CSCI 599
Geospatial Data Integration
Spring 2011

Instructors: Craig Knoblock (knoblock@isi.edu) and Yao-Yi Chiang (yaoyichi@isi.edu)

Meeting Time: Tuesdays 3:30-6:20pm

Location: THH 217

Office Hours:

Professor Knoblock
- Tuesday 6:20-6:50pm (THH 217, if available)
- Friday 1:30-2pm (ISI 922 or by phone: 310-448-8786)
- Or by appointment
  - On campus Tuesdays 2:45-3:15pm
  - At ISI or by phone: 310-448-8786 other times

Course Web Page: USC Blackboard (blackboard.usc.edu)

There is an ever-increasing amount of geospatial data available, including satellite imagery, aerial imagery, maps, vector layers, elevation data, photos, etc.. There is also a huge amount of information that can be linked to location and integrated with the geospatial layers. This course will focus on the problem of how to integrate the diverse sources of geospatial data. The course will cover a wide variety of topics within this area, including building geospatial mashups, geospatial source discovery, geospatial mediation, geospatial semantic web, geocoding, registering and aligning geospatial layers, extracting layers from maps, linking documents to locations, integrating data for the mobile phone, and open-source GIS systems.

The class will be run as a lecture course with lots student participation, student presentations, and hands-on experience. The class will cover the latest research papers, software, tools, and results on the various topics. Each student will present a research paper in class and develop and build a geospatial data integration project based on the research and tools covered in the class.

Prerequisites: None

Recommended Courses:

CSCI561 -- Introduction to AI
CSCI585 – Database Systems

CSCI587— Geospatial Information Management

Grading:

Course project -- 35%
Quizzes – 25%
Homeworks – 20%
Class Presentation/Participation – 20%

Books: There is no required textbook. We will read technical papers on each topic.

Course Syllabus and Schedule

- **January 11**
  - **Topic: Introduction and Geospatial Basics (Professor Knoblock)**
  - **Readings:**
    - Getting Started with Geographic Information Systems, Keith C. Clarke, Prentice Hall, 2010, Chapters 2 & 3.
- **January 18**
  - **Topic: Geographic Information Systems: ArcGIS (Dr. Chiang)**
  - **Readings:**
    - **Advances in Geographic Information Systems**, Keith C. Clarke, Computers, Environment and Urban Systems, Volume 10, Issues 3-4, 1986, Pages 175-184.
      Paper
    - **The Geographic Info System - A Map for the Future**, Dansby, Bishop, Onsrud, Harlan J., Probate and Property 1989 , Pages 20 – 27.
      Paper

- **January 25**
  - **Topic: Building Geospatial Mashups (Professor Knoblock)**
  - **Readings:**
    - **Making Mashups with Marmite: Towards End-User Programming for the Web**
      Paper
    - **Intel Mashmaker**
      Paper
    - **Building Mashups by Example**, Rattapoom Tuchinda, Pedro Szekely, and Craig A. Knoblock, In Proceedings of the International Conference on Intelligent User Interface 2008
      Paper
    - **Building geospatial mashups to visualize information for crisis**

management, Shubham Gupta and Craig A. Knoblock. In Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management, 2010.
Paper

- **Suggested Readings:**
  - **Mashroom: End-User Mashup Programming Using Nested Tables.** Guiling Wang, Shaohua Yang, and Yanbo Han, In Proceedings of the World Wide Web Conference, 2009.
    Paper
  - **Spatio-textual spreadsheets: Geotagging via spatial coherence.** M. D. Lieberman, H. Samet, J. Sankaranarayanan, and J. Sperling. In GIS'09: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 524-527, Seattle, WA, November 2009.
    Paper

- **February 1**
  - **Topic: Geographic Information Systems (Dr. Chiang)**
  - **Readings:**
    - **An introduction to spatial database systems**, Ralf Hartmut Guting. 1994. The VLDB Journal 3, 4 (October 1994), 357-399.
      Paper
    - **Geographical information systems and location science**, Richard L. Church, Computers & Operations Research, Volume 29, Issue 6, May 2002, Pages 541-562
      Paper
  - **Suggested Readings:**
    - **Using ontologies for integrated geographic information systems**, Frederico T. Fonseca, Max J. Egenhofer, Peggy Agouris, Gilberto Câmara3, Transactions in GIS, Volume 6, Issue 3, pages 231–257, June 2002
      Paper
    - **Using geographic information system analyses to monitor large-scale distribution of nicotine replacement therapy in New York City**, Karen Davis Czarneckia, Chris Goransonb, Jennifer A. Ellisa, Laura E. Vichinskya, Micaela H. Coadya, and Sarah B. Perla, Preventive Medicine, Volume 50, Issues 5-6, May-June 2010, Pages 288-296
      Paper

- **February 8**
  - **Topic: Geospatial Semantic Web (Professor Knoblock)**
  - **Readings:**
    - **The Semantic Web in Breath** by Aaron Swartz
      Paper
    - **The Semantic Web: An Introduction**
      Paper
    - **The Geospatial Semantic Web** by Frederico Fonseca
      Paper (Follow the "Open URL", Read pages 367-376 in NetLibrary)
    - **Toward the Semantic Geospatial Web** by Max J. Egenhofer
      Paper
  - **Suggested Readings:**
    - **Geospatial Semantics: Why, of What, and How?**
      Paper
    - **Exploring the Geospatial Semantic Web with DBpedia Mobile**
      Paper

- **February 15**
  - **Topic: Mapping Addresses to Locations (Geocoding) (Professor Goldberg)**
  - **Readings:**
    - **Exploiting Online Sources to Accurately Geocode Addresses.** Bakshi, R., C.A. Knoblock, and S. Thakkar, 2004, In D. Pfoser, I. F. Cruz, and M. Ronthaler (Eds.), ACM-GIS '04: Proceedings of the 12th ACM International Symposium on Advances in Geographic Information Systems, Washington DC, USA, November, 2004, 194–203.
      Paper
    - **Improving geocode accuracy with candidate selection criteria.** Goldberg, D. W., Cockburn, M. G. (2010). Transactions in GIS. Vol. 14 (S1), pp. 129-146.
      Paper
    - **Toward Quantitative Geocode Accuracy Metrics.** Goldberg, D. W., Wilson, J. P., Cockburn M. G. (2010) In Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. pp. 329-332 Leicester, UK.
      Paper
    - **From text to geographic coordinates: The current state of geocoding. Journal of the Urban and Regional Information Systems Association.** Goldberg, D. W., Knoblock, C. A., Wilson, J. P. (2007). Vol. 19 (1), pp. 33-46.
      Paper
  - **Suggested Readings:**

        - **A comparison of address point, parcel and street geocoding techniques**, Paul A. Zandbergen, Computers, Environment and Urban Systems 32 (2008) 214–232
          Paper
        - **A Flexible Addressing System for Approximate Geocoding**, Davis et al.
          Paper

- **February 22**
    - **Topic: Linking Text Documents to Location (Professor Knoblock)**
    - **Readings:**
        - **STEWARD: Architecture of a spatio-textual search engine.** M. D. Lieberman, H. Samet, J. Sankaranarayanan, and J. Sperling. In *Proceedings of the 15th ACM International Symposium on Geographic Information Systems (ACM GIS'07)*, pages 186-193, Seattle, WA, November 2007.
          Paper
        - **Web-a-where: Geotagging Web Content.** Amitay E., Har'El N., Sivan R., Soffer A. (2004). ACM **SIGIR** 2004.
          Paper
        - **A confidence-based framework for disambiguating geographic terms.** E. Rauch, M. Bukatin, and K. Baker. In Proceedings of the HLT-NAACL 2003 Workshop on Analysis of Geographic References, pages 50-54, Edmonton, CA, May 2003.
          Paper
    - **Suggested Readings:**
        - **Geospatial Mapping and Navigation of the Web** Kevin S. McCurley, Proceedings of the World Wide Web Conference, 2001.
          Paper
        - **Determining the spatial reader scopes of news sources using local lexicons.** G. Quercini, H. Samet, J. Sankaranarayanan, M. D. Lieberman, In A. El Abbadi, D. Agrawal, M. Mokbel, and P. Zhang, editors, Proceedings of the 18th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 43-52, San Jose, CA, November 2010.
          Paper
        - **Geotagging: Using proximity, sibling, and prominence clues to understand comma groups.** M. D. Lieberman, H. Samet, J. Sankaranarayanan In R. Purves, C. Jones, and P. Clough, editors, Proceedings of 6th Workshop on Geographic Information Retrieval, Zurich, Switzerland, February 2010.
          Paper

- **March 1**
    - **Topic: Integrating data for mobile phones (location-based services) (Dr. Chiang)**

- **Readings:**
  - **Location-based services**. Iris A. Junglas and Richard T. Watson. 2008. Commun. ACM 51, 3 (March 2008), 65-69.
    Paper
  - **Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns**, Louise Barkuus, and Anind Dey, Proceedings of the INTERACT 2003, 9th IFIP TC13 International Conference on Human-Computer Interaction
    Paper
  - **SPIRAL: A Scalable Private Information Retrieval Approach to Location Privacy**. A. Khoshgozaran, H. Shirani-Mehr, C. Shahabi. 2008. *The 2nd International Workshop on Privacy-Aware Location-based Mobile Services (PALMS)*
    Paper
- **Suggested Readings:**
  - **Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking**. Marco Gruteser and Dirk Grunwald. 2003. In Proceedings of the 1st international conference on Mobile systems, applications and services (MobiSys '03). ACM, New York, NY, USA, 31-42.
    Paper
  - **Private queries in location based services: anonymizers are not necessary.** Gabriel Ghinita, Panos Kalnis, Ali Khoshgozaran, Cyrus Shahabi, and Kian-Lee Tan. 2008. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (SIGMOD '08). ACM, New York, NY, USA, 121-132.
    Paper

- **March 8**
  - **Topic: Geospatial Reasoning and Fusion (Professor Knoblock)**
  - **Readings:**
    - **A Framework for Integrating and Reasoning about Geospatial Data.** Gupta, S., and Knoblock, C. A. 2010.In *Extended Abstracts of the Sixth International Conference on Geographic Information Science (GIScience)*.
      Paper
    - **A Constraint Satisfaction Approach to Geospatial Reasoning.** Michalowski, M., and Knoblock, C. A. 2005. In *Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05).*
      Paper
    - **Exploiting Automatically Inferred Constraint Models for Building Identification in Satellite Imagery.** Michalowski, M.; Knoblock, C. A.; Bayer, K. M.; and Choueiry, B. Y. 2007. In *Proceedings of the 15th ACM International Symposium on Advances in Geographic Information Systems*

(ACMGIS 07), 35-42.
Paper

- **Suggested Readings:**
  - **Information Fusion for Feature Extraction and the Development of Geospatial Information.** Michael A. O'Brien and John M. Irvine
    Paper
  - **Merging of Heterogeneous Data for Emergency Mapping: Data Integration or Data Fusion.** Florin Savopol and Costas Armenakis
    Paper

- **March 14-18**
  - **Spring Break!**

- **March 22**
  - **Topic: Registering and Aligning Geospatial Layers (Dr. Chiang)**
  - **Readings:**
    - **Automatically conflating road vector data with orthoimagery**, Ching-Chien Chen, Craig A. Knoblock, and Cyrus Shahabi. *Geoinformatica*, 10(4):495--530, December 2006.
      Paper
    - **Automatically and Accurately Conflating Raster Maps with Orthoimagery,** Chen, C.; Knoblock, C. A.; and Shahabi, C. 2008. *Geoinformatica*, 12(3):377--410.
      Paper
    - **Automatic alignment of large-scale aerial rasters to road-maps. X. Wu, R. Carceroni**, H. Fang, S. Zelinka, and A. Kirmse. , In Proceedings of the 15th ACM International Symposium on Advances in geographic information systems, pages 1–8, 2007.
      Paper
  - **Suggested Readings:**
    - **Design of a conceptual framework and approaches for geo-object data conflation**, Li, Linna, Ph.D., UNIVERSITY OF CALIFORNIA, SANTA BARBARA, Chapter 2: Geo-Object Data Conflation: Review and Overview
      Thesis
    - **Image registration methods: a survey**, B. Zitova, Image and Vision Computing, Vol. 21, No. 11, 2003, pp. 977-1000.
      Paper

- **March 29**
  - **Topic: Building 3D Models from LIDAR (Professor Neumann)**
  - **Readings:**
    - **2.5D Dual Contouring: A Robust Approach to Creating Building Models from Aerial LiDAR Point Clouds**, Q. Zhou and U. Neumann,

In proceeding of 11th European Conference on Computer Vision (ECCV), Greece, September 5-11, 2010 (Oral Paper)
Paper

- **A Robust Approach for Automatic Registration of Aerial Images with Untextured Aerial LiDAR Data**, L. Wang and U. Neumann, In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, USA, 2009.
  Paper
- **Approaches to Large-Scale Urban Modeling**, J. Hu, S. You, U. Neumann, IEEE Computer Graphics and Applications, Vol. 23, No. 6, pp. 62-69, November 2003.
  Paper

- **Suggested Readings:**
  - **3D model generation for cities using aerial photographs and ground level laser scans**, C. Früh and A. Zakhor, In *Computer Vision and Pattern Recognition Conference*, Kauai, Hawaii, December 2001.
    Paper
  - **Fast 3D model generation in urban environments**, C. Früh and A. Zakhor, In *International Conference on Multisensor Fusion and Integration for Intelligent Systems 2001*, Baden-Baden, Germany, August 2001, p. 165-170.
    Paper

- **April 5**
  - **Topic: Extracting Layers from Maps (Dr. Chiang)**
  - **Readings:**
    - **Harvesting Geographic Features from Heterogeneous Raster Maps**, Y.-Y. Chiang, Ph.D. Thesis, Department of Computer Science, University of Southern California. Chapter 2, pages 12-57
      Thesis
    - **Integrated text and line-art extraction from a topographic map**, L. Li, G. Nagy, A. Samal, S. C. Seth, and Y. Xu. International Journal of Document Analysis and Recognition, 2(4):177-185, 2000.
      Paper
    - **Reviving legacy population maps with object-oriented image processing techniques**, N. Kerle and J. de Leeuw. IEEE Transactions on Geoscience and Remote Sensing, 47(7):2392-2402, 2009.
      Paper
  - **Suggested Readings:**
    - **Toponym recognition in scanned color topo- graphic maps**, J. Pouderoux, J. C. Gonzato, A. Pereira, and P. Guitton. In Proceedings of the Ninth International Conference on Document Analysis and Recog- nition, volume 1, pages 531–535, Sept. 2007.
      Paper
    - **Colors of the past: color image segmentation in historical topographic maps based on homogeneity**, S. Leyk and R.

Boesch. GeoInformatica, 14(1):1-21, 2010.
Paper

- **April 12**
  - **Topic: Geospatial Source Discovery (Professor Knoblock)**
    Readings:
    - **Geospatial search service**. Procházka, David and Motyčka, Arnošt, Information Society. Ljubljana, Slovenia, 2008. vol. A, pp. 227--230. ISSN 1581-9973.
      Paper
    - **Identifying Maps on the World Wide Web.** Michelson, M.; Goel, A.; and Knoblock, C. A. In *Proceedings of the 5th International Conference on GIScience, LNCS 5266*, 249--260, Springer, New York.
      Paper
    - **A Data Integration Approach to Dynamically Fusing Geospatial Sources.** Thakkar, S. Ph.D. Thesis, Department of Computer Science, University of Southern California. Chapter 3, pages 42-80
      Thesis
  - **Suggested Readings:**
    - **Semantic Web Service Based Geospatial Knowledge Discovery.** P. Zhao and L. Di, Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006.
      Paper

- **April 19**
  - **Project Presentations**

- **April 26**
  - **Project Presentations**

Statement for Students with Disabilities
Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

Statement on Academic Integrity
USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the

expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. *Scampus,* the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://www.usc.edu/dept/publications/SCAMPUS/gov/. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: http://www.usc.edu/student-affairs/SJACS/.