Global Software Development in the FreeBSD Project

Diomidis Spinellis

Athens University of Economics and Business

Outline

- **FreeBSD**
- **Methodology**
- **Findings**
  - Global development
  - Productivity and quality
  - Human interactions

FreeBSD

- Complete operating system
  - High performance application servers
  - Networked embedded systems
- Derived from BSD Unix
- Focus on stability and reliability
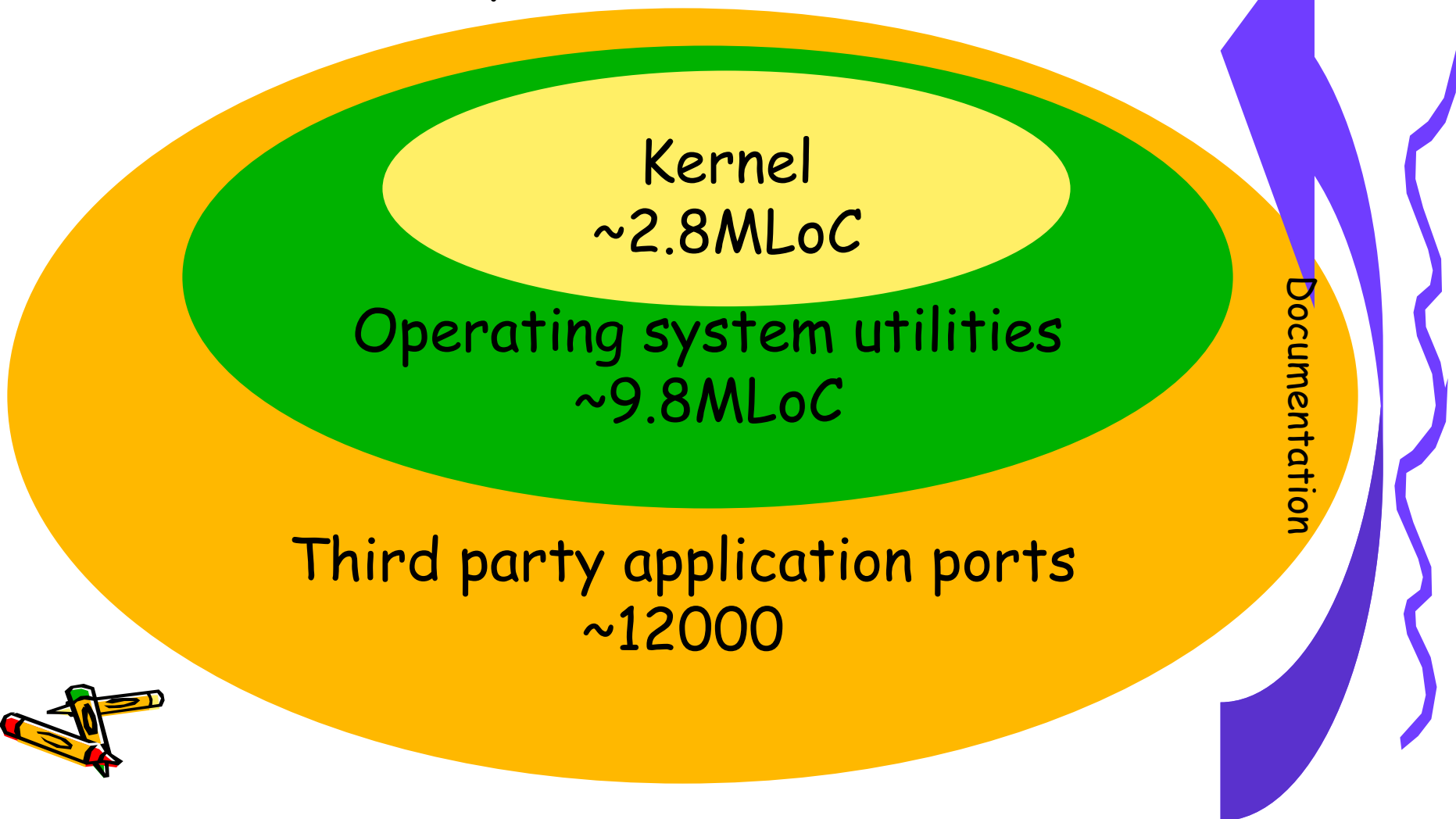 - Yahoo's servers
  - Part of Apple's Mac OS X

Development

- ~350 developers

Kernel
~2.8MLoC

Operating system utilities
~9.8MLoC

Third party application ports
~12000

Documentation

Development Facilities

- CVS (10 years of historical data; >938k)
- Problem reports (>95k)
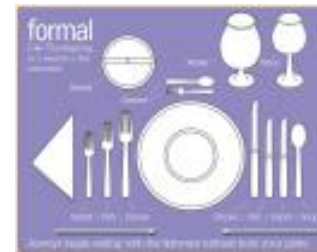- Mailing lists (>100)
- Tinderbox
- Web site
  - Developer's handbook
  - Release engineering
  - Browsable CVS
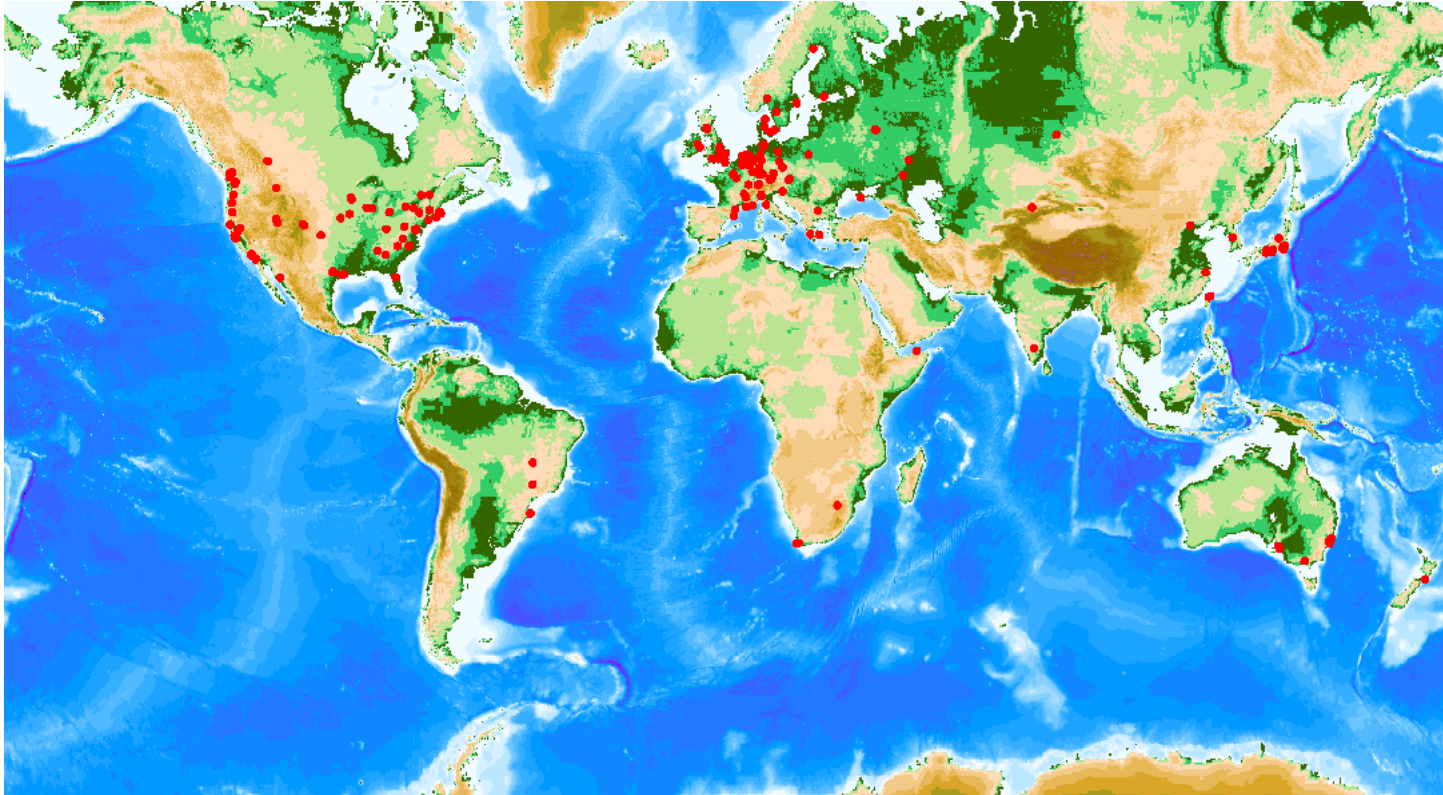 - Mailing lists
  - PRs

Development Process

- Unpaid volunteers
- Teams
  - Core
  - Release engineering
  - Security
  - Ports
  - Donations
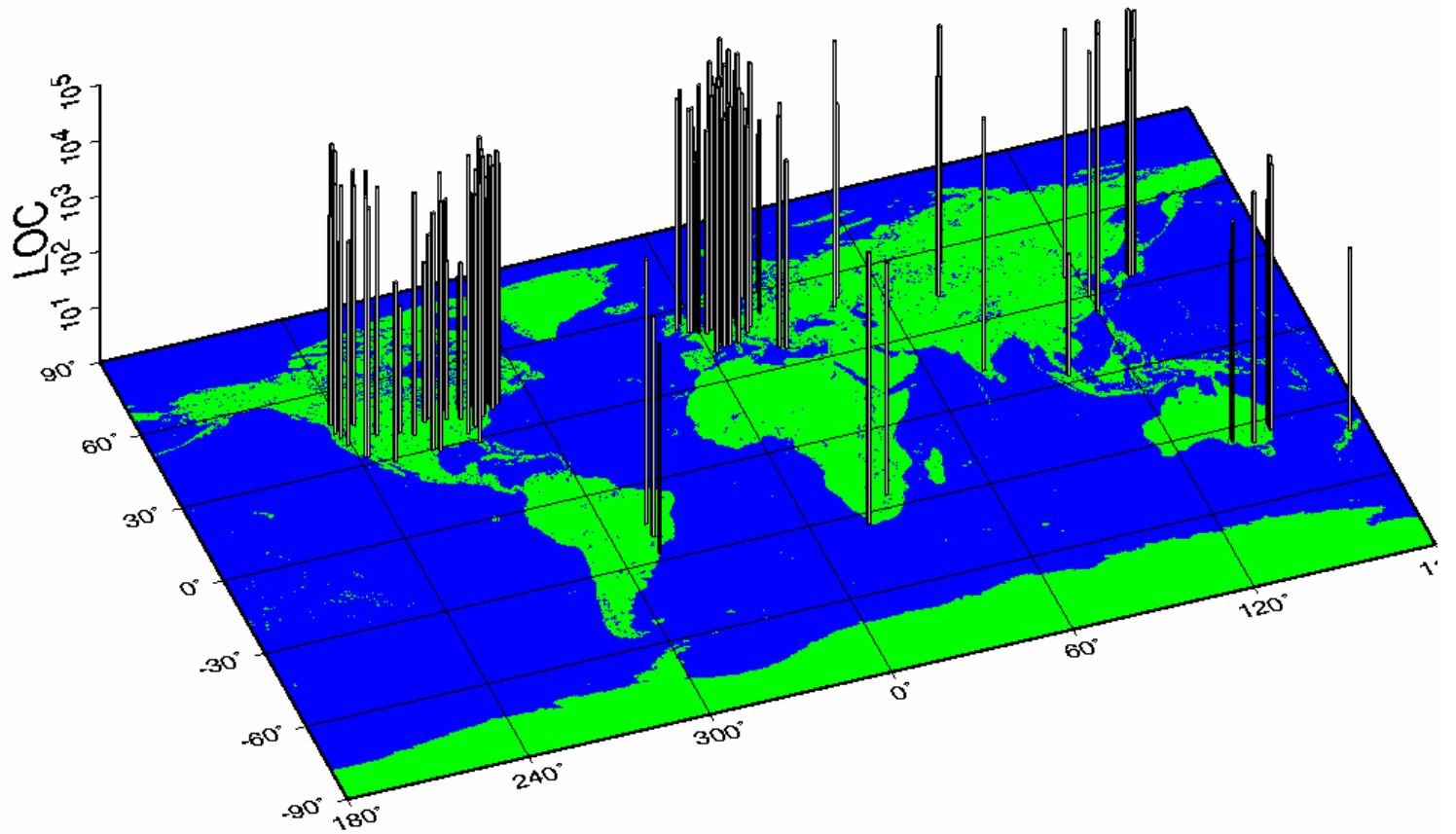- Committer etiquette
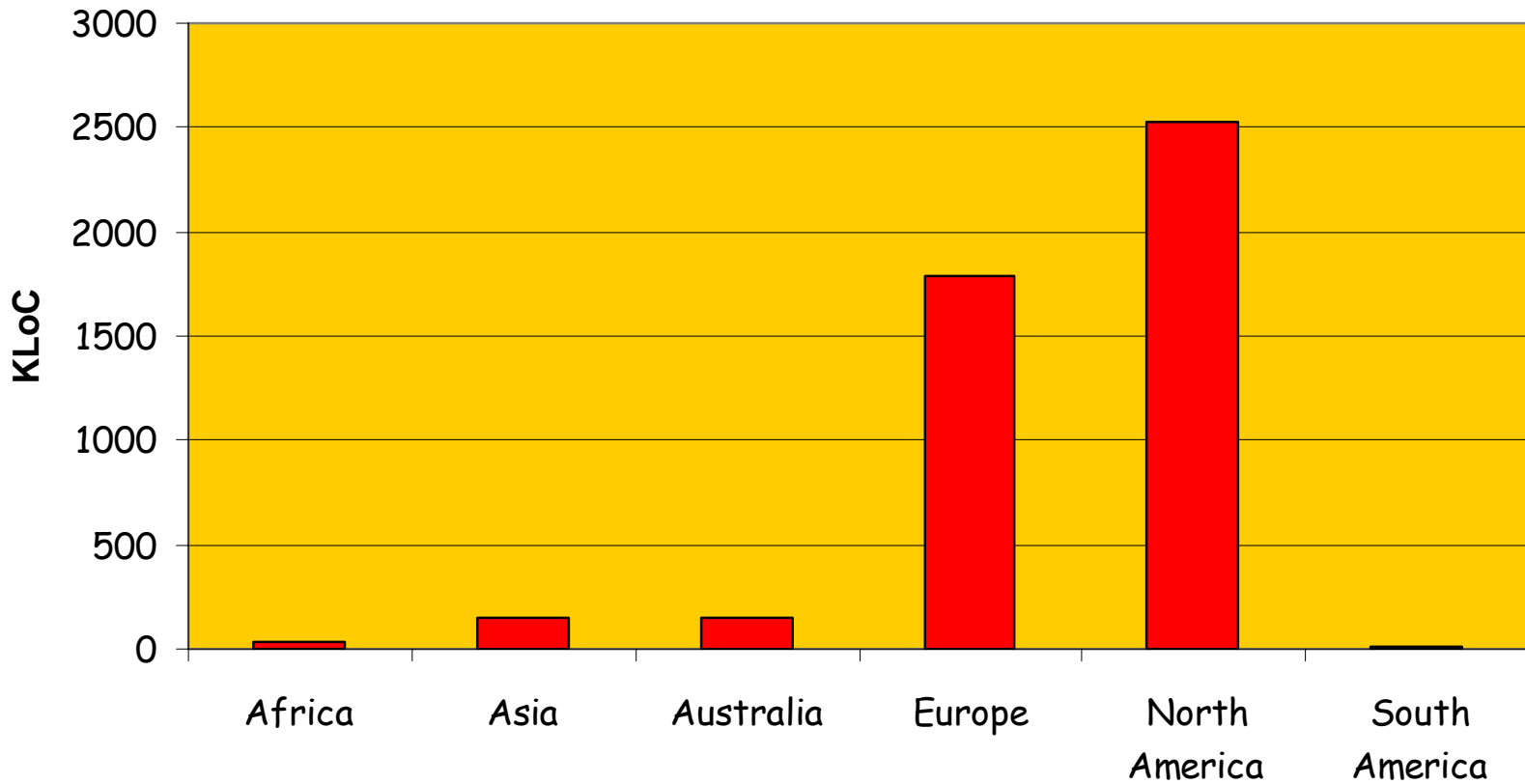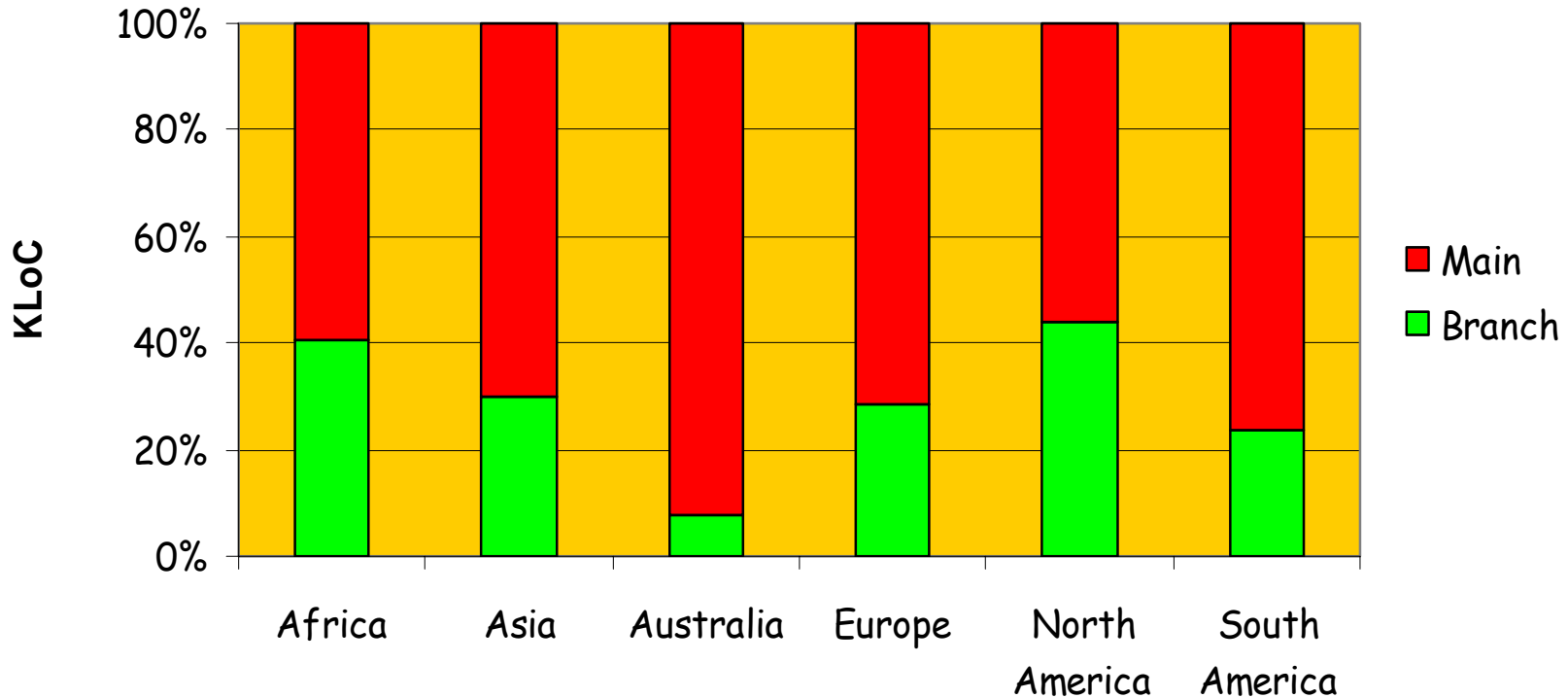- Mentoring

Global Development

Around the Globe

Division of Work

Hold and Advance

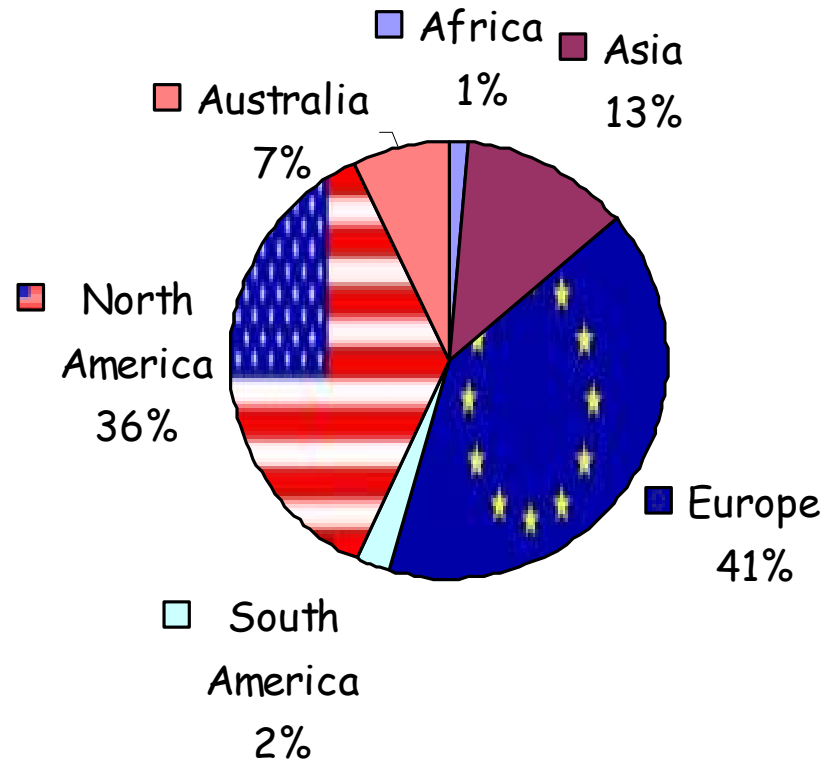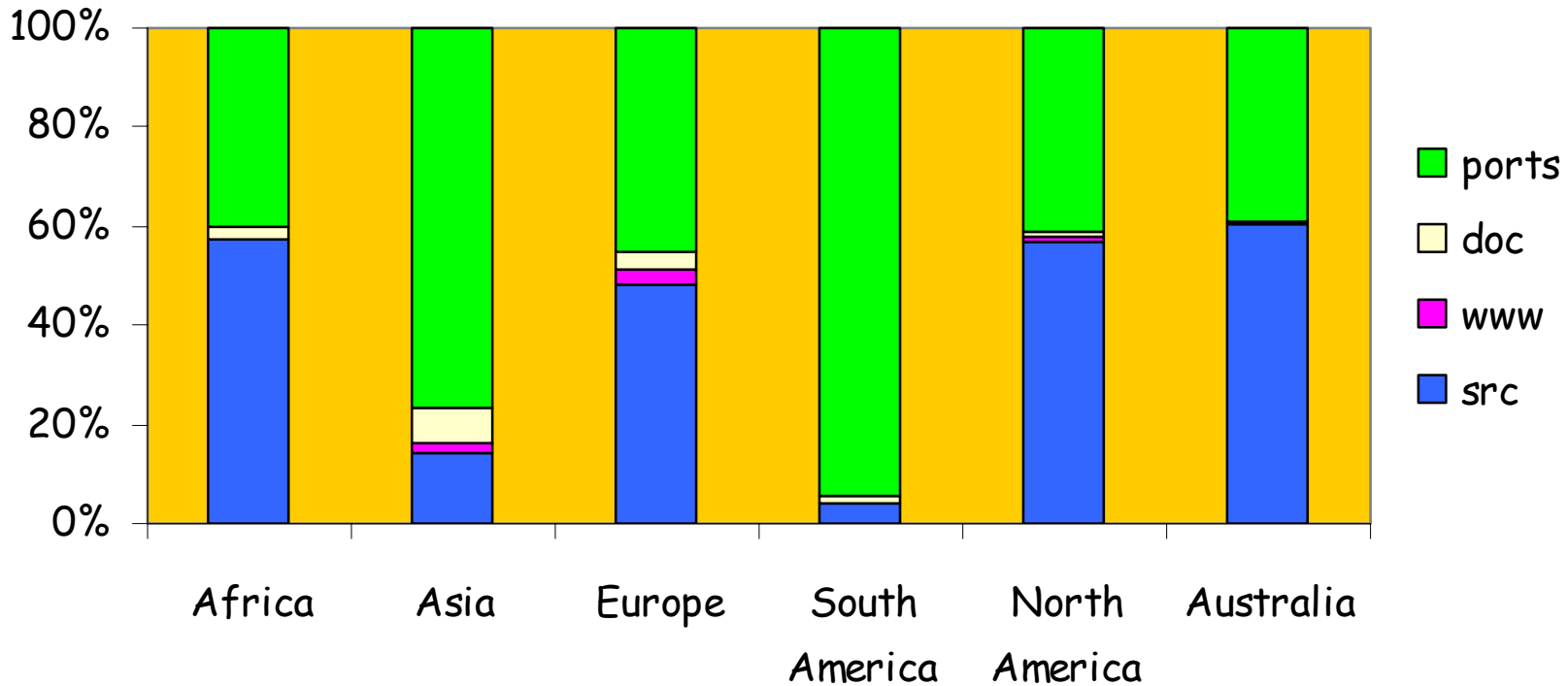
Work type per area

… and Fix

Closed GNATS entries per area

Roads and Buildings

Work type per area
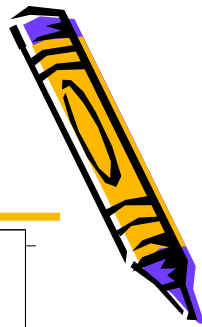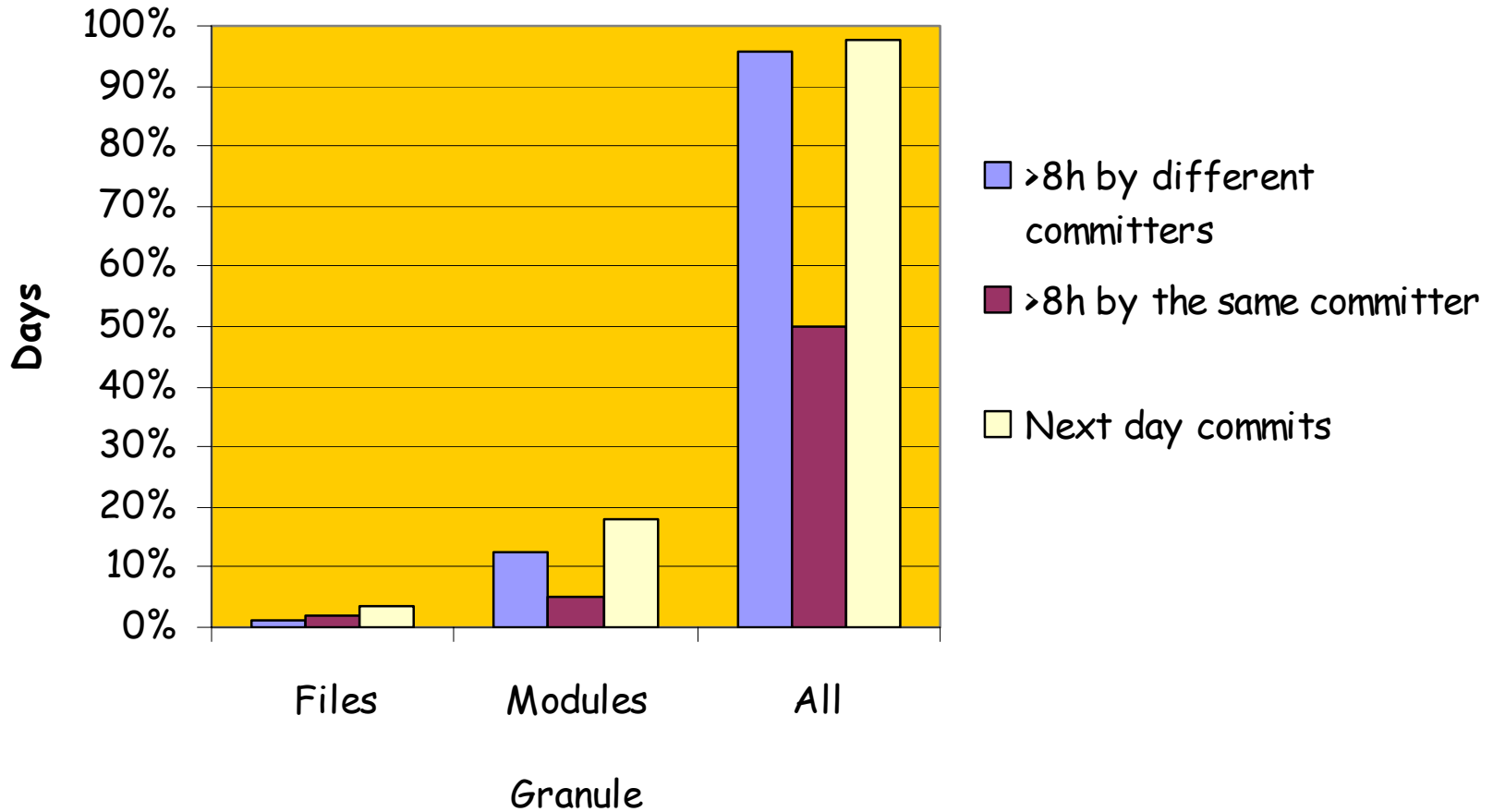
Round the Clock

Passing Work Around

- >8h by different committers
- >8h by the same committer
- Next day commits
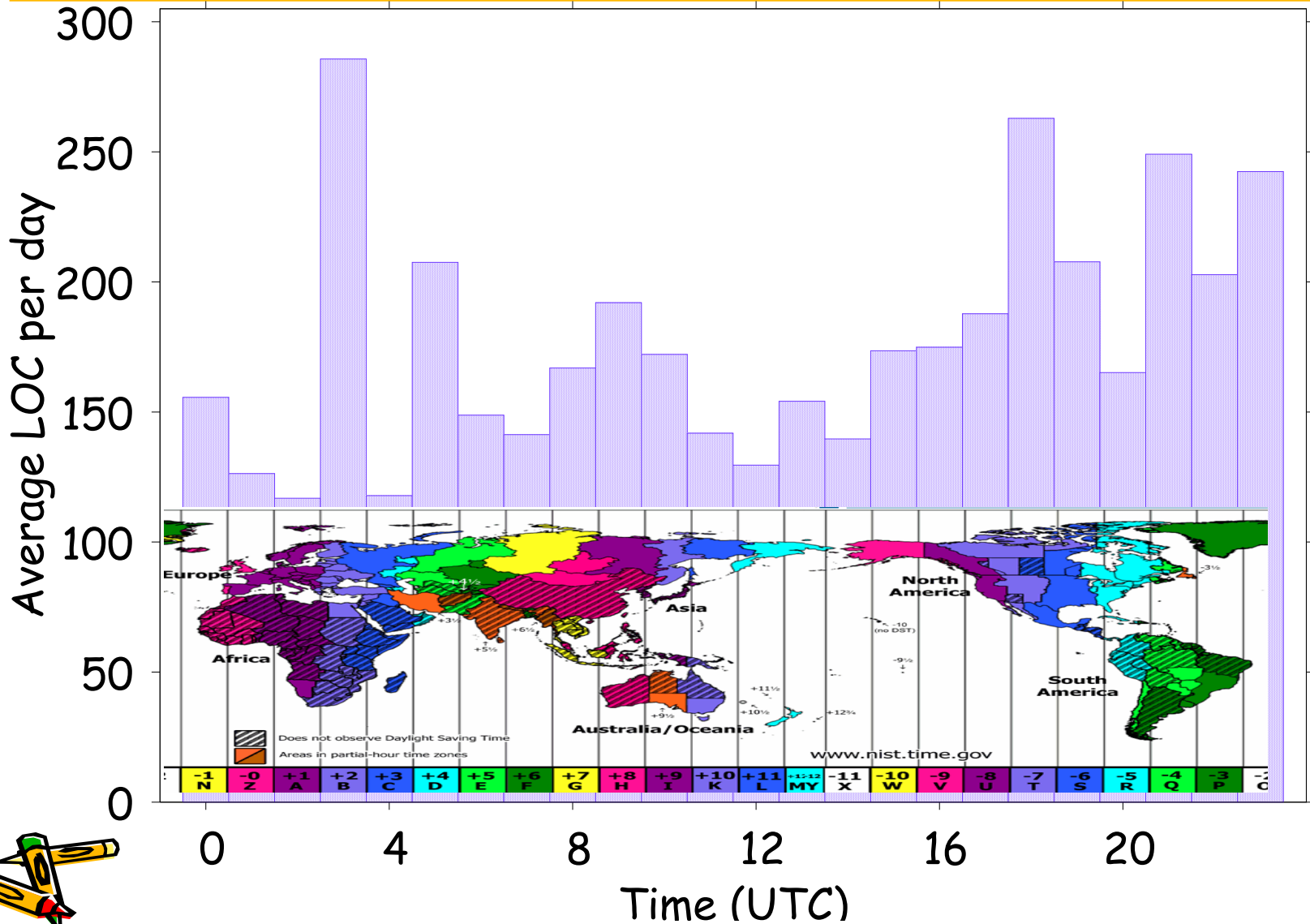
Productivity and Quality

+ Round the clock development
+ Software and hardware diversity
− Lack of face-to-face communication
− Asynchronous communication
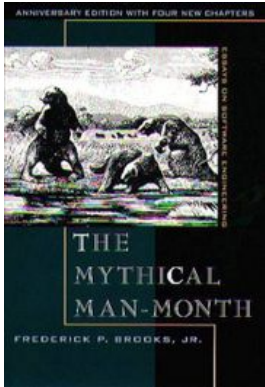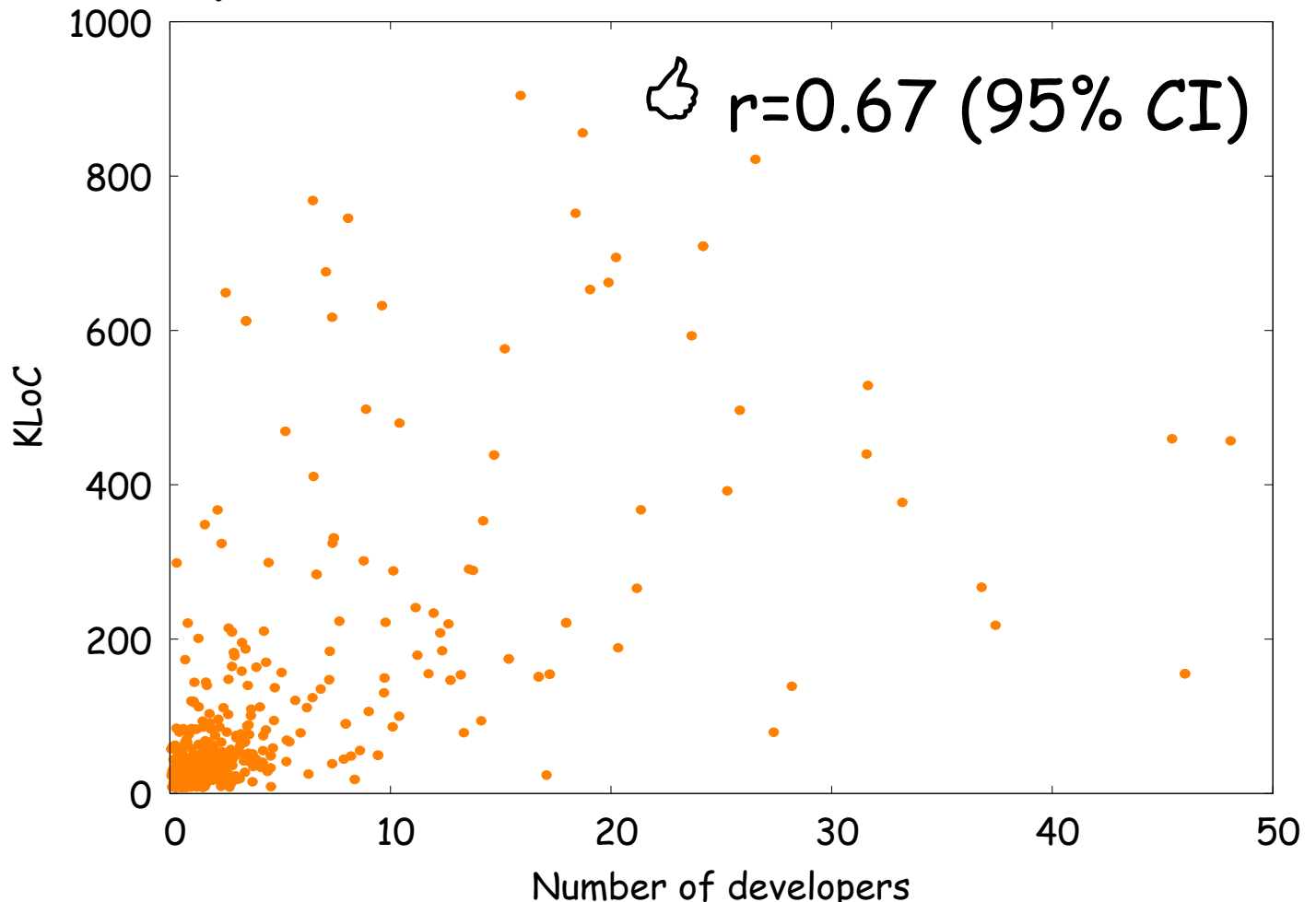? A more rigorous process is required
? Cultural diversity

Productivity

- More developers produce more software (base case)

Scatter plot with "Number of developers" on the x-axis (0 to 50) and "KLoC" on the y-axis (0 to 1000). Annotation: 👍 r=0.67 (95% CI)

Productivity

- ... but developer distance doesn't affect productivity.

Scatter plot of KLoC (y-axis, 0 to 1000) versus Developer distance (x-axis, 0 to 18000). r=-0.14 (95% CI)

Code Style

r=0.05 (95% CI)

% of non-adhering lines (y-axis)

Number of developers in the annotated file (x-axis)

Defect Density

👎 r=0.07 (95% CI)

Number of developers (x-axis, 0 to 200)

PRs per commit (y-axis, 0 to 1)
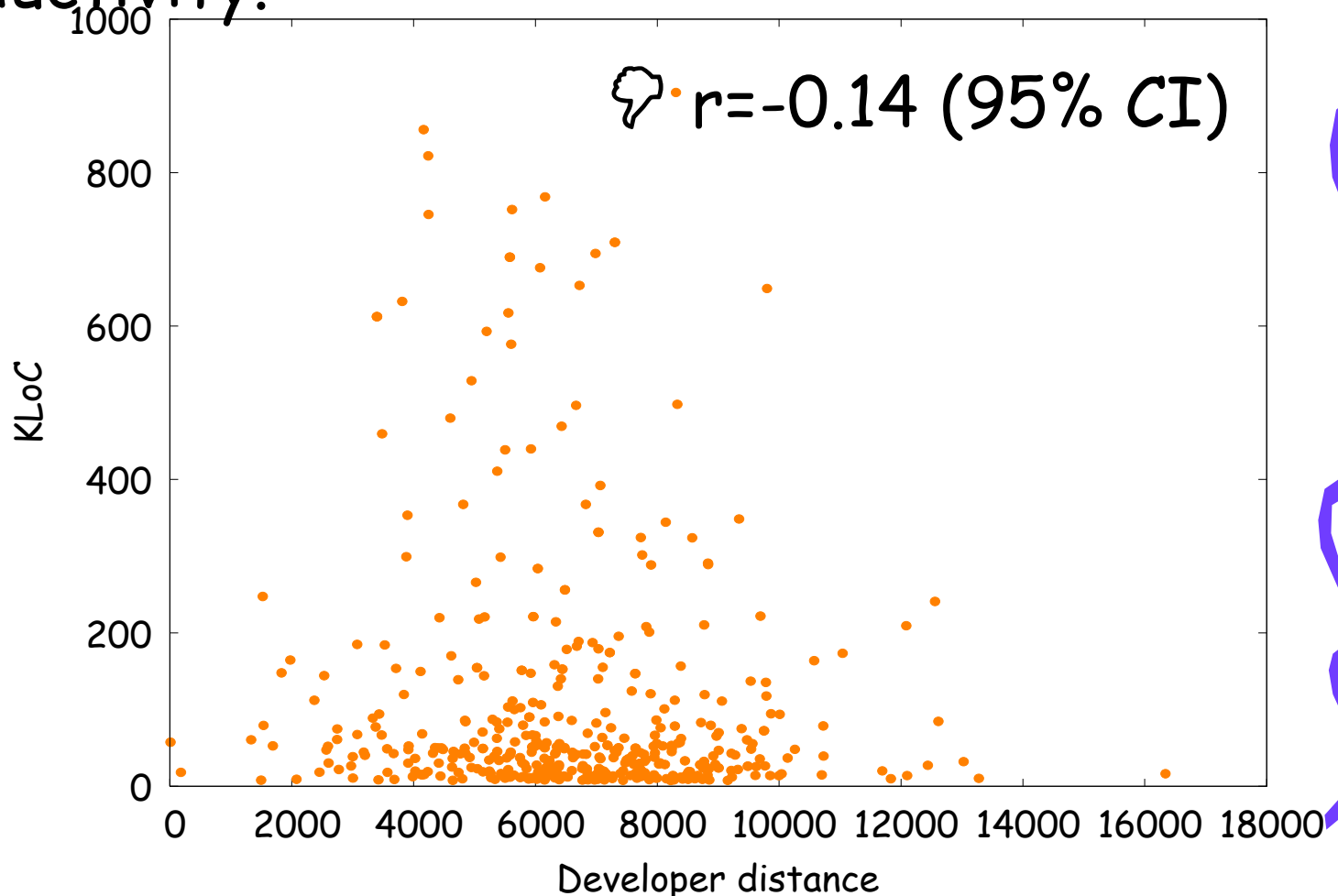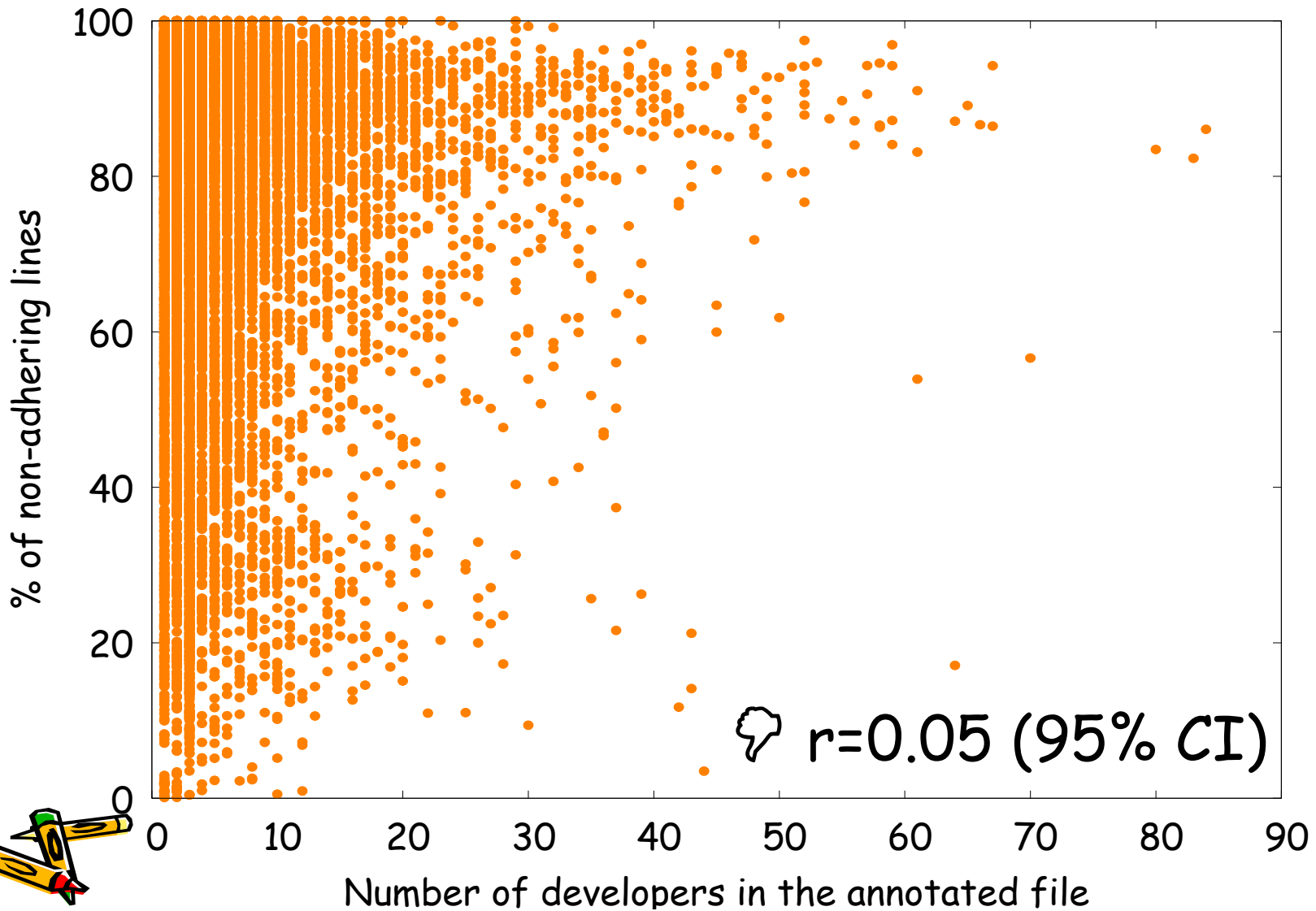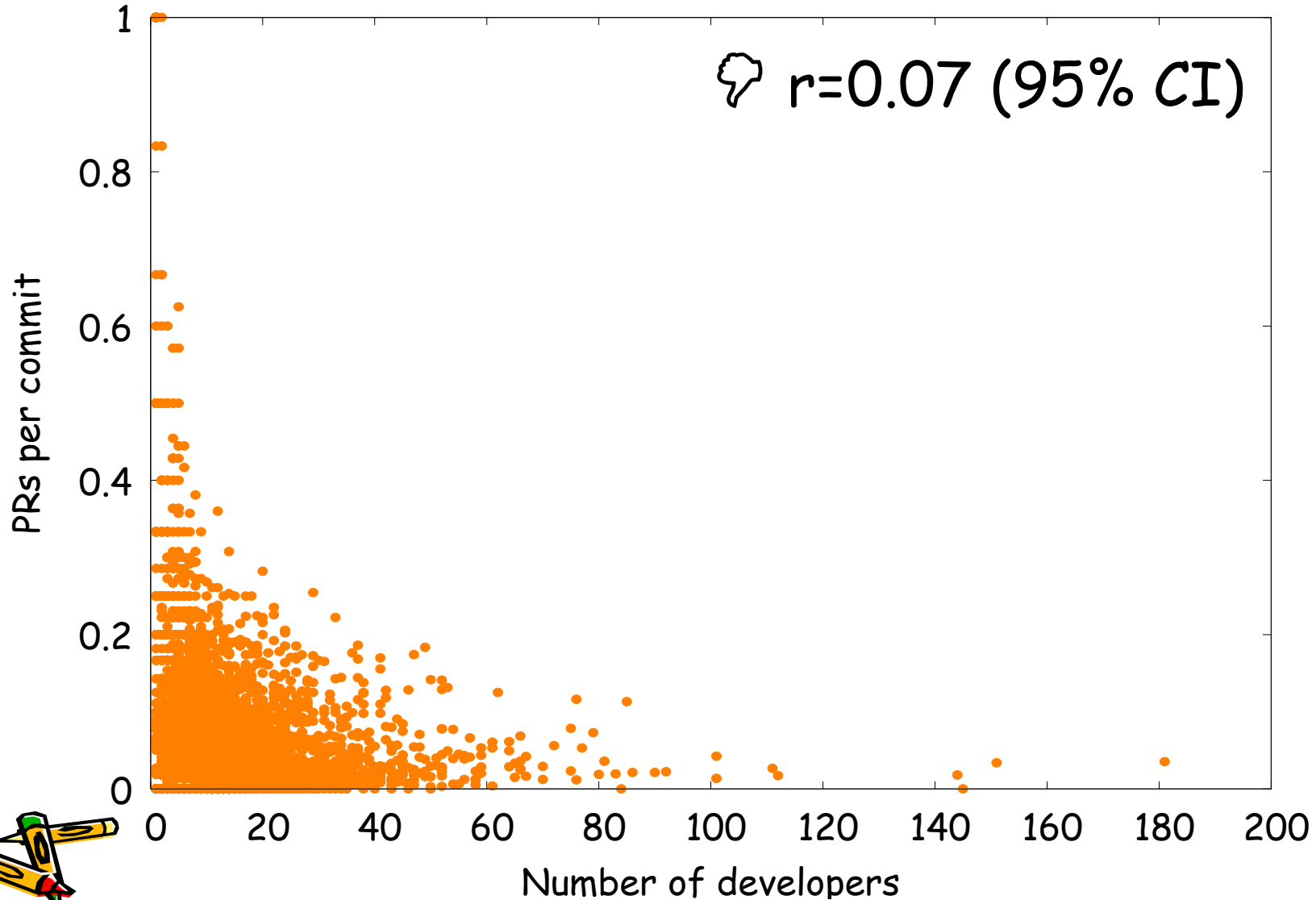
Human Interactions

- **Look for neighbors**
  - Cooperating developers
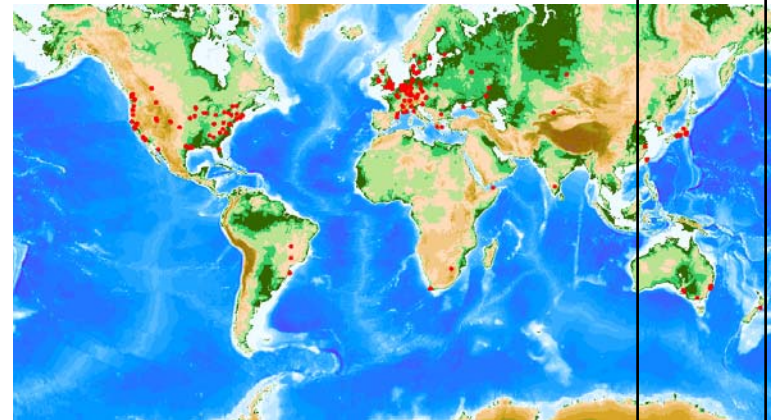 - Mentors and mentees
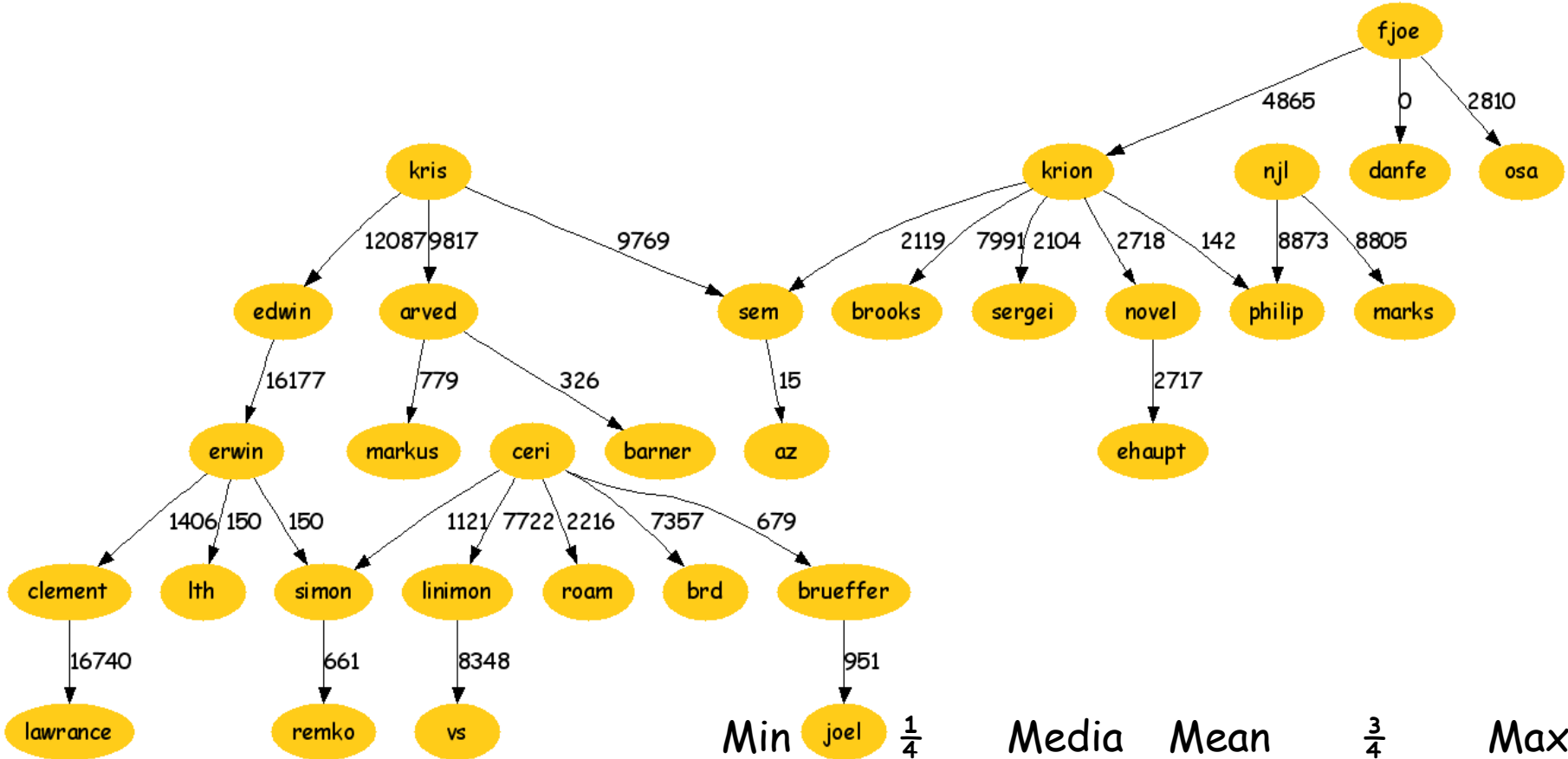

Jan Van Haasteren: **Neighbours**

Cooperating developers

- **Mean distance between**
  - any two FreeBSD developers: 6,701km
- **4010 instances of commits**
  - Same file
  - Same day
  - Different developers
  - Mean distance: 6,489km

Mentors and mentees

	Min	$\frac{1}{4}$	Media	Mean	$\frac{3}{4}$	Max
Any	0	2.2	7.8	6.7	9.3	19
M/M	0	0.7	3.8	5.0	8.8	18

GSD Works

- Round the clock development
- No ill effects on
  - Productivity
  - Code quality
  - Bug density
- Distance
  - Doesn't show up in ad hoc cooperation
  - Appears to matter in mentoring