

# Yan Sheng

#Yangzhou University, JiangSu province  
Institute of Information Engineering

13813147410  
shengyan1985-at-gmail.com  
<http://liz.appspot.com>

## Basic Info

- Sex: Female
- Origin/Birthday: SuZhou, China/1985, August
- Health status: Good
- Graduate institutions: Institute of Information Engineering, [Yangzhou University](#)
- Professional/Direction: Computer Application Technology/Data Mining
- Qualifications: Master/graduated in June 2010
- Mobile Phone: 13813147410
- Website: [liz.appspot.com](http://liz.appspot.com)
- Email: [shengyan1985@gmail.com](mailto:shengyan1985@gmail.com)
- Communication Address: 400173#, Institute of Information Engineering, Yangzhou University, JiangSu Province

## Education

- **Master**  
*2007.09 - now*
  - studying in Institute of Information Engineering, Yangzhou University and doing some research about data mining;
  - The courses I learned are: Matrix theory, data warehouse, data mining, distributed database theory, parallel processing, parallel algorithm design and analysis, computer architecture, computer network technology, optimization theory, etc;
  - Do some research about data mining, concept lattice and be familiar with many data mining algorithms, such as: the classic Apriori, Decision Tree classifier, naive Bayesian classifier, KNN classifier, KMeans clustering, Rough Set Theory, Fuzzy Set Theory and Godin, Bordat algorithm in Formal Concept Analysis;
  - Published many papers, and in mid-August 2009, participate in ICNC-FSKD'09 international academic conference.
- **Bachelor**  
*2003.09 - 2007.07*
  - learned lots of things in Institute of Information Engineering, Yangzhou University and obtained the bachelor's degree of Computer Science and Technology;

## Skills

- Language: Python, C/C++, Java, Pascal, L<sup>A</sup>T<sub>E</sub>X;
- OS: Linux (Ubuntu, Fedora);
- Web Technology: Familiar with the Web application development; Master Html, CSS, Javascript and other Web technologies, and use J<sup>Q</sup>uery skilled, also have Ajax development experience; Become adroit at the D<sup>j</sup>ango Web development framework and know about K<sup>a</sup>rrigell, GAE;
- Database: Familiar with the use of M<sup>y</sup>SQ<sup>L</sup>, S<sup>q</sup>lite and other databases;
- System Maintain: Familiar with Apache, L<sup>i</sup>ghttpd, ftp, svn, trac, samba;
- Other: Photoshop, Illustrator software;
- English: CET-4, have the good ability of English reading, writing, listening and speaking. Do well in reading or translating English literature of computer science and have published scholarly English papers.

## Project Experience

- **Galicia**  
*2007.9 - 2008.3*
  - Describe: Galicia is an open source project to implement concept lattice construction algorithms.
  - Duties: Improve the Godin algorithms to obtain better time and space complexity.
- **Openbookproject**  
*2008.4 - 2009.6*
  - Describe: The project is about the translation or creatation of Chinese Pythonic books which is hold on <http://code.google.com/p/openbookproject>. LovelyPython is a original book which aimed to introduce the python to reader in the most understandable way.
  - Duties: Took part in the writting process of LovelyPython.
- **Construction of Avian virus genome bioinformatics analysis platform**  
*2008.11 - 2009.1*
  - Describe: Since major international Bioinformatics Center provide many anaysis software for gene and gene/nucleic acid database, such as, BLAST retrieval system(The Basic Local Alignment Search Tool) and NCBI database(National Center for Biotechnology Information), SMS2(The Sequence Manipulation Suite 2, for analyzing the shorter DNA or Protein sequence), Clustalx-2.0.10(Used for DNA or multiple protein sequence alignment tool), these analysis tools must be localized. Further more, some specail function need be extented, for instance, selecting the Avian virus genome from the complete database, then updating and other maintenance.
  - Duties: We installed RedHat Enterprise Linux 4.0 AS as server operating systems, Apache2.2 as Web server and other utilities. Then after localizing the BLAST, SMS2 and Clustalx, we integrated three systems to one complete system. The main extension is selecting the Avian virus genome to build a sub-database which updates with the NCBI database.
- **eachdn Domain Name Trading Platform**  
*2009.5 - 2009.9*

- Describe: An e-commerce website for domain name transactions, the website is <http://www.each.dn>.
- Duties: Responsible for the overall website's backend part, including needs analysis, database design, architecture design, including js. The main technologies and framework: Django, jQuery, Orbited, Xapian.

### • Collaborative information retrieval research using semantic info

• 2009.3 - now

- Describe: This project was supported by Jiangsu Province Universities Research and Innovation Fund in 2009. The project is to design a information retrieval method using semantic info. The main contents contains: 1, design a new feature selection method; 2, advance an concept similarity measure based on concept lattice and rough set theory; 3, build a framework for collaborative information retrieval; 4, build a user interest model using WordNet and re-rank search results;
- Duties: Read a number of papers, study of the text feature extraction methods, and semantic ontology-related knowledge, presents a new collaborative information retrieval method.

## Papers

- 2009 Yun Li, Yan Sheng, Luan Luan, Lianglei Sun and Ling Chen. A Personalized Search Results Ranking Method Based on WordNet. The 5th International Conference on Natural Computation(ICNC'09) and The 6th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'09), Tianjin, China;
- 2009 Li Yun, Sheng Yan, Luan Luan. A Text Classification Method with an Effective Feature Extraction based on Category Analysis. The 5th International Conference on Natural Computation(ICNC'09) and The 6th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'09), Tianjin, China;
- 2008 SHENG Yan, LI Yun, TIAN Su-fang, LUAN Luan. A Rough Concept Lattice Model of Variable Precision. IEEE International Symposium on Intelligent Information Technology Application 2008 (IITA'08), Shanghai City, China;
- 2008 Shengyan, Li yun, Li tuo, Luanluan. Ontology mapping based on the lattice model. Jiangsu Computer Conference 2008, JSCC 2008;
- 2008 SHENG Yan, LI Yun, LI Tuo, YUAN Yun-hao. A method for ontology merging based on the concept lattice. DPCS2008;

## Interests

- Life: Like classical movies, listening English songs, drawing and playing badminton;
- Computer: Also like image processing software and services, such as PhotoShop, Illustrator, GIMP, Picasa. Concerned about the IT technologies;
- Research: interested in Intelligent Information Retrieval, NLP methods and techniques.