

## EDUCATION

**School of Computer Science, Carnegie Mellon University (CMU), Pittsburgh**

*Master of Science in Very Large Information Systems, Expected December 2013*

- **Core Curriculum:**

- Multimedia Database and Data Mining
- Distributed Systems
- Studio Data Systems (Hadoop, BigTable etc.)
- Machine Learning
- Design & Engineering of Intelligent Information Systems (IBM Watson)

**University of Toronto (U of T), Toronto**

*Honours Bachelor of Science with High Distinction in Computer Science, June 2012*

- Awards: Honours List; Summer Research Experience Award
- Leadership: VP Technology of University of Toronto China Conference

## PROFESSIONAL EXPERIENCE

**IBM Toronto Software Lab, Toronto, ON**

Oct. 2010 – Dec. 2011

*Co-op DB2 Quality Assurance Developer*

- Enhanced the defect detection tool by implementing a message token comparison algorithm and “Omni-Search” extension based on grep.
- Awarded IBM Certified DB2 Academic Associate

**Caseware International Inc., Toronto, ON**

May 2009 – Dec. 2009

*Co-op Application Developer*

- Developed the client side of a web-based ERP System by using Google Web Toolkit (GWT), GXT and JavaScript.
- Main features include draggable file explorer, text editor, file upload etc.

## PROJECTS

**Parallel Data Mining from System Logs using Hadoop, CMU**

Fall 2012

- Proposed and implemented a data mining algorithm to discover users’ usage pattern from terabytes of data by using Hadoop and MapReduce

**Parallel Graph Mining using Hadoop, CMU**

Fall 2012

- Designed and developed an efficient algorithm to perform *Belief Propagation* on large graphs using Pig on Hadoop and conducted performance analysis

**Biomedical Question Answering System, CMU**

Fall 2012

- Produced a QA System that can answer biomedical questions in terms of genes based on data sources on a server running Apache Lucene and Solr. Some techniques, such as “query expansion”, were used to improve the correctness.

**SyncGallery+ (Android Cloud-based Applicaton), U of T**

Spring 2012

- Significantly reduced the complexity of editing and sharing photos among different platforms by utilizing the features of Dropbox, Aviary and Bump (<https://www.sunyangkobe.com/cscd08>)

**SmartShare Carpooling Reputation System, U of T**

Summer 2011

- Created a Map based carpooling reputation system by using “WAMP”, Mootools, hMailServer etc. and imported SNS features, such as “friends list”, into the system. (<https://www.sunyangkobe.com/smartshare/>)

## TECHNICAL SKILLS

Java, C, Python, Perl, SQL, Linux Shell Scripting, PHP, AJAX, JSP, GWT

Github, SVN, Maven, Hadoop, Accumulo, Pig, Linux, Windows