INFO5992 Understanding IT Innovations

Week 5: Distributed Innovation:

Crowdsourcing, Open Data, Free and Open Source Software

Semester 1, 2025

THE UNIVERSITY OF SYDNEY

Acknowledgement of Country

I would like to acknowledge the Traditional Owners of Australia and recognise their continuing connection to land, water and culture. I pay my respects to the first nations people and their Elders, past, present and emerging.

Copyright warning

COMMONWEALTH OF AUSTRALIA

Copyright Regulations 1969

WARNING

This material has been reproduced and communicated to you by or on behalf of the University of Sydney pursuant to Part VB of the Copyright Act 1968 (**the Act**).

The material in this communication may be subject to copyright under the Act. Any further copying or communication of this material by you may be the subject of copyright protection under the Act.

Do not remove this notice.

UoS Semester Outline

Week	Learning Outcomes	Lectures
Module 2: Innovation Framework		
Week 01	L01, LO2, LO3	Unit of Study Introduction, Administrivia, Definition of IT Innovation, Importance of Innovation to a Country, General Purpose Technologies, Overview of Emerging Technologies
Week 02	LO4, LO5	Innovation Frameworks I: Dynamics of IT Innovation, Dominant Design
Week 03	LO6	Innovation Frameworks II: Disruptive Innovation, Innovator's Dilemma, Value Chain & Value Network
Module 2: Development of Key Intellectual Property in the Modern Age		
Week 04		Introduction to Open Innovation and Closed Innovation Distributed Innovation I: Product Platforms, Web APIs
Week 05	LO7	Distributed Innovation II: Crowdsourcing, Free and Open- Source Software, Open Data
Week 06		Distributed Innovation III: Platform Ecosystems, User Innovation
Module 3: Commercialisation Process and Business Strategies for Emerging Technologies		
Week 07	LO8	Commercialisation I: Startup vs Traditional Companies, Lean Startup Methodology and Agile Development
Week 08		Commercialisation II: Customer Development Process, Value Proposition Canvas
Mid semester break		
Week 09	LO8, LO9	Commercialisation III: Innovation Management, Business Model Canvas Commercialisation IV: Capital & Fundraising for IT Innovation
Week 10	LO11, LO12	Organisational Cultures and Structures Supporting Innovation, Judging IT Innovation
Module 4: Innovation At-Scale		
Week 11	LO10	Innovation Ecosystem: Silicon Valley and Australia
Week 12	N/A	Course Review | Innovation Pitch Presentation
Week 13	N/A	Innovation Pitch Presentation
Final Exam		

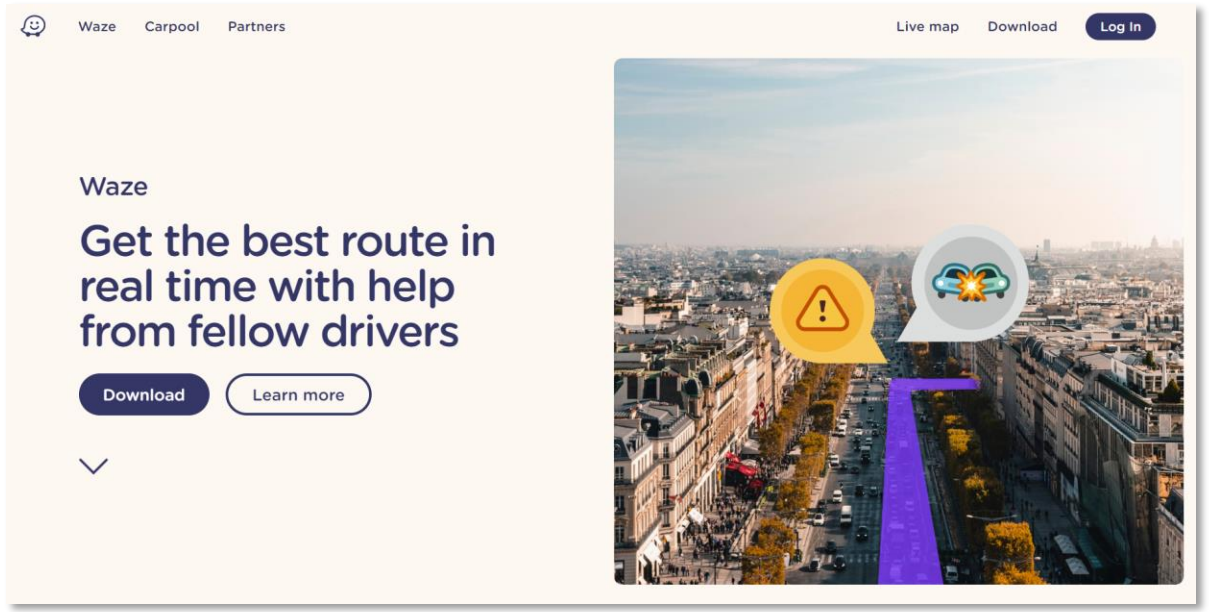
Agenda – Week 04

These are some approaches companies use to get external companies/individuals involved in their innovation:

A. Product platforms
B. Web APIs
C. Crowdsourcing innovation / Crowdfunding Innovation
D. Releasing data sets "Open data"
E. Free and Open Source Software
F. User innovation
G. Platform ecosystems
H. Accelerators, investment and others

C. Crowdsourcing (including crowdfunding)

[Wikipedia, the free encyclopedia]



[Driving directions, live traffic and road conditions update - Waze]

[Minecraft]

Crowdsourcing: What is it?

- Original definition (from 2006)
- = Crowd + Outsourcing
- Crowdsourcing represents the act of a company or institution **taking a function once performed by employees and outsourcing** it to an undefined (and generally large) network of people in the form of an open call.

  The crucial prerequisite is the use of the **open call format and the wide network** of potential laborers.

- (Jeff Howe, Wired Magazine, 2006)

Crowdsourcing: Newer definition

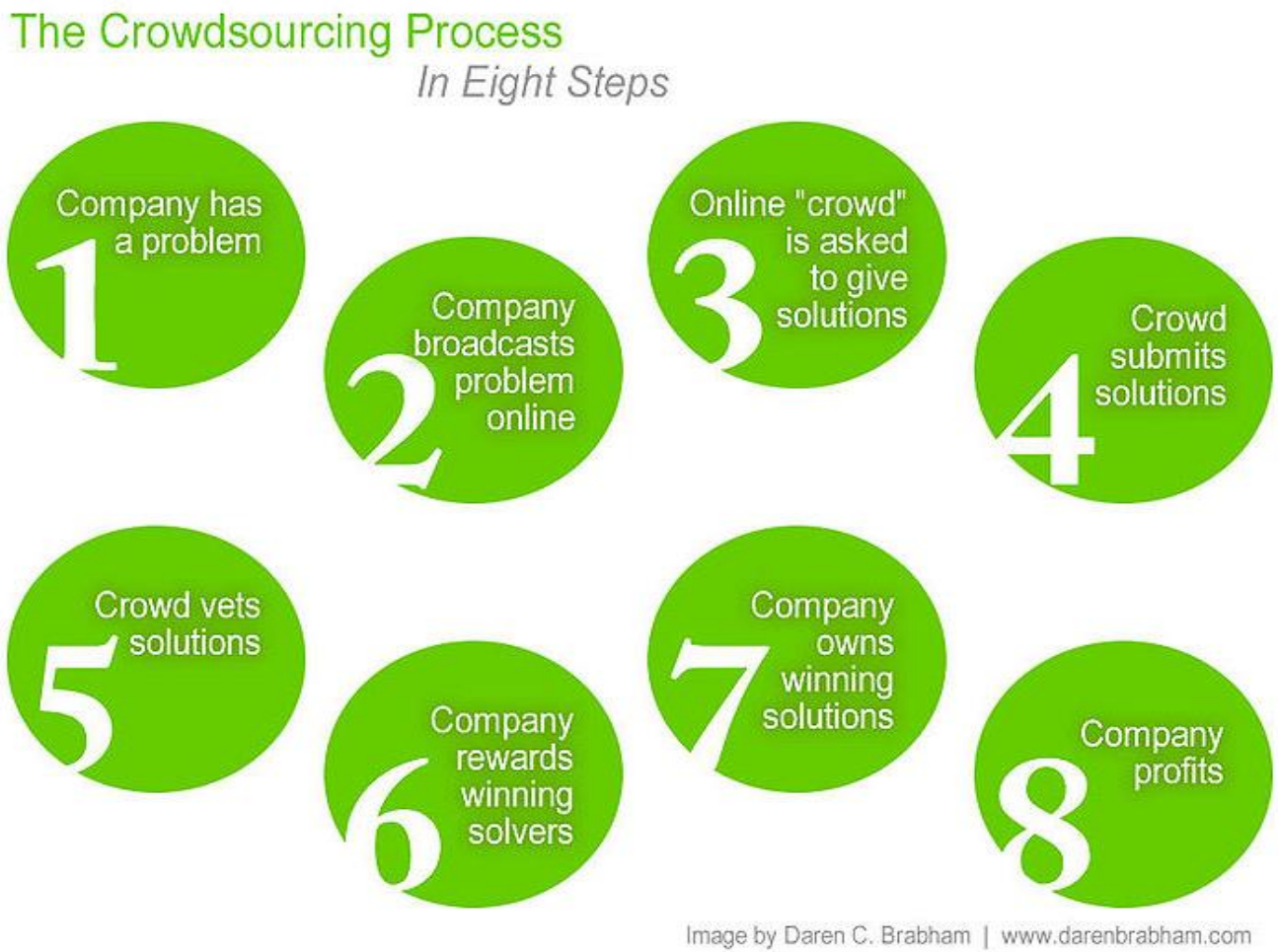
- Crowdsourcing is a type of **participative online activity** in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the **voluntary undertaking of a task**.

- The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, **always entails mutual benefit**.

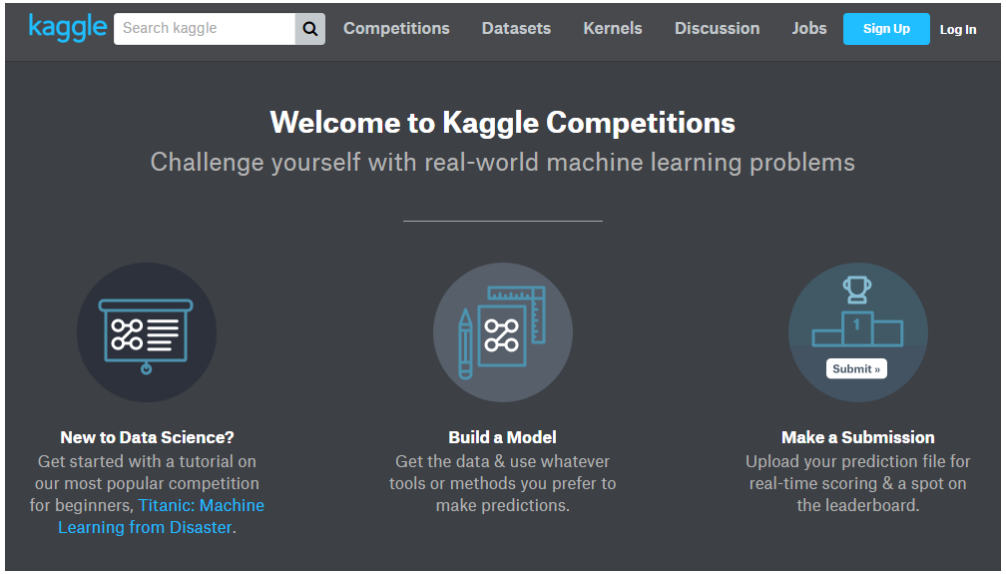
- The **user will receive the satisfaction** of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, **while the crowd-source organization** will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science*, *38*(2), 189-200.

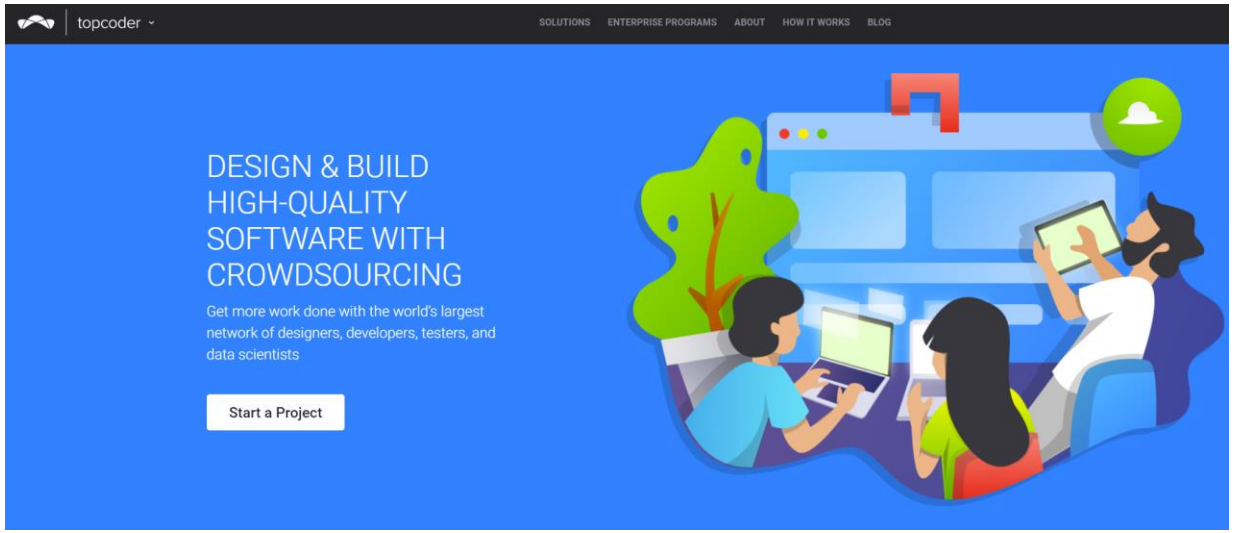
Crowdsourcing: The typical crowdsourcing process



Other Examples – crowdsourcing for developments



https://www.kaggle.com/competitions



https://www.topcoder.com/

Crowdfunding

- Crowdsourcing is the sourcing of anything from a crowd
- Crowdfunding is the sourcing of funds from a crowd - a specific type of crowdsourcing.

Platform	Total Raised	Supporters	Platform Fee	Payment Fee	Important to Know
gofundme GoFundMe	$25B	50M	0%	2.9% + $0.30	✓ Can quickly set up withdraws and deposits take an average of 2-5 business days ✓ Coaching and account support throughout the fundraising and donation process ✓ Easy to use fundraising tools make setup fast (e.g., mobile app and superior add beneficiary feature) ✓ The GoFundMe Giving Guarantee – in the very rare case that something isn't right with a fundraiser, donors may be eligible for a 100% refund of their donation
INDIEGOGO Indiegogo	$1.5B	10M	5%	3.0% + $0.30	✓ Offers "flexible funding" ✓ Specializes in technology and hardware product launches ✓ Regular email support hours; marketing and campaign strategy support
KICKSTARTER Kickstarter	$3B	15M	5%	3.0% + $0.20	✓ Specializes in creative projects with robust reward level feature ✗ 14-day wait to withdraw and deposits take 5-7 business days ✗ Limited email support hours ✗ Requires Kickstarter approval to launch a fundraiser
FUNDLY Fundly	$330M	NA	4.9%	2.9% + $0.30	✓ Can withdraw immediately and deposits take 2-5 business days ✗ No donor guarantee policy for fraud protection ✗ Limited email support hours
JustGiving JustGiving	NA	22M	Nonprofits: 0-5% Personal: 0%	Nonprofits: 2.9% Personal: 2.9% + $0.30	✓ Supports UK gift aid ✗ 14-day wait to withdraw and deposits take 6-10 business days ✗ No fraud protection offered ✗ Limited email support hours

http://www.crowdfunding.com/

Why do people engage with crowdsourcing?

Many reasons including:

- "the desire to earn money;
- to develop one's creative skills;
- to network with other creative professionals;
- to build a portfolio for future employment;
- to challenge oneself to solve a tough problem;
- to socialize and make friends;
- to pass the time when bored;
- to contribute to a large project of common interest;
- to share with others; and
- to have fun."

Brabham, D. C. (2012). Crowdsourcing: A model for leveraging online communities. In *The participatory cultures handbook* (pp. 120-129). Routledge.

Types of Crowdsourcing (Brabham, 2011)

Type	How it Works	Kinds of Problems	Examples
Knowledge Discovery and Management	Organization tasks crowd with finding and collecting information into a common location and format	Ideal for information gathering, organization, and reporting problems, such as the creation of collective resources	Peer-to-Patent *peertopatent.org* SeeClickFix *seeclickfix.com*
Broadcast Search	Organization tasks crowd with solving empirical problems	Ideal for ideation problems with empirically provable solutions, such as scientific problems	InnoCentive *innocentive.com* Goldcorp Challenge *Defunct*
Peer-Vetted Creative Production	Organization tasks crowd with creating and selecting creative ideas	Ideal for ideation problems where solutions are matters of taste or market support, such as design or aesthetic problems	Threadless *threadless.com* Doritos Crash the Super Bowl Contest *crashthesuperbowl.com* Next Stop Design *nextstopdesign.com*
Distributed Human Intelligence Tasking	Organization tasks crowd with analyzing large amounts of information	Ideal for large-scale data analysis where human intelligence is more efficient or effective than computer analysis	Amazon Mechanical Turk *mturk.com* Subvert and Profit *subvertandprofit.com*

SeeClickFix | 311 Request and Work Management Software

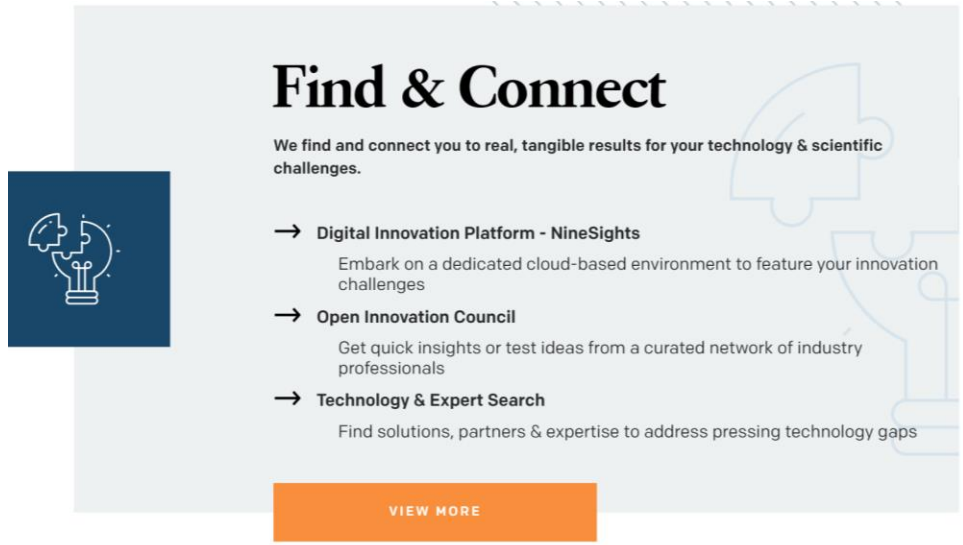
Challenges – Wazoku (now Wazoku)

Threadless

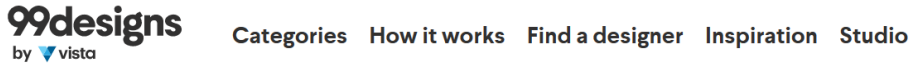
Amazon Mechanical Turk

Crowdsourcing for innovation: Another typology

- **Intermediary platforms**
  - Research & Development platforms (e.g. Innocentive, NineSigma)
  - Marketing, Design & Idea platforms (e.g. 99designs)
  - Collective intelligence & Prediction platforms (e.g. Kaggle)
  - HR and Freelancers platforms (e.g. TopCoder, Amazon Mechanical Turk)
  - Open innovation software
  - Intermediary open innovation services



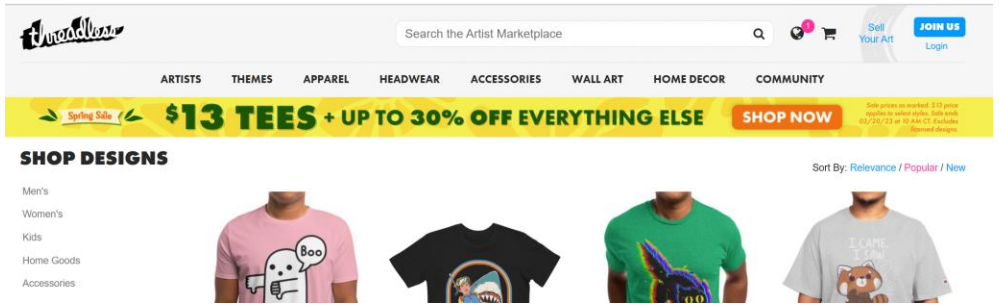
Contribute to open innovation - We make innovation happen | NineSigma



https://99designs.com.au/

Crowdsourcing for innovation: Another typology (Cont.)

- **Creative co-creation**
  - E.g. **Threadless,** Custom Tattoo Design Contests & Tattoo Ideas | CreateMyTattoo.com

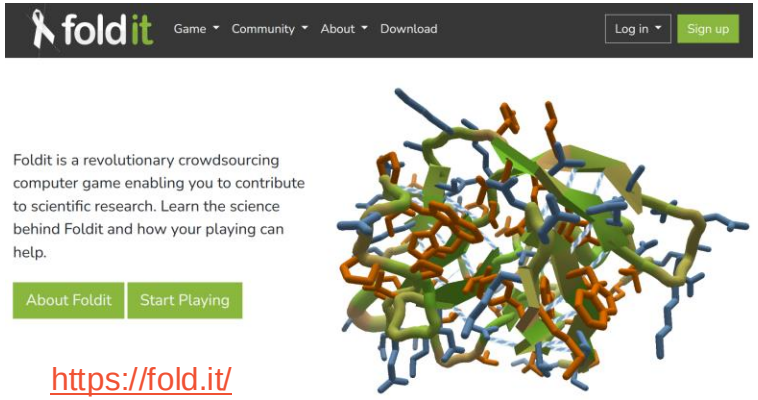


Browse Designs | Threadless

- **Corporate initiatives**
  - Product ideas crowdsourcing (e.g. IBM InnovationJam IBM InnovationJam® - Overview | IBM)
    - *What began as an internal experiment in 2001… is now a proven management tool for driving innovation and collaboration. IBM's InnovationJam® offering is ideal for companies and enterprises looking to kick-start a transformation or change program through a transparent 'conversation'.*
  - Branding and Design crowdsourcing (e.g. Fluevog)

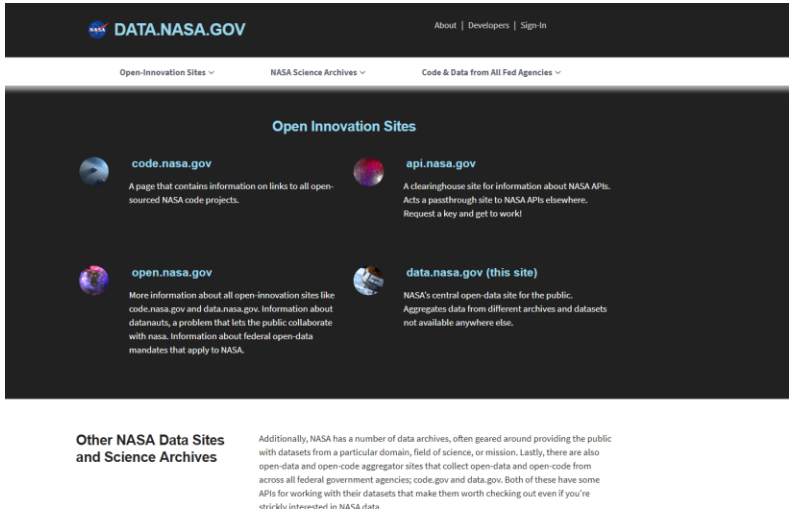
Crowdsourcing for innovation: Another typology (Cont.)

- **Peer production**
  - E.g. Linux, Wikipedia
- **Public crowdsourcing**
  - E.g.
  - Fold it - *a revolutionary crowdsourcing computer game enabling you to contribute to scientific research. Learn the science behind Foldit and how your playing can help.*



https://fold.it/

  - Nasa Open Innovation

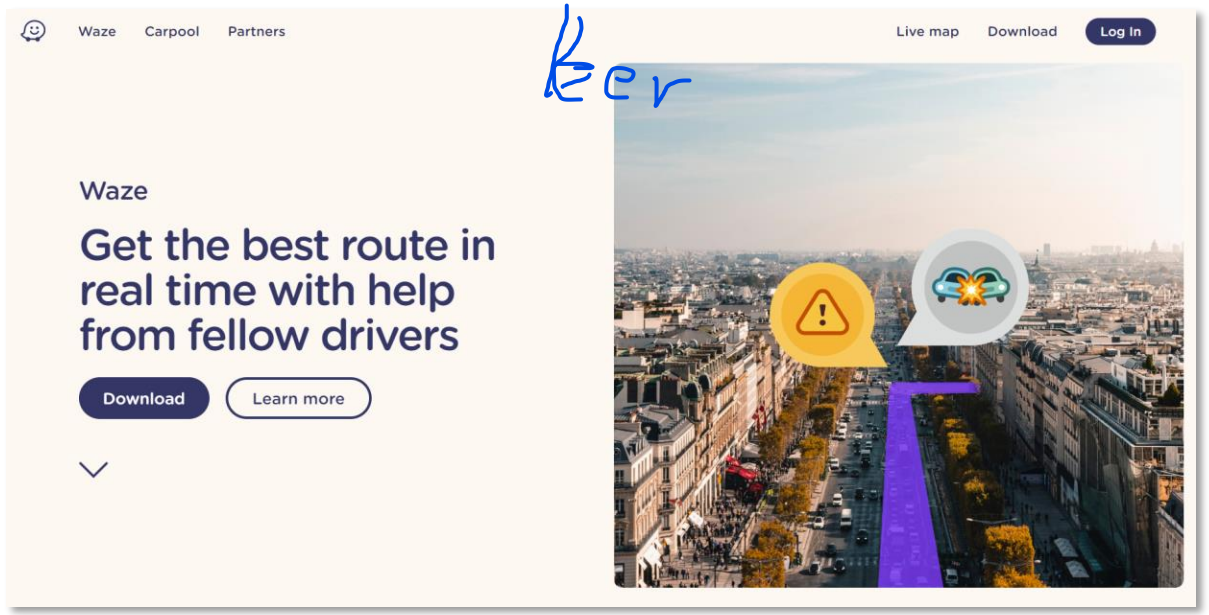


Open Innovation | NASA

Why Crowdsourcing Initiatives?

- Crowdsourcing is the collective intelligence and creativity of a diverse community
- Innovation can be significantly accelerated and enriched
- Companies can demonstrate open collaboration, not just internal R&D, to generate breakthrough ideas, solve complex problems, and create value that benefits both the contributors and the organisation
- A company could take the following **initiatives:**
  - Open up the innovation process, such as by creating competitions, hackathons, or open calls for proposals
  - Encourage voluntary collaboration, such as in the Peer Production
  - Implement public crowdsourcing initiatives
  - Incentivise participation

Revisted…



Wikipedia, the free encyclopedia

Driving directions, live traffic and road conditions update - Waze

Minecraft

- What type of crowdsourcing are they?
- Why do people engage??
- Could they exist without crowdsourcing? Could they change their crowdsourcing type??

Crowdsourcing: What is it?

https://www.youtube.com/watch?v=WCPfxPSpEsI

D. Releasing data sets “Open data”

What is 'Open Data'?

"Open means **anyone** can **freely access, use, modify, and share** for **any purpose** (subject, at most, to requirements that preserve provenance and openness)."

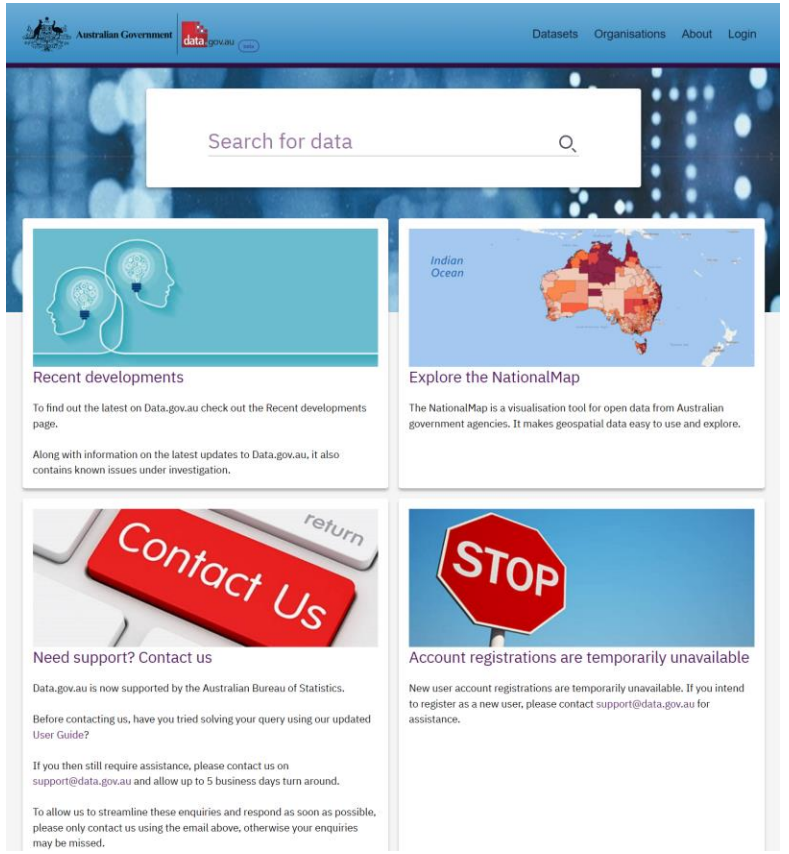
Put most concisely:

"Open data and content can be freely used, modified, and shared by anyone for any purpose"

https://opendefinition.org/

Releasing data sets

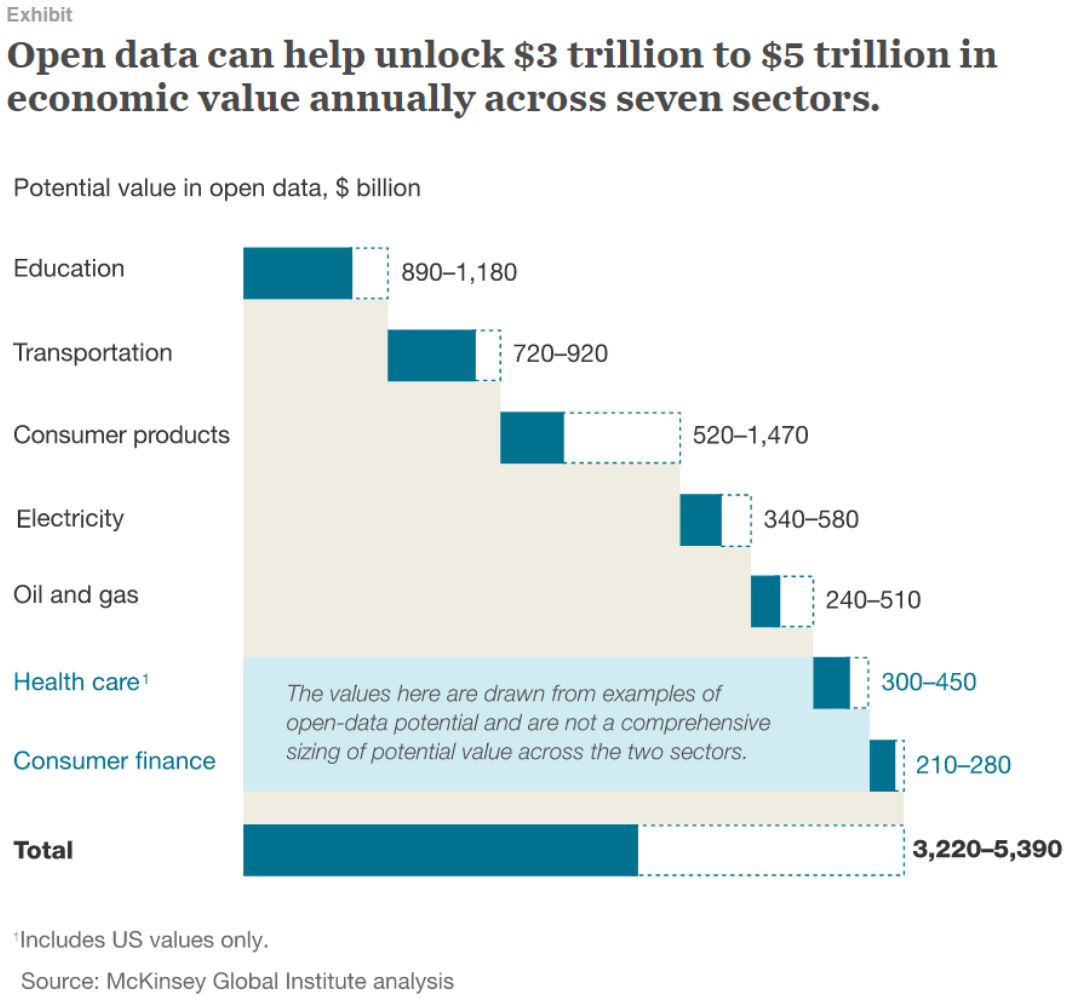
- Many governments have opened up government data ("open data")
  - In some cases, static data (e.g. tables of static data)
  - In some cases, live data feeds (e.g. an RSS feed or data service)
  - The Australian federal government http://data.gov.au includes:
    - Electoral boundaries
    - Crime data, census data
  - NSW Government http://data.nsw.gov.au/ includes:
    - Bus stop data, Electricity consumption data, pollution, education etc.



Releasing data sets

- Many communities are also building open data sets
  - E.g. www.openstreetmap.org, www.openaddresses.io
- Some companies are encouraging users to develop applications using their data – more later

Value in Open Data

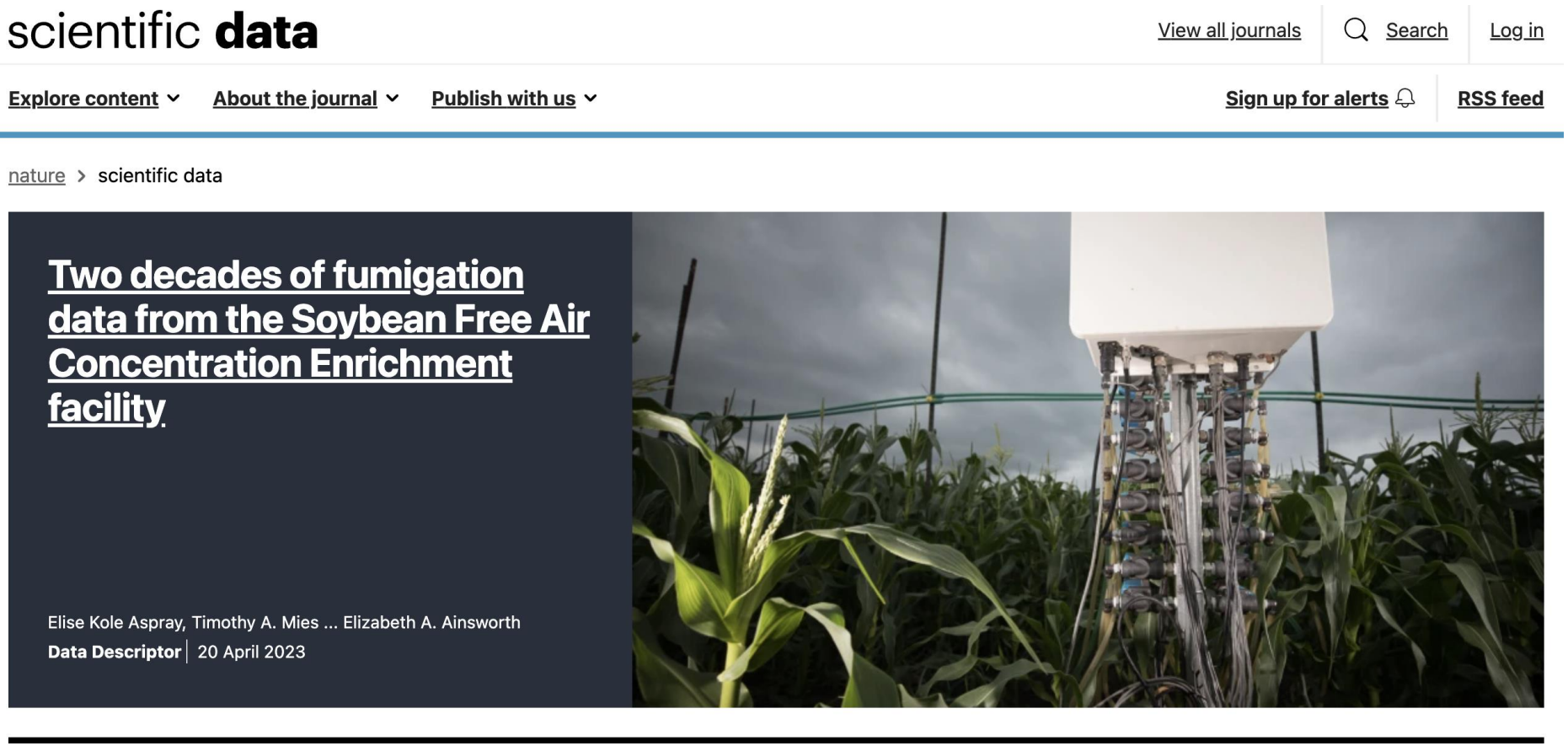


- **Economic value** e.g. increased efficiency, new products and services, and a consumer surplus (cost savings, convenience, better products)
- **Big data's impact** e.g., replacing or supporting human decision making
- **Business opportunities** e.g., new products and services
- **Governments to play a central role**

https://opendatatoolkit.worldbank.org/en/data/opendatatoolkit/starting (Mar'25)

Value in Open Data – Self-reinforcing cycle

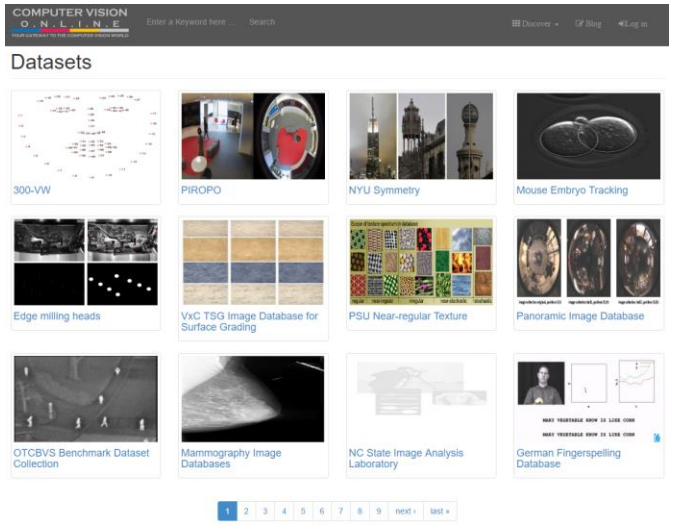
- The benefits of open data can be self-reinforcing: they will increase as **individuals perceive the advantages and help to improve the accuracy** and detail of the information available.
- However, this cycle can gather momentum only if **private industries and public agencies cultivate a vibrant open-data ecosystem** and implement policies to protect stakeholders.
- For companies, that means putting in place the **technologies and talent to collect and analyze data.**
- For individuals—as both consumers and citizens – it means **being vigilant, savvy providers and users of open data.**

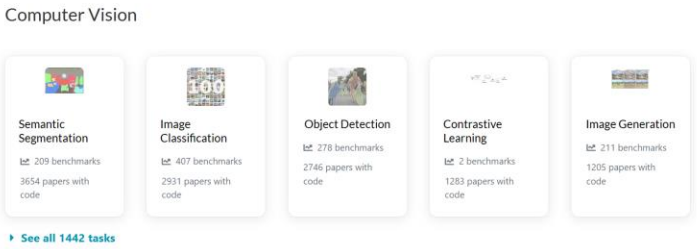


Scientific Data (nature.com)

Example: Computer Vision / Medical Imaging Community

- In Research community, open source and open data is important for
  - Benchmarking / Dissemination
  - Replication / Citations
  - Reputation / State of the Art
  - Building a community

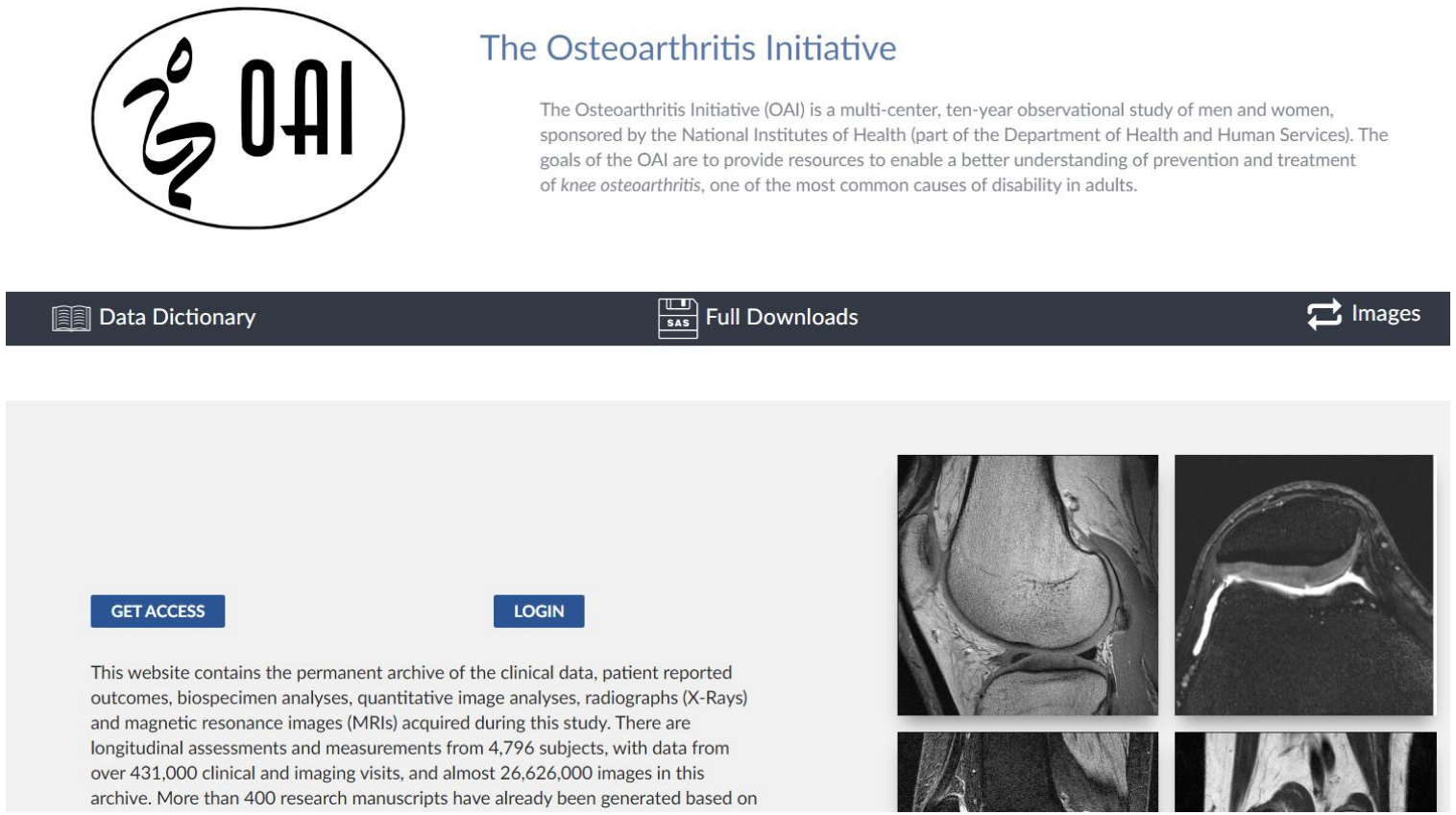




[Browse the State-of-the-Art in Machine Learning | Papers With Code](#) (Mar'25)

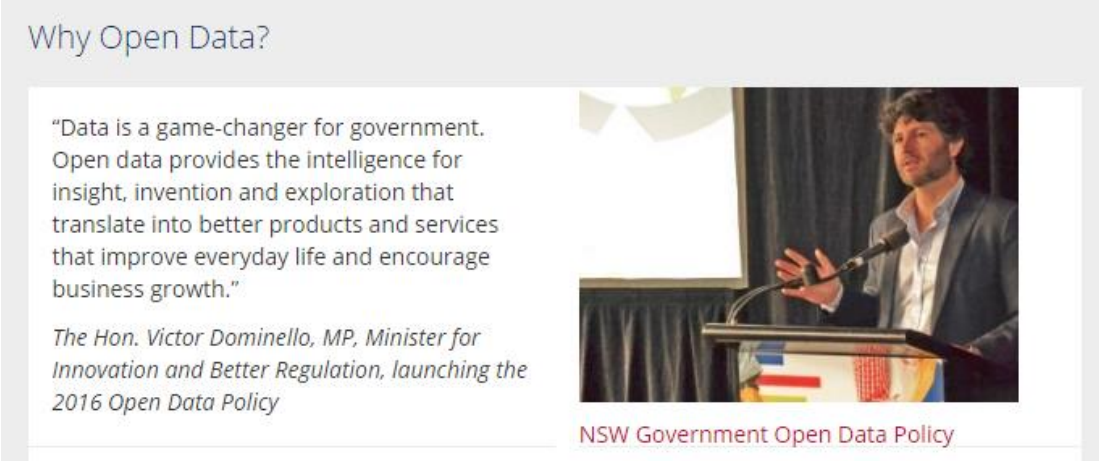
Example: The Osteoarthritis Initiative

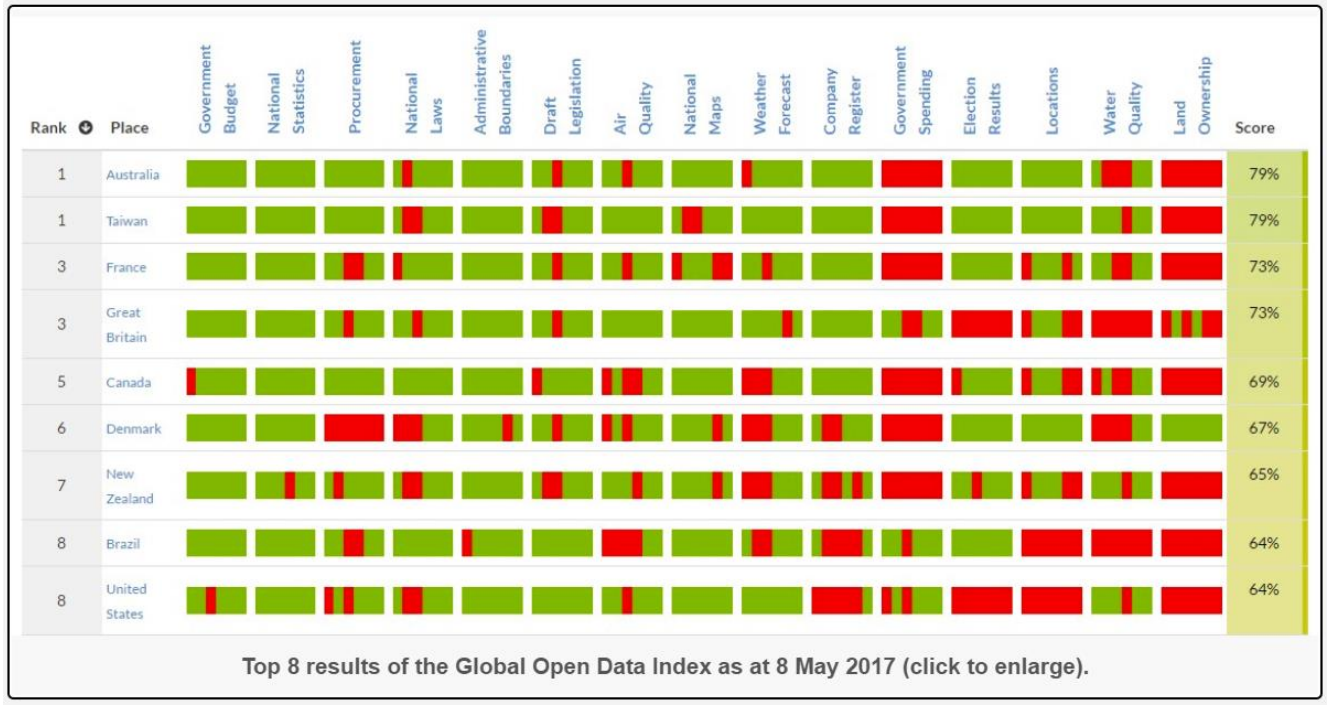
- The Osteoarthritis Initiative (OAI) is a multi-center, ten-year observational study of men and women, sponsored by the National Institutes of Health (part of the Department of Health and Human Services).
- **The goals of the OAI are to provide resources to enable a better understanding of prevention and treatment of *knee osteoarthritis*, one of the most common causes of disability in adults.**



https://nda.nih.gov/oai/

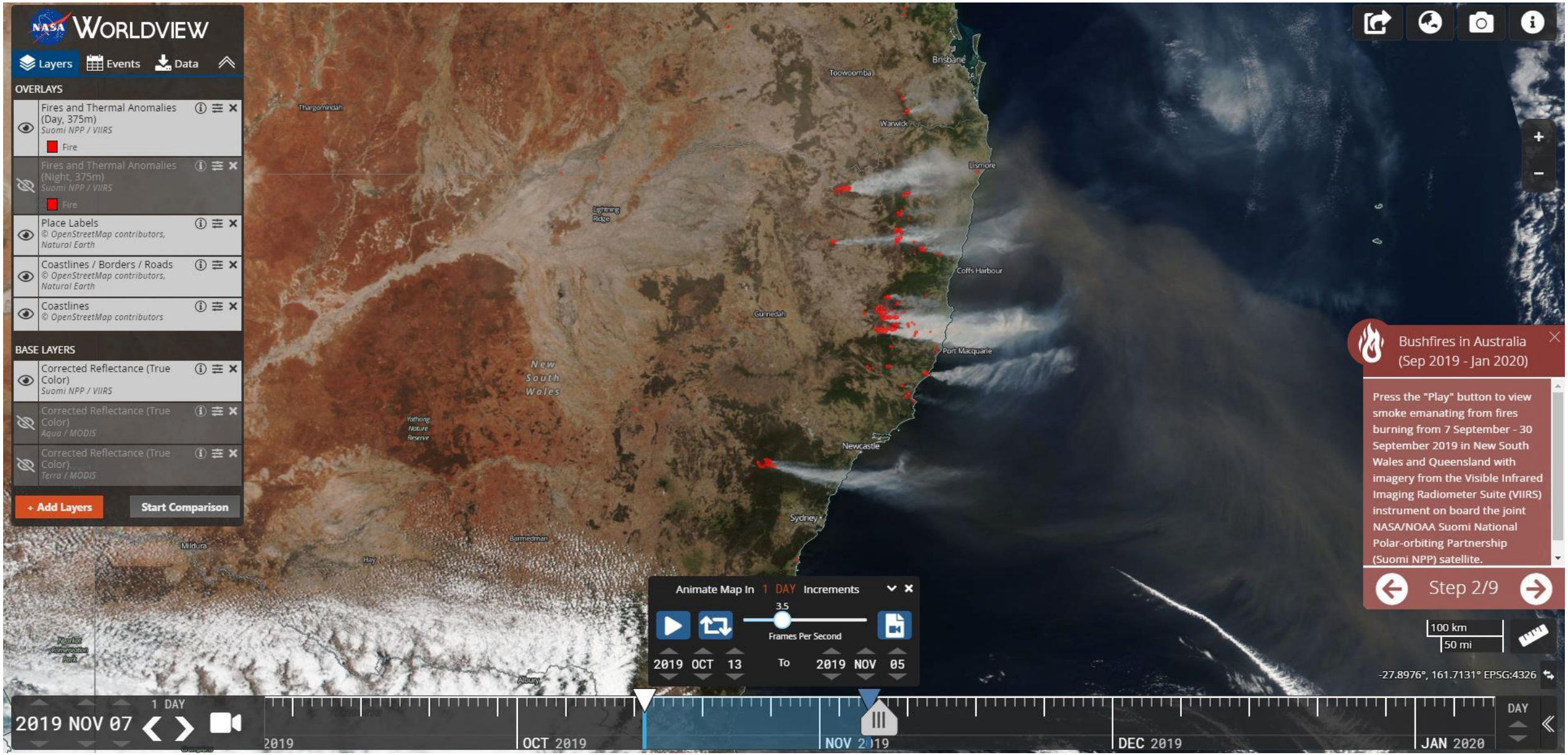
Example: Australian Government





https://www.spatialsource.com.au/gis-data/australia-leads-world-open-data (Mar'25)

http://data.nsw.gov.au/
https://www.ipc.nsw.gov.au/open-data-infographic (Mar'25)



EOSDIS Worldview (nasa.gov)

Open collaborative innovation

E. Free and open source software

Proprietary software vs Free and open source software

- Traditionally, most software was proprietary.
- Proprietary software is software built by or for a specific person, organisation or group of organisations where:
  - The owner holds intellectual property rights over the software; and
  - The owner has total control over the software and how it is used
- In free and open source software:
  - The source code is made available
  - Source code can be changed and redistributed by others
  - (more precise definitions coming later)

Some examples

Proprietary software

SAP ERP

IBM DB2

Oracle Database

Internet Explorer

(Many of these also use open source within them)

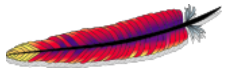
Free and open source software

Android

Chrome

Linux

Firefox



Apache HTTP Server

Alfresco

Hadoop

Tomcat

(Not all versions of the above are open source)

Free and open source software: Example GNU

- Richard Stallman was a programmer at MIT AI Lab
- In the 1970s, many manufacturers (e.g. copiers) supplied source code (e.g. drivers)
- Stallman (and others) modified software to meet their needs
- Companies started to be more protective of their source code.
- In 1980, Stallman and others were refused access to the source code for the software of the first **laser printer** (from Xerox)
- Stallman was not able to modify the software, and this was inconvenient for him
- This helped convince Stallman that people should be free to modify all software.

Richard Stallman

"technical means to a social end"

What is Free Software?
(using Free Software Foundation definition)

- Free Software Foundation (**FSF**)
- Started by Richard Stallman in 1985
- “Free” as in “free speech” not as in “free beer”

Stallman's Argument

- Computer software is becoming more and more critical for the running of a free society
- If companies or governments control that software, the software can be used to restrict or monitor people
- So, source code must be available for all software
- If source code were not freely available, a limited number of mighty people would dominate computing

Free Software Foundation http://www.fsf.org/about/
and
DiBona, Chris (ed) (1999). *Open Sources: Voices from the Open Source Revolution.*
O'Reilly & Associates: Sebastopol. pp.2.

What is Free Software?
(using Free Software Foundation definition)

- "Free software is a matter of the users' freedom to run, copy, distribute, study, change and improve the software. More precisely, it means that the program's users have the four essential freedoms:
  - The freedom to run the program, for any purpose (freedom 0).
  - The freedom to study how the program works, and change it to make it do what you wish (freedom 1). Access to the source code is a precondition for this.
  - The freedom to redistribute copies so you can help your neighbour (freedom 2).
  - The freedom to distribute copies of your modified versions to others (freedom 3). By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this. "

Source: http://www.gnu.org/philosophy/free-sw.html

"Copyleft"

- Play on word "copyright"
- "Copyleft is a general method for making a program (or other work) free, and requiring all modified and extended versions of the program to be free as well." (Free Software Foundation)
- Example of a copyleft licence is the GNU Public License (GPL)
  - More in later lectures

http://www.gnu.org/copyleft/

Copyright is a legal concept that provides exclusive rights to creators, while Copyleft is a licensing method that encourages the unrestricted sharing, modification, and utilisation of creative works.

Copyleft symbol

Open Source Software

- Concept of Open Source Software started by Eric Raymond
- Launched in 1998
- He said that the term/concept "free software" was:
  - Ambiguous (e.g. confusion with freely downloadable binaries and shareware)
  - Intimidating to companies (due to ideological approach)
- He focussed on the pragmatic aspects of providing access to source code, rather than the "rights" of users.

Eric Raymond

"Given enough eyeballs, all bugs are shallow"

What is Open Source Software (OSS)? (using Open Source Initiative definition)

- To be classified as OSS, the software must be (according to its licence):
  - Freely redistributable
  - Source code must be available for free or at reasonable reproduction cost
  - Modifications and derived works must be allowed and be distributable under same terms
  - Can protect integrity of author's source code as long as allow source code patches
  - No discrimination against people/groups
  - No discrimination against fields of endeavour
  - Must not be restricted to use with a specific product
  - Must not place restrictions on other software distributed with it
  - Must be technology-neutral

Source: http://www.opensource.org/docs/osd

Difference between Free Software and Open Source Software

- According to Stallman, "Open source is a development methodology; free software is a social movement.“
- Open Source covers a wider range of licence types
- More ability to mix Open Source software with proprietary software than is the case for free software
- The Open Source concept was developed to bring major software businesses and other high-tech industries into the mix.
- When avoiding distinguishing between these, people use the terms:
  - FOSS (Free and Open Source Software); or
  - FLOSS (Free/Libre and Open Source Software)

Source: http://www.gnu.org/philosophy/open-source-misses-the-point.html

Free software and open source software: Examples

- **OSS and copyleft** (changes to the source must be made available to others)
  - The Linux kernel
  - MariaDB (database software based on MySQL codebase)
  - Eucalyptus (for building private clouds – company bought by HP)
- **OSS and not copyleft** (changes to the source do not need to be made available to others)
  - Apache web server
  - OpenCV (Computer Vision library originally by Intel)
  - Chromium (the core of Google Chrome web browser)

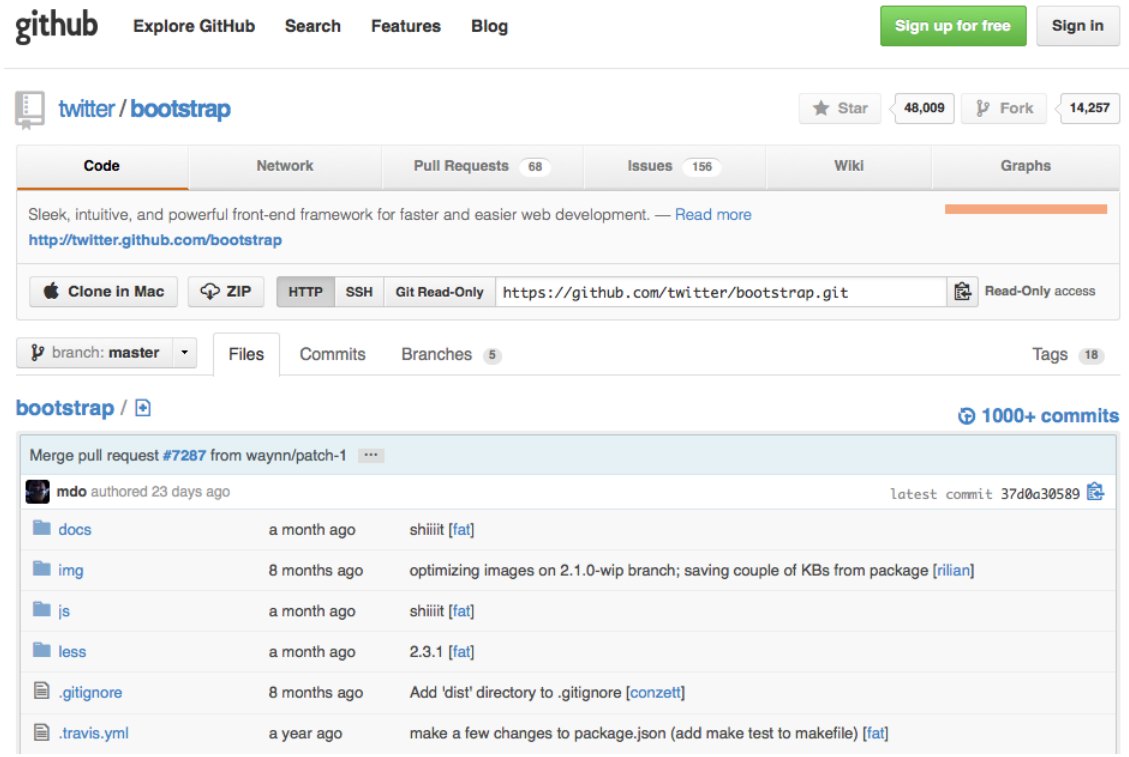
Note: It's a bit more complicated than this as some of this software is available under multiple licences. More later on OSS licences.

Open source hosting sites

- Offer hosting, version control, issue tracking, wikis, download support, etc.
- Some support code reviews, etc.
- Examples:
  - Github (>420 million repositories)
  - Sourceforge (>500k projects)
  - Google Code (250k projects) (closed down Jan 2016)

Github

- Github (>500 million projects):
  - E.g. Ruby on Rails (web application framework), jquery (JavaScript query engine), node.js (evented i/o for javaScript), Diaspora (distributed social networking)
- GitHub is free to use for public and open source projects. Work together across unlimited private repositories with a paid plan.



New Model for FOSS?

elementary OS community logo ELEMENTARY

- Inspired by an established business tactic in the gaming space, Foré and his team flipped a somewhat controversial switch that led to a massive increase in the company's income: they simply started charging money for it.
- 'pay what you want' – the price floor is zero so people can still go in and get it for free.
- This practice -- and its outcome -- seems to contradict a belief many Linux and FOSS advocates cling to: that donations to their favorite Linux distribution or open source projects do matter, but paying for them up front is a line that shouldn't be crossed.

https://www.forbes.com/sites/jasonevangelho/2019/09/03/paying-for-linux-distros-and-foss-software-successful-elementary-os/#3ce63a5f3b3a (Mar'25)

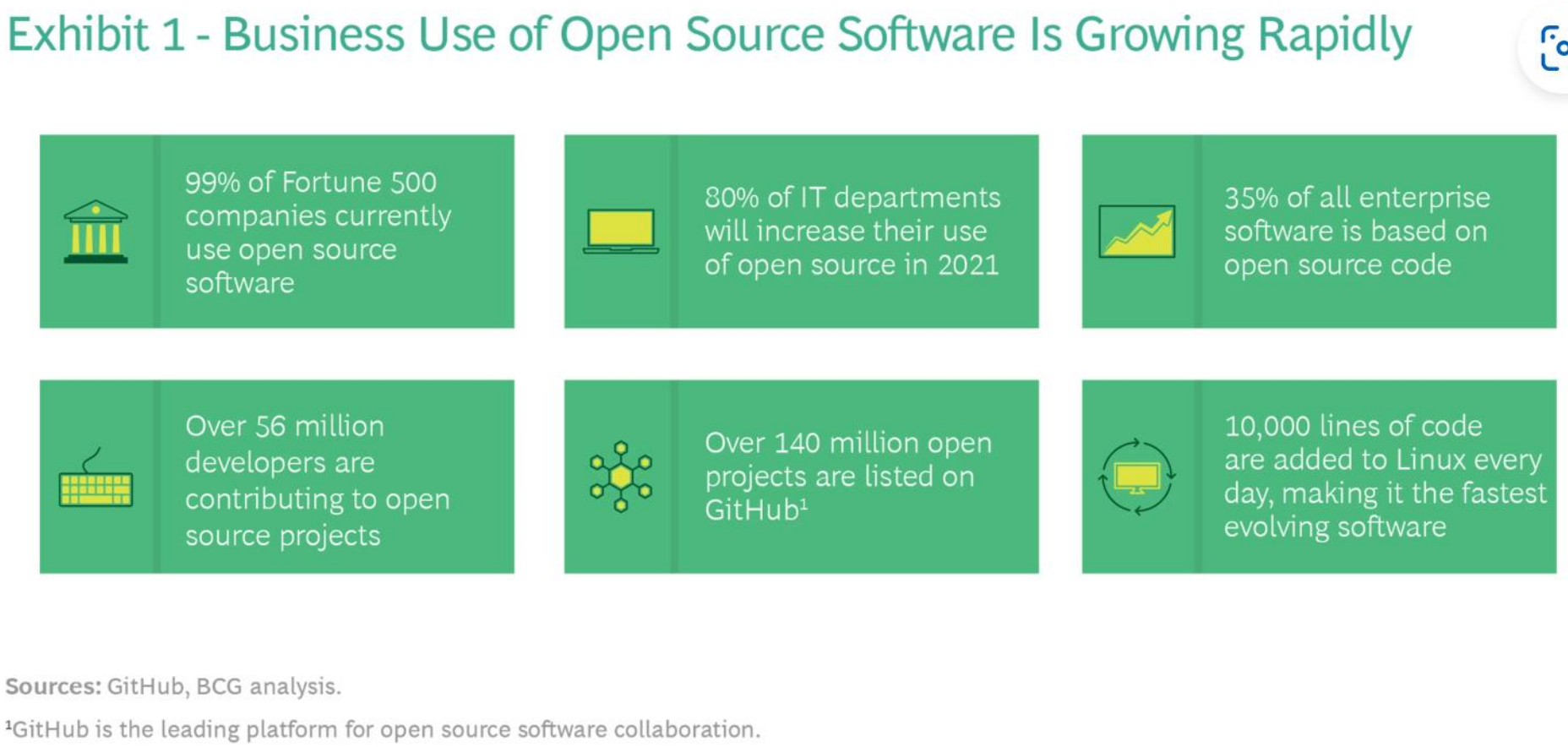
Importance of Free and Open Source Software (FOSS) for innovation

Growing importance of free and open source software

- Many companies use open source software such as for:
  - Internal IT infrastructure (e.g. Linux)
  - Building and running web services (e.g. Apache, Apache Tomcat, JBoss)
  - Building software for redistribution

- Open-source software allows companies to innovate their infrastructure and services rapidly

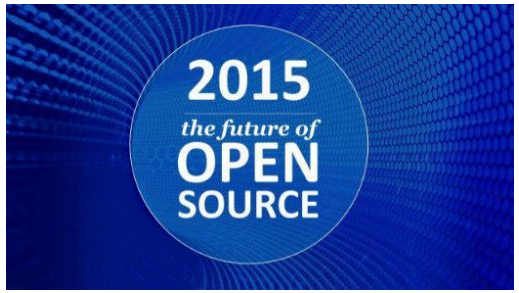
Why You Need an Open Source Software Strategy



Why You Need an Open Source Software Strategy | BCG (Mar'25)

Open Source usage in Enterprises

- **Survey** by North Bridge and Black Duck Software
- Input from 1300 senior IT professionals, in 2015
- 78% of companies run all or part of their operations on open source software
  - Compared to 42% in 2010
- 3% of companies said they don't use open source software
- 64% say that their company participates in open source projects
- 66% say that they consider open source software before considering proprietary software



Importance of FOSS

- Why we need open-source science innovation — not patents and paywalls (theconversation.com)
- Google Open Source
- Open source is more important than ever, say developers. Here's what's driving adoption | ZDNET

Importance of FOSS in R&D and startups

- Most infrastructure used in R&D and startups uses FOSS:
  - Operating systems (e.g. Linux)
  - Containers (e.g. Docker)
  - System configuration management (e.g. Puppet, Chef)
- Most new software is built using FOSS:
  - Software platforms (e.g. Java, Scala, Python, Ruby on Rails, node.js)
  - Software libraries/frameworks (e.g. Spring framework, glibc)
  - Software build and test automation (e.g. Jenkins, Cucumber)
- Most new software contains FOSS:
  - To reduce the time and cost of development
  - To reduce testing and maintenance costs (assuming using stable FOSS)
  - To provide compatibility with other software
  - To focus on the core differentiator of your own software

Some open source business models

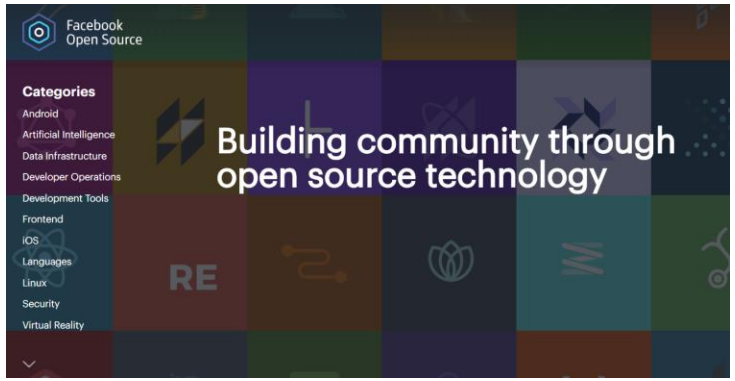
- Sell support and services
  - Example: Canonical (with Ubuntu)
- Sell certified version (with support and services)
  - Example: Cloudera (with Hadoop)
- Sell "enterprise edition" (effectively proprietary software)
  - Example: MySQL "standard edition" (not "community edition")
- Dual licensing (copyleft so need commercial license if modify source)
  - Example: Digia (with Qt)
- Other advantages to the company
  - Example: Google (with Android)

Video: Tim O'Reilly on Open Source Business Models

http://www.youtube.com/watch?v=GYarQ1r2yZo

Example: Facebook and Open Source

- Accelerates innovation in the world. Users can build apps more quickly and Facebook can benefit from the improvements that others make to their code.
- If the company knows something will be open from the start, it just builds it better so that it can be more accessible and dependable because it's going to be used in the outside world.
- Open source provides opportunities to share challenges. It attracts the interest of people who want to work on these challenges, and as a result it helps improve the quality of the company's staff.

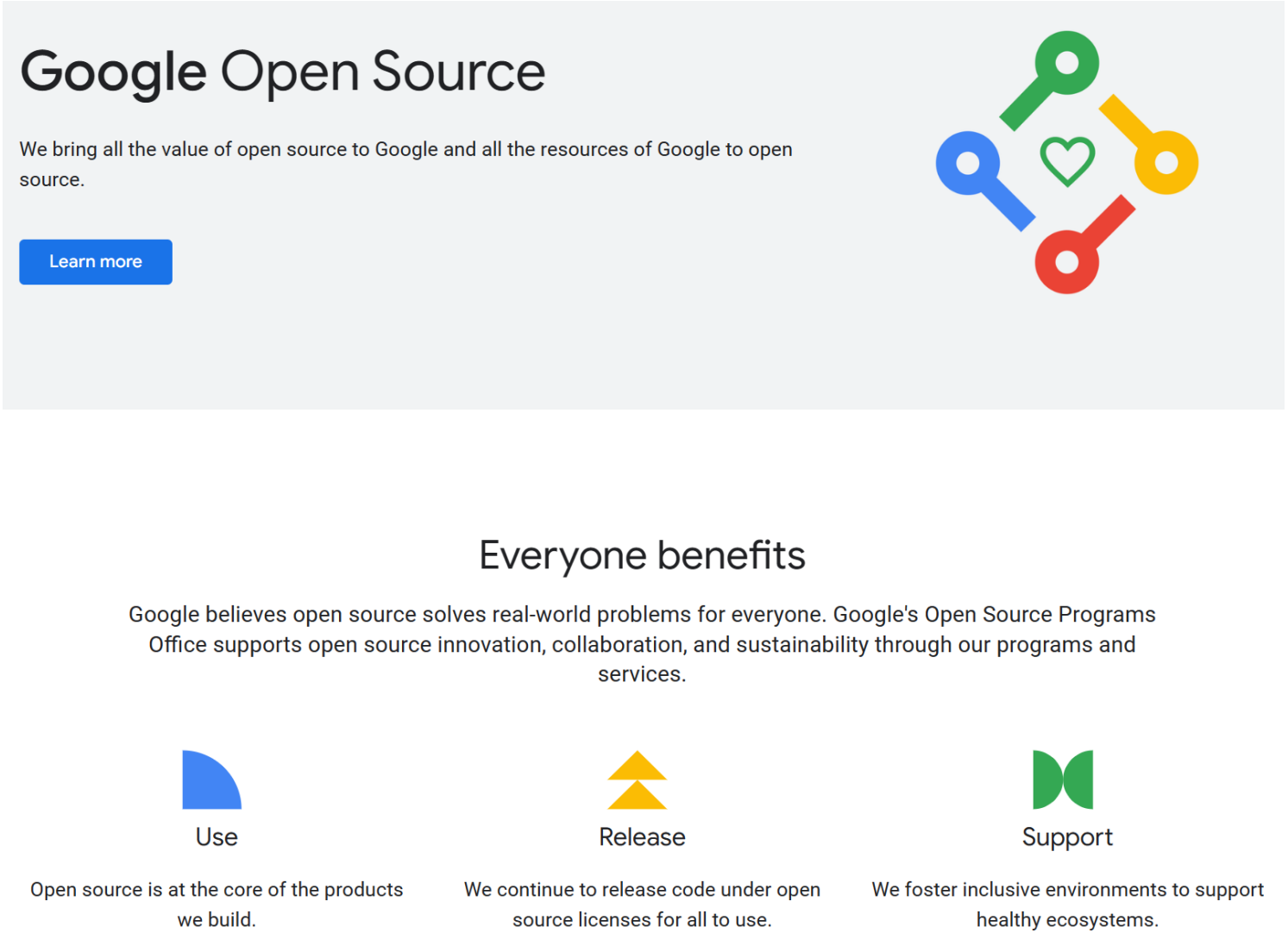


https://opensource.com/business/15/7/keynote-oscon-james-pearce-facebook (Mar'25)

https://opensource.fb.com/

Google Open Source

- "Open source makes it possible for us to work together." —Google OSPO Founder Chris DiBona
- A key part of OSPO's mission is helping Google-led projects thrive and grow. Some of Google's most notable open-source projects were released in 2008, including Android



Example: Apple and Swift programming language

- Apple is among the most closed tech companies, yet it is doing open source with its **Swift** programming language
- Swift is open source, and they want you to help make it the best general-purpose programming language available everywhere.

- The success of a software platform is proportional to the number of developers that use it.
- Apple wants to attract the best developers in the world to create new apps and desktop applications that showcase its latest iPhones, Apple Watches and MacBook devices.

http://www.computerweekly.com/news/450296755/Why-Apple-is-wooing-open-source-developers-with-Swift (Mar'25)

https://swift.org/

Example: Microsoft and .Net

- Microsoft is porting its server-side .NET stack to Linux and Mac OS X, and is making more of that stack available as open source. With its engineers involved in more than 2,000 open source projects, you'd have to agree that open source has more than a foothold at Microsoft these days.
- Microsoft also wants to bring technologies to Linux, in large part because of Azure.
- Running a cloud platform gives Microsoft an interest in Linux that goes far beyond the open source contributions the Windows Server team has been making to the Linux kernel
- As of September 2015, more than 20 percent of the virtual machines running on Azure IaaS (Infrastructure as a service) were Linux.
  - "As we pursue our vision of the fabric and the cloud anywhere, that is as much a story about supporting Linux workloads as it is Windows workloads," says lead architect for Windows Server, Jeffery Snover.

http://www.zdnet.com/article/microsoft-to-open-source-more-of-net-and-bring-it-to-linux-mac-os-x/ (Mar'25)

Open-source lab model (e.g. Amplab at UC Berkeley)

- Berkeley Lab: Multi-year collaborative effort at UC Berkeley
- Sponsors: Amazon, Google, IBM, SAP + 19 more
- Open Source Software developed:
  - Spark (cluster computing framework), Mesos, Tachyon, GraphX, MLBase
- Companies formed:
  - Mesosphere ($122.25M invested), Databricks ($47M invested), Tachyon Networks ($17M invested)

https://amplab.cs.berkeley.edu/

Amplab: benefits to participants

- What do the companies get:
  - Deep knowledge of technology as it is developed
  - Influence technology direction and outcome
  - Protection from disruption by other companies
  - Access to university talent pool
- What does the university get:
  - Clear focus for computer science research
  - Additional income to fund activities
  - Close interaction with market needs
  - Incubation of new businesses
  - Stronger innovation ecosystem

Challenges in using FOSS in products and services

- Meeting obligations of software licenses (ensuring appropriate notices, etc.)
- Possibility of accidentally "contaminating code"
  - E.g. a programmer introduces some GPL (General Public Licence) code from the Internet into some proprietary product code and then the product is released
    - legally, the company should release the proprietary source code
- Ensuring adequate quality of the final product if it includes some open source software of unknown quality
- Avoiding security vulnerabilities in underlying code (that may already be known to hackers)

How can companies address the challenges

- Companies developing products (hardware or software) or services and using open-source software should have an open-source policy and controls to ensure good governance.

- According to a Gartner report, <50% of Global 2000 IT Companies planned to implement an open-source governance program by 2014.
- In a related survey, only 1/3 of the companies surveyed had an OSS policy

Source: *http://blog.blackducksoftware.com/2011/07/29/policy-precedes-controls/* (Mar'25)

Tools for managing open source software

- Most open source analysis tools provide ways for checking if open source software is used and check if usage conforms to a policy (e.g. a company OSS policy)
  - Black Duck Software (http://www.blackducksoftware.com/)
  - Software tools and online Knowledgebase containing open-source software
- Palamida (http://www.palamida.com/)
- FOSSology (http://www.fossology.org/)
  - Developed by HP – released as open source software
  - Openhub.net (https://openhub.net/)
  - Now owned by Black Duck Software
  - Free online service for searching open source code
  - Searching of over 21b lines of FOSS code

Case Study

Innovations with Maps

https://commons.wikimedia.org/wiki/File:World_Map_1689.JPG

THE UNIVERSITY OF SYDNEY

Geographic Information System (GIS)

- With the advent of data mining, machine learning, mobile applications, the Internet of Things, social media, and other recent additions to the computing landscape, there are orders of magnitude more data available across every discipline, and the race to do useful and interesting things is in full throttle.
- Once limited to more traditional realms like land records, agriculture, natural resources, and urban planning, GIS now permeates practically every field.

https://opensource.com/life/15/11/getting-started-web-mapping (Mar'25)

Pymble, Sydney, Ku-ring-gai Council, New S

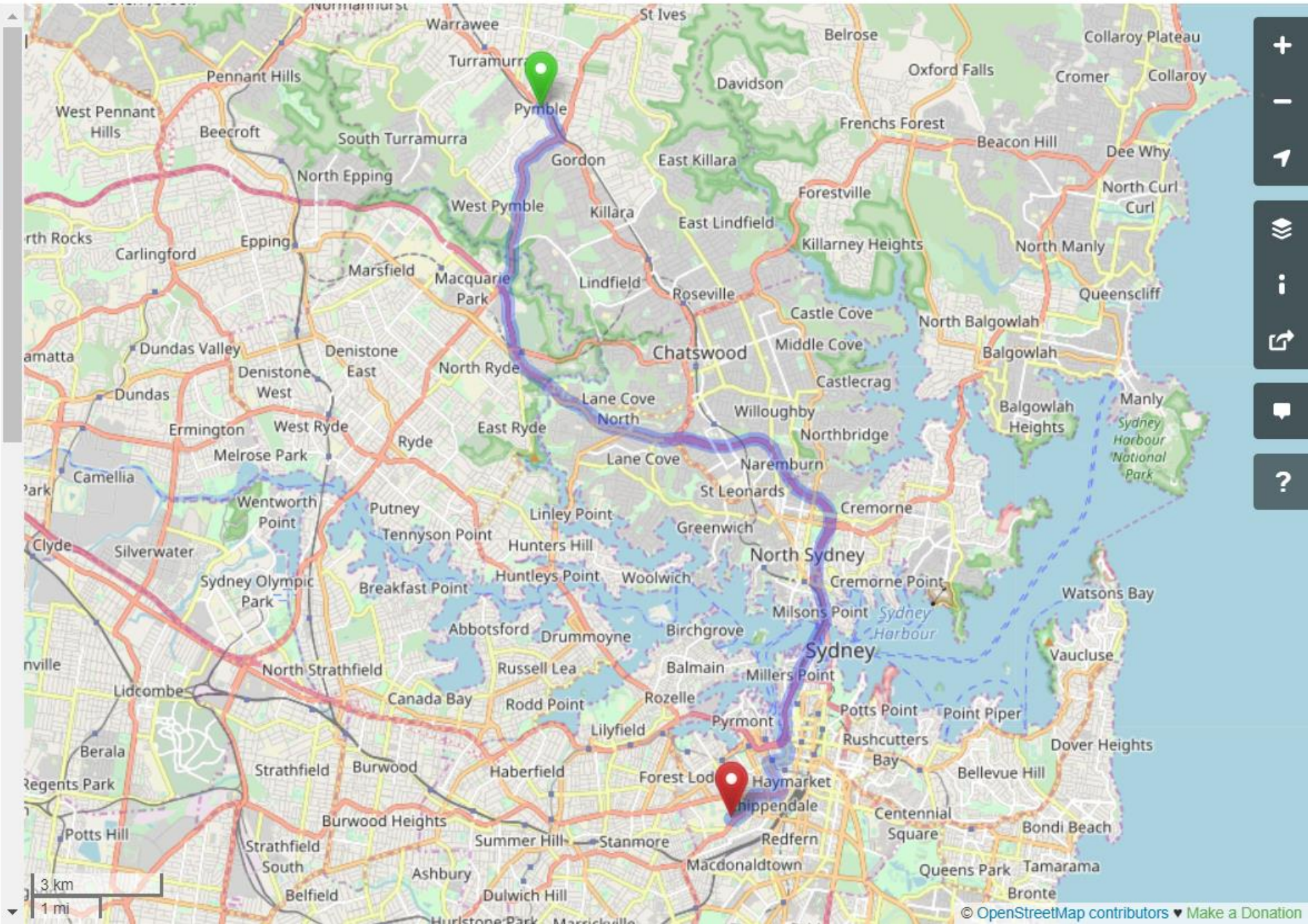
The University of Sydney, Fisher Road, Cam

Car (OSRM) ▾ Go

Directions

Distance: 24km. Time: 0:26.

1.	Start at end of unnamed road	20m
2.	Turn left onto **Pacific Highway**	700m
3.	Take the sliproad on the right onto unnamed road	200m
4.	Continue on **Ryde Road Onramp**	200m
5.	Merge right onto **Ryde Road**	2.9km
6.	Continue on **Lane Cove Road**	700m
7.	Take the sliproad on the left onto **M2 Hills Motorway Onramp**	50m
8.	At the fork turn left onto **M2 Hills Motorway Onramp**	200m
9.	Merge right onto **M2 Hills Motorway**	2.6km
10.	Continue on **Lane Cove Tunnel**	2.5km
11.	At the fork turn right onto **Lane Cove Tunnel**	1.8km



OpenStreetMap

OpenStreetMap

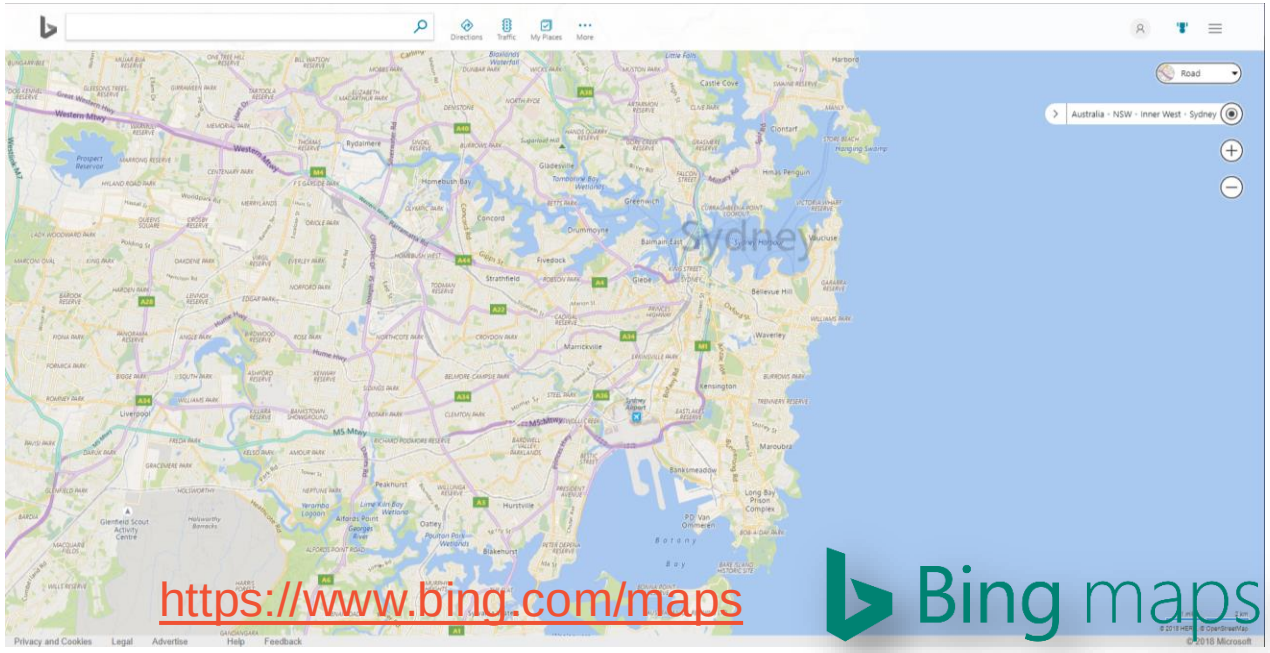
https://www.openstreetmap.org/

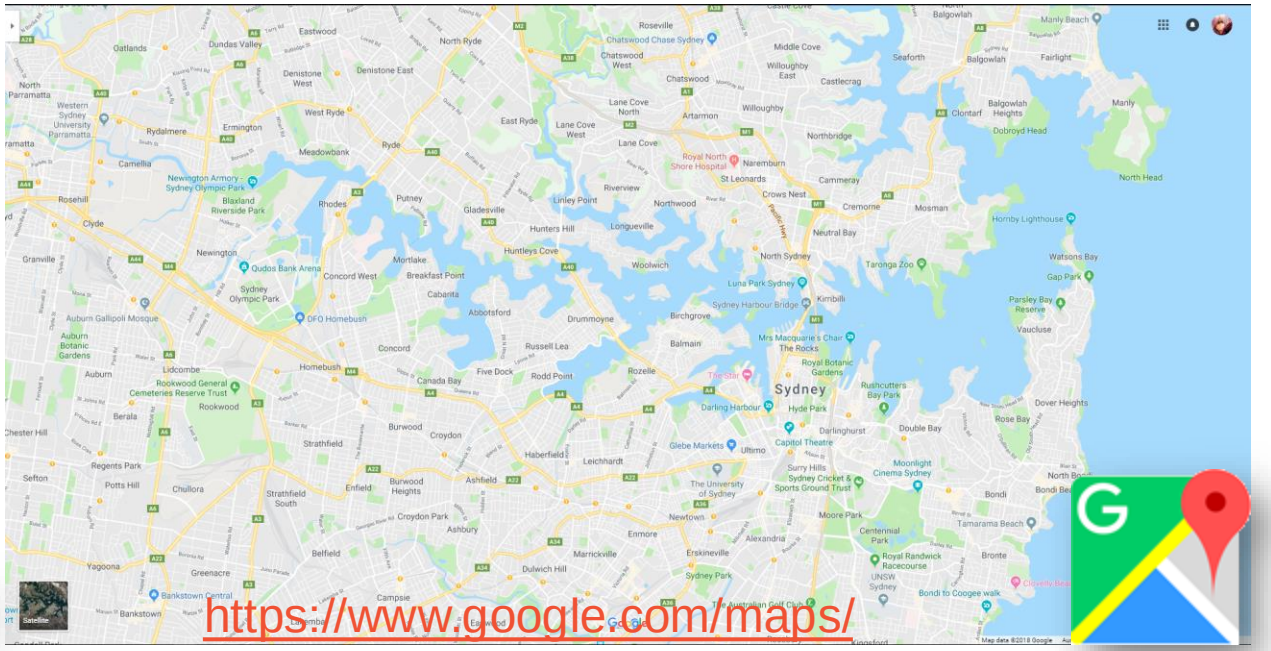
- Crowd sourcing
- Open source
- User innovation

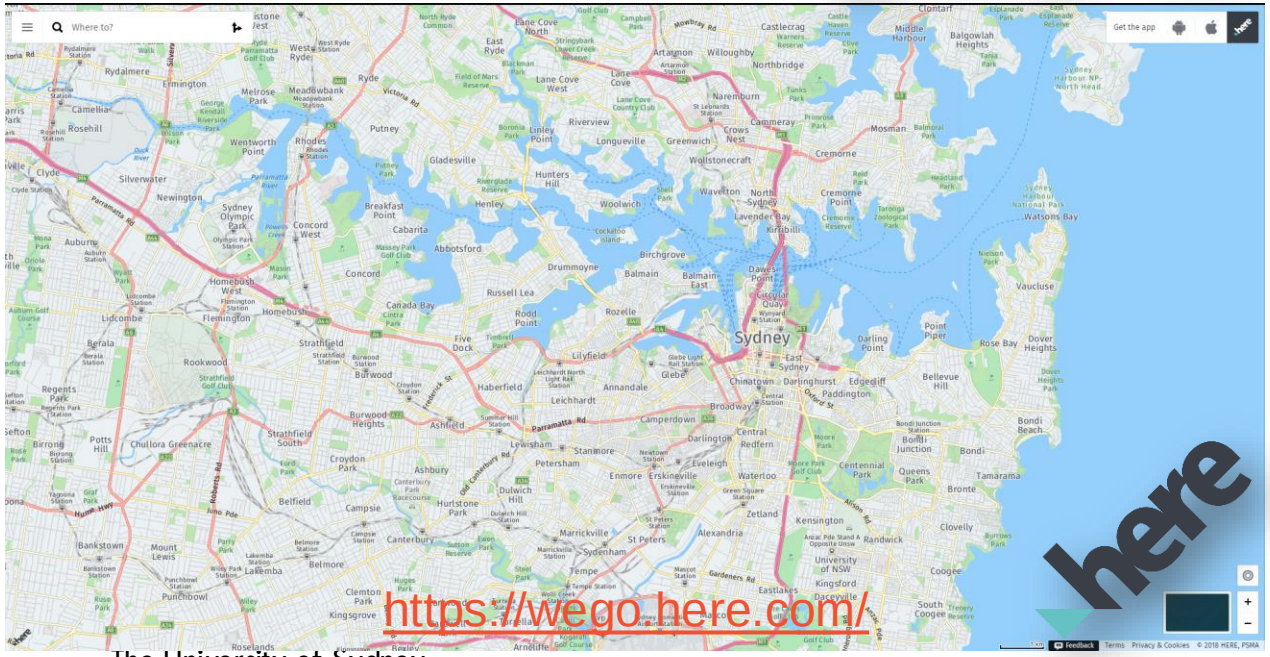
OSRM

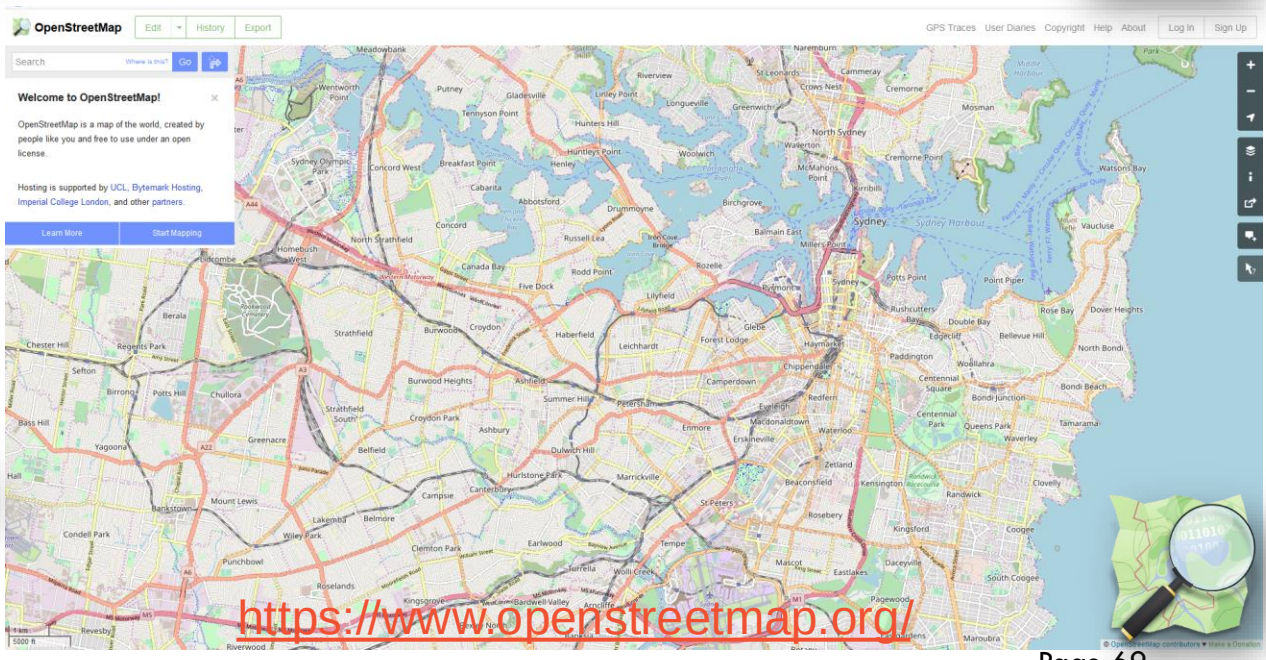
OPEN SOURCE ROUTING MACHINE

http://project-osrm.org/


https://www.bing.com/maps
Bing maps

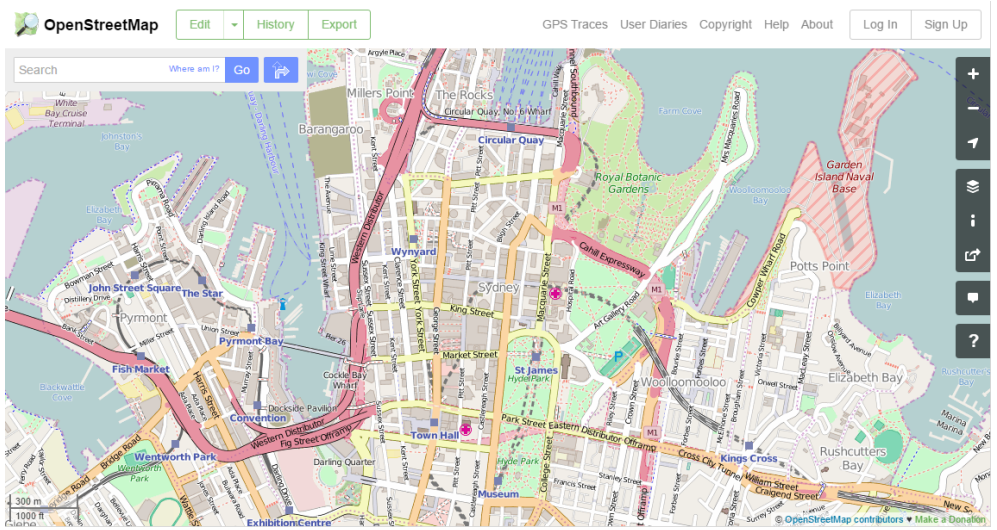

https://www.google.com/maps/


https://wego.here.com/
here


https://www.openstreetmap.org/
Welcome to OpenStreetMap!

An open source map

- Open Street Map



http://www.openstreetmap.org/#map=15/-33.8680/151.2100

M Haklay, and P Weber, “Openstreetmap: User-generated street maps”, *IEEE Pervasive Computing*, 2008 - ieeexplore.ieee.org

OpenStreetMap Foundation

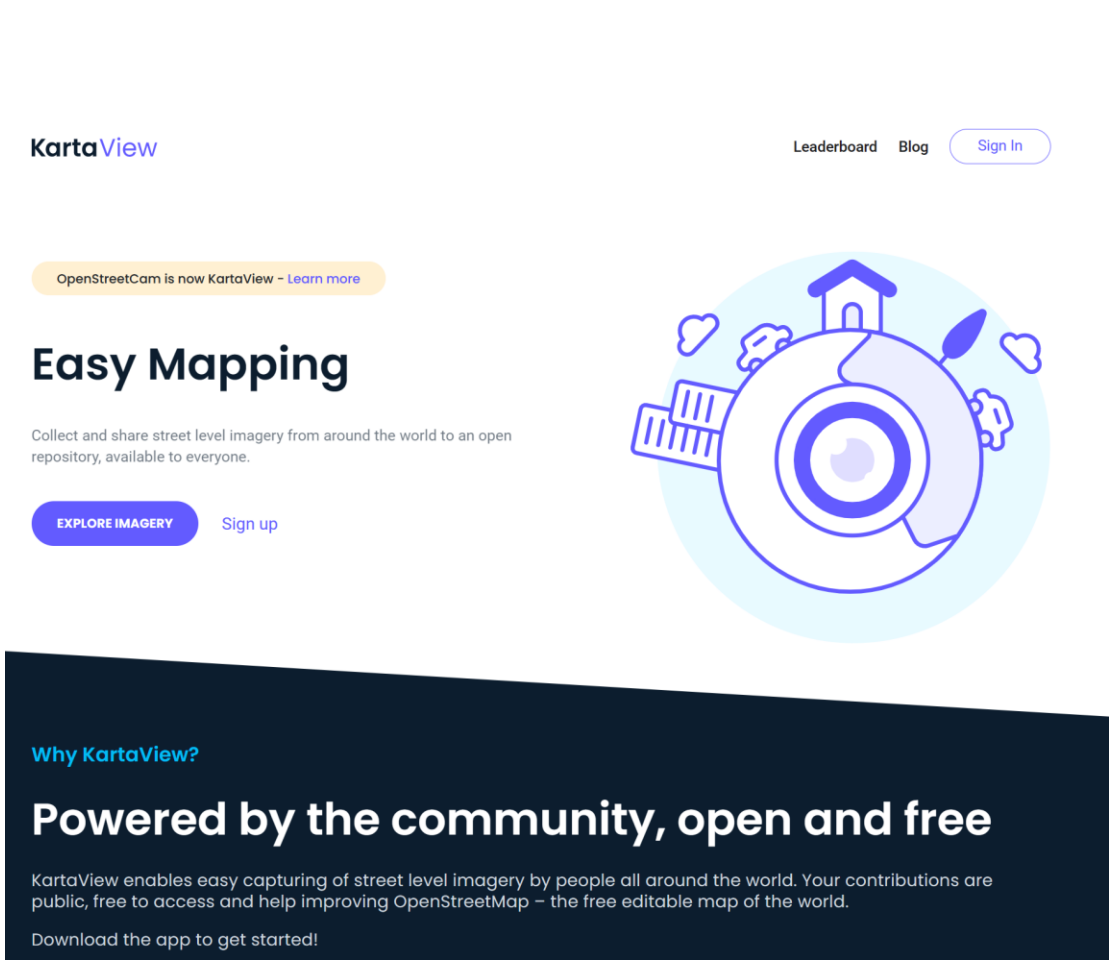
- OpenStreetMap is an initiative to create and provide free geographic data, such as street maps, to anyone.
- The OpenStreetMap Foundation is an international not-for-profit organization supporting, but not controlling, the OpenStreetMap Project.
- It is dedicated to encouraging the growth, development and distribution of free geospatial data and to providing geospatial data for anyone to use and share.

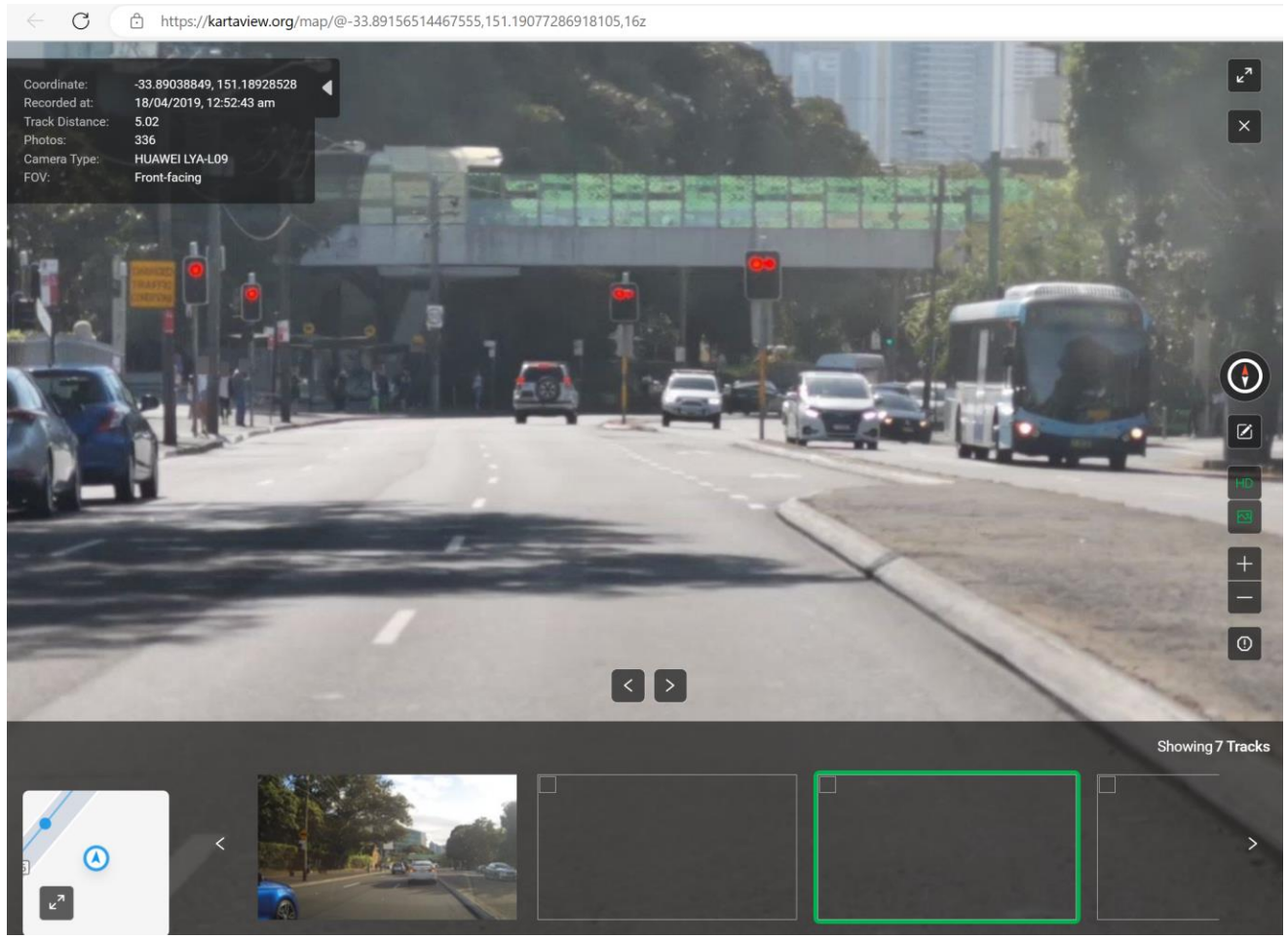
http://wiki.osmfoundation.org/wiki/Main_Page (Mar'25)

OpenStreetMap

- Open source – enable innovations to happen
  - introduce services / businesses
  - for non-profit communities
- Crowd sourcing – over 2M users collecting data using manual survey, GPS and other free sources
  - Government and 'out of copyright' data
  - Innovating through the power of the crowd
- Open data license – share left

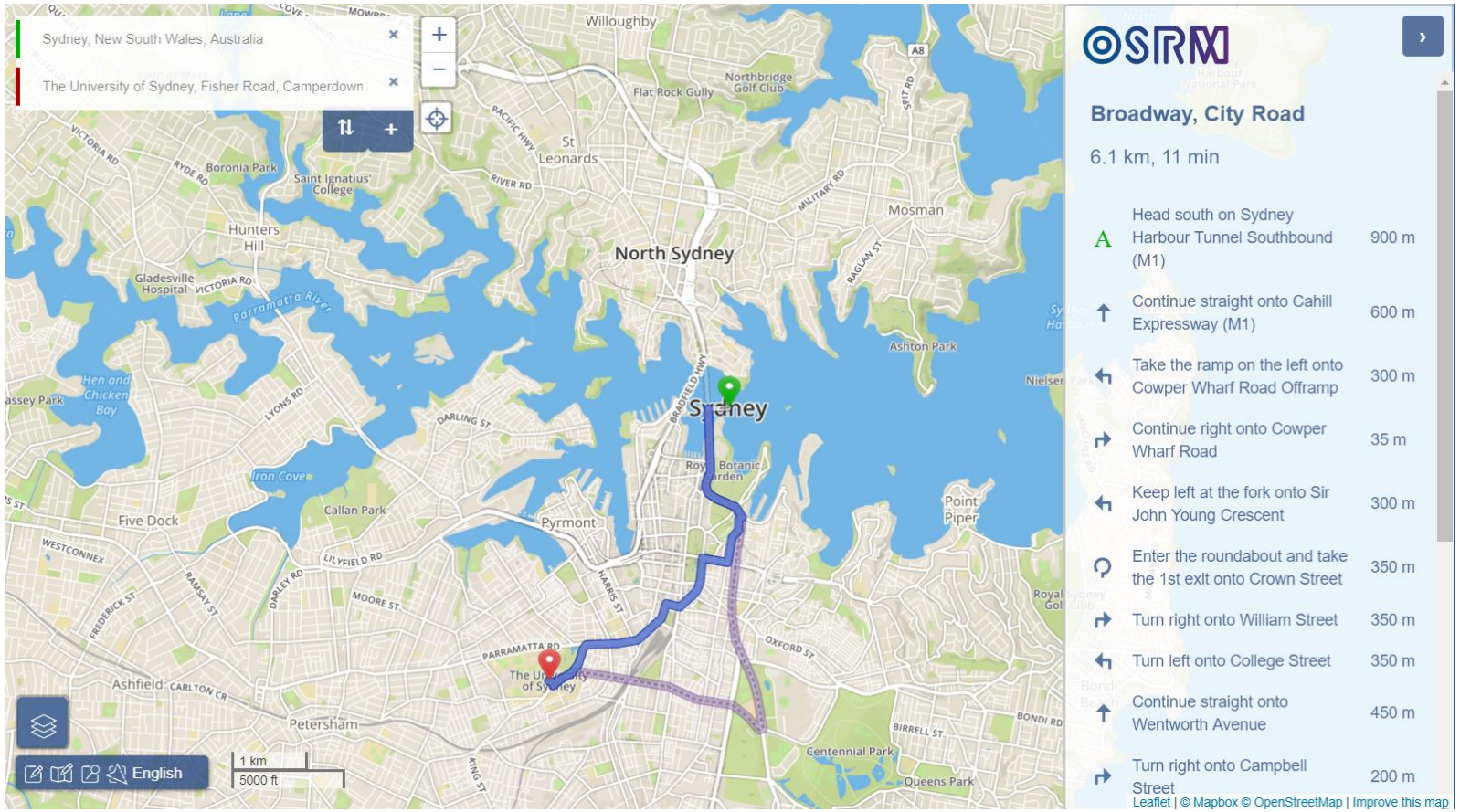
- Leading to many new innovations!


KartaView
Leaderboard
Blog
Sign In
OpenStreetCam is now KartaView - Learn more
Easy Mapping
Collect and share street level imagery from around the world to an open repository, available to everyone.
EXPLORE IMAGERY
Sign up
Why KartaView?
Powered by the community, open and free
KartaView enables easy capturing of street level imagery by people all around the world. Your contributions are public, free to access and help improving OpenStreetMap – the free editable map of the world.
Download the app to get started!


https://kartaview.org/map/@-33.89156514467555,151.19077286918105,16z
Coordinate: -33.89038849, 151.18928528
Recorded at: 18/04/2019, 12:52:43 am
Track Distance: 5.02
Photos: 336
Camera Type: HUAWEI LYA-L09
FOV: Front-facing
Showing 7 Tracks

KartaView

Open Source Routing Machine



http://project-osrm.org/

Pokemon GO

Users of Open Source Maps….

- Pokémon Go is a mobile augmented reality game by Niantic in which the player catches and collects Pokémon that "spawn" at locations all over the map. The game was released for Android and Apple iOS in July 2016.
- Niantic switched the base map from Google Maps data to OpenStreetMap data in December 2017.
- Since its initial release, many Pokémon Go players have edited OpenStreetMap in an attempt to improve their gameplay. Many players have made good-faith edits to OpenStreetMap, such as adding buildings and roads that reflect reality on the ground. On the other hand, there have also been many cases of Pokémon Go players adding spurious parks or deleting schools in an attempt to essentially tag for the renderer.
- Pokémon Go players are strongly encouraged to improve the map but are expected to follow good practice like any other mappers.

https://wiki.openstreetmap.org/wiki/Pok%C3%A9mon_Go

Openpilot — open source advanced driver assistance system

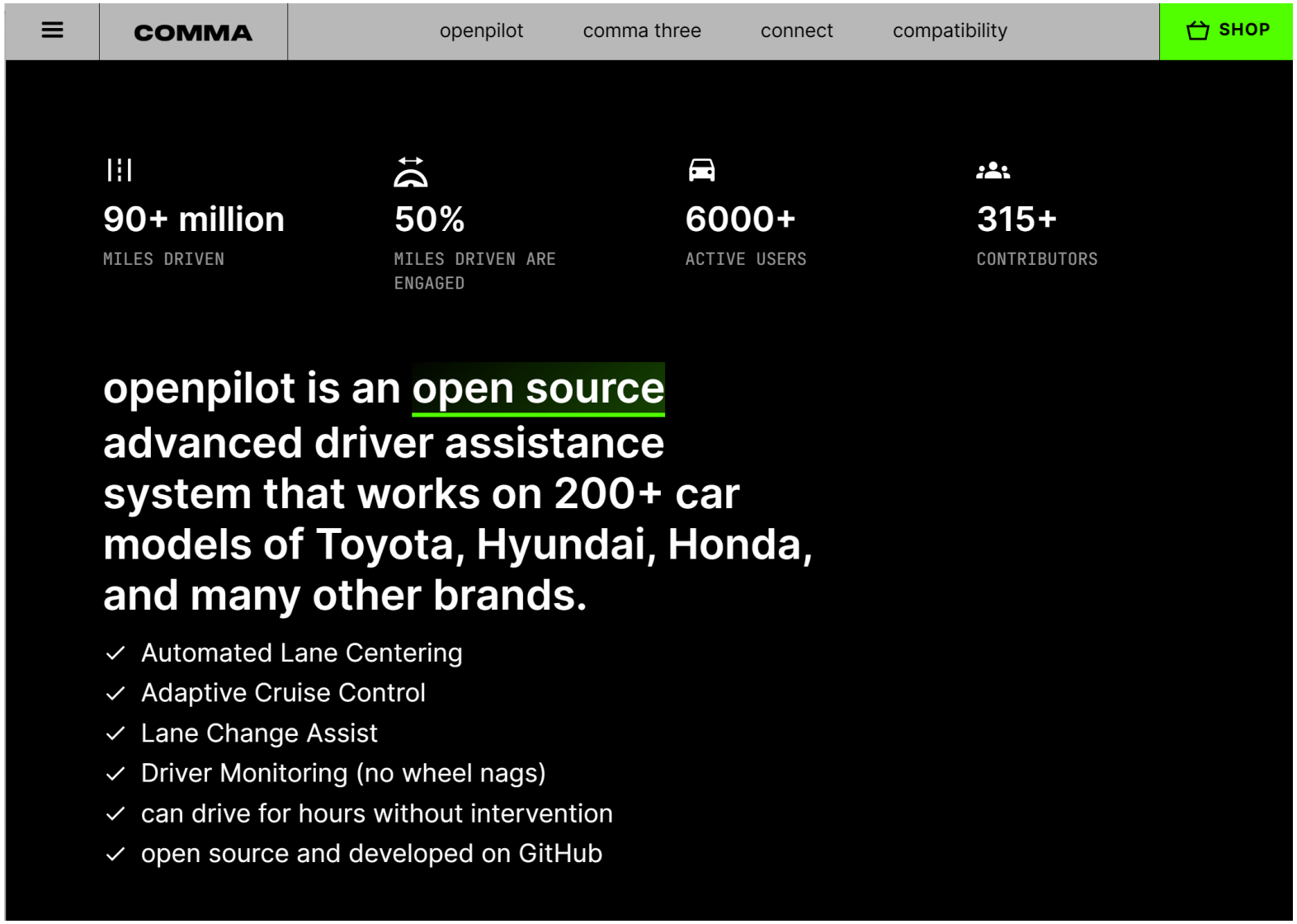
openpilot — open source advanced driver assistance system (comma.ai)

04

open source and community supported

openpilot is developed by comma and by users like you. We welcome both pull requests and issues on GitHub. Bug fixes and new car ports are strongly encouraged. Check out the contributing docs.

VIEW ON GITHUB →



FOSS Licenses

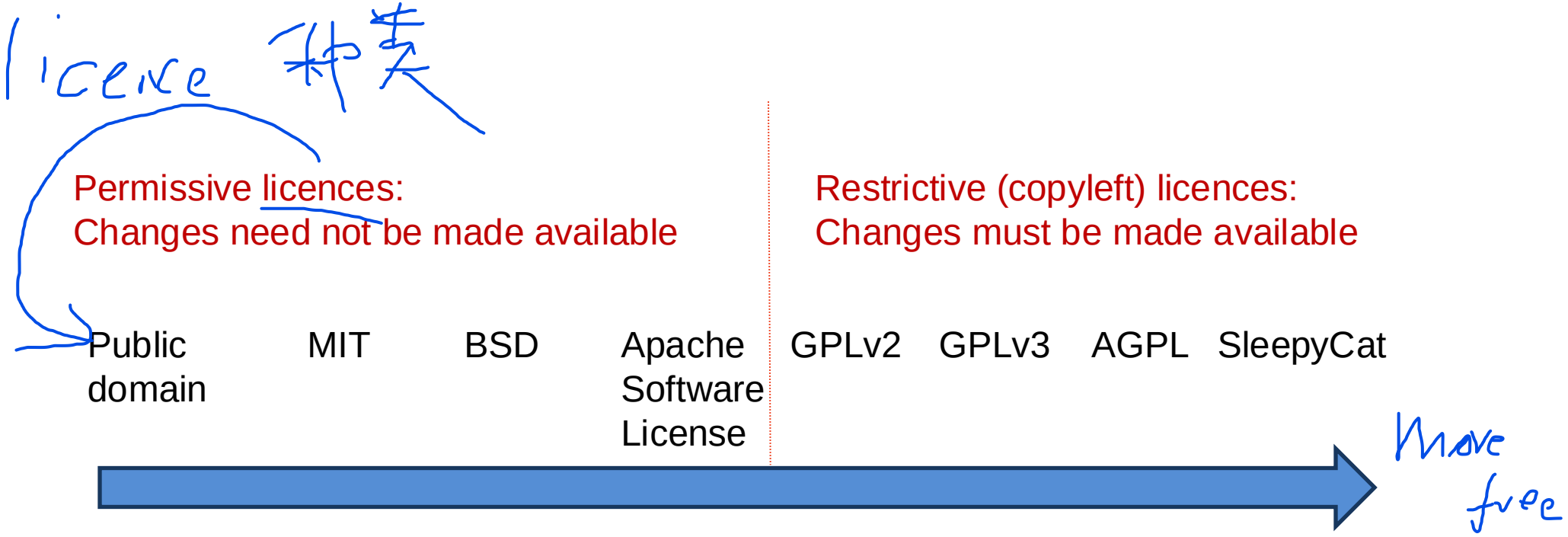
Usage and Licenses

THE UNIVERSITY OF SYDNEY

Obligations when using open source software

- The obligations depend on the actual software licence used by the software
- Your obligations may include:
  - Nothing (i.e. no special obligations); or
  - If you redistribute the open source software in your software:
    - Mentioning that you have used it; or
    - Redistributing any changes you made to it; or
    - Not suing other companies in relation to patents you may hold related to the features of the open source software; etc

Summary of Main open source licences



- More restrictions/conditions on users
- More assurances of software staying free

Public domain

- Work in the public domain does not have intellectual property rights
  - (eg the right has expired or has been deliberately placed in the public domain)
- Examples: the English language, Shakespeare's works, Beethoven's music, many old photos for which copyright has expired
- Not commonly used for software because:
  - As software development is a recent activity, copyright hasn't expired yet
  - Author can't make disclaimer (unlike open source licences)

Massachusetts Institute of Technology Licence (MIT License)

- User can do anything with the software…
- But they must make sure that the copyright of the original author is maintained
- No warranty

Copyright (C) <year> by <copyright holders>

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Berkeley Software Distribution Licences (BSD Licences)

- Similar to MIT Licence but, if redistribute software is using it, it must acknowledge its use
- 4-clause (original), 3-clause ("modified") and 2-clause ("simplified") versions exist
- 3-clause version:

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

• Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

• Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

Removed in simplified version

• Neither the name of the <organization> nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL <COPYRIGHT HOLDER> BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

GNU General Public Licence (GPL)

- More restrictive than MIT and BSD – it is copyleft
- You can use the code and change it, but you must release all modified code under the same licence and any other code of yours that touches it
- 2 main versions - GPL v2 and GPL v3
- http://www.gnu.org/licenses/gpl.html
- Clause 5 of GPL v3:
- You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:
- a) The work must carry prominent notices stating that you modified it, and giving a relevant date.
- b) The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to "keep intact all notices".
- c) You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not invalidate such permission if you have separately received it.
- d) If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.
- A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an "aggregate" if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation's users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate.

When creating open source software: How do you know what licence to use?

- Or use "Dual-licensing" – this is now very common
- Build up the market first and then provide services
- Eg:
  - Software can be licensed as GPL or proprietary licence
  - If a company doesn't want to make their changes available, they can come to you to negotiate a proprietary licence

When creating open source software: How do you know what licence to use?

Permissive licences:
Changes need not be made available

Restrictive (copyleft) licences:
Changes must be made available

Public domain | MIT | BSD | Apache Software License | GPLv2 | GPLv3 | AGPL | SleepyCat

If:
- You want a lot of companies to adopt your software in their products/services, and
- You don't care if they make their changes available (eg as you just want the code to be used or you have deep enough knowledge & expertise that they will come back to you):

=> use a permissive licence (eg BSD, Apache)

If:
- You want to ensure that companies (using your software in their products) make their changes available (so you and others can get them):

=>use a restrictive licence (eg GPLv3)

Example: Open CV

- OpenCV is released under a BSD license and hence it's free for both academic and commercial use.

OpenCV license

By downloading, copying, installing or using the software you agree to this license.
If you do not agree to this license, do not download, install, copy or use the software.

License Agreement
For Open Source Computer Vision Library
(3-clause BSD License)

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

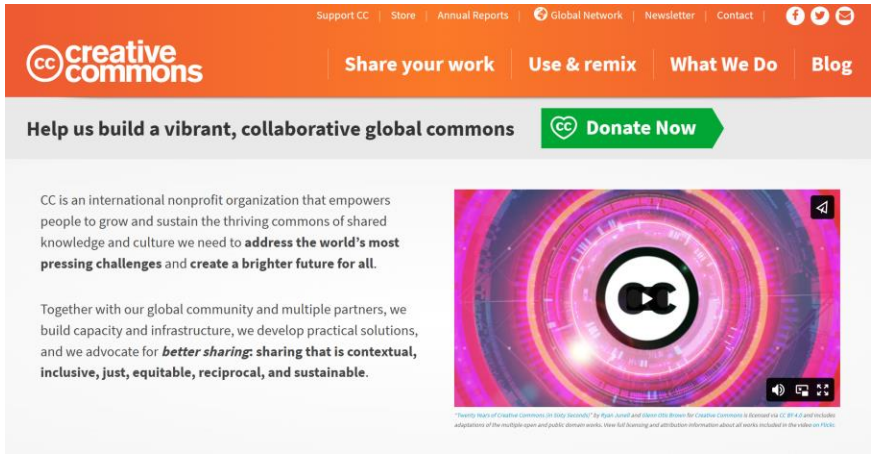
- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the names of the copyright holders nor the names of the contributors may be used to endorse or promote products derived from this software without specific prior written permission.

This software is provided by the copyright holders and contributors "as is" and any express or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall copyright holders or contributors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

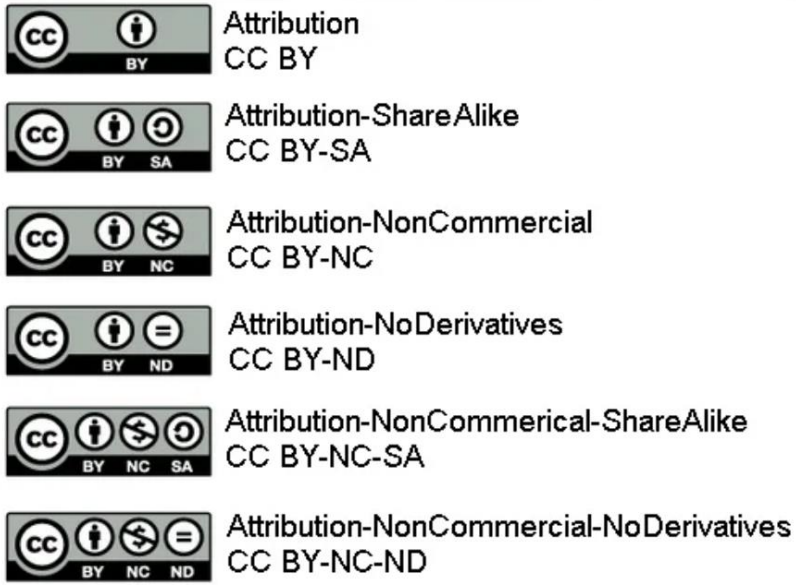
Creative Commons

- Provide Creative Commons licenses and public domain tools that give every person and organization in the world a **free, simple, and standardized way to grant copyright permissions for creative and academic works; ensure proper attribution; and allow others to copy, distribute, and make use of those works**

- Creative Commons licenses give everyone from individual creators to large institutions a standardized way to grant the public permission to use their creative work under copyright law. From the reuser's perspective, the presence of a Creative Commons license on a copyrighted work answers the question, "What can I do with this work?"

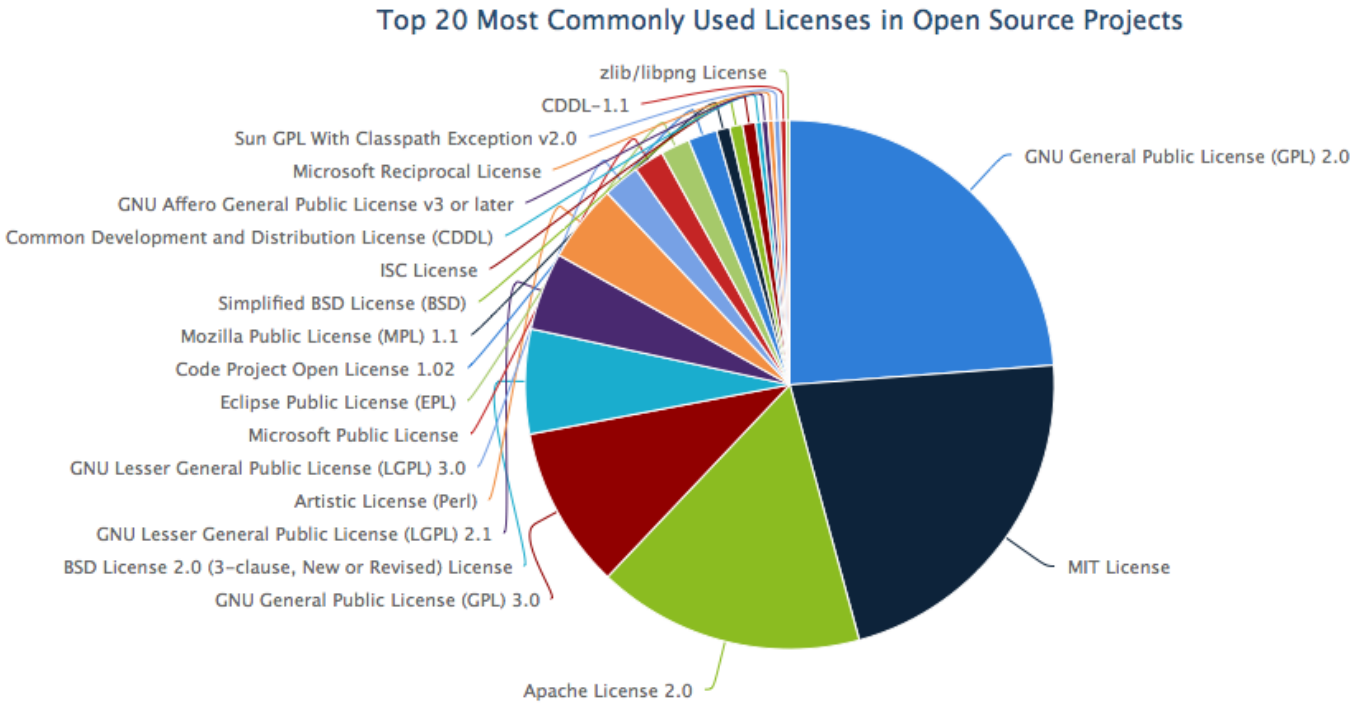


When we share, everyone wins - Creative Commons



Most common open source licences

- The Black Duck Software Knowledgebase contains OSS packages from:
  - Over 1.1 million open source projects from more than 8,500 sites
  - Using 2,400 different software licences



Source: http://www.blackducksoftware.com/ (Mar'25)