Si Yin

Doha, Qatar

github.com/zyzyis

zyzyis@gmail.com

+974 6641 9808

Work Experience

- **Senior Research Software Engineer** in **QCRI**   2013.4 - present (Doha)

    - Be the initial member to develop NADEEF, the first-kind end-to-end data cleaning system. I experienced the whole process of transforming an idea into a working product, and growing the team into a company.
    - Contribute more than 90% of the open-source version of NADEEF. Develop the data pipeline engine, the distributed architecture of NADEEF dashboard, and HTML5 front-end.
    - Develop highly scalable version of NADEEF based on Apache Spark.
    - Evanglize the product in enormous of data conferences (e.g. SIGMOD, VLDB).

- **Software Engineer** in **Philips**   2012.1 - 2013.4 (Eindhoven)

    - Deliver Imaging system SDK for Philips Healthcare products (MR, CT, iXR).
    - Involve in full-cycle product development activities including architecture design, coding workflow, and testing.
    - Mainly work on the Rendering / AppHosting components using technology including C#, .Net, GPGPU, MultiThread/MultiProcess concurrency, and DirectX.
    - Work in a global development team ($\approx$ 200 engineers) on millions lines codebase with daily scrum / agile methodology.

- **Software Consultant** in **IBM**   2008.8 - 2011.12 (Amsterdam)

    - Work in the IBM OpenCL team to develop special software component which provides 40x speed-up for client's HPC application.
    - Work in the IBM Cognos team to provide Real-time monitoring, data visualisation / reporting, and asset management solutions to clients from Water and Energy sectors.
    - Experience working with global outsource partners from India and China.
    - Mentor students from IBM Extreme Blue program and their master thesis projects.

- **ExtremeBlue Intern** in **IBM**   2007 Summer (Amsterdam)

    - Engage, brainstorm, and whiteboard ideas with customer ABNAMRO bank.
    - Develop NFC (Near Field Communication) software on secure bank card.
    - Be the key presenter in the IBM EMEA Extreme Blue Expo 2007.

- **Research Assistant** in **Vrije Universiteit Amsterdam**   2006 Summer (Amsterdam)

    - Develop web application framework (PANORAMA) for gaming and education using PHP, Google GWT, JQuery, and VRML/X3D technology.

Education

Vrije Universiteit Amsterdam   The Netherlands
MSc with *Cum Laude*, Computer Science, 2005 - 2008
Thesis: Map/Reduce based ray tracer on IBM Cell B.E. processor
Supervised by Thilo Kielmann and Henri Bal
(The paper and source code can be found in https://github.com/zyzyis/GoldenMonkey)

Sichuan University   China
BSc with *Honor*, Computer Science, 2001 - 2005
Thesis: The Research and Implementation of Query-by-humming system
Supervised by Changjie Tang

Engineering Portfolio

- Core Stack: C/Java/Python/C++/C#/OOP/Clojure
- Web Stack: Javascript/JQuery/D3JS/RequireJS/Bootstrap/Jetty/MongoDB/RDBMS
- Visualization technology (D3JS/OpenGL/Realistic Rendering/OpenCV)
- Parallelization technology (MultiThreading/OpenCL/SIMD/GPGPU)
- Familar with Data Analytics/Machine Learning
- Experienced in Agile/Scrum/Waterfall Methodologies
- **Excellent in Communication and Presentation**

Publication

- Ahmed Elmagarmid, Ihab F. Ilyas, Mourad Ouzzani, Jorge-Arnulfo Quian-Ruiz, Nan Tang, Si Yin: NADEEFER: Generic and Interactive Entity Resolution. **SIGMOD 2014**

- Amr Ebaid, Ahmed K. Elmagarmid, Ihab F. Ilyas, Mourad Ouzzani, Jorge-Arnulfo Quian-Ruiz, Nan Tang, Si Yin: NADEEF: A Generalized Data Cleaning System. **PVLDB 6(12)**: 1218-1221 (2013)

Extra Sugar

- **Data Analysis and Statistical Inference, Coursera, Duke University**, 2014.
- **Machine Learning, Coursera, Stanford University, Score: 100/100**, 2013.
- **IBM Client-Value Outstanding Technical Achievement Award**, 2011.
- **IBM Technical Talent Program**, 2011 - 2012.
- **Cum Laude, M.Sc**, Vrije Universiteit Amsterdam, 2007.
- **V.H Rutgers scholarship**, Vrije Universiteit Amsterdam, 2005.
- **ACM ICPC Asia/Northwest Europe**, 2002, 2003, and 2005.
- **Best Graduation Thesis**, Sichuan University, 2005.