



# Who is behind open source?

---

Gartner Open Source Summit,

Barcelona, June 14, 2006

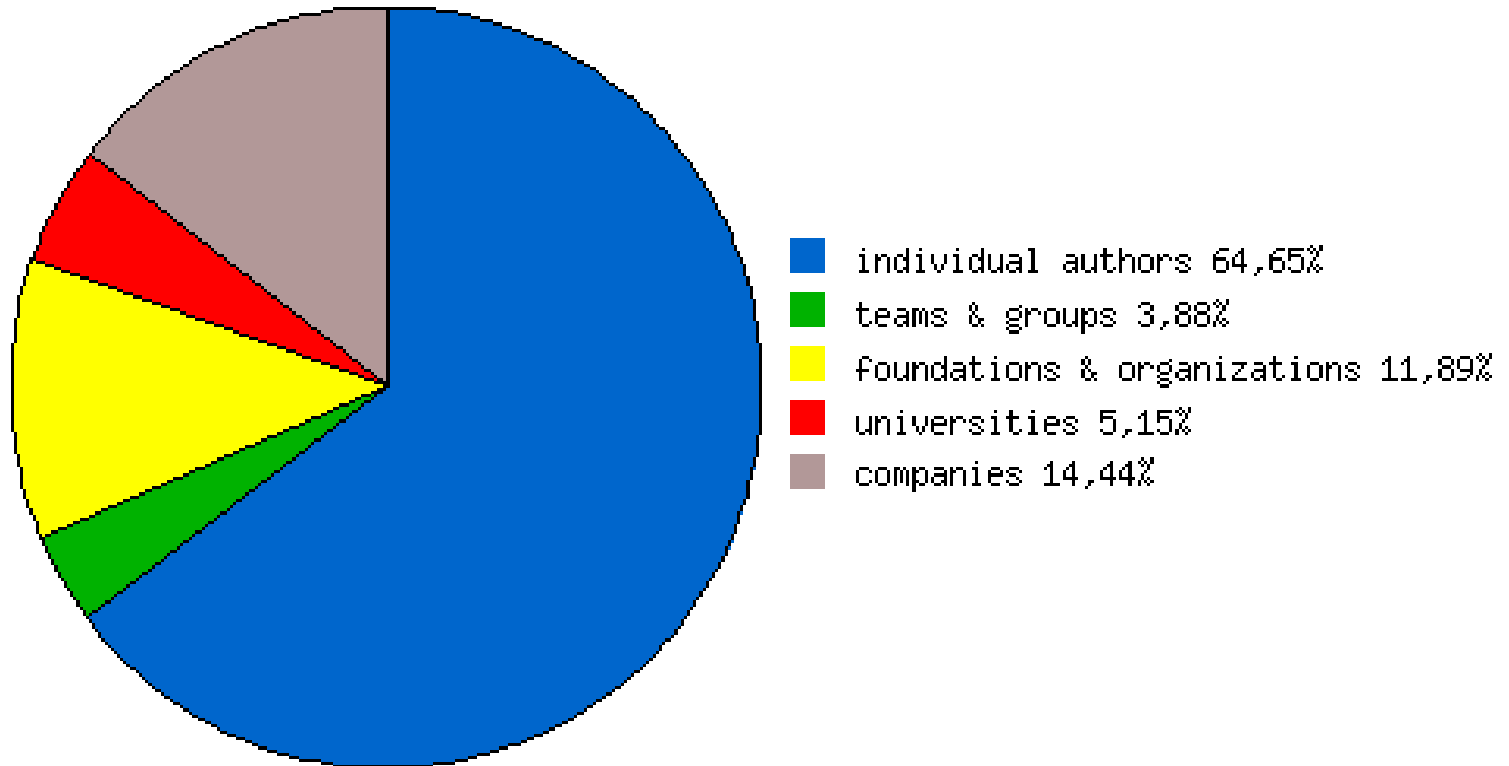
Rishab Aiyer Ghosh

[rishab@dxm.org](mailto:rishab@dxm.org)

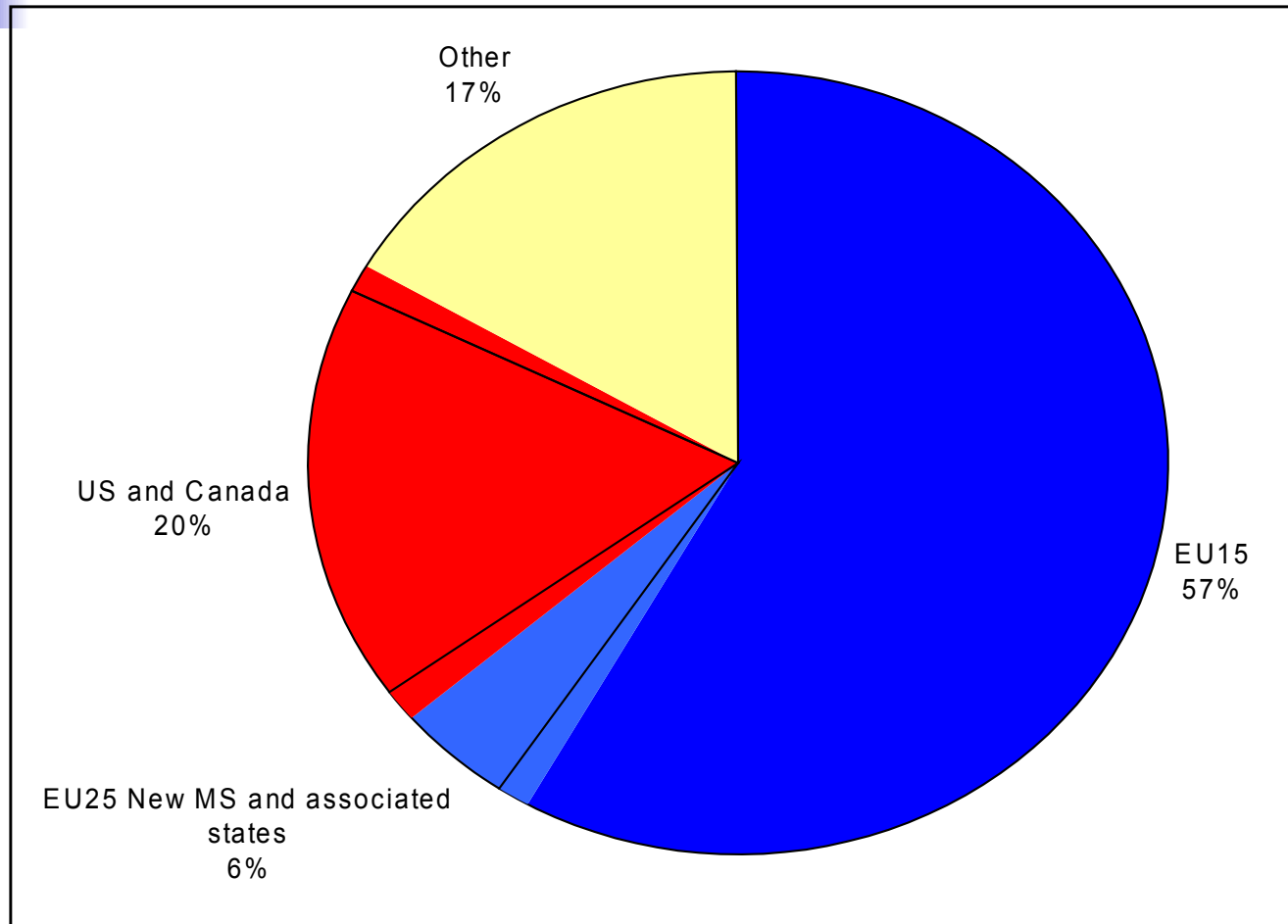
UNU-MERIT

This presentation is based on work conducted in the projects CALIBRE, funded by the European Union's FP6 IST programme

# Who creates open source

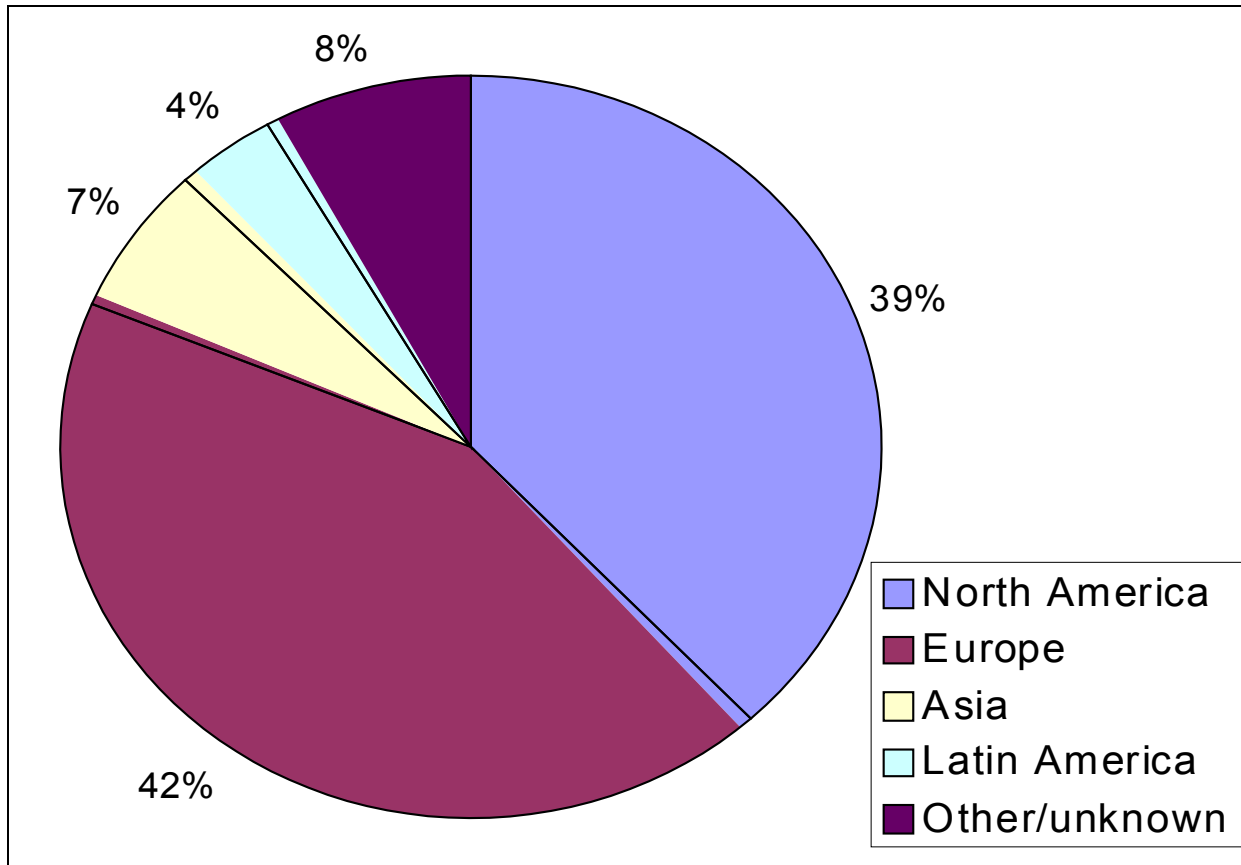


# Where developers live



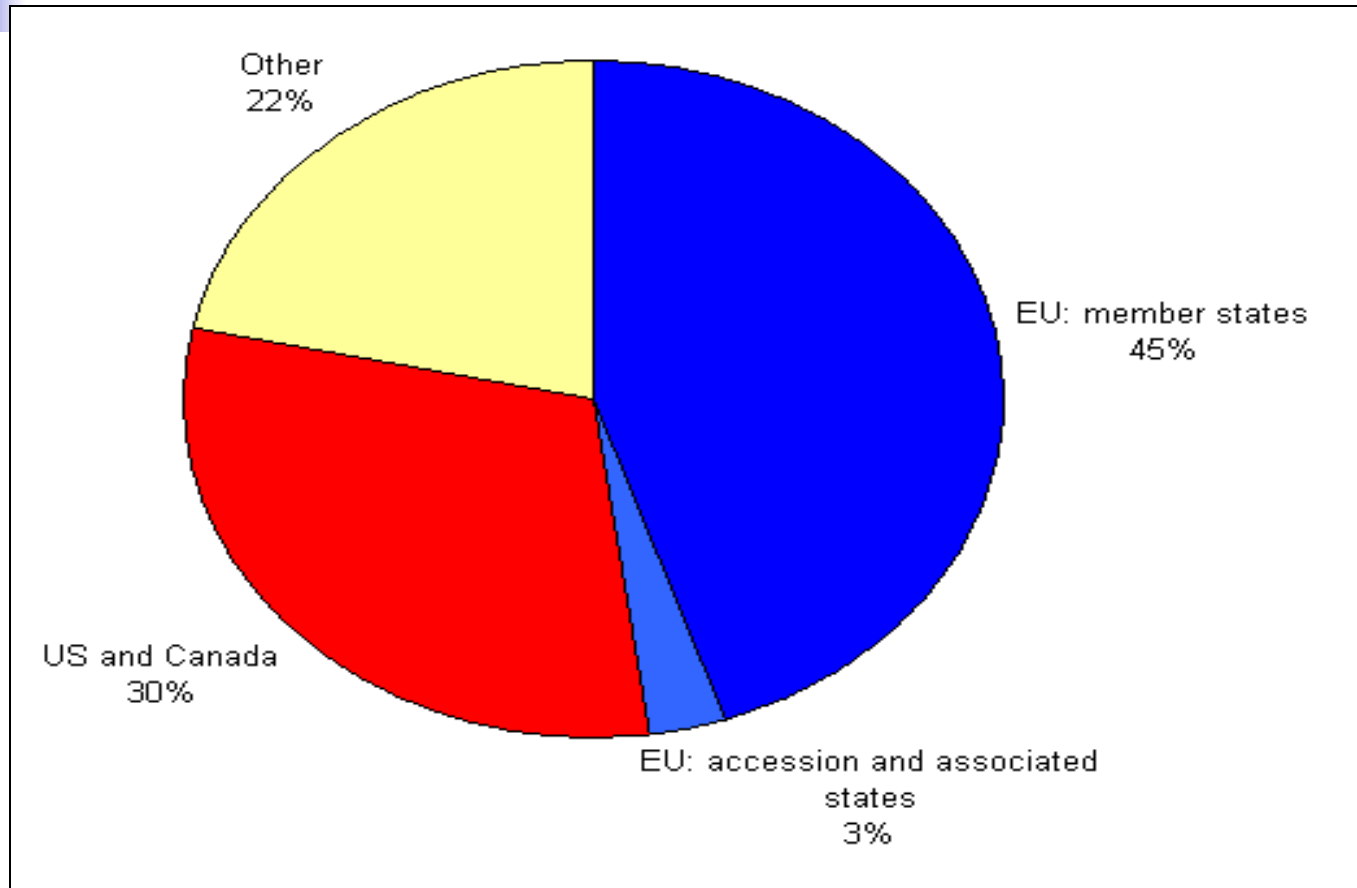
(Surveys totalling 5000+ developers)

# Where developers live



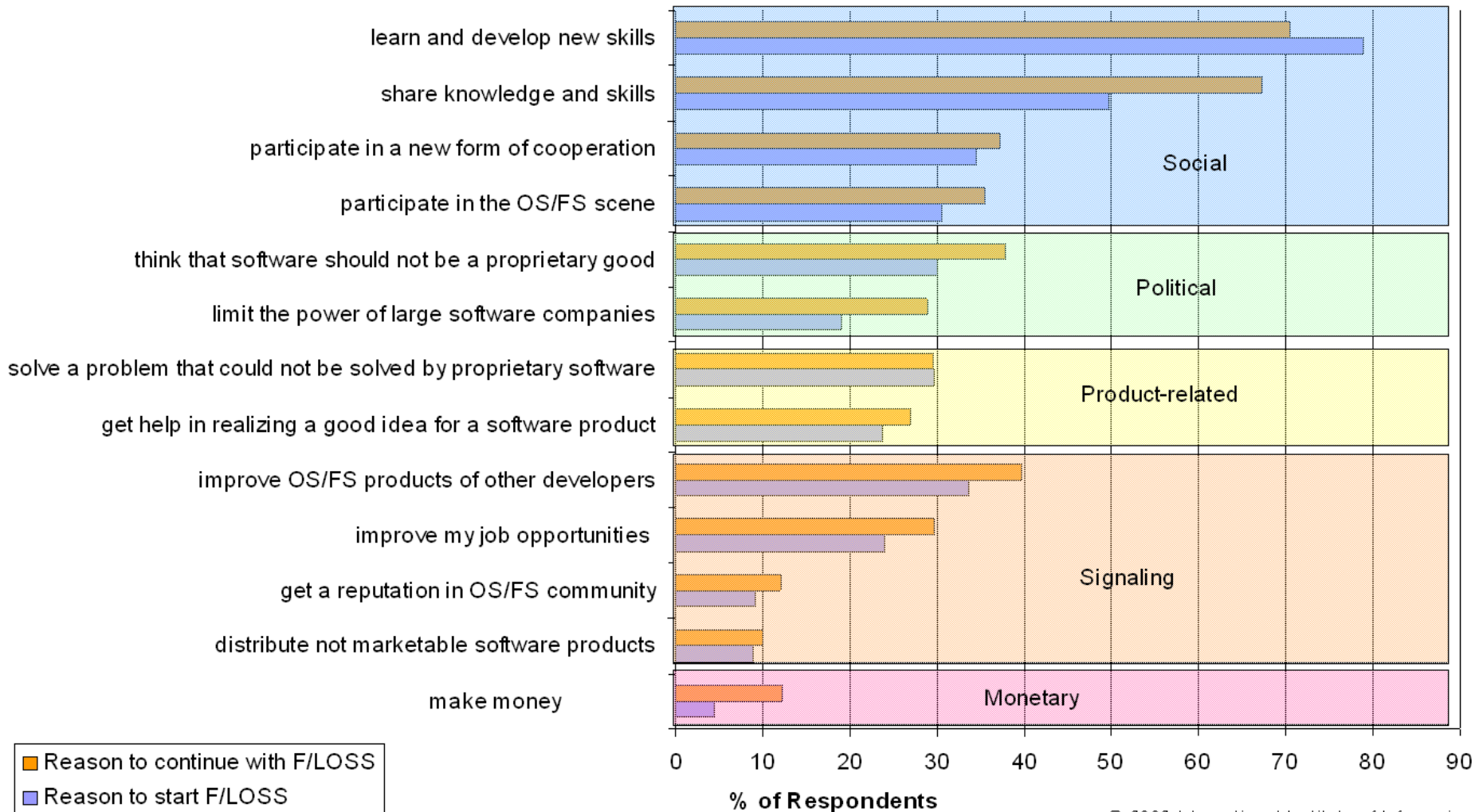
(Developers on Sourceforge)

# Where developers lead from

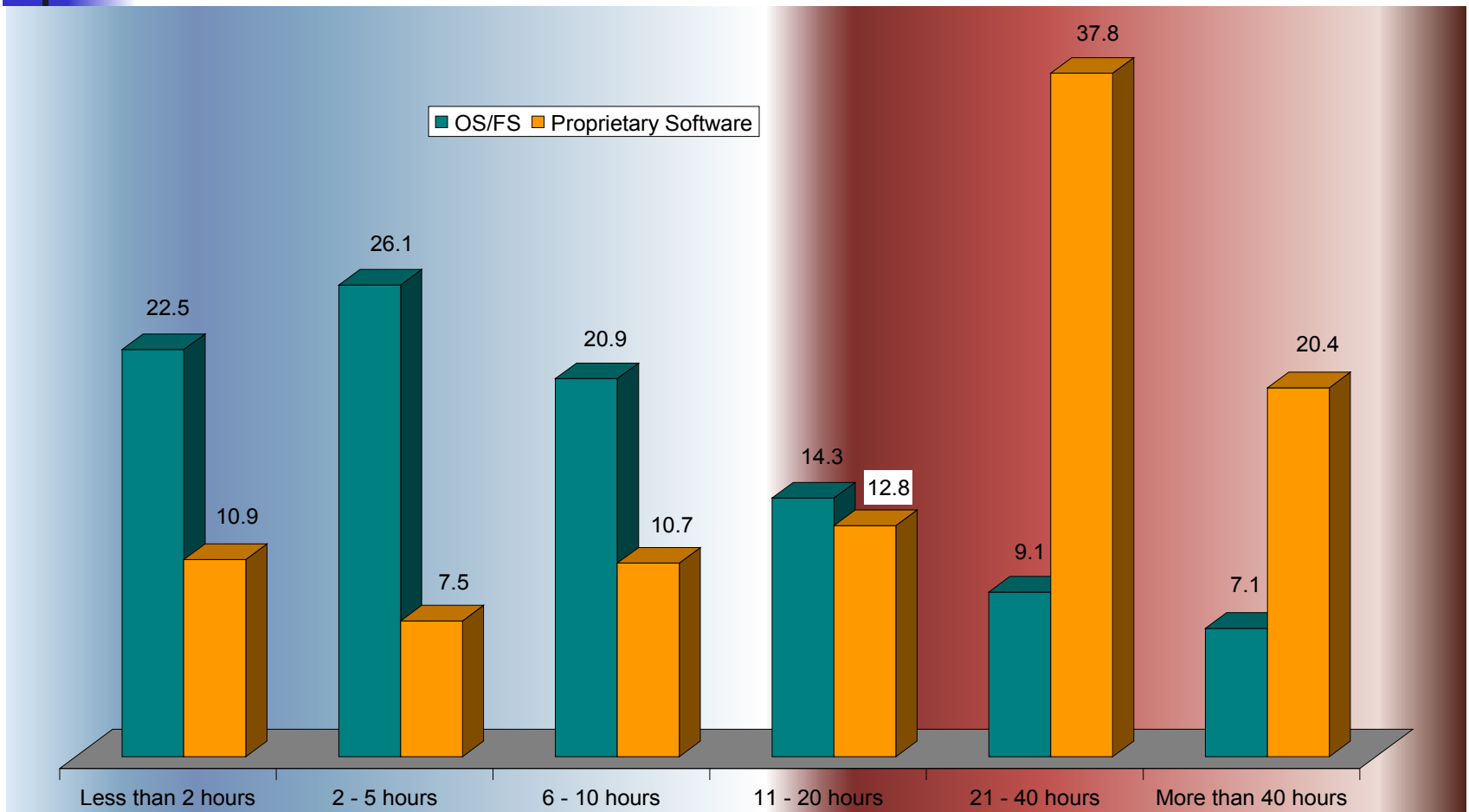


(leaders and maintainers of projects included in debian)

# Why they do it

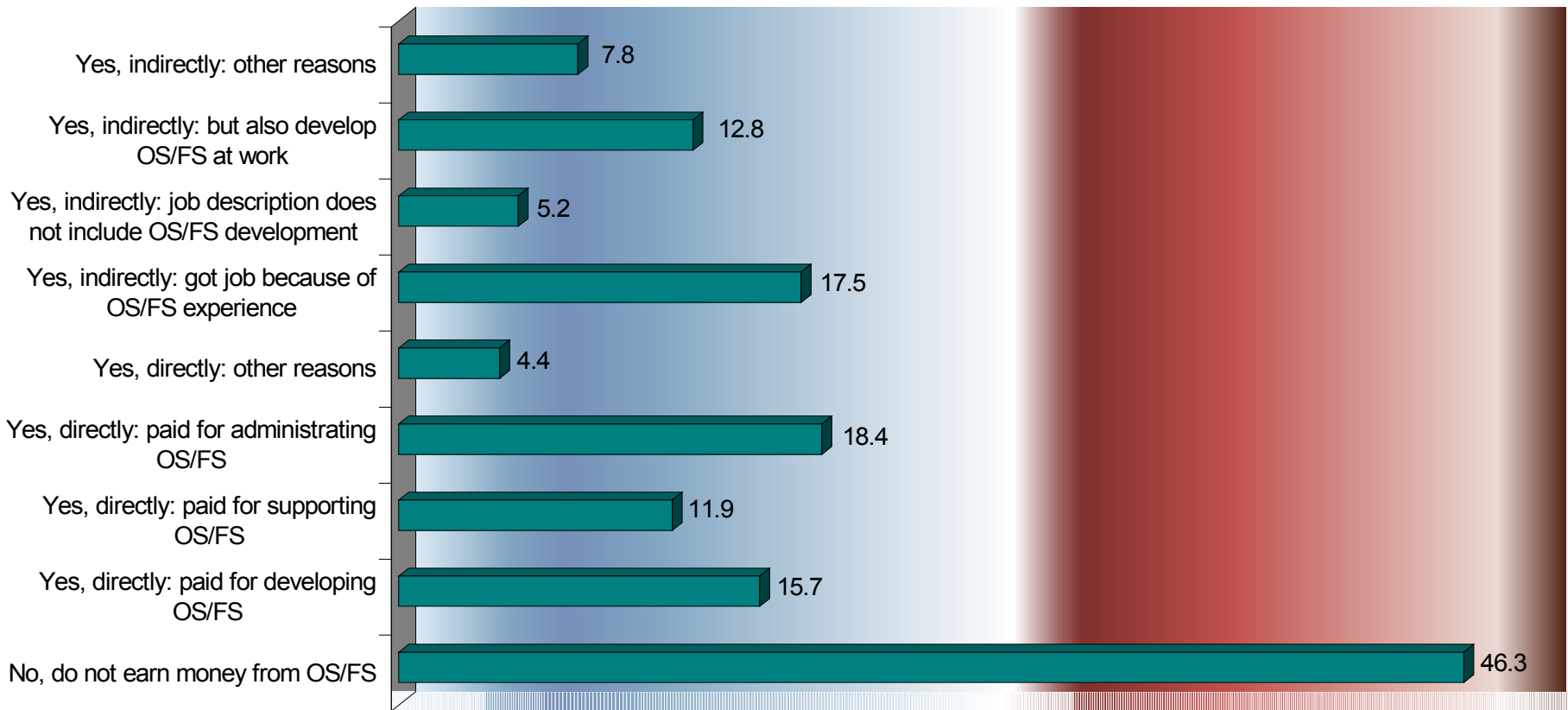


# Time spent





# ... is rewarded





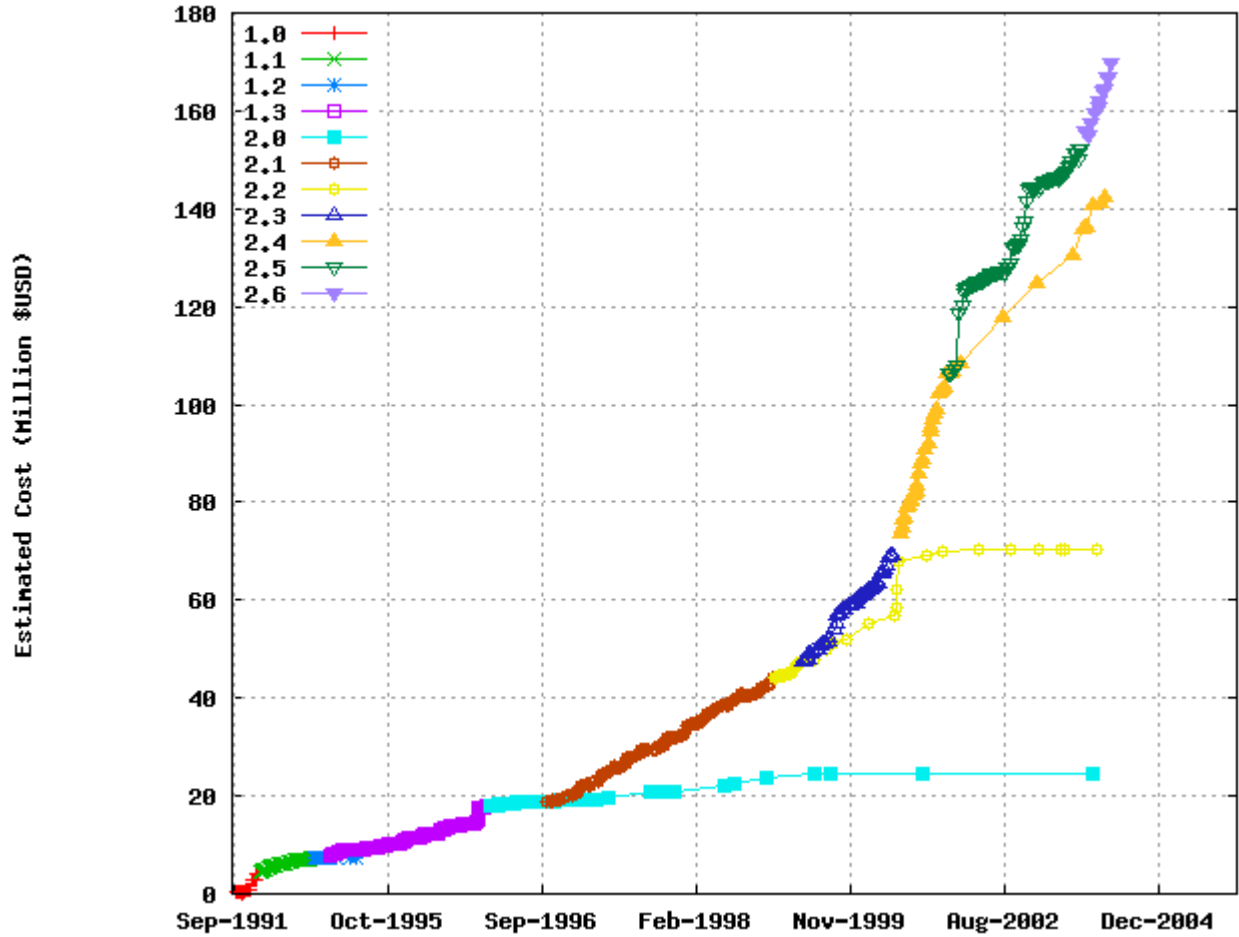
# Valuing FLOSS

---

- Example: Debian 2.2 GNU/Linux (2001)
  - Source lines of code: 55,201,526 (of which the Linux kernel forms under 6%)
- If Debian was written in a software company:
  - Estimated effort: 14,005 person years
  - Estimated schedule: 6.04 years (team of 2,318!)
  - Development cost: US\$ 1,891,990,000

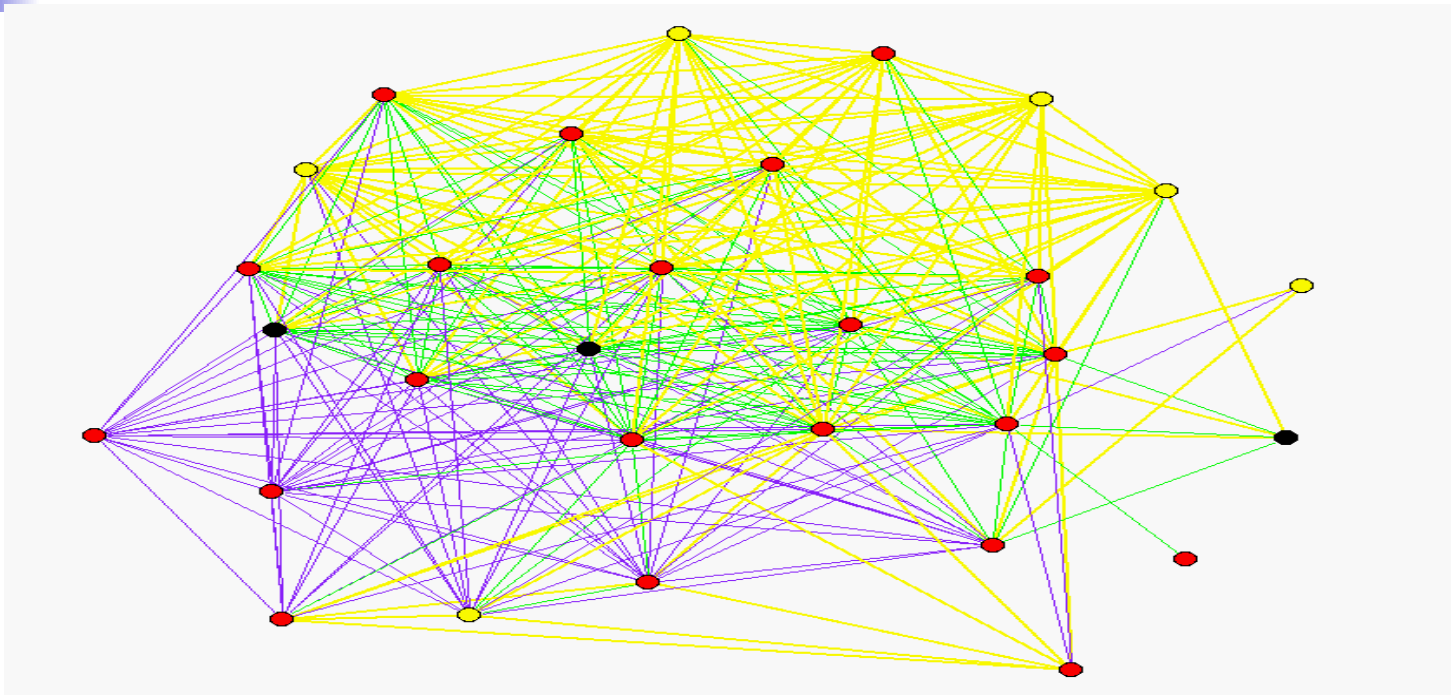
(Source: "Counting potatoes" by Gonzalez-Barahona et al)

# Linux kernel substitution cost



Linux kernel versions 1.0 to 2.6. Source: URJC

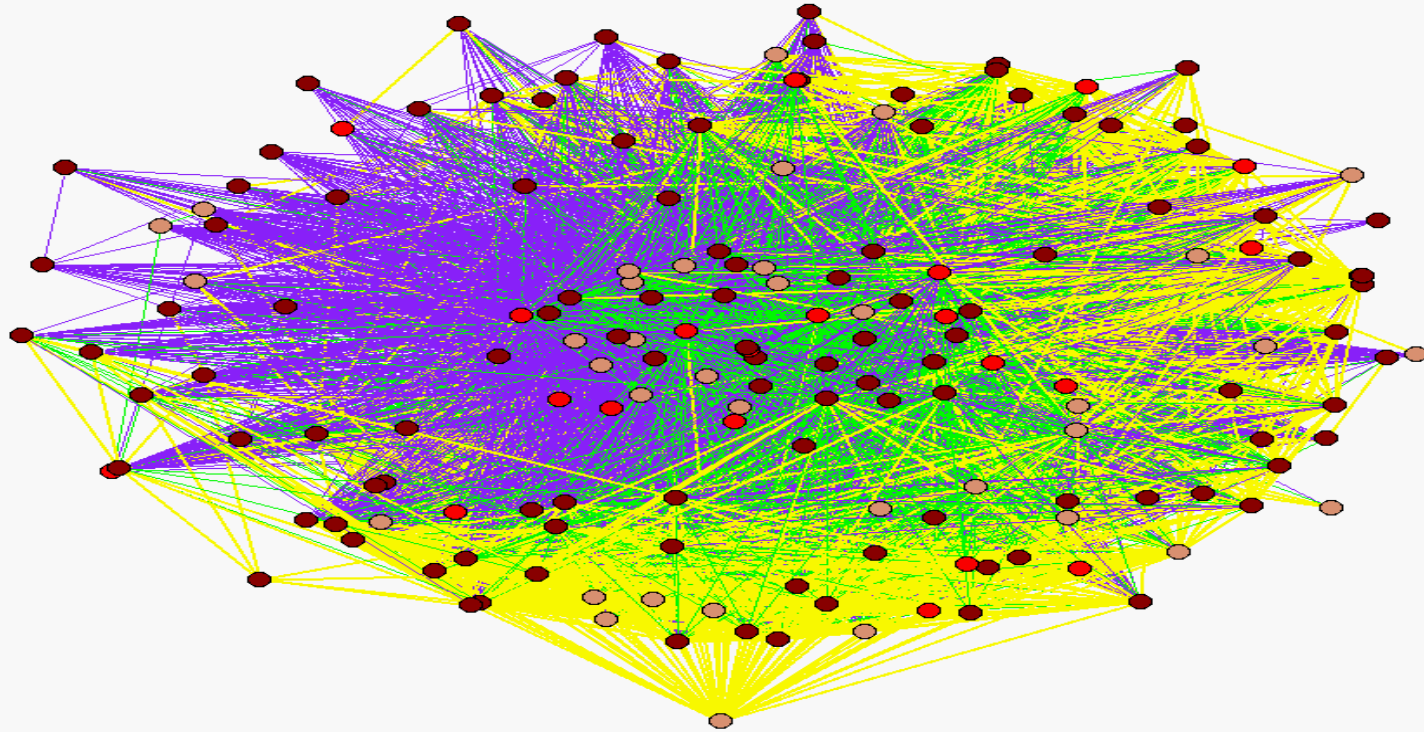
# Linux kernel v1.0: 158 authors



1994. Nodes are 30 modules. Arcs represent **common authors**, **code dependencies**, or **both**

(Source: "Nature and composition... ", Rishab Ghosh & Paul David)

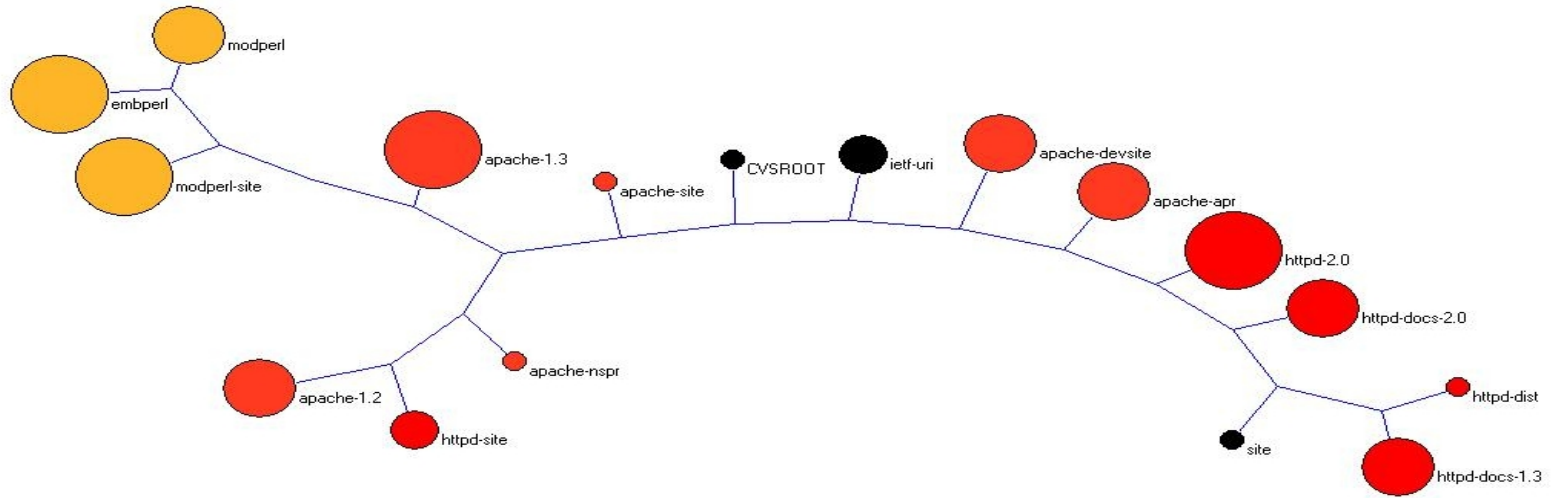
# Linux kernel v2.5.25: 2263 authors



2002. Nodes are 169 modules. Arcs represent **common authors**, **code dependencies**, or **both**

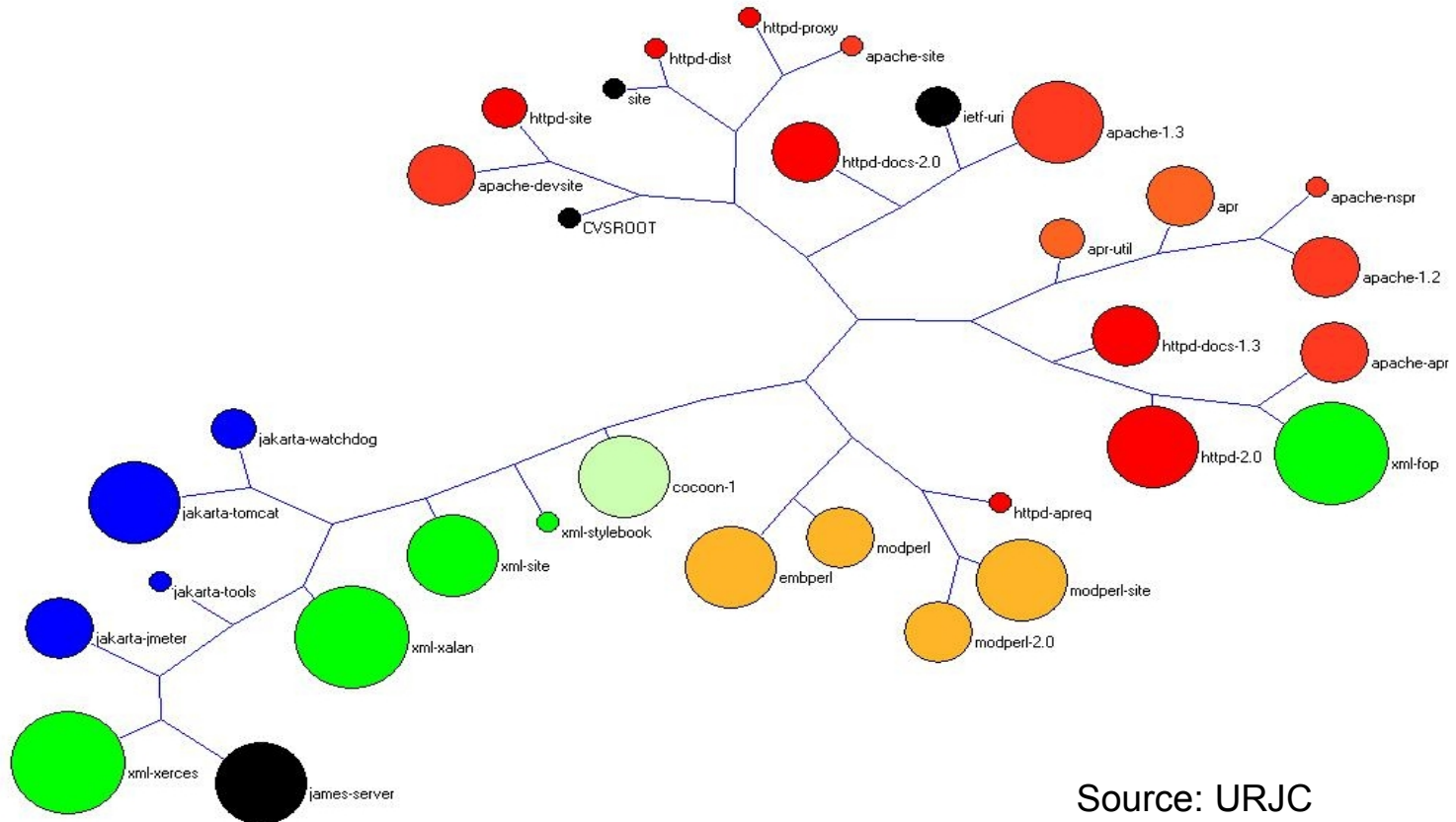
(Source: "Nature and composition... ", Rishab Ghosh & Paul David)

# Social networks: Apache mods 1999



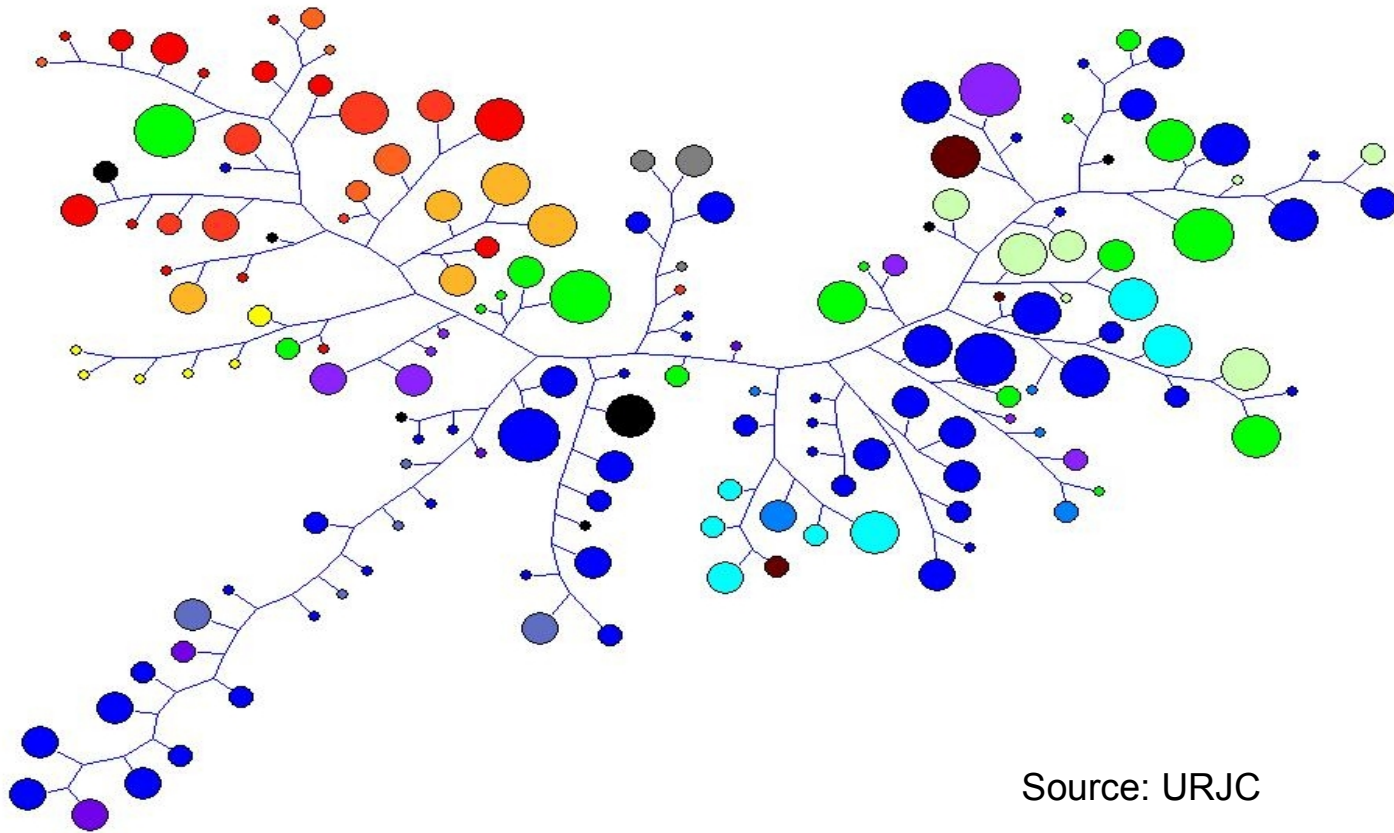
Source: URJC

# Social networks: Apache mods 2000



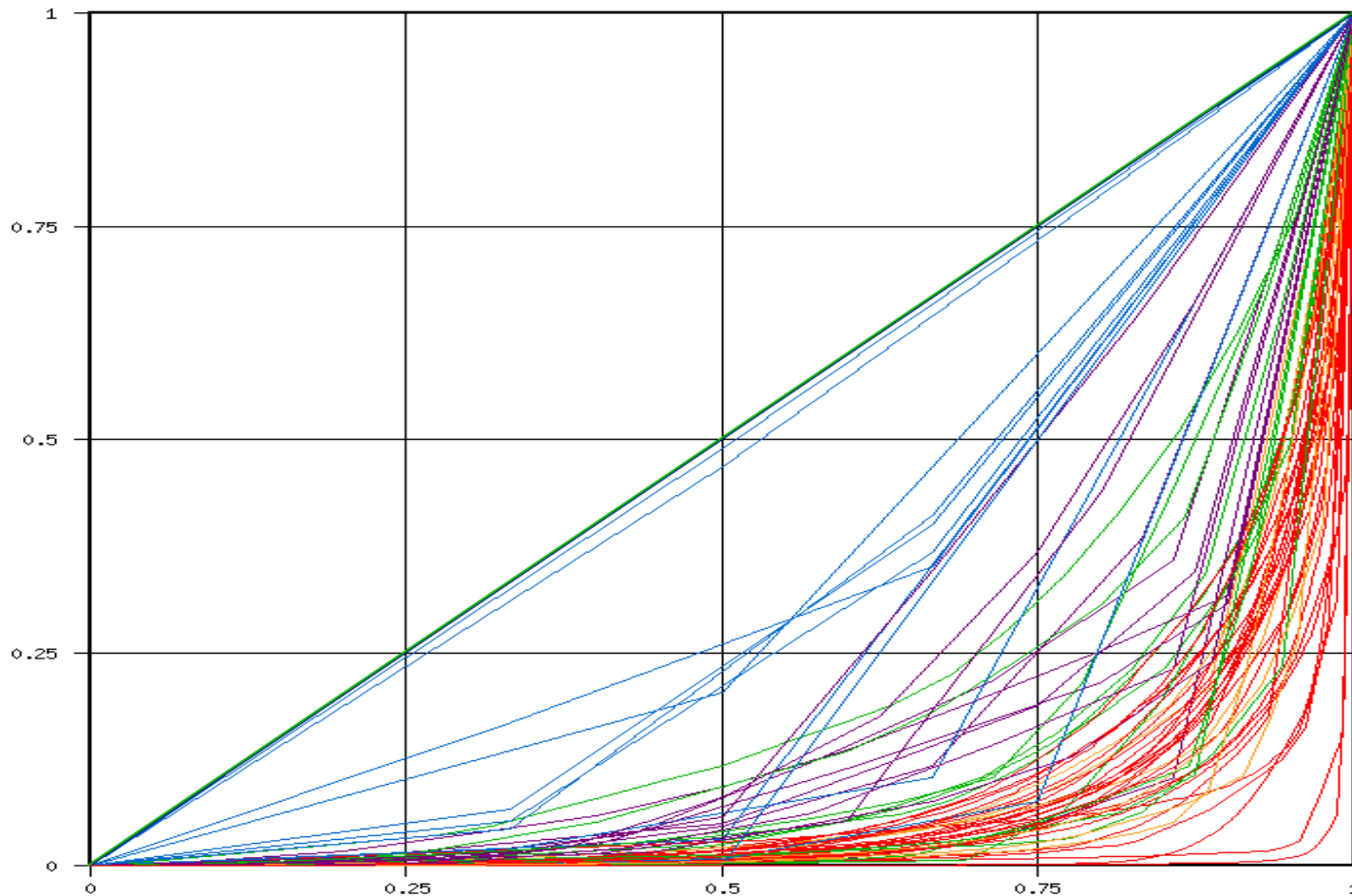
Source: URJC

# Social networks: Apache mods 2004



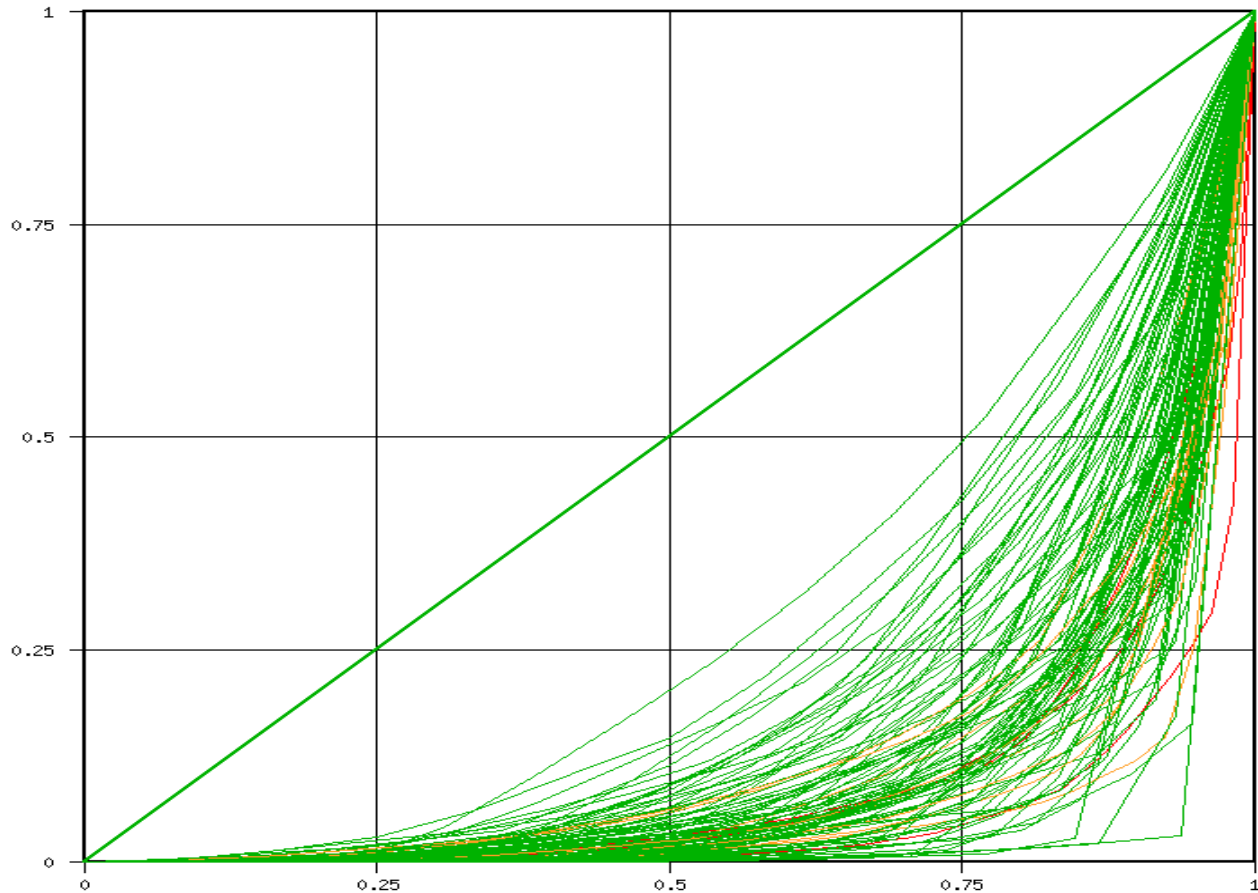
Source: URJC

# Distribution of "authorship"



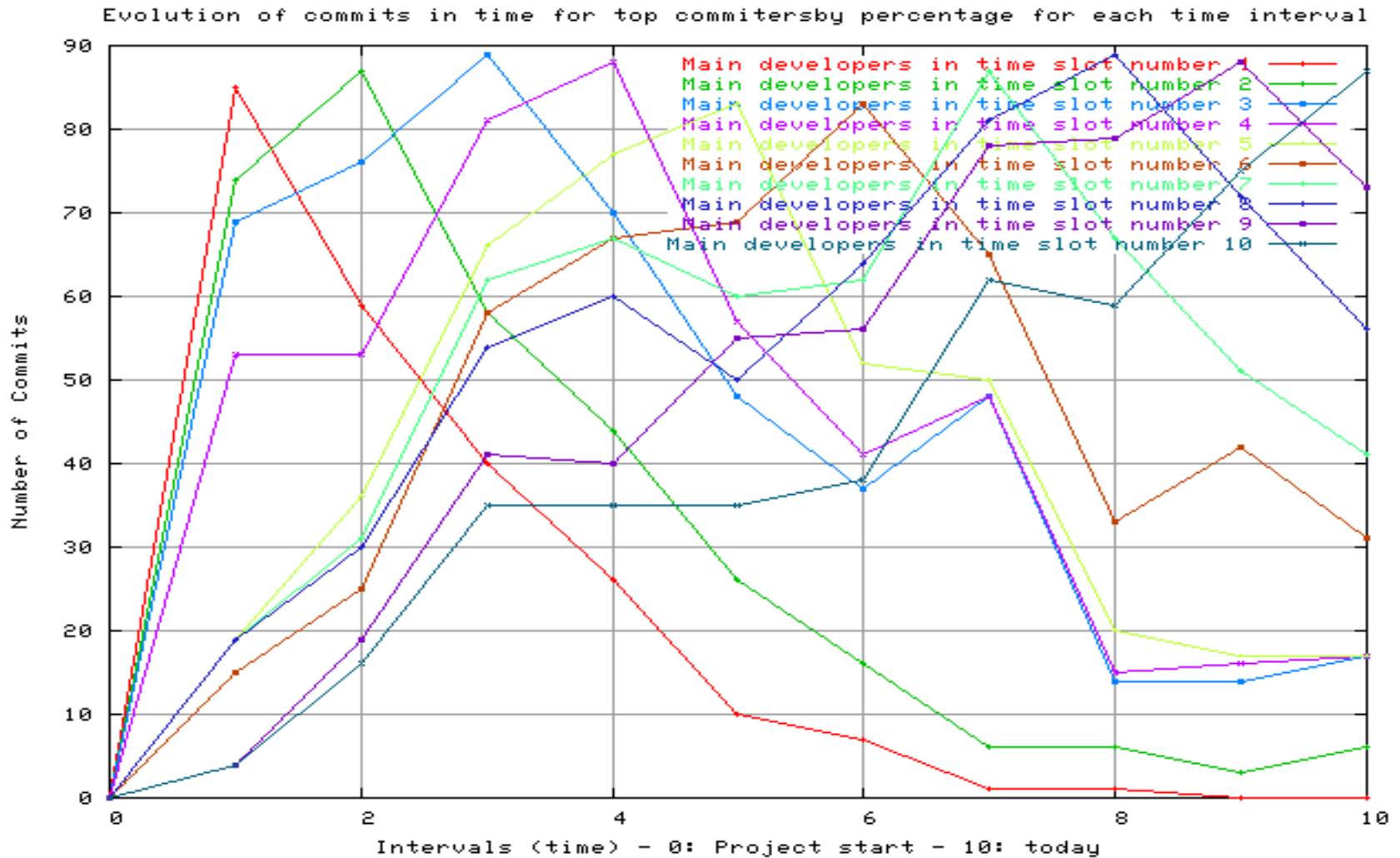
Lorenz curves for KDE modules: "hot" colours are for modules with more developers  
(Ghosh & Robles)

# Distribution of "authorship"



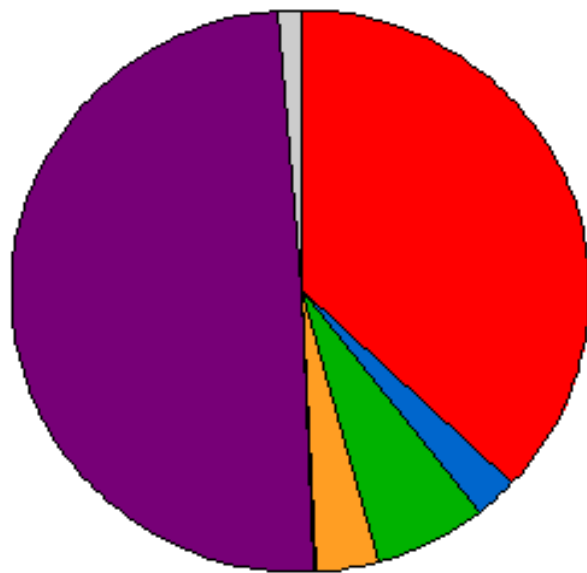
Lorenz curves for SourceForge projects with **>11** and **>50** developers  
(Ghosh & Robles)

# Generations of "core" developers: FreeBSD

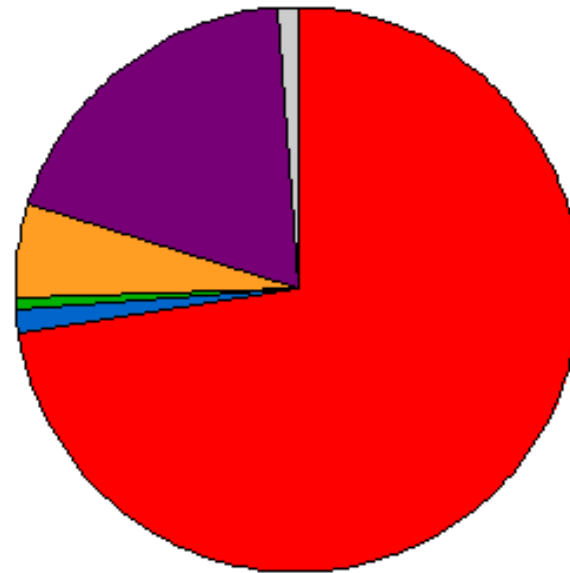


Over 10 years. Source: Barahona & Robles 2003, "Code Gods..."

# What is "development"? filetypes: KDE



KDEbase



KOffice

- development
- documentation
- i18n
- images
- sound
- ui
- unknown



## More information

---

CALIBRE: <http://www.calibre.ie>

FLOSSPOLS: <http://flosspols.org>

References:

- FLOSS report: <http://flossproject.org/report/>
- “Nature and composition...” (Ghosh & David):  
<http://dxm.org/papers/licks1/>
- Other papers: <http://libresoft.urjc.es>