/translate?hl=es&sl=en&u=http://ras-mol.org/&ei=eF_AT8fnNdL3ggfbqo3HCQ&sa=X&oi=translate&ct=result&resnum=2&ved=0CGYQ7gEwAQ&prev=/search%3Fq%3Drasmol%26hl%3Des%26biw%3D1366%26bih).
4. WorlWide Protein Data Bank [En línea ] [Citado el: 15 de marzo de 2013.] <http://www.wwpdb.org>.
5. GARCÍA, E. M. (2001): Ciencia, Tecnología y Sociedad: una aproximación conceptual.
6. LIBERA, B. E.: Impacto, impacto social y evaluación del impacto.
7. JOVER, J. N. (2003): La Ciencia y la Tecnología como procesos sociales. Lo que la educación científica no debería olvidar, La Habana.
8. JOVER, J. N.: Ética, Ciencia y tecnología: sobre la función social de la tecnología.
9. MARQUÈS GRAELLS, P. (2008): Las TIC y sus aportaciones a la Sociedad.



[www.sociedadelainformacion.com](http://www.sociedadelainformacion.com)

Edita:



Director: José Ángel Ruiz Felipe  
Jefe de publicaciones: Antero Soria Luján  
D.L.: AB 293-2001  
ISSN: 1578-326x