CS499 - OPEN SOURCE SOFTWARE DEVELOPMENT

Lecture #02: Basics about open source software

Dr. Igor Steinmacher

e-mail: igorfs@utfpr.edu.br

Twitter: @igorsteinmacher

OSD Criteria

1. **Free Redistribution**

2. Source Code

3. **Derived Works**

4. Integrity of The Author's Source Code

5. **No Discrimination Against Persons or Groups**

6. No Discrimination Against Fields of Endeavor

7. **Distribution of License**

8. License Must Not Be Specific to a Product

9. **License Must Not Restrict Other Software**

10. License Must Be Technology-Neutral

https://opensource.org/osd

QUESTIONS

- What is Open Source about?

  – **Price?**

QUESTIONS

- What is Open Source about?

    – **Price?**

    – **License?**

QUESTIONS

- What is Open Source about?

  – **Price?**

  – **License?**

  – **Social Movement?**

QUESTIONS

- What is Open Source about?

  – **Price?**

  – **License?**

  – **Social Movement?**

  – **Hobby?**

QUESTIONS

- What is Open Source about?

  - **Price?**

  - **License?**

  - **Social Movement?**

  - **Hobby?**

  - **Market trend?**

- Software *vs.* hardware: what's the difference?

  - **Engineering *vs.* Software Engineering**

  - **Tool *vs.* Knowledge**
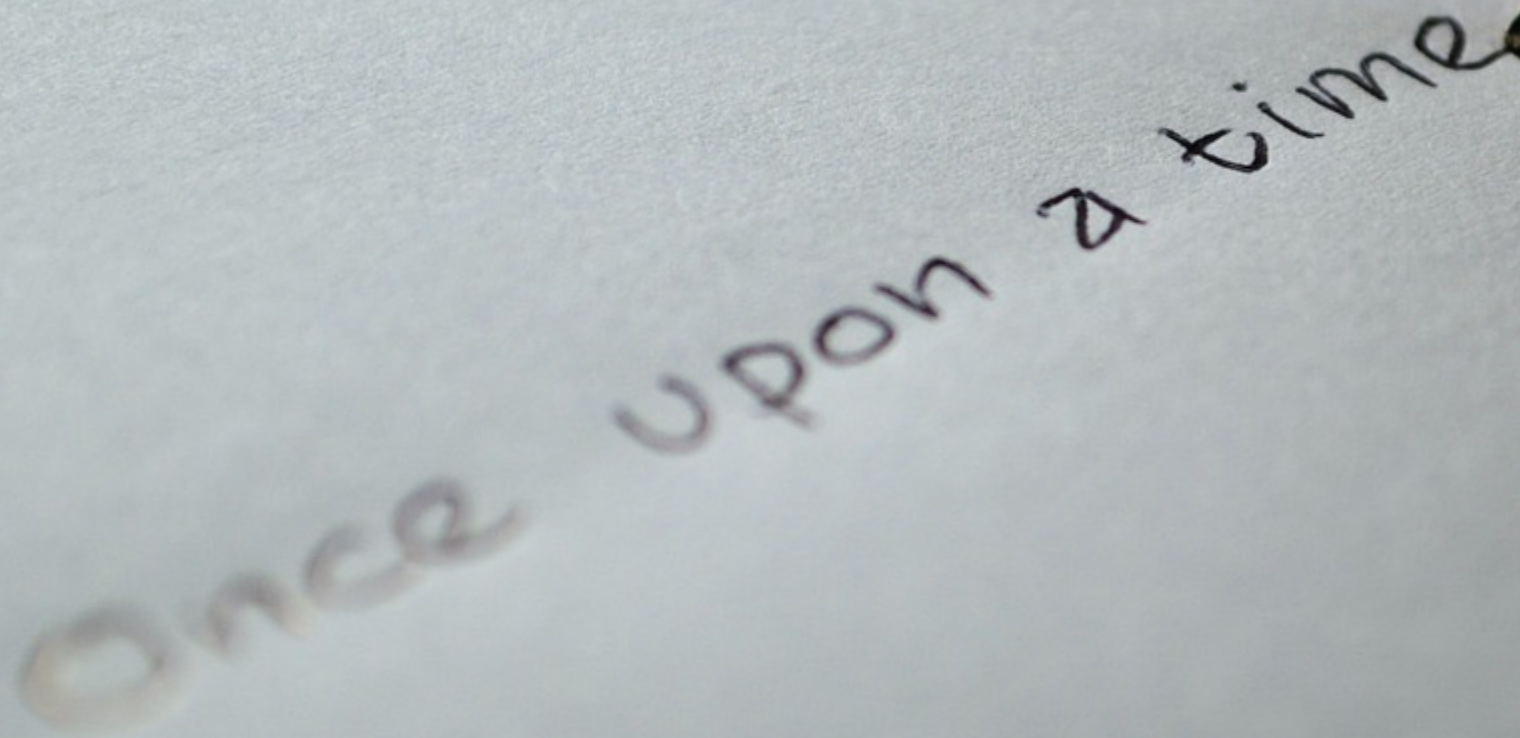
  - **Product *vs.* Service**

So…

- Open Source is about…

  - **Freedom**

  - **License**

  - **Code as a way to share knowledge**

  - **Community!**

PRE-HISTORIC…

- Commercial settings:

  – **software is usually part of a bigger solution (consulting, hardware, maintenance, etc.)**
  – **not a solution *per se***

- Academic settings:
  – **Software is informally shared as a way to share knowledge**

- Hobbyists
  – **Hacking for fun**
  – **Garage programmers community**

AND THE HISTORY…

- 1976 – Bill Gates' "Open Letter to Hobbyists"
  - **Copyright infringement + commercial power**
  - *"Will quality software be written for the hobby market?"*
  - *"One thing you do do is prevent good software from being written."*

- 1980 – Stallman problem with Xerox printer
  - **Stallman refused access to the source code for the printer driver**

- 1981 – IBM signs a deal with "Micro-Soft"
  - **MS-DOS embedded in each and every IBM machine**

- 1984 – Stallman launches GNU project
  - **1985 – Free Software Foundation (FSF) was formed to house funding for the GNU Project**

AND THE HISTORY…

- 1991 – Linus Torvalds creates (or kicks off) Linux
    - **Personal/informal project**
    - **Message to a newsgroup at Usenet**
    - **V 1.0.0 released in 1994 with 170+ KLoC**

- 1995 – Internet Boom
    - **Linux is distributed in large-scale**

- 1997 – Eric Raymond writes "*The Cathedral and the Bazaar*"
    - **Advantages of Open Source**
    - **Development in a decentralized way**

AND THE HISTORY…

- 1998 – Netscape open-sourced
  - **Fight with MS Explorer?**

- 1998 – Eric Raymond, Linus Torvalds and others launch the Open Source Initiative (OSI)
  - **Technical reasons to Open-source**
  - **Use of "Open Source" instead of "Free software"**

- 1999 – Sourceforge is launched

- 2000 – OpenOffice is launched

- 2001 – IBM invests $1 billion on Linux

AND THE HISTORY…

- 2005 – Sun launches Solaris 10 as Open Source

- 2007 – Sun launches JDK under GPL

- GitHub launches in 2008
  - **GitHub BOOM ~2010–2011 (2 million repos in 2011)**

GitHub ✔ @github · 16m

We're thrilled to announce that we've entered into an acquisition agreement with @Microsoft! news.microsoft.com/?p=406917

- … MANY COMPANIES JOINING

WHY OSS???

WHY OPEN SOURCE?

- Incentive to collaboration → human relationship

- Technical education

- "Low-cost" access to "high-end" technology

- Decentralizes the "power" of the software

- Government → knowledge and transparency
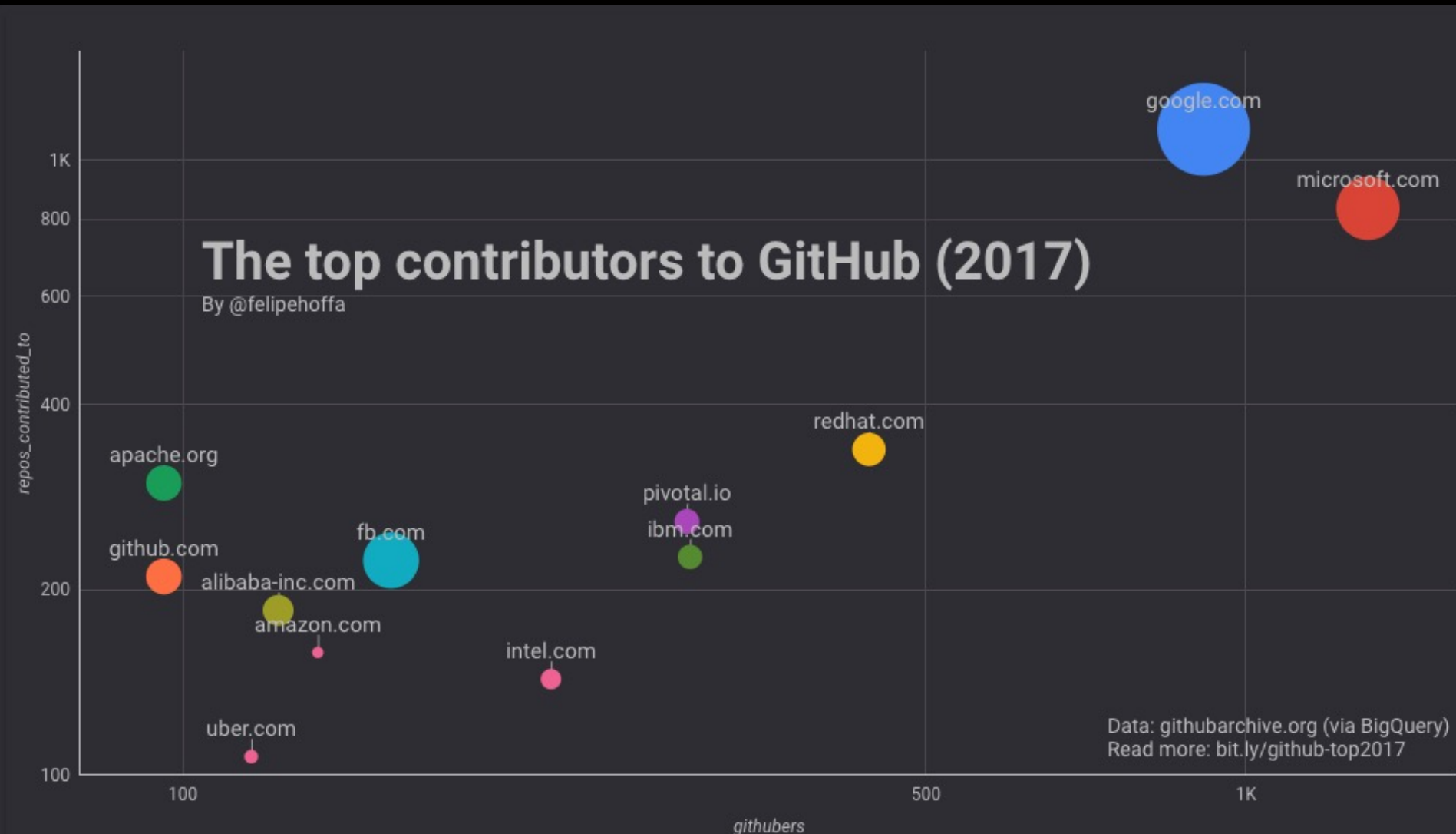
WHY OPEN SOURCE?

- Less cost (is that a thing? When?)

- Less bugs → More quality
  - *"Given enough eye-balls all bugs are shallow"*

- Customization and improvements

- Competition for support

"Problems"

- Quality assessment of the solutions

- Reputation (FUD)

- Sustainability

- "Intellectual property"

FINAL FACTS

The top contributors to GitHub (2017)
By @felipehoffa

Data: githubarchive.org (via BigQuery)
Read more: bit.ly/github-top2017

RANK	ORGANIZATION	ACTIVE CONTRIBUTORS		TOTAL COMMUNITY		
1	Google	2912	+843	6955	+1245	>
2	Microsoft	2644	+719	6327	+1528	>
3	Red Hat	2426	+567	3955	+583	>
4	Intel	1142	+350	2813	+566	>
5	IBM	994	+314	2660	+678	>
6	Amazon	887	+287	2460	+699	>
7	Facebook	706	+250	2518	+712	>
8	VMware	584	+169	1133	+231	>
9	GitHub	536	+145	1473	+375	>
10	SAP	484	+158	1035	+221	>

< 1 2 3 ... 28 29 30 >

291 records

Top languages

TypeScript	15.0%
Python	14.6%
JavaScript	12.2%
Java	9.1%
Go	8.8%
C	7.3%

License usage

4%

37%

30%

29%

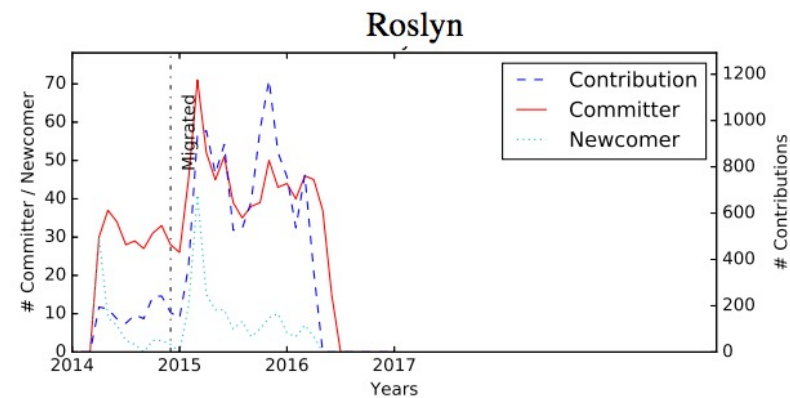MIT — Apache-2.0 — GPL-3.0 — Other
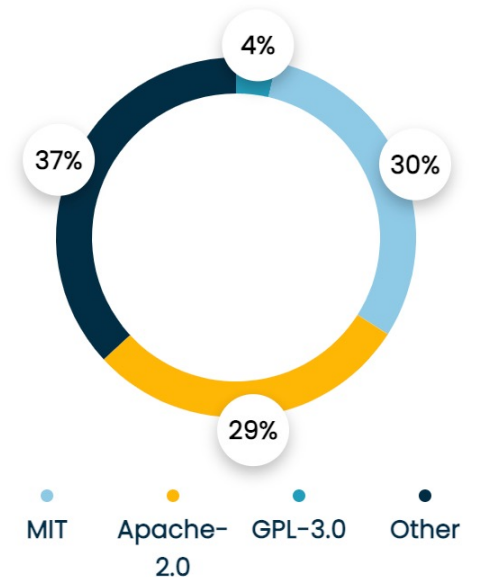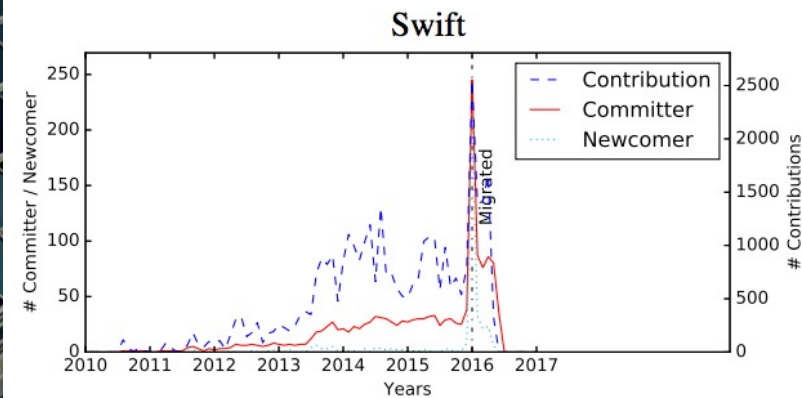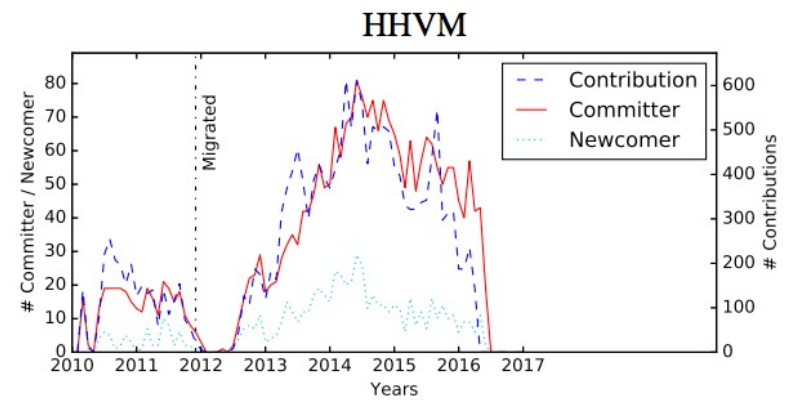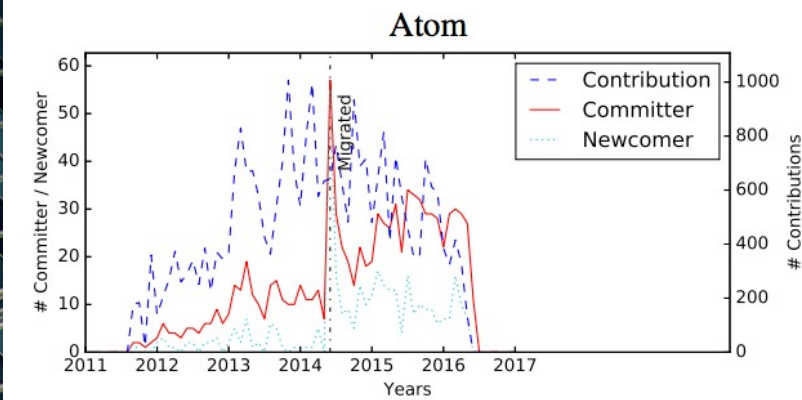
LEVERAGING THE CROWD

"The World Wide Web became a tool for bringing together the small contributions of millions of people and making them matter"

Collaboration in large-scale

Pinto, Steinmacher, Dias and Gerosa, On the Challenges of Open-Sourcing Prop
Empirical Software Engineering, 2018
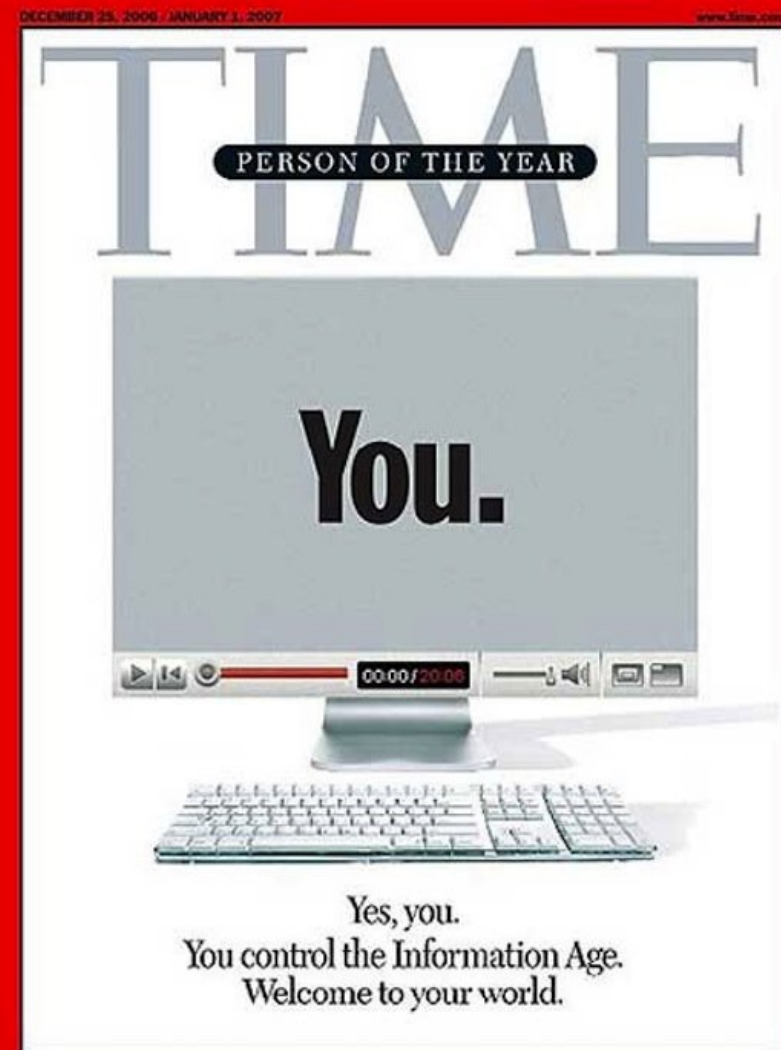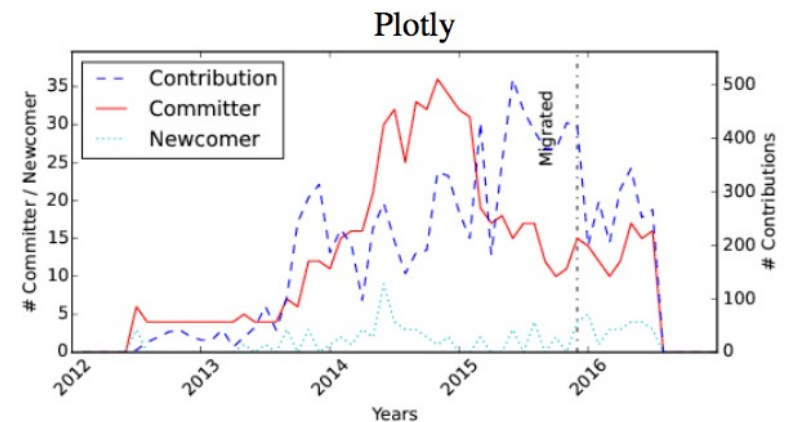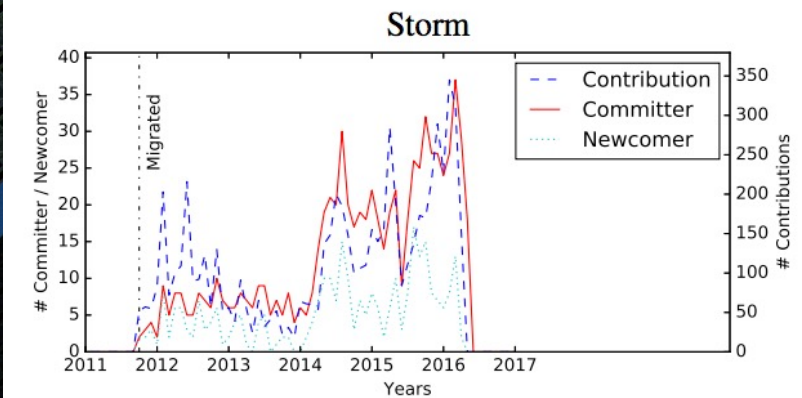
ALL ABOUT COMMUNITY!

- OpenStack:
  - **12.9M lines of code**
  - **37 programming languages (mainly Python)**
  - **11.6K+ code contributors**

- Mozilla Firefox:
  - **24.5M lines of code**
  - **48 programming languages**
  - **8.5K+ contributors (1K+ in the last 12 months)**

- Swift:
  - **800K lines of code**
  - **1K+ contributors**

https://www.openhub.net/p/openstack
https://www.openhub.net/p/firefox
https://www.openhub.net/p/apple_swift

ASSIGNMENT!

- **TRABALHO EM GRUPO**
- **Prazo:** 14/03/2022 (Segunda) antes da aula
- **Como submeter**: Enviar um email para igorfs@utfpr.edu.br, com assunto "SL2022: Tarefa 1"
  - Buscar referências sobre modelos de negócio e meios de ganhar dinheiro com software livre
  - Produzir um documento que apresente diferentes maneiras de criar negócios que se baseiem em software livre. Ressaltar os pontos importantes relativos a cada uma das maneiras encontradas e citar exemplos.
  - Escolher um dos modelos de negócio e analisar com mais profundidade. Pensar em como seria possível que vocês criassem uma empresa em torno disso. Pense no público alvo, nos tipos de serviços oferecidos, nos desafios e nos benefícios que tal empresa pode ter.

FOR MY DEAR PPGCC STUDENTS

- Cada aluno deverá apresentar **DOIS** seminários de pesquisa, discutindo artigos científicos dentro da área de comunidades de software livre (gestão de comunidades, licenças, criação, manutenção, entrada de pessoas, diversidade, etc.)

- As apresentações estão agendadas para **04/04/2022 e 06/06/2022**

- Os alunos devem escrever uma **resenha** de cada artigo e enviar como parte do trabalho

- As apresentações devem ter entre **10 e 12 minutos**