

Curriculum Vitae
Indra Neil Sarkar

40 River Road – Apt 10 R
New York, NY 10044 USA
Phone: 212 935 6646
Electronic mail: sarkar@mac.com

PERSONAL DATA:

Security Clearance DoD SECRET
Military None
Birthplace/Citizenship US/US

ACADEMIC TRAINING:

<u>From</u>	<u>To</u>	<u>School / Institution</u>	<u>Major, Subject</u>
09 / 1999	09 / 2004	Graduate School of Arts & Sciences College of Physicians & Surgeons <i>Columbia University</i> New York, New York	Ph.D. (2004) Biomedical Informatics M.Phil. (2002) Biomedical Informatics
08 / 1995	05 / 1999	Lyman Briggs School of Science College of Natural Science <i>Michigan State University</i> East Lansing, Michigan	B.Sc. (1999) Microbiology w/Foci in Computer Science & Philosophy of Science
09 / 1991	01 / 1995	Concord-Carlisle Regional High School Concord, Massachusetts	High School Diploma (1995)

CURRENT APPOINTMENTS & AFFILIATIONS:

<u>From</u>	<u>To</u>	<u>School / Institution</u>	<u>Position</u>
09 / 2004	(present)	Divisions of Invertebrate Zoology & Library Services <i>American Museum of Natural History</i> New York, New York	<i>Bioinformatics Associate</i>
11 / 2004	(present)	Department of Biomedical Informatics College of Physicians & Surgeons <i>Columbia University</i> New York, New York	<i>Affiliated Research Scientist</i>

HONORS, DISTINCTIONS, AND AWARDS:

<u>Date</u>	<u>Description</u>
2006	Donald A.B. Lindberg Research Fellowship, Medical Library Association
2005	Guest Editor, Journal of Biomedical Informatics: Special Issue on Phylogenetic Inferencing
2003	International Lipari School Scholarship, Italian Ministry of Research & Universities
2002	Who's Who in America, Marquis Publication Group
2001	Summer Research Fellowship, Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health
1999	Medical Informatics Training Fellowship, National Library of Medicine, National Institutes of Health

PREVIOUS RESEARCH EXPERIENCE:

<u>From</u>	<u>To</u>	<u>School / Institution</u>	<u>Position</u>
09 / 1999	09 / 2004	Department of Biomedical Informatics College of Physicians & Surgeons Columbia University New York, New York	<i>Graduate Research Fellow</i>
05 / 2001	08 / 2001	Lister Hill National Center for Biomedical Communications National Library of Medicine Bethesda, Maryland	<i>Research Assistant</i>
05 / 1996	08 / 1996	Research Computing Division Genetics Institute Cambridge, Massachusetts	<i>Research Intern</i>
09 / 1998	07 / 1999	Department of Physiology Michigan State University East Lansing, Michigan	<i>Research Assistant</i>
09 / 1996	08 / 1998	Department of Microbiology Michigan State University East Lansing, Michigan	<i>Research Assistant</i>
01 / 1995	05 / 1995	Institut Fuer Klinische und Molekulare Virologie Frederik-Alexandr Universitaat Erlangen-Nurenburg Erlangen, Bavaria, Germany	<i>Lab Practicum</i>

PREVIOUS TEACHING EXPERIENCE:

<u>From</u>	<u>To</u>	<u>School / Institution</u>	<u>Position, Subject(s)</u>
09 / 2004	05 / 2006	Downstate Medical Center <i>State University of New York</i> New York, New York	Instructor 1. <i>Introduction to Medical Informatics</i> 2. <i>Clinical Decision Support Systems</i> 3. <i>Internet Integration into Health Care</i>
08 / 2001	05 / 2002	Department of Biomedical Informatics College of Physicians & Surgeons <i>Columbia University</i> New York, New York	Teaching Assistant 1. <i>Introduction to Genomics (2001)</i> 2. <i>Representation & Coding of Medical Data (2002)</i>
01 / 1997	05 / 1999	Lyman Briggs School College of Natural Science <i>Michigan State University</i> East Lansing, Michigan	Teaching Assistant & Student Instructor 1. <i>C++ w/ Biological Applications (1997)</i> 2. <i>Introduction to Computers (1998)</i> 3. <i>Biocomputing (1997-1999)</i>
06 / 1994	09 / 1994	Education Studies Program <i>Massachusetts Institute of Technology</i> Cambridge, Massachusetts	Instructor 1. <i>Chemistry</i> 2. <i>Calculus I(AB)</i>

PROFESSIONAL EXPERIENCE, SERVICE, AND LEADERSHIP:

<u>From</u>	<u>To</u>	<u>School / Institution</u>	<u>Position</u>
06 / 1999	(present)	Board of Directors <i>Lyman Briggs Alumni Association</i> East Lansing, Michigan	Vice-President (2004 – Present) Assistant Treasurer (2003 – 2004) Director (1999 – Present)
01 / 1991	(present)	Board of Directors <i>Mosaic Data Systems</i> Bedford, Massachusetts	Director (1999 – Present) Technical Consultant (1991 – 1997)
02 / 2004	05 / 2006	Education Programs <i>New York Academy of Sciences</i> New York, NY	NYCSEF Judge (2004-2006) JSHS Judge (2005) SRTP Mentor (2005-2006)
09 / 2004	05 / 2006	Townsend High School <i>CUNY - Queens College</i> New York, NY	Mentor (2004 - 2006)
04 / 2002	12 / 2005	Advisory Committee <i>John Wiley InterScience</i> Hoboken, New Jersey	Member
12 / 2002	06 / 2003	Advisory Committee Columbia Speakers Fund GSAS at Columbia University New York, New York	Member

09 / 2000	07 / 2002	Department of Biomedical Informatics College of Physicians & Surgeons Columbia University New York, New York	<i>Student Representative (2000-2002)</i>
09 / 1999	12 / 2003	Graduate Student Advisory Council Columbia University New York, New York	<i>Chair (2002); Ex-officio (2003) Vice-Chair (2001) Member (1999-2002)</i>
05 / 1996	09 / 1999	Board of Directors Interdream Bedford, Massachusetts	<i>President & Co-Founder</i>

SPOKEN LANGUAGES:

<u>Language</u>	<u>Ability</u>
English	Fluent
Bengali	Fluent
French	Competent
German	Basic
Italian	Basic

CERTIFICATIONS AND MISCELLANEOUS TRAINING:

<u>Training / Certification</u>	<u>Year of Initial Certification</u>
GCP - Patient/Clinical	2004
HIPAA	2004
ORCBS Radiological	1996
ORCBS Chemical Hygiene	1996
Biohazard Level II & III	1995
Biohazard Level I	1994

PROFESSIONAL AFFILIATIONS:

<u>From</u>	<u>To</u>	<u>Organization</u>
2006	(present)	Willi Hennig Society
2005	(present)	Life Sciences Society
2005	(present)	Association for Computing Machinery
2005	(present)	Institute of Electrical and Electronics Engineers
2004	(present)	International Society for Infectious Diseases
2000	(present)	International Society for Computational Biologists Member, Conferences Committee (2005-present)
2000	(present)	American Medical Informatics Association
1995	(present)	American Association for the Advancement of Science
1996	2006	New York Academy of Sciences
1996	2001	American Institute of Biological Sciences

FUNDED RESEARCH:

<u>Dates</u>	<u>Funding Agency/Title</u>	<u>Role</u>
03/01/2004 -	National Science Foundation (IIS)	Co-PI
02/28/2007	Development of new digital library applications in the context of a basic ontology for biosystematics information using the literature of entomology (Ants).	
07/01/2006 -	Medical Library Association (DAB Lindberg Fellowship)	PI
06/30/2007	Bridging Biomedical and Biodiversity Knowledge via Taxonomic Information	

PENDING FUNDED RESEARCH:

<u>Dates</u>	<u>Funding Agency/Title</u>	<u>Role</u>
12/01/2006 -	National Institutes of Health (NLM)	PI
11/30/2008	TaxonLink: Bridging Biodiversity and Biomedical Knowledge Through Literature	
12/01/2006 -	National Institutes of Health (NLM)	Co-Inv
11/30/2008	Taxonomic Organization of Biomedical Literature.	
02/01/2007 -	National Institutes of Health (NIGMS)	Co-Inv
01/31/2011	Genetic Transfer in the Evolution of Infectious Disease	
08/01/2006 -	National Science Foundation (CEO)	Co-PI
07/31/2009	Name-Based Tools as Cyberinfrastructure for Biology	
04/01/2007 -	National Science Foundation (DBI)	PI
03/31/2010	OHM: Ontology for Homologous Morphology	

SOFTWARE APPLICATIONS:

1. **Sarkar IN**. *ORFcurator*. (2004)
2. Cantor M, **Sarkar IN**, Lussier YA. *GenesTrace*. (2003)
3. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *Phylogenomenclature*. (2002)
4. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *CAOS*. (2002)

INTELLECTUAL PROPERTY:

1. Lussier YA, **Sarkar IN**, Cantor MN. *In Silico Clinical Genomics*. Provisional Patent: Columbia University Science and Technology Ventures, Health Sciences Division; Priority: November 2002.
2. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. *CAOS: Method and System for Recognizing One or More Characteristic Attributes of One or More Groups of Information Items*. Patent-Pending: PCT/US02/03540; 60/267,972; Priority: February 2001.

PEER-REVIEWED JOURNAL PUBLICATIONS:

1. Perkins SL, **Sarkar IN**, Carter R. "The Phylogeny of Rodent Malaria Parasites: Simultaneous Analysis Across Three Genomes." *Infection, Genetics and Evolution* 2006 (in press).
2. Chiu JC, Lee EK, Egan MG, **Sarkar IN**, Coruzzi GM, DeSalle R. "OrthologID: Automation of Genome Scale Ortholog Identification within a Parsimony Framework." *Bioinformatics* 2006 22(6):699-707.

3. **Sarkar IN**. "Phylogenetics in the Modern Era." *Journal of Biomedical Informatics* Jan/Feb;39(1):3-5. 2006.
4. Planet PJ, **Sarkar IN**. "mLLD: A Tool for Constructing and Analyzing Matrices of Pairwise Character Incongruence Tests." *Bioinformatics* 21:4423-4224. 2005.
5. Koning D, **Sarkar IN**, Moritz TD. "TaxonGrab: Extracting Taxon Names from Text." *Journal of Biodiversity Informatics* 2;79-82. 2005.
6. Rosenfeld JA, **Sarkar IN**, Planet PJ, Figurski DH, DeSalle R. "ORFcurator: Molecular Curation of Genes and Gene Clusters in Prokaryotic Organisms." *Bioinformatics* 20: 3462-3465. 2004.
7. Cantor MN, **Sarkar IN**, Hartel F, Bodenreider O, Lussier YA. "An Evaluation of Hybrid Methods for Matching Biomedical Terminologies: Mapping the Gene Ontology to the UMLS." *Stud Health Technol Inform* 95:62-7. 2003.
8. **Sarkar IN**, Planet PJ, Bael TE, Stanley SE, Siddall M, DeSalle R, Figurski DH. "Characteristic Attributes in Cancer Microarrays." *Journal of Biomedical Informatics* Apr/May;35(2):111-122. 2002.
9. **Sarkar IN**, Thornton J, Planet PJ, Schierwater B, and DeSalle R. "A systematic method for classification of novel homeoboxes." *Molecular Phylogenetics and Evolution* Sep;24(3):388-399. 2002.
10. **Sarkar IN** and Starren JB. "Desiderata for Personal Electronic Messaging in Clinical Systems." *Journal of the American Medical Informatics Association* May/Jun;9:209-216. 2002.
11. Planet PJ, DeSalle R, Sidall ME, Bael T, **Sarkar IN**, Stanley SE. "Systematic Analysis of DNA Microarray Data: Ordering and Interpreting Patterns of Gene Expression." *Genome Research* Jul;11(7):1149-1155. 2001.

PEER-REVIEWED CONFERENCE PROCEEDINGS:

1. **Sarkar IN**, Agrawal A. "Literature Based Discovery of Gene Clusters Using Phylogenetic Methods." *AMIA Annual Meeting*, Washington, DC. 2006 (in press).
2. Cantor M, **Sarkar IN**, Bodenreider O, Lussier YA. "GenesTrace: Phenomic Knowledge Discovery Via Structured Terminology." *Pacific Symposium on Biocomputing*, Big Island, HI: 10:103-114. 2005.
3. **Sarkar IN**, Cantor MN, Gelman R, Hartel F, Lussier YA. "Linking Biomedical Language Information and Knowledge Resources: GO and UMLS." *Pacific Symposium Biocomputing*, Kua'i, HI: 8:427-450. 2003.
4. Lussier YA, **Sarkar IN**, Cantor MN. "An Integrative Model for In-Silico Clinical Genomics Discovery Science." *AMIA Annual Meeting*, Washington, DC: 469-473. 2002.
5. **Sarkar IN** and Rindfleisch TC. "Discovering Protein Similarity Using Natural Language Processing." *AMIA Annual Meeting*, Washington, DC: 677-681. 2002.

CONFERENCE ABSTRACTS:

1. Umejei C, **Sarkar IN**. "zMedline: A Medline-Based Content Management System." *AMIA Annual Meeting*, Washington, DC: 469-473. 2006 (in press).
2. **Sarkar IN**, Remsen DP. "Enabling Biological Knowledge Integration Through Scientific Nomenclature." *International Society for Phylogenetic Nomenclature Conference*, Yale University, New Haven, CT. 2006 (in press).
3. **Sarkar IN**, DeSalle R. "Automated Barcoding Using the Characteristic Attribute Organization System." *DIMACS*, National Museum of Natural History, Paris, France. 2006 (in press).
4. **Sarkar IN**, Planet PJ, DeSalle R. "TEPC: Total Evidence Phylogenetic Correlation of Microbial Phenotypes and Genotypes." *ISMB 2005: 13th Annual International Conference on Intelligent Systems for Molecular Biology*, Detroit, MI. 2005.

5. **Sarkar IN**, Cantor MN, Bodenreider O, Lussier YA. "GenesTrace: Biological Knowledge Discovery via Structured Terminology." *MedInfo 2004: 11th World Congress on Medical Informatics*, San Francisco, CA. 2004.
6. Cantor MN, **Sarkar IN**, Gelman R, Lussier YA. "Hybrid Lexical Methods for Mapping the Gene Ontology to the UMLS." *Pacific Symposium on Biocomputing*, Kaua'i, HI. 2003.
7. **Sarkar IN**, Cantor MN, DeSalle R, Lussier YA. "Exploring SNOMED Using Phylogenetic Analysis Tools." *AMIA Annual Meeting*, Washington, DC. 2002.
8. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. "Knowledge Acquisition of Organized Sets (KAOS) of Clinical Data." *AMIA Annual Meeting*, Washington, DC. 2001.
9. **Sarkar IN**, Planet PJ, DeSalle R, Figurski DH. "Knowledge Aggregation and Separation Through a Novel Classification Algorithm." *AAAS Annual Meeting and Innovation Exposition*, San Francisco, California. 2001.
10. **Sarkar IN**, Malesewski JJ, Luckie DB. "Biocomputing in the Undergraduate Classroom." *AAAS Annual Meeting and Innovation Exposition*, San Francisco, California. 2001.
11. **Sarkar IN**, Luckie DB. "Biocomputing: Learning How to Write Programs to Solve Problems." *Michigan State University Undergraduate Research Forum*, East Lansing, Michigan. 1998.
12. **Sarkar IN**, Koslowsky DJ. "Mitochondrial Cleavage mRNA Cleavage/Polyadenylation Regulation in Pathogenic African Trypanosomes." *Michigan State University 1997 Undergraduate Research Forum*, East Lansing, Michigan. 1997.

PRESENTATIONS:

1. **Session Chair.** Pacific Symposium on Biocomputing (PSB 2007). "Biodiversity Informatics: Managing Knowledge Beyond Humans and Model Organisms." Maui, HI. 3-7 January 2007. (Invited)
2. **Conference Presentation.** American Medical Informatics Association Annual Meeting (AMIA 2006). Washington, DC. 11-15 November 2006. (Invited)
3. **Invited Presentation/Workshop Participant.** National Center for Biomedical Ontology Anatomy Ontology Workshop. Seattle, WA. 8-9 September 2006. (Invited)
4. **Invited Presentation.** Alumni Speaker. Columbia University Department of Biomedical Informatics, Department Retreat. Hudson Valley, NY. 28-30 August 2006. (Invited)
5. **Invited Presentation/Workshop Participant.** "Automated Barcoding Using the Characteristic Attribute Organization System." DIMACS, National Museum of Natural History, Paris, France. 7 July 2006 (Invited).
6. **Conference Presentation.** International Society for Phylogenetic Nomenclature Conference (PhyloCode). "Enabling Biological Knowledge Integration Through Scientific Nomenclature." Yale University, New Haven, CT. 28 June - 02 July 2006. (Invited)
7. **Invited Presentation/Workshop Participant.** "Bridging Biomedical and Biodiversity Knowledge Through Literature." TeleNature Workshop on Refactoring Natural History Literature. University of Illinois Urbana-Champaign, Urbana-Champaign, IL. 17-18 April 2006.
8. **Symposium Chair.** American Association for the Advancement of Science Annual Symposium (AAAS 2006). "Using Ontologies to Teach Computers Biology." St. Louis, MO. 20 February 2006.
9. **Tutorial.** IEEE Computational Systems Bioinformatics Conference (CSBICON 2005). "Organizing and Understanding the Biological Data Deluge through Phylogenetics." Stanford, CA. 8 August 2005.
10. **Conference Presentation.** Annual International conference on Intelligent Systems for Molecular Biology (ISMB 2005). "TEPC: Total Evidence Phylogenetic Correlation of Microbial Phenotypes and Genotypes." Detroit, MI. 28 June 2005.
11. **Invited Presentation.** ACM Joint Conference on Digital Libraries (JCDL 2005); Networked Knowledge Organization Systems/Services (NKOS) Workshop. "Using the Existing

Biomedical Infrastructure to Guide the Design and Development of Biological Ontologies.” Denver, CO. 10 June 2005.

12. **Invited Presentation.** Quantitative Biology (QuBi) Symposium at Hunter College, City University of New York. “Contemporary Topics in Bioinformatics: Organizing and Curating the Genome Deluge.” New York, NY. 6 May 2005.
13. **Invited Panel.** Pacific Symposium on Biocomputing (PSB 05). Panelist Speaker: “Enabling Resources for Biomedical Discovery.” Big Island, HI. 6 January 2005.
14. **Conference Presentation.** Pacific Symposium on Biocomputing (PSB 05). “GenesTrace: Phenomic Knowledge Discovery Via Structured Terminology.” Big Island, HI. 6 January 2005
15. **Conference Presentation.** National Library of Medicine. Training Directors’ Meeting. “The Characteristic Attribute Organization System.” Washington, DC. 9 July 2003.
16. **Conference Presentation.** Pacific Symposium on Biocomputing (PSB 03). “Linking Biomedical Language Information and Knowledge Resources: GO and UMLS.” Kaua’i, HI. 7 January 2003.
17. **Invited Presentation.** Michigan State University - Lyman Briggs School. 2002 Fall Commencement. Alumni Speaker. East Lansing, MI. 7 December 2002.
18. **Invited Presentation.** Michigan State University - Lyman Briggs School. 35th Anniversary Alumni Guest Speaker. “Physiognomonics.” East Lansing, MI. 26 September 2002.