

# Crowdsourcing and AI

CUI, spring 2022

## Lecture 3

François Grey and Jose Luis Fernandez Marquez

Citizen Cyberlab

UNIGE + UNITAR + CERN

# Classwork

BOINC (Berkeley Open Infrastructure for Network Computing) is an open source software (or more precisely, middleware) for sharing computing power.

*What licence does BOINC use.*

*In your opinion, why did the developer choose this option?*

Load BOINC on your machine.

Choose a BOINC project and join it.

*Science: what did you learn from trying to participate in one?*

*Technology: what was your experience with the interface?*

*Society: How many volunteers? How do they communicate?*

# Homework

Read the PNAS paper about “crowd science”

Crowd science user contribution patterns and their implications  
Henry Sauermann and Chiara Franzoni

PNAS January 20, 2015 112 (3) 679-684;  
first published January 5, 2015

What challenges do the authors identify for crowd science?

How do the results inform current science policy discussions?

What aspects of their analysis were not clear to you, and why?

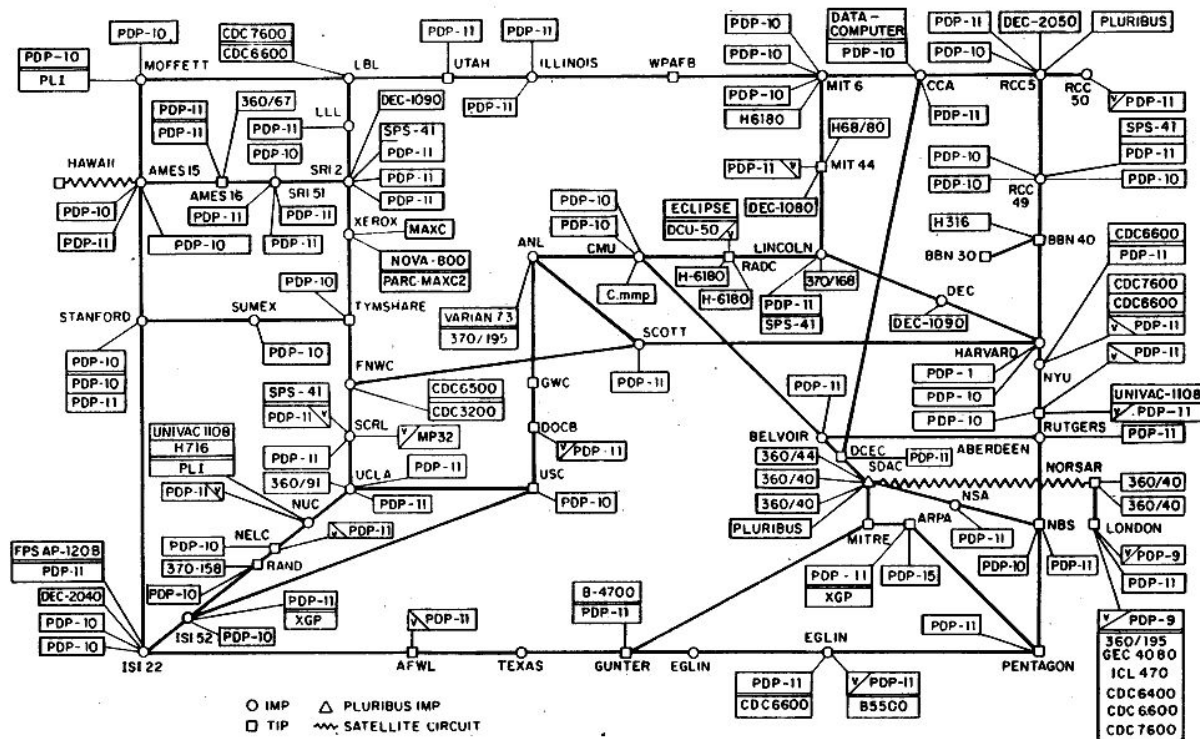
# Crowdsourcing and Open Source

# 1969: IBM “unbundles” software and services from hardware

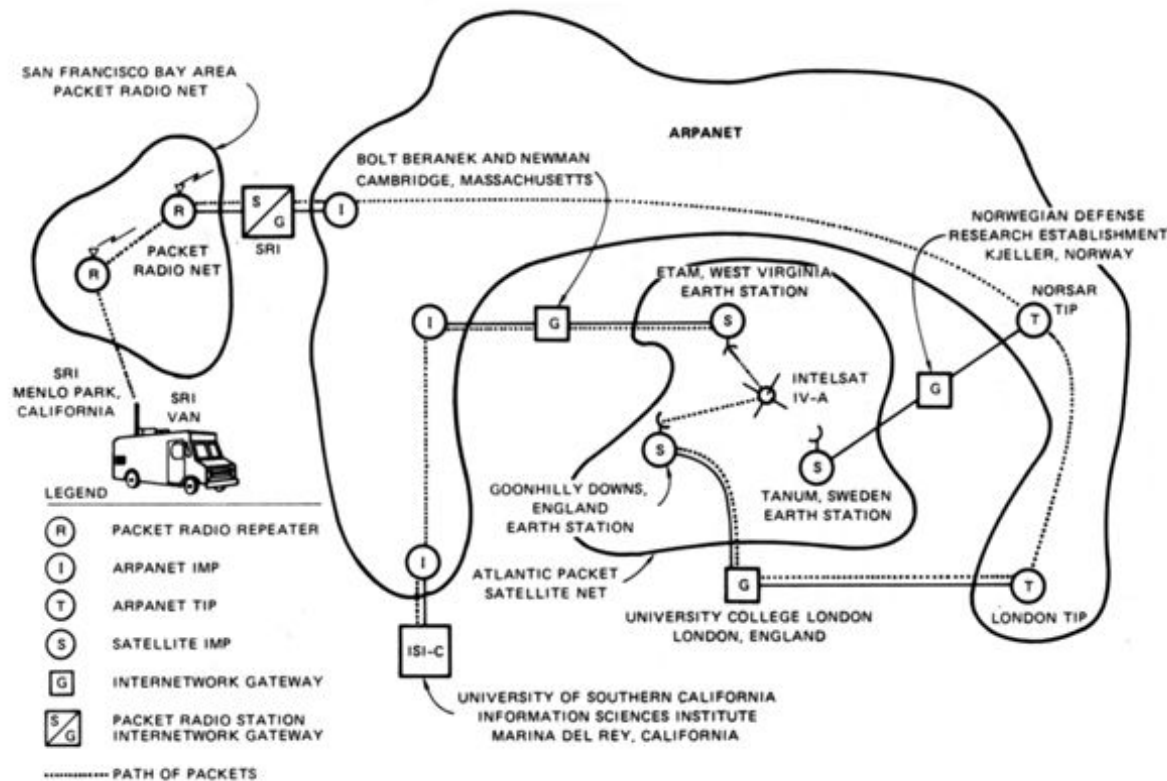


# 1969: First computers connected on the ARPANET

ARPANET LOGICAL MAP, MARCH 1977



# 1973-83: From ARPANET to Internet



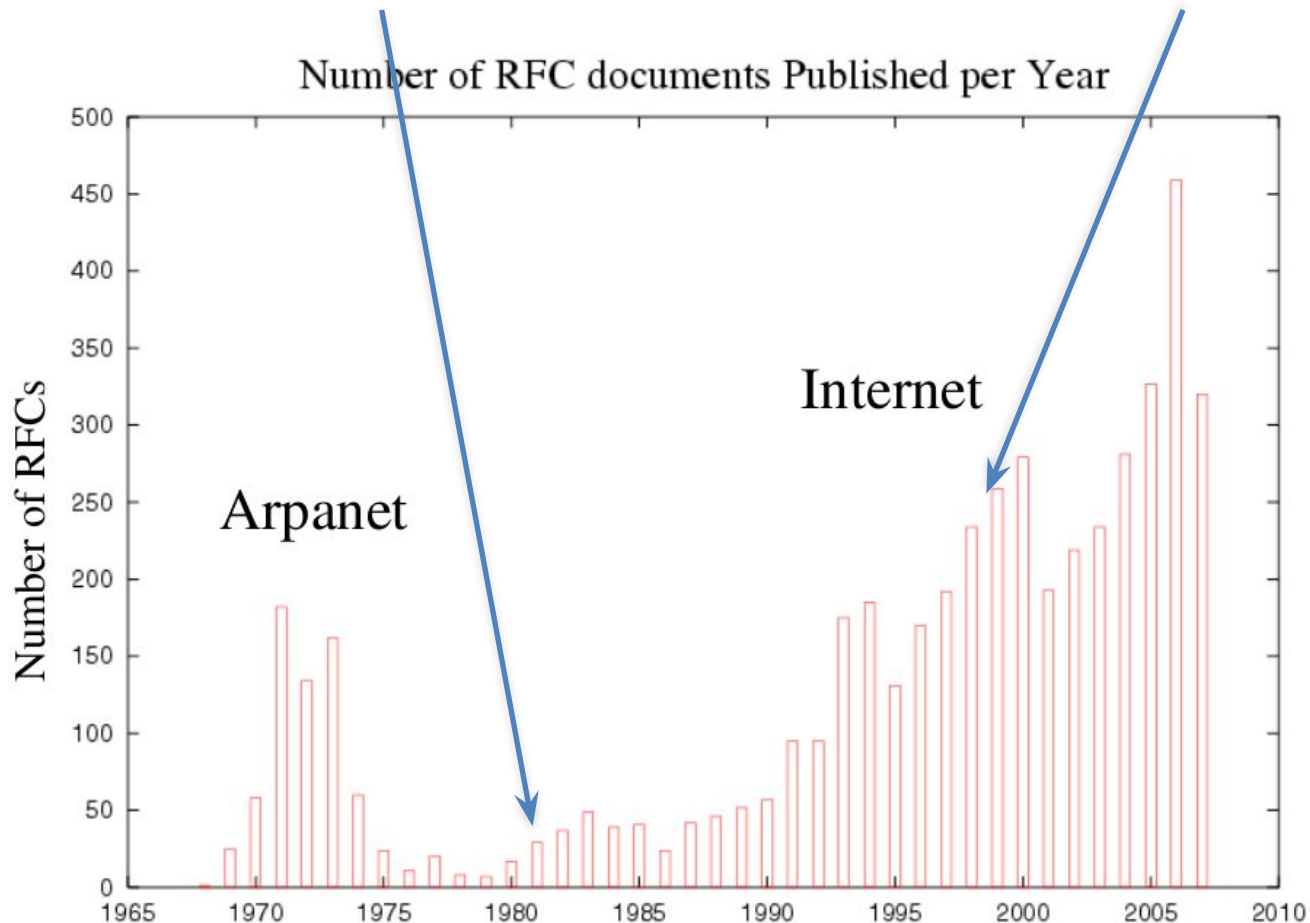
1973: Cerf and Khan propose “internetwork protocol”

1983: TCP/IP becomes protocol of ARPANET

# Request for Comments (RFC): a precursor of open source

IPV4 = RFC 791

HTTP = RFC 2616





# 1971: AT&T launches UNIX operating system



1970s: UNIX free to government, researchers  
1980s: scientists must pay for a licence

What is intellectual property?

# Definition

“Intellectual property refers to creations of the mind: inventions, literary and artistic works, and symbols, names, images, and designs used in commerce.”

What is Intellectual Property?,

World Intellectual Property Organization (WIPO)

<http://www.wipo.int/about-ip/en/>

# Types of intellectual property

1. Expressions (copyright)
2. Inventions (patent)
3. Commercial identifiers (trademark)

Contrast with Public Domain: there is no owner, belongs to the public as a whole.

# What is copyright?

“Original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.” (US Law)

A copyright holder has **the exclusive right**:

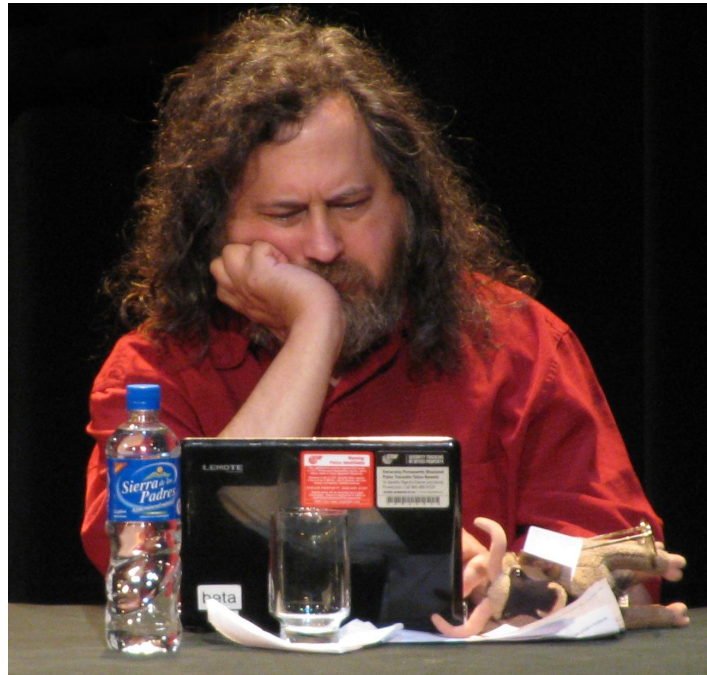
- To make copies of the work
- To prepare derivative works
- To distribute copies of the original work or derivative works
- For certain types of works, to perform the work
- For certain types of works, to perform the work

# What is a licence?

- Grants permission to use a copyrighted work
- Can grant any or all of the rights associated with copyright
- Can impose other restrictions, such as type or place or usage, or duration of the license
- Does not transfer ownership of the copyright
- Open Source licenses grant the right to make copies free of charge and distribute them

# The origins of Open Source

# 1983: Richard Stallman starts the GNU project



GNU objective: an operating system free of restrictions  
GNU meaning: GNU's Not UNIX



# Free vs. Open

1986: Free Software Foundation (Richard Stallman)

- promotes the universal freedom to study, distribute, create and modify computer software
- popularizes the concept of “copyleft” and “all rights reversed”

1989: GNU General Public Licence (GPL) published

1991: Linus Torvalds publishes the Linux Kernel

1992: Linux becomes GPL, GNU Project + Linux = first free OS

# Linux: role model for building Open Source communities

From: [torvalds@klaava.Helsinki.FI](mailto:torvalds@klaava.Helsinki.FI) (Linus Benedict Torvalds)  
Newsgroups: comp.os.minix  
Subject: What would you like to see most in minix?  
Summary: small poll for my new operating system  
Message-ID: <[1991Aug25.205708.9541@klaava.Helsinki.FI](mailto:1991Aug25.205708.9541@klaava.Helsinki.FI)>  
Date: 25 Aug 91 20:57:08 GMT  
Organization: University of Helsinki

Hello everybody out there using minix -  
I'm doing a (free) operating system (just a hobby, won't be big and professional like gnu) for 386(486) AT clones. This has been brewing since april, and is starting to get ready. I'd like any feedback on things people like/dislike in minix, as my OS resembles it somewhat (same physical layout of the file-system (due to practical reasons) among other things).

I've currently ported bash(1.08) and gcc(1.40), and things seem to work. This implies that I'll get something practical within a few months, and I'd like to know what features most people would want. Any suggestions are welcome, but I won't promise I'll implement them  
Linus ([torvalds@kruuna.helsinki.fi](mailto:torvalds@kruuna.helsinki.fi))

PS. Yes - it's free of any minix code, and it has a multi-threaded fs. It is NOT protable (uses 386 task switching etc), and it probably never will support anything other than AT-harddisks, as that's all I have :-(

# Evolution of Open Source

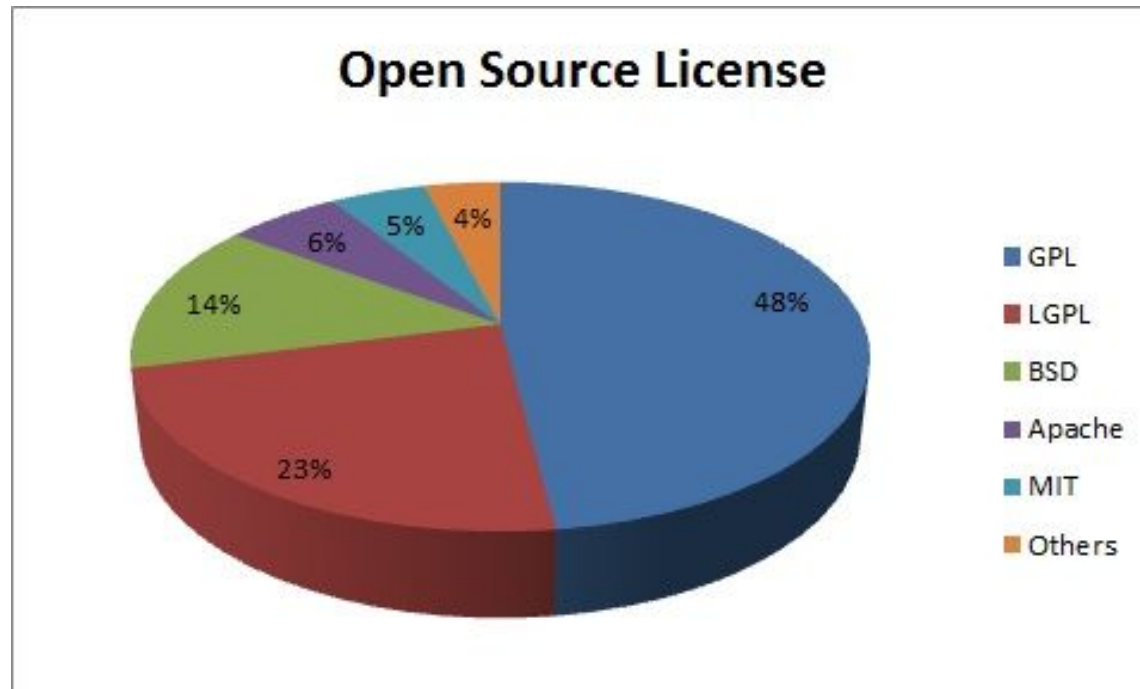
1997: Essay by Eric Raymond “The Cathedral and the Bazaar” contrasts code that is controlled and with code developed by a group on the Internet (Linux)

1998:

- Netscape releases browser Netscape Communicator as free software (basis of Firefox) using Netscape Public Licence (now Mozilla Public Licence).
- Raymond and others propose term Open Source to make free software more appealing to corporations.
- Freeware Summit by O'Reilly Publishing becomes “Open Source Summit”
- **Open Source Initiative** is created as an organization to promote Open Source Software

# Open Source today

- Over 30 million open source projects
- Over 2 billion lines of code



# Why does Open Source matter for Open Science?

“Scientific communication relies on evidence that cannot be entirely included in publications... We argue that [...] anything less than the release of source programs is intolerable for results that depend on computation. “

*The Case for Open Computer Programs*

*Ince et al., Nature 2012*

# Classwork

Amazon Mechanical Turk is the most well-known commercial crowdsourcing platform to date.

*What steps are necessary to become a requester?*

*What steps are necessary to become a worker?*

# Homework

Read Chapter 4 of Reinventing Discovery (Book in Moodle)

“Patterns of Online Collaboration” and answer the questions:

1) What example of Open Source innovation does the author give, that is inspired by open source software?

2) What unique features of open source collaboration does the author distinguish?