

Google and Open Source

Chris DiBona

April 10, 2007

Flourish, Chicagoland



Open Source Programs Manager, Google Inc.

- Compliance
- Code Release
- The Summer of Code
- APIs and such

Former Editor: Slashdot

Long-Time Linux Advocate & Writer about F/OSS

- Co-Editor Open Sources & Open Source 2.0

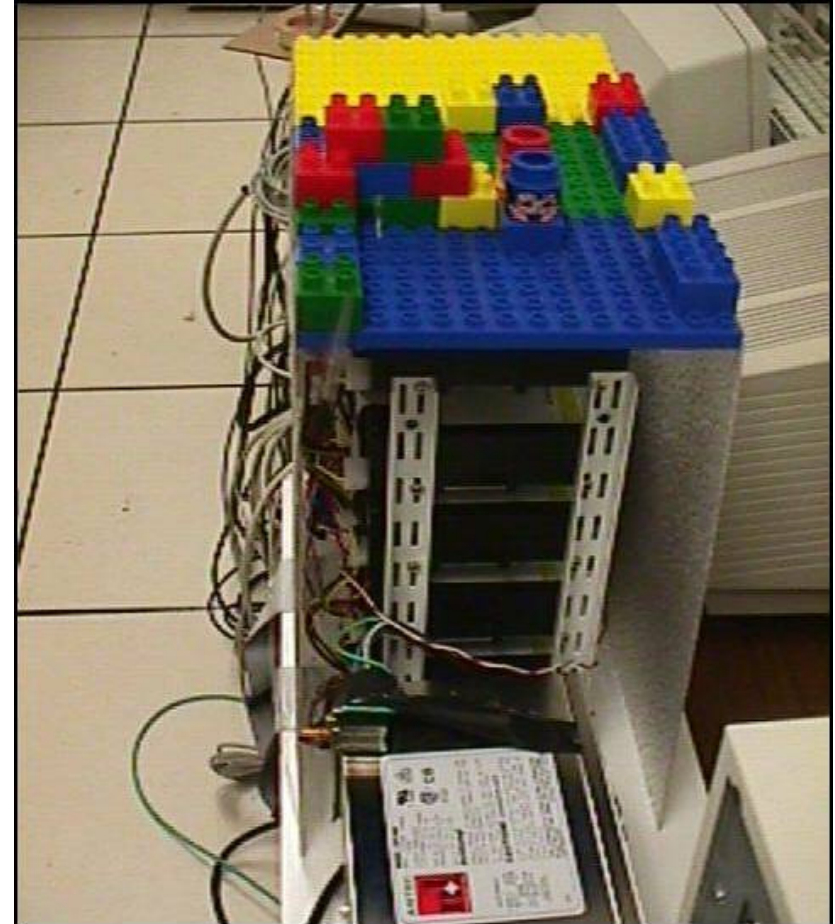
Free Software Developer

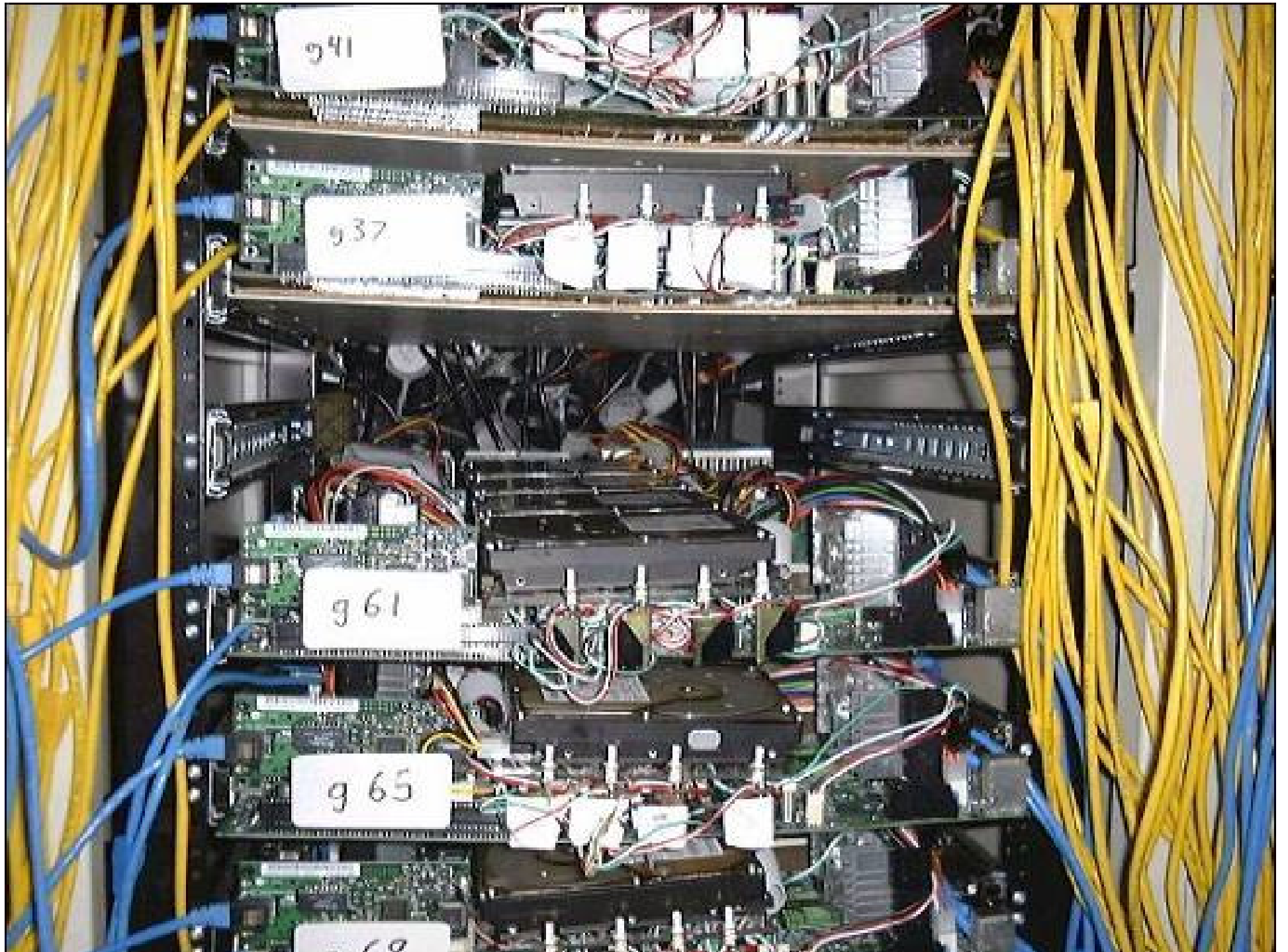
Floss Weekly Podcast

Agendas, you see, destroy suspense.



Lego Disc Case





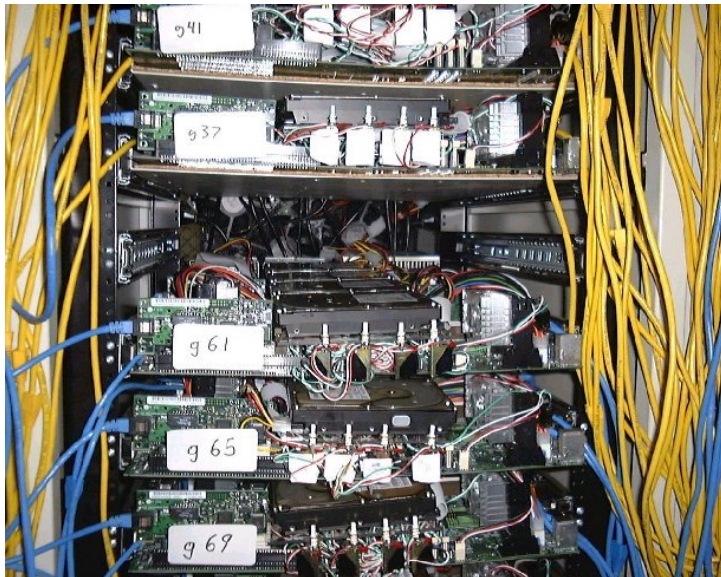


Reliability/Fault Tolerance

PCs are unreliable, especially if you have thousands

But they are cheap and fast

Strategy: Exploit processing power of off-the-shelf PC hardware, make it reliable in software







But **First!**



What am I actually here to talk about?

“A **license** can ruin a perfectly good piece
of software”

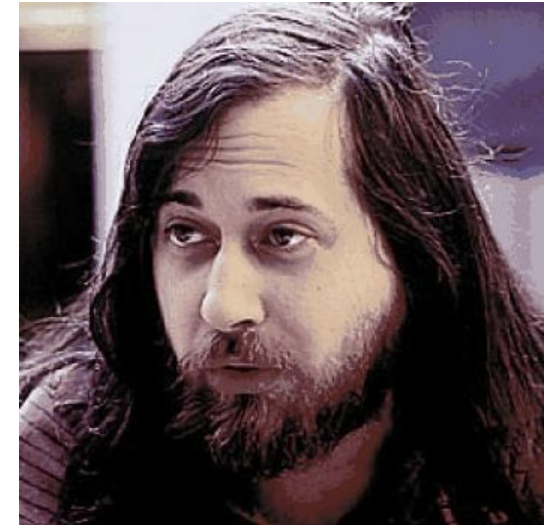
– *Jon Stevens*

Where Did **Free Software** Come From?



Richard Stallman had a broken printer

- The FSF
- The GPL
- Emacs/GCC
- Now the GPL covers ~50% of all open source software



Why am I telling you this?

- This experience guides Richard's life.
- Indeed it guides all of the GPLv3 process.

Where did **Open Source** Come From?

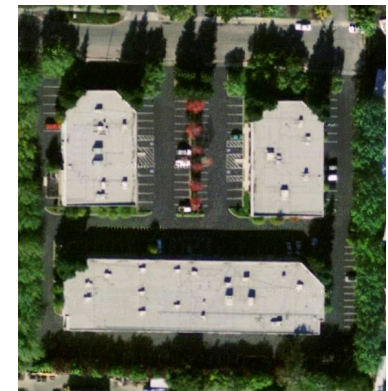


Free Software wasn't perfect for everyone.

- Free was weird.
- Other Licenses were growing
 - BSD
 - Apache
 - MIT
 - MPL/NPL

How do you sell Free?

Tim O'Reilly called together a number of people to find out what to do.



A license that is approved by the Open Source Initiative and complies with the Open Source Definition

Definition Problems:

- “Open Source” is not trademarked and thus, not statically defined.
- Many firms call themselves open source ,when in fact they are not.

Other Problems:

- 60+ Approved to date.
- Misuse of the term.
- Scare tactics from vendors.

The important parts:

- No field of use restrictions
- The software must be distributed with no royalties.
 - Save for media costs.
- The program must be available as source code
- No discrimination against people or groups
- Code must be redistributable

The Confusing parts:

- License must not be specific to a product
- License must not restrict other software
- License must be technology neutral

Licenses:

- Apache Software License: Feel free to use it, closed or open.
- LGPL/MPL/CDDL: If you use it, you have to keep the stuff you use freely available, but you can keep your stuff closed.
- FSF GPL: If you use this library, than everything you link it to must be free as well.

All requirements are keyed on ***distribution***, meaning shipping or allowing the code to be downloaded and ***linking*** meaning the inclusion of the code in some manner.

Yes, but **why** do people **release** code at all?



A BCG/OSDN Study found that open source developers are motivated by:

- Intellectual Curiosity/Stimulation (44.9%)
- Skills Improvement (41.3%)
- Work needs (33.8 %)
- Open Source Idealism (33.1%)
- Non-Work Functionality (29.7%)
- Professional Status (17.5%)
- Reputation building within Open Source (11%)
- Dislike/distrust of proprietary software (11.1%)

<http://osdn.com/bsg>

OSS Developers understand licenses.

They know they are “giving away” their software.

They know people will use it.

The vast majority expect nothing for the work.

If they did, they’d just do it as a commercial offering.

That said, Every 6 months, there is a problem with someone not getting paid for their free software. Most recent was OpenBSD.

How does Google use Open Source?



As Infrastructure

- Linux kernel
- Apache Tomcat and many other Apache tools
- SSH, and other system management tools
- The Languages and compilers that we use.
- Engineers and others running Linux (Goobuntu)

As building blocks

- The //depot/third_party directory has hundreds of libraries.
- Put it in third_party, don't cut and paste

Control and Ownership

- Maintain our independence from external software companies

Adaptability and Flexibility

- We can drill down to repair and enhance our services
- If we want to do something out of the ordinary, we can do so without showing our hand

Roots

- Appeals to the Google ethic

Every time you use Google you are using Open Source

How does Google take part?

Patching and Code Release
Donations

Internal License Discipline

API/Protocol License Discipline

Industry Trade Group Participation

Open Source Infrastructure & Leadership

Summer of Code

About a million lines of code released to date

Developer Tools:

- Perftools
- Google Web Toolkit
- Javascript encryption library
- And dozens more.
 - See <http://code.google.com>

Hosting on Google Code which provides:

- Version Control
- Issue/Bug Tracking
- Wikis
- Project Management

Thousands of projects, 2nd Largest Repository Online

Andrew Morton

- 2.6 Kernel Maintainer

Guido Van Rossum

- Python BDFL

Greg Stein

- Apache Chairman

Jeremy Allison

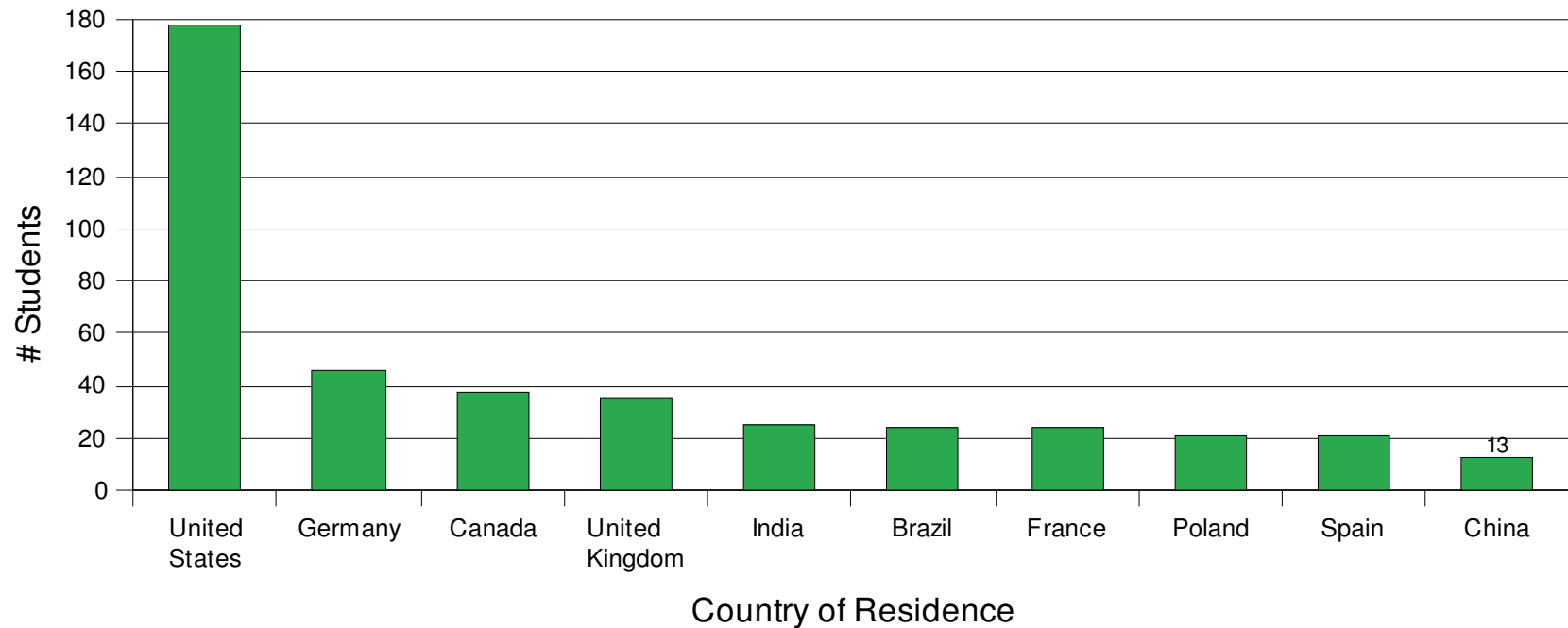
- Samba Lead Developer

And Others....

6338 SoC **Applications** *from*
3044 SoC **Applicants** *to*
102 **Open Source Orgs** *with*
1260 **Mentors** *selecting*
630 **Students** *from*
456 **Schools** *in*
90 **Countries** *for*
One Summer of Code

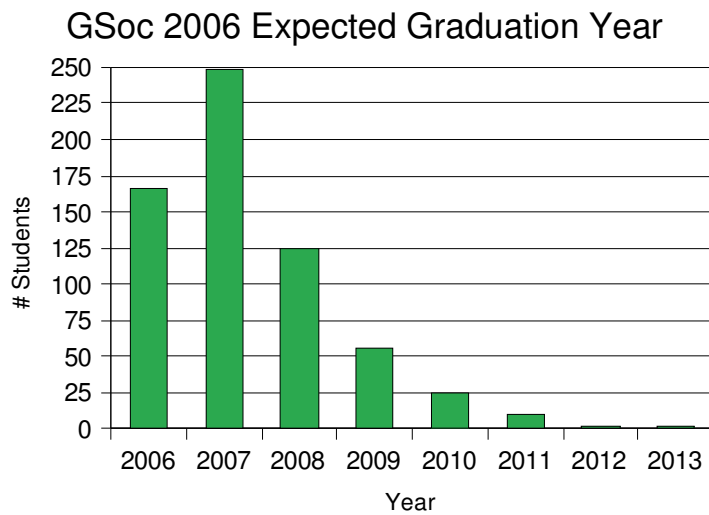
- 1) Student applies to work on a project for, say, the ASF
- 2) ASF mentors review the application and accept the student
- 3) Google pays student \$500 for being accepted
- 4) Student works with an ASF mentor to create project
- 5) Mid-Term, good students are paid \$2000 by Google
- 6) Student continues work with mentor to create project
- 7) Final: If the student has completed the goals set forth in the application and student is paid \$2000 by Google
- 8) Successful student gets t-shirt, certificate and valuable experience along with the money.

GSoc 2006 Geographic Distribution: Accepted Students (Top 10 Countries)



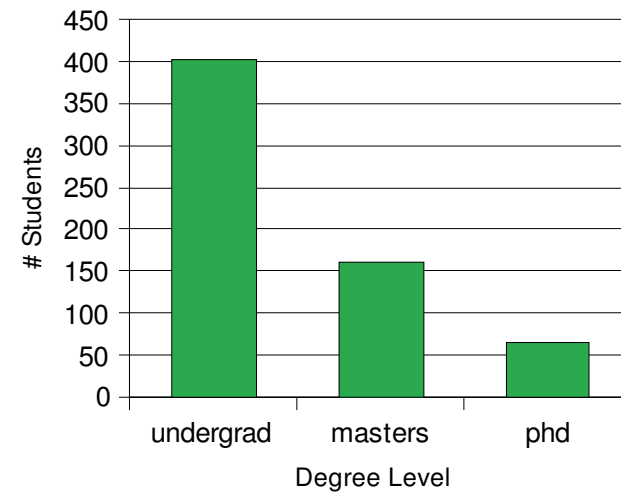
Students Per Country	
Min:	1
Max:	178
Sum:	630
Mean:	10.68
Median:	3
Stdev:	24.26

Student Educational Distribution



Average graduation year: 2007

GSoc 2006 Student Level Distribution



Undergraduates most common participants

Not Just Computer Scientists



Mechanical Engineering

English Literature

Interior Design

Urban Planning

Astronomy

Cartography

Genetics

Developmental Psychology

Why do we do all this funding?



Keep students coding

- *Flip Bits, not Burgers*

Ensure a healthy open source ecosystem

- Good for the Internet
- Good for Google

A great way to meet students all over the world

A great way for open source organizations to grow and thrive.

A lot of Googlers come from the Open Source World

- Pre-training for our kind of engineer.
- Lead into internships.

Giving Back

The Summer of Code, on Google Earth



http://code.google.com/soc/soc_map2.kml

More **Students**, More **Organizations**, More **Time**

6336 applications

800+ students

131 Organizations

1400+ Mentors

Announcing the winning 800 **students** next week.

The End



Questions?

chris@dibona.com

<http://dibona.com>

<http://code.google.com>

Backup Slides Follow



Backup Slides Follow



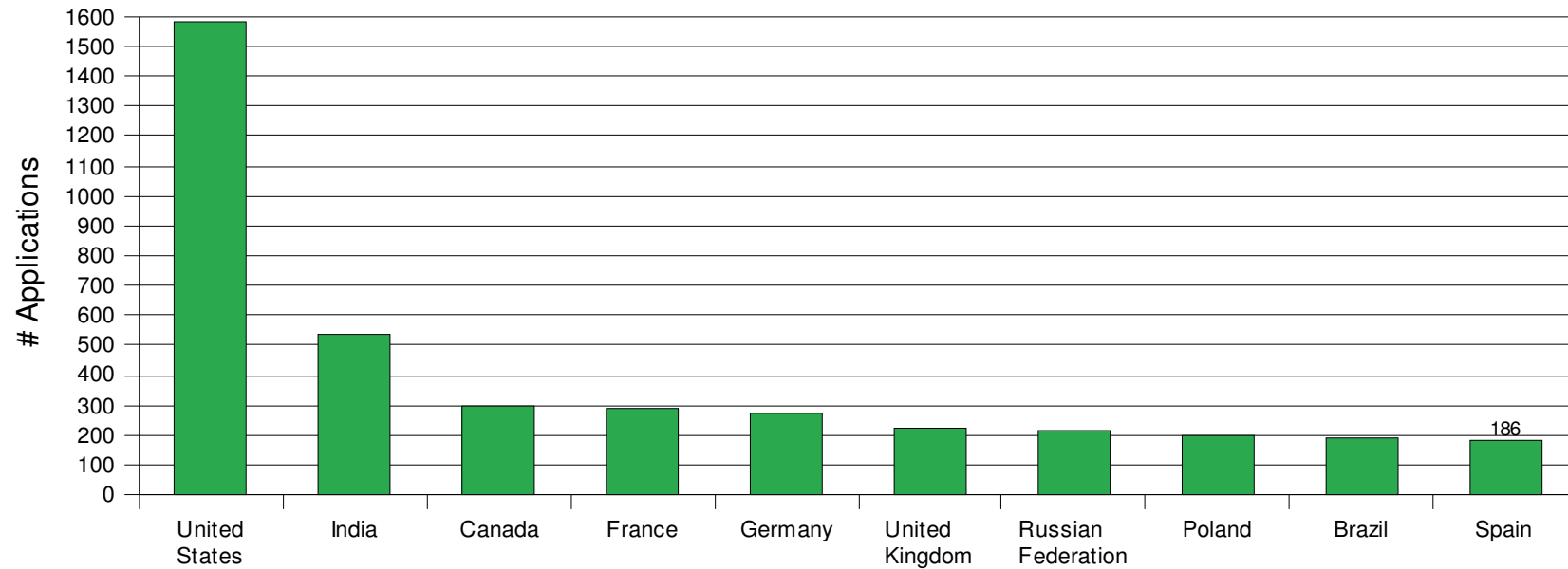
Backup Slides Follow



The Applicants



Applications Per Country (Top 10 Countries)



Applicants Per Country	
Min:	1
Max:	1587
Sum:	6338
Mean:	70.42
Median:	19.5
Stdev:	183.03