

Google

SUMMER

of

CODE

TENTH

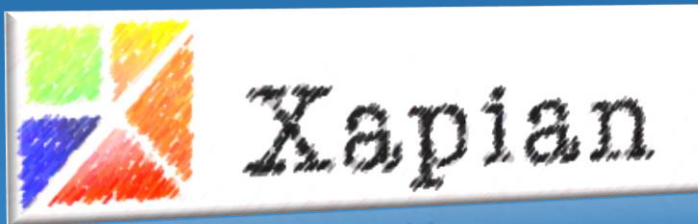
2014

YEAR

# My GSoC 2014 Experience

— Learning to Rank in Xapian

Hanxiao Sun | Nov, 14 2014



# Self Introduction

I am a third year computer science master student at ICT, CAS

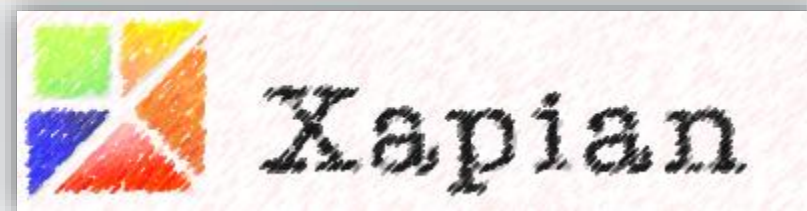
My research interests include IR, ML, RecSys

I like sports. I have ridden my bike to Qinghai Lake, Hanoi, Hainan, Taiwan and participated in Beijing Marathon and Beijing Triathlon.

Find me at

<http://about.me/sunhanxiao>





## ● History

- ❑ Xapian has evolved out of more than 30 years
- ❑ Replacement for Muscat software(1984)
- ❑ Released under the GPL in 2000
- ❑ Name changed to Xapian in 2001

## ● So, what is it?

- ❑ Toolkit for IR
- ❑ Index documents and search them
- ❑ Associated tools / applications
- ❑ C++ core, bindings for many languages(C#, Erlang, Java, Lua, Perl, PHP, Python, Ruby, Node.js)

## ● User



**GMANE**

*grokbase*

UNIVERSITY OF  
Southampton

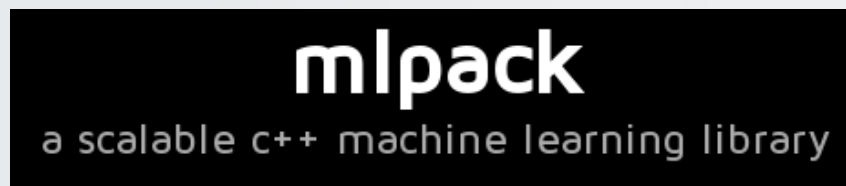
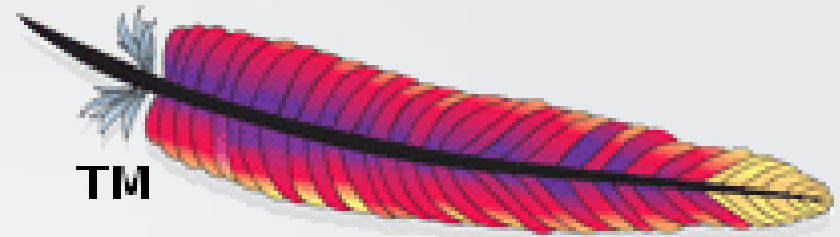
## ● GSoC

□ Xapian has been involved with Google Summer of Code (GSoC for short) in the following years:

- 2009 SWIG-based Perl bindings (Mentored under SWIG)
- 2011 (Mentoring Org)
- 2012 (Mentoring Org)
- 2013 Divergence from Randomness Weighting (Mentored under Debian)
- 2014 (Mentoring Org)

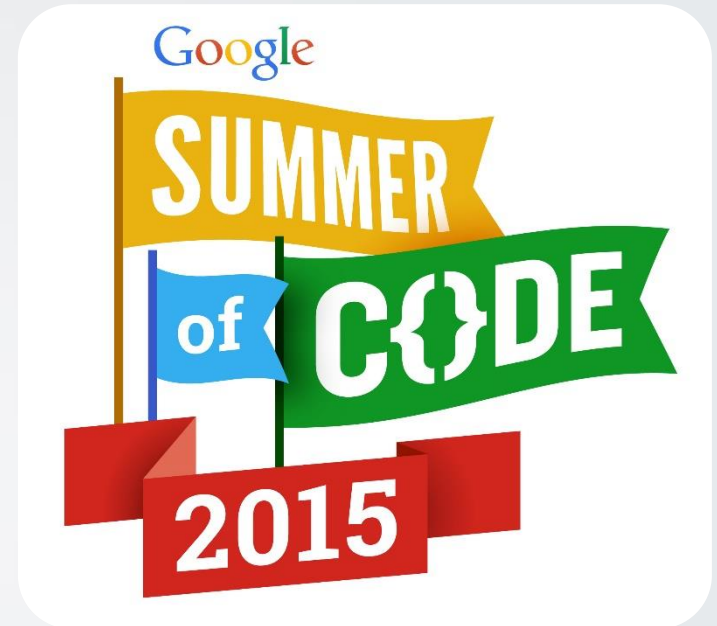
- GSoC 2011:
  - Parth Gupta, PhD student in Universidad Politécnica de Valencia
  - Finish the first algorithm in Xapian——SVM-rank
- GSoC 2012:
  - Rishabh Mehrotra
  - Quit in the midterm
- GSoC 2014:
  - Me, Hanxiao Sun
  - Finish the ListNet, ListMLE, Adarank in Xapian-letor

- Why Xapian ?



- How to apply
  - Choosing an idea to work on
  - Checking out and building the code
  - Reading the Resources for your chosen Project Idea
  - Learn about Xapian's API
  - Get familiar with the code
  - Writing your proposal

- Tips :
  - ❑ The sooner the better
  - ❑ Find your passion
  - ❑ Do not be shy
  - ❑ Keep your promise
  - ❑ Enjoy your summer
  - ❑ Hope to see you in the next year !





Thank you