

## **Introducción a la informática de la biodiversidad**

### **Programa**

#### **Contenidos teóricos y metodológicos**

##### **Introducción general**

Tipos de datos de biodiversidad: Taxonómicos, geográficos, ambientales, cronológicos, anatómicos, desarrollo, parentescos.

Origen de datos de biodiversidad: Colecciones biológicas, observaciones, multimedia.

Datos y metadatos. Especificaciones utilizadas en biodiversidad y disciplinas relacionadas.

##### **Introducción a las bases de datos**

Concepto de normalización e integridad referencial. El modelo entidad-relación.

Bases relacionales vs. bases orientadas a objetos, estado del arte.

Las bases de datos heterogéneas distribuidas, su aplicación a la biodiversidad.

Repositorios de datos.

##### **Introducción a los metadatos**

Datos y metadatos.

Especificaciones utilizadas en biodiversidad y disciplinas relacionadas.

Darwin Core: Versiones e implementaciones.

##### **Metadatos de conjuntos de datos**

Catálogos de metadatos de recursos de biodiversidad.

##### **Metadatos de ubicación geoespacial**

Método punto-radio.

Protocolos de georreferenciación.

##### **Metadatos de imágenes biológicas**

Metadatos embebidos.

Vistas estandarizadas.

Repositorios.

Digitalización de ejemplares mediante imágenes.

##### **Datos taxonómicos**

Catálogos de autoridad taxonómica.

Agregadores de datos taxonómicos.

Estimas de diversidad con taxonomía incompleta.

##### **Identificadores únicos**

Importancia y problemas de identificadores únicos.

URIs, LSIDs.

##### **Ontologías biológicas**

Conceptos, relaciones y razonamientos automáticos.

Web semántica, repositorios RDF.

Anotaciones utilizando ontologías.

##### **Herramientas de adquisición de datos de colecciones biológicas**

Funciones y problemática general.

Modelos de datos usuales.



### **Proveedores de datos**

Función y tecnología empleada.  
Proveedores TAPIR, IPT

### **Portales de datos**

Función y tecnología empleada.  
Portales de iniciativas de biodiversidad.

### **Aplicaciones de datos de biodiversidad**

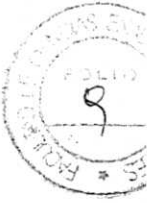
Cybertaxonomía.  
Identificación taxonómica automática: *DNA Barcodes* y sistemas basados en imágenes.  
Modelado de distribuciones predictivas.  
Áreas de endemismo.  
Inventarios y estimas de riqueza.

### **Trabajos prácticos**

Búsquedas y procesamiento de datos de portales: GBIF, SNDB, BOLD, Morphbank.  
Consultas en bases de datos relacionales.  
Mapeo de campos de colecciones biológicas a Darwin Core.  
Validación de colecciones de datos. Problemas generales, métodos manuales y automáticos.  
Georreferenciación de localidades geográficas.  
Uso de servicios web.

### **Bibliografía general**

- Chapman, A. D. 2005. Uses of Primary Species-Occurrence Data, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen.
- Chapman, A. D. 2005. Principles of Data Quality, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen.
- Chapman, A. D. 2005. Principles and Methods of Data Cleaning – Primary Species and Species-Occurrence Data, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen.
- Date, C.J. 2001. Introducción a los sistemas de bases de datos, 7ª edición. Prentice Hall.
- Döring, M. & R. De Giovanni. 2004. GBIF Data Access and Database Interoperability: A unified protocol for search and retrieval of distributed data. Report for the Global Biodiversity Information Facility, Copenhagen.
- Funk, V.A., P.C. Hoch, L.A. Prather & W.L. Wagner. 2005. The importance of vouchers. *Taxon*, 54: 127-129.
- Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.*, 19, 497-503.
- Gwinn, N.E. & C. Rinaldo. 2009. The Biodiversity Heritage Library: sharing biodiversity literature with the world. *IFLA Journal*, 35: 25-34.
- Hebert, P. D., A. Cywinska, S. L. Ball y J. R. deWard. 2002. Biological identifications through DNA barcodes. *Proc Biol Sci.* 270:313-21.
- Jones, M.B., Schildhauer, M.P., Reichman, O.J. and Bowers, S. (2006) The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology Evolution and Systematics* 37, 519-544.
- Korth, H.F., A. Silberschatz & S. Sudarshan. 2006. Fundamentos de bases de datos, 5ª edición. McGraw-Hill.



- Monk, R.R., y R. J. Baker. 2001. e-vouchers and the use of digital imagery in natural history collections. *Museology, Museum of Texas Tech University*, 10: 1-11.
- Page, R. D. M. 2005 A Taxonomic Search Engine: Federating taxonomic databases using web services. *BMC Bioinformatics*, 6:48, 8pp. (leído en seminarios 2008).
- Patterson, D.J., D. Remsen, W.A. Marino & C. Norton. 2006. Taxonomic Indexing—Extending the Role of Taxonomy. *Systematic Biology*, 55:367–373.
- Patterson, D.J., S. Faulwetter & A. Shipunov. 2008. Principles for a names-based cyberinfrastructure to serve all of biology. *Zootaxa*, 1950: 153–163.
- Penev, L., T. Erwin, J. Miller, V. Chavan, T. Moritz, C. Griswold. 2009. Publication and dissemination of datasets in taxonomy: ZooKeys working example. *ZooKeys* 11: 1-8. doi: 10.3897/zookeys.11.210.
- Pyke, G.H. & Ehrlich, P.R. 2010. Biological collections and ecological/environmental research: a review, some observations and a look to the future. *Biological Reviews*, 85: 247–266.
- Pyle, R. y E. Michel. 2008. ZooBank: Developing a nomenclatural tool for unifying 250 years of biological information. *Zootaxa* 1950: 39–50.
- Ramírez, M. J., J. A. Coddington, W. P. Maddison, P. E. Midford, L. Prendini, J. Miller, C. E. Griswold, G. Hormiga, P. Sierwald, N. Scharff, S. P. Benjamin, W. C. Wheeler. 2007. Linking of digital images to phylogenetic data matrices using a morphological ontology. *Systematic Biology*, 56, 283–294.
- Ratnasingham, S. & P.D.N. Hebert. 2007. BOLD: The Barcode of Life Data System ([www.barcodinglife.org](http://www.barcodinglife.org)). *Molecular Ecology Notes* 7: 355–364.
- Schlick-Steiner, B.C., F.M. Steiner, B. Seifert, C. Stauffer, E. Christian y R.H. Crozier. 2009. Integrative Taxonomy: A Multisource Approach to Exploring Biodiversity. *Annual Review of Entomology*, 55:421–38.
- Shotton, D., K. Portwin, G. Klyne & A. Miles. 2009. Adventures in semantic publishing: exemplar semantic enhancements of a research article. *PLoS Computational Biology* 5:e1000361.
- Smith, V., S.D. Rycroft, K.T. Harman, B. Scott & D. Roberts. 2009. Scratchpads: a data-publishing framework to build, share and manage information on the diversity of life. *BMC Bioinformatics* 2009, 10:S.
- Soberón, J.S. & T. Peterson. 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society of London B* 359, 689–698.
- Soberón, J.S., J. Golubov y J.S. Sarukhan. 2001. The importance of *Opuntia* in Mexico and routes of invasion and impact of *Cactoblastis cactorum* (Lepidoptera: Pyralidae). *Florida Entomologist*, 84: 486-492.
- Stein, B. and Wieczorek, J. (2004) Mammals of the World: MaNIS an example of data integration in a distributed network environment. *Biodiversity Informatics*, 1, 14-22.
- Szumik, C., F. Cuezco, P. A. Goloboff, A. E. Chalup. 2002. An optimality criterion to determine areas of endemism. *Systematic Biology* 51:806–816.
- Wheeler, Q. D. (editor). 2008. *The New Taxonomy*. The Systematics Association Special Volume Series 76. CRC Press, New York. xi + 237 pp.
- Wieczorek, J., Guo Q. y Hijmans, R. 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *Int. J. Geographical Information Science*, 18: 745–767.
- Wilson, E.O. 2003. The encyclopedia of life. *Trends in Ecology & Evolution*, 18: 77-80.

### Links

Introduction to the GBIF portal. <http://data.gbif.org/tutorial/tutorial>  
Distributed Generic Information Retrieval (DiGIR). <http://digir.net/>  
TAPIR (TDWG Access Protocol for Information Retrieval).  
<http://www.tdwg.org/activities/tapir/>  
Morphbank. <http://www.morphbank.net/>



Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales

Ref. Expte. N° 498.257/2010

Buenos Aires, 17 MAY 2010

**VISTO:**

La nota de fecha 04/05/2010, presentada por la Dra. María Busch, Directora del Departamento de Ecología Genética y Evolución, mediante la cual eleva la Información del Curso de Posgrado **INTRODUCCION A LA INFORMATICA DE LA BIODIVERSIDAD**, que será dictado entre el 2 de agosto y el 13 de agosto por el Dr. Martín J. Ramirez con la colaboración de Renato Mazzanti

La nota de la Directora del Departamento de Graduados Dra. Silvia Moreno del 07/05/2010

**CONSIDERANDO:**

lo actuado por la Comisión de Doctorado de esta Facultad el 07/05/2010,  
lo actuado por la Comisión de Enseñanza, Programas, Planes de Estudio y Postgrado,  
lo actuado en la Comisión de Presupuesto y Administración,  
lo actuado por este cuerpo en Sesión Ordinaria realizada en el día de la fecha,  
en uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD  
DE CIENCIAS EXACTAS Y NATURALES  
RESUELVE:**

**Artículo 1°:** Autorizar el dictado del Curso de Posgrado **INTRODUCCION A LA INFORMATICA DE LA BIODIVERSIDAD**, de 60 horas de duración.

**Artículo 2°:** Aprobar el Programa (obrante a fs 5 a 9) del Curso de Posgrado **INTRODUCCION A LA INFORMATICA DE LA BIODIVERSIDAD**.

**Artículo 3°:** Aprobar un Puntaje Máximo de tres (3) puntos para la Carrera del Doctorado.

**Artículo 4°:** Aprobar un Arancel de 80 Módulos. Disponer que los montos recaudados serán utilizados conforme a los dispuesto por Resolución CD N° 072/03.

**Artículo 5°:** Comuníquese a la Dirección del Departamento de Ecología, Genética y Evolución, a la Subsecretaría de Postgrado y a la Biblioteca de la FCEN (con fotocopia del Programa incluida). Comuníquese a la dirección del departamento de Alumnos (sin fotocopia del Programa). Cumplido, archívese

RESOLUCION CD N° 1105 ==  
SP-med-07/05/2010