JOHAN LINÅKER (RISE)

Open Source Software and AI

- The What, the Why, the How, and the When

First, what's Open Source Software?

Liberally Licensed, Collaboratively Developed Software

Photo by Billy Huynh | https://unsplash.com/photos/v9bnfMCyKbg

Open Source Software (OSS) today

- Approximately…
  - 95+ % of all software contains OSS components
  - Some estimates showing 90% of code in applications consists of open source code
  - 150 million developers, and 4 million organizations collaborate on OSS projects on GitHub.
  - Collaboration in and between verticals, including Energy, Automotive, Telco, Health

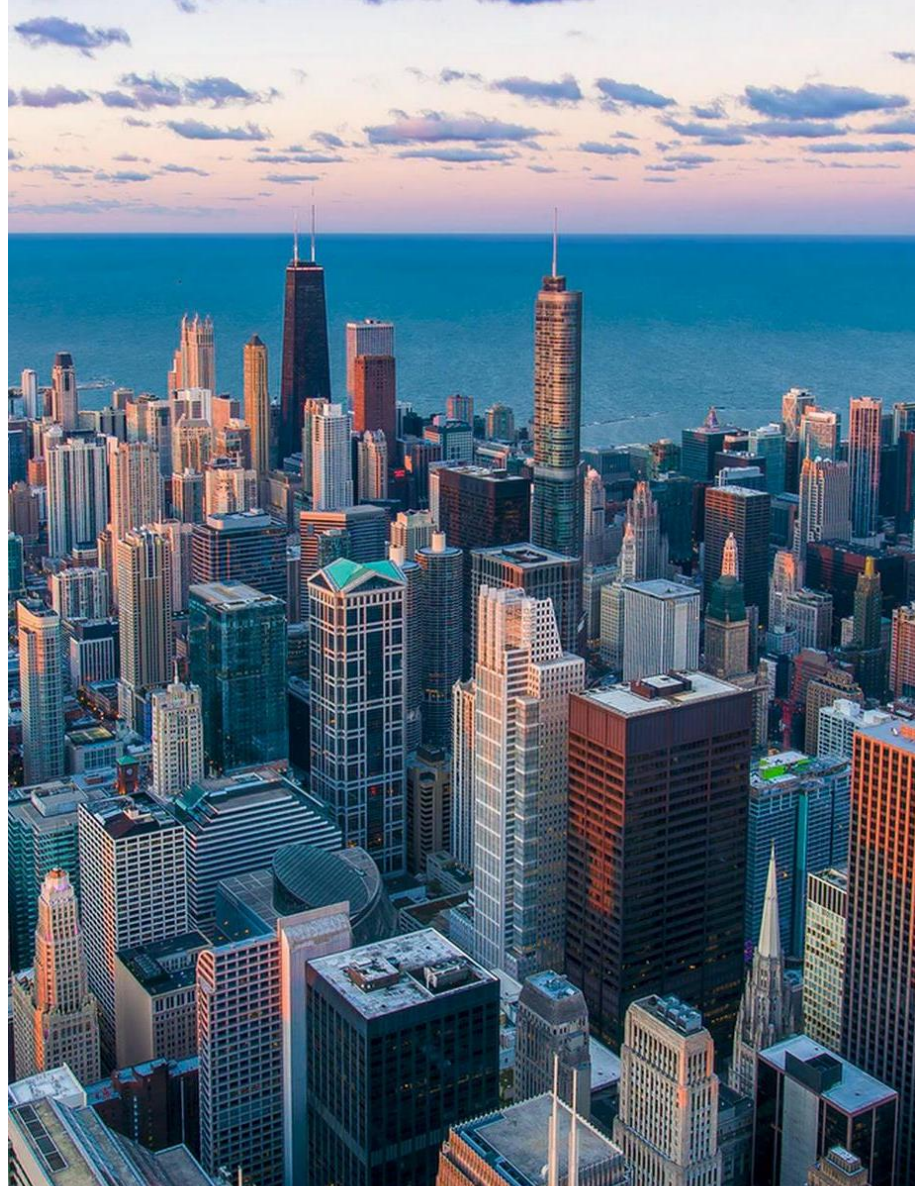


Photo by Pedro Lastra | https://unsplash.com/photos/Nyvq2juw4_o

In other words, it’s everywhere

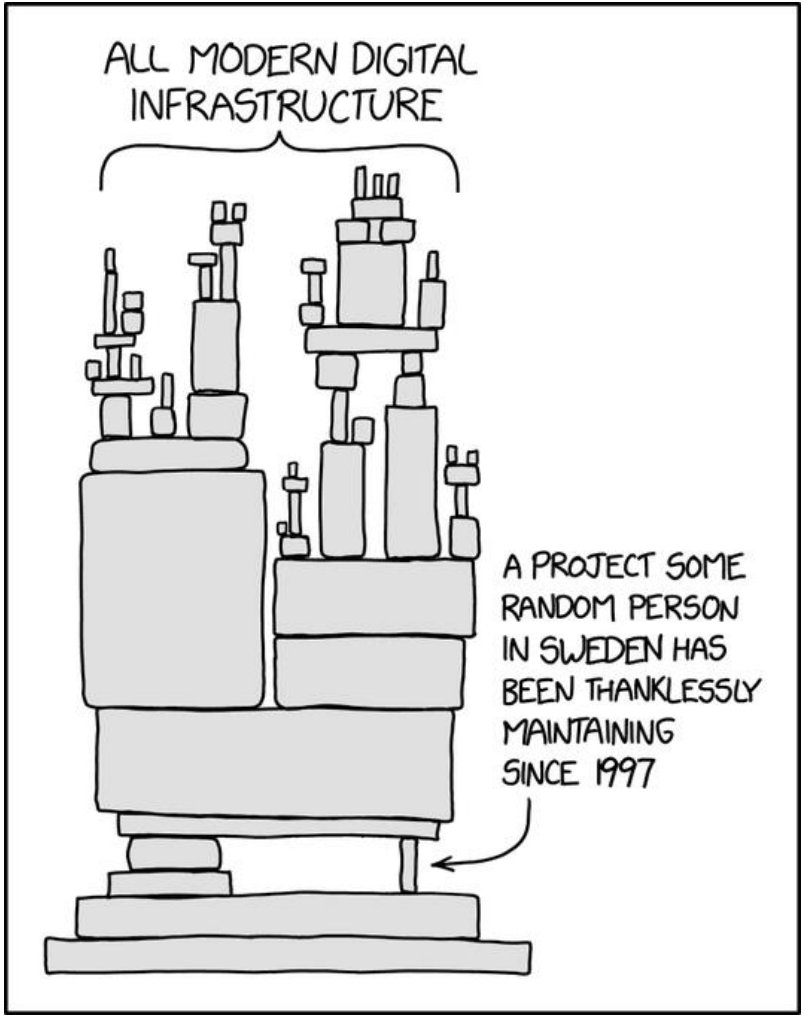


Photo by @tsjost and XKCD | https://twitter.com/tsjost/status/1295662118113619968

Increasing risk of exploitation

- 86% of applications audited in 2025 contained 1+ vulnerable OSS components
  - 81% high- or critical-risk vulnerabilities
- 91% of applications used outdated OSS,
  - 90% more than 10 versions behind latest release
- Over 512,000 malicious open source packages were discovered in 2024—a 156% annual increase
- Several high-profile vulnerabilities including SolarWinds, Log4j, and XZ Utils

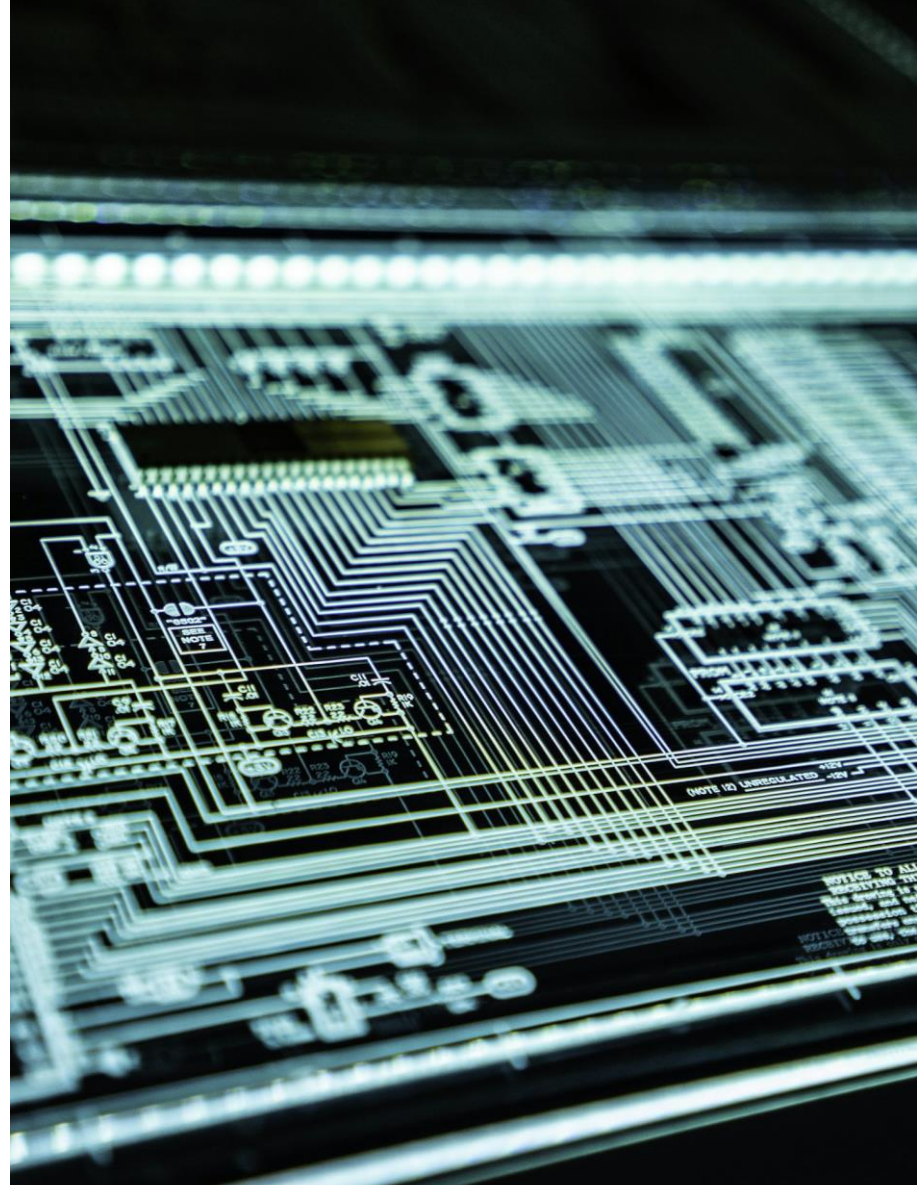
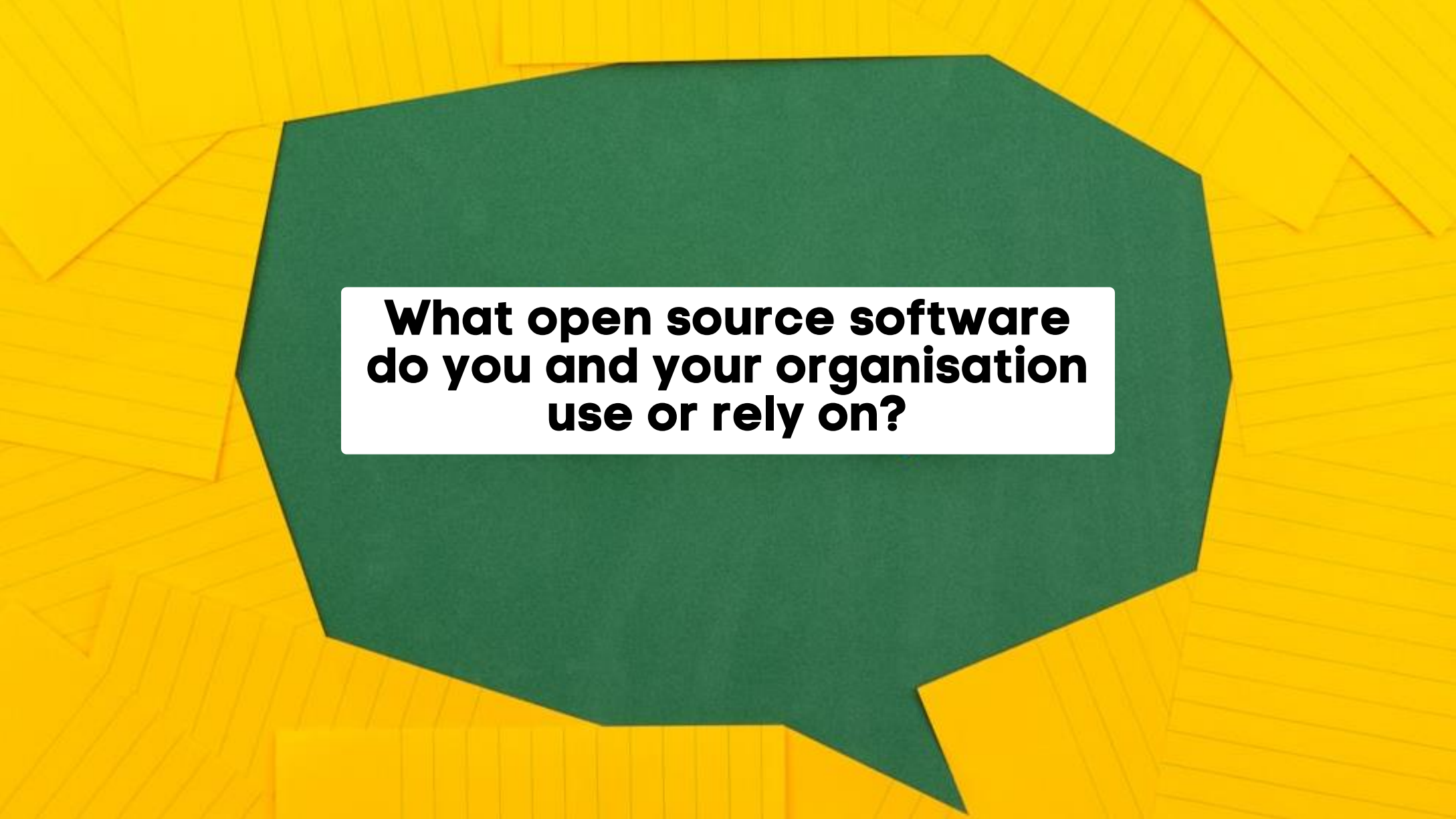


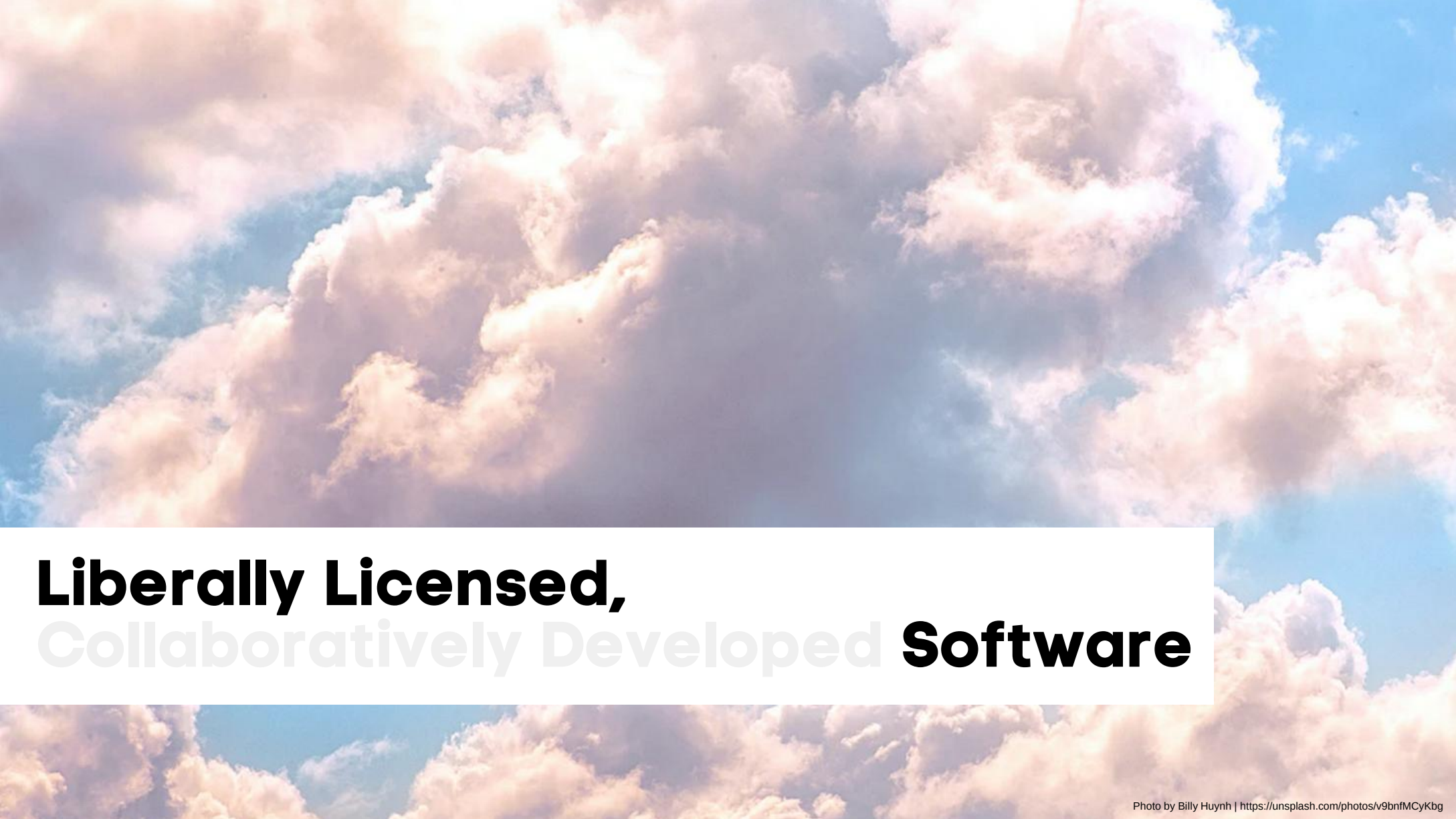
Photo by Adi Goldstein | https://unsplash.com/photos/teal-led-panel-EUsVwEOsblE

Increasing regulatory requirements

- Executive Order (EO) 14028 mandates the adoption and self-attestation of security practices to improve supply-chain security
- European Union's Cyber Resilience Act (CRA), approved in March 2024, mandates common cybersecurity standards for digital products
- Standardization and organization from industry and community ongoing

Photo by Frederic Köberl | https://unsplash.com/photos/aerial-view-photography-of-room-x_0hW-KaCgI


What open source software do you and your organisation use or rely on?



Liberally Licensed, Collaboratively Developed Software

Photo by Billy Huynh | https://unsplash.com/photos/v9bnfMCyKbg

Liberally licensed software

- Software available under an OSS license
- License follows the Open Source Definition and approved by Open Source Initiative (http://opensource.org)
- Who ever, for what ever reason may inspect, use, modify and redistribute the software
- Further conditions may vary between licenses

Photo by Fabian Fauth | https://unsplash.com/photos/1_EedIBc6jY

Permissive vs. Copyleft licenses

- Permissive licenses – do whatever you want, as long as you recognize the copyright holder
  - E.g., MIT, BSD, Apache
- Copyleft licenses – Above + share any modifications, additions and connecting code under same license.
  - GPL 2, GPL 3, AGPL
- Permissive common for standardiszations and collaboration on non-differentiating software
- Copyleft common when copyright holder wants to capture value back

Photo by Fabian Fauth | https://unsplash.com/photos/1_EedIBc6jY

Liberally Licensed,

Collaboratively Developed Software

Photo by Billy Huynh | https://unsplash.com/photos/v9bnfMCyKbg

Collaboratively Developed Software

- Software developed as projects by networks of individuals and organizations, aka. Open Source Communities
- ”Members” of the community commonly both users and developers
- Are united by a common vision and goal around the Open Source Software.

Photo by Annie Spratt | https://unsplash.com/photos/QckxruozjRg

Incentives for going open source

- Individuals:
  - Sense of belonging,
  - Recognition for contributions,
  - Solves paintpoint,
  - Build CV
- Organizations:
  - Lower costs,
  - Increased innovation,
  - Branding and PR,
  - Strategic tool

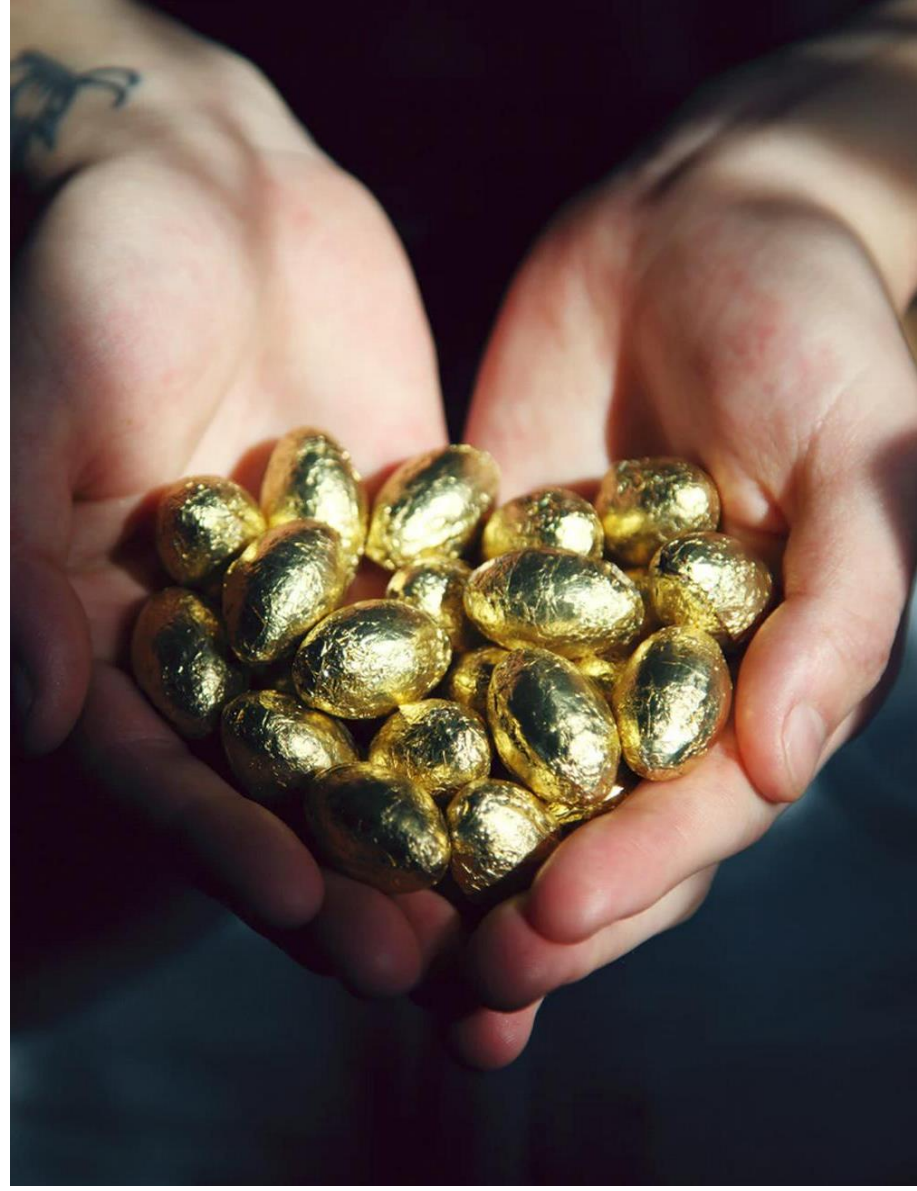


Photo by Annie Spratt | https://unsplash.com/photos/QckxruozjRg

Incentives for going open source

- Public policy:
  - Open strategic autonomy/ Digital sovereignty
  - Competition
  - Economic growth
- Researchers:
  - Disseminate research outputs
  - Sustain OSS development between project
  - Collaborate with partners and scientific community
  - Enable reproducibility

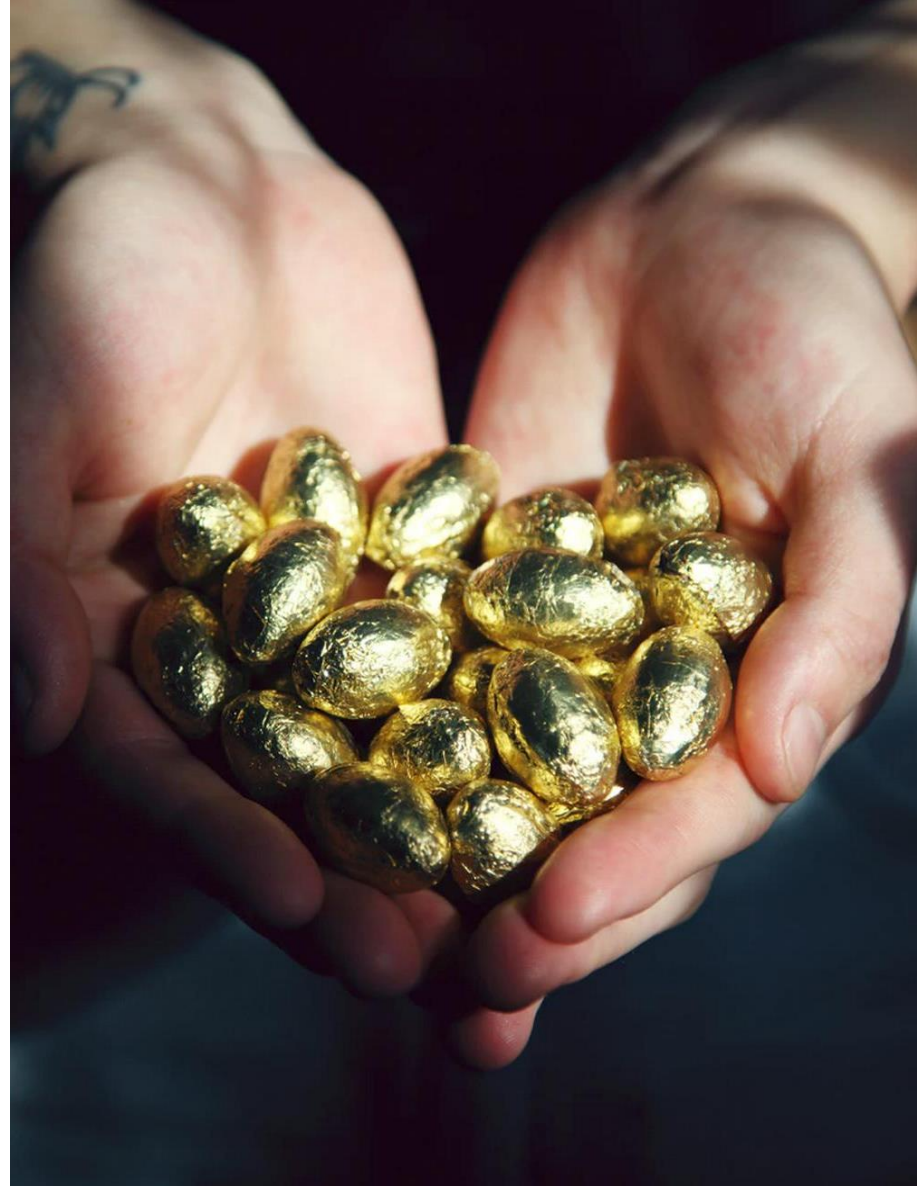


Photo by Annie Spratt | https://unsplash.com/photos/QckxruozjRg

Risks, costs and complexities

- Companies:
  - Differentiating functionality, competitive edge and commoditization
  - Sensitive IPR and patents
- Public administrations
  - Compete with industry
  - Ethical aspects and responsibility
  - Integrity and confidentiality
- General:
  - Internal budget and resource constraints
  - Modularity and technical architecture

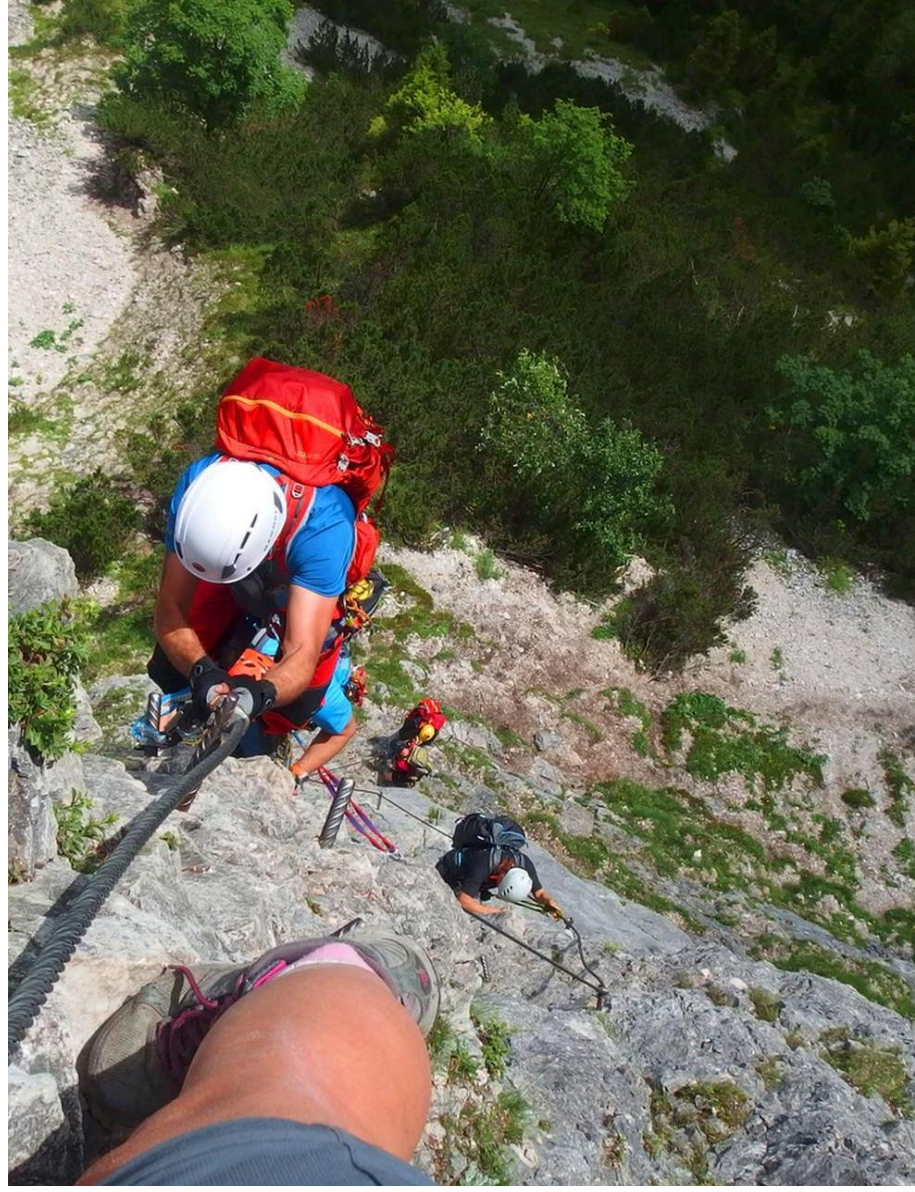
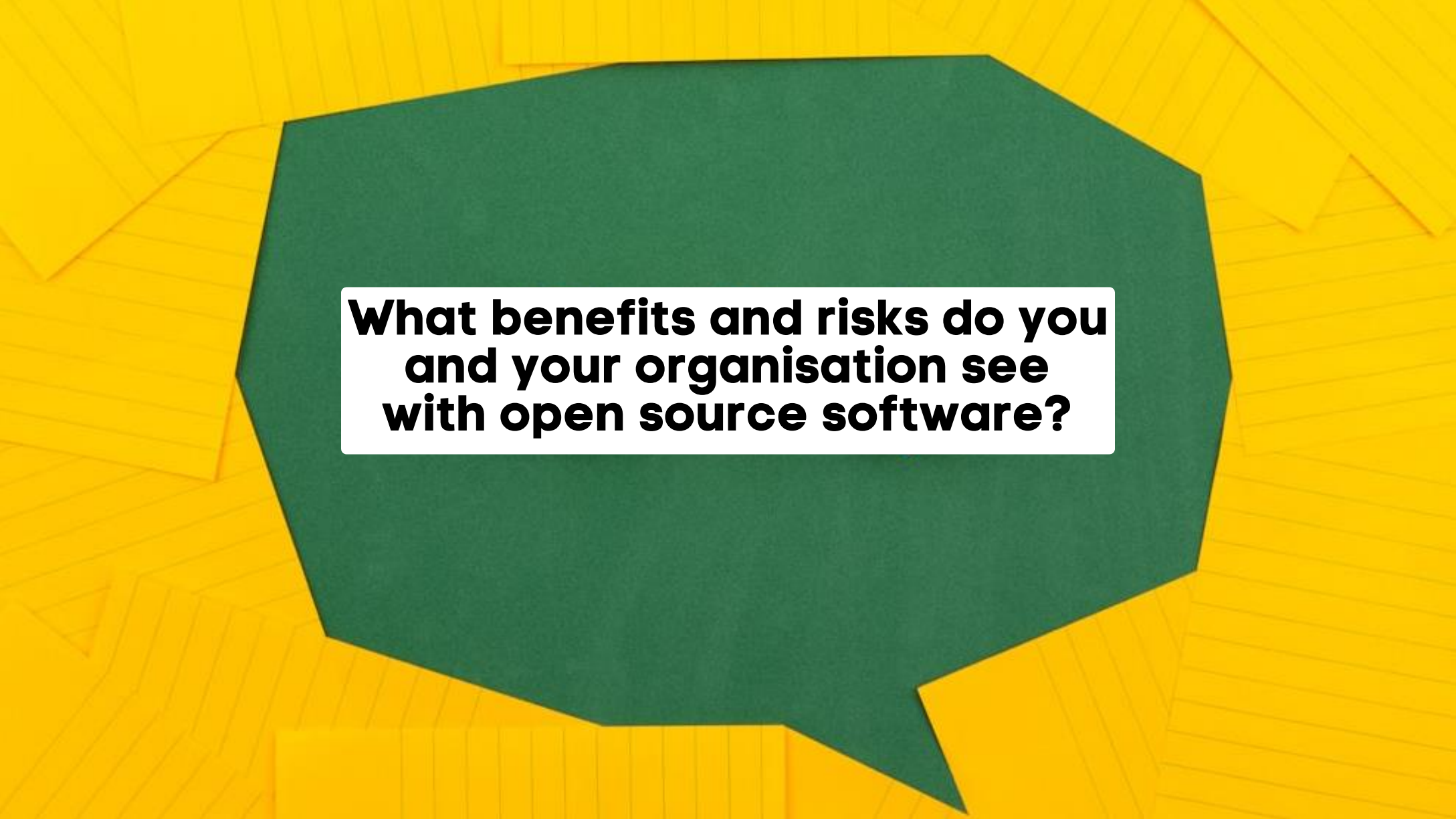


Photo by Maja Kochanowska | https://unsplash.com/photos/E-v6j5Y9xAs


What benefits and risks do you and your organisation see with open source software?

Technical and non-technical contributions

- Development of new functionality and bug fixes
- Requirements identification, analysis, and prioritization
- Testning and quality assurance
- Documentation, marketing and community management
- Financial and infrastructural support
- ...

Photo by Amélie Mourichon | https://unsplash.com/photos/hcLSA1MiU3A

Open development process

- Informal structure, dependent the community
- Focus on openness
  - Whomever can contribute
  - Influence through merit
  - Self-appointment of tasks
- Traditional development
  - Structured in silos
  - Influence through hierarchy
  - Appointed tasks

Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Open development process

- Meaning, you cannot...
  - Expect quick and professional support
  - Expect to get your feature requests implemented
  - Order individuals to act according to your agenda

Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Open development process

- Transparent and open discussions on bug reports, features, and road map
- Conversations and information persisted in an open infrastructure
- Requirements fragmented and decentralized in various ”informalisms” (Scacchi), e.g., bug reports, mail threads, code commits, etc.
- Formality typically dependent on corporate interest
- Community full of (un)known stakeholders, all with their own agendas

Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk


Have you or any in your organisation contributed to open source software projects?

Governance for OSS projects

- Means and processes for
  - Deciding on requirements, i.e., the technical direction of the OSS project, and
  - How the collaboration should be coordinated to enable this direction.

Photo by krakenimages | https://unsplash.com/photos/Y5bvRlcCx8k

Community Structure and Governance

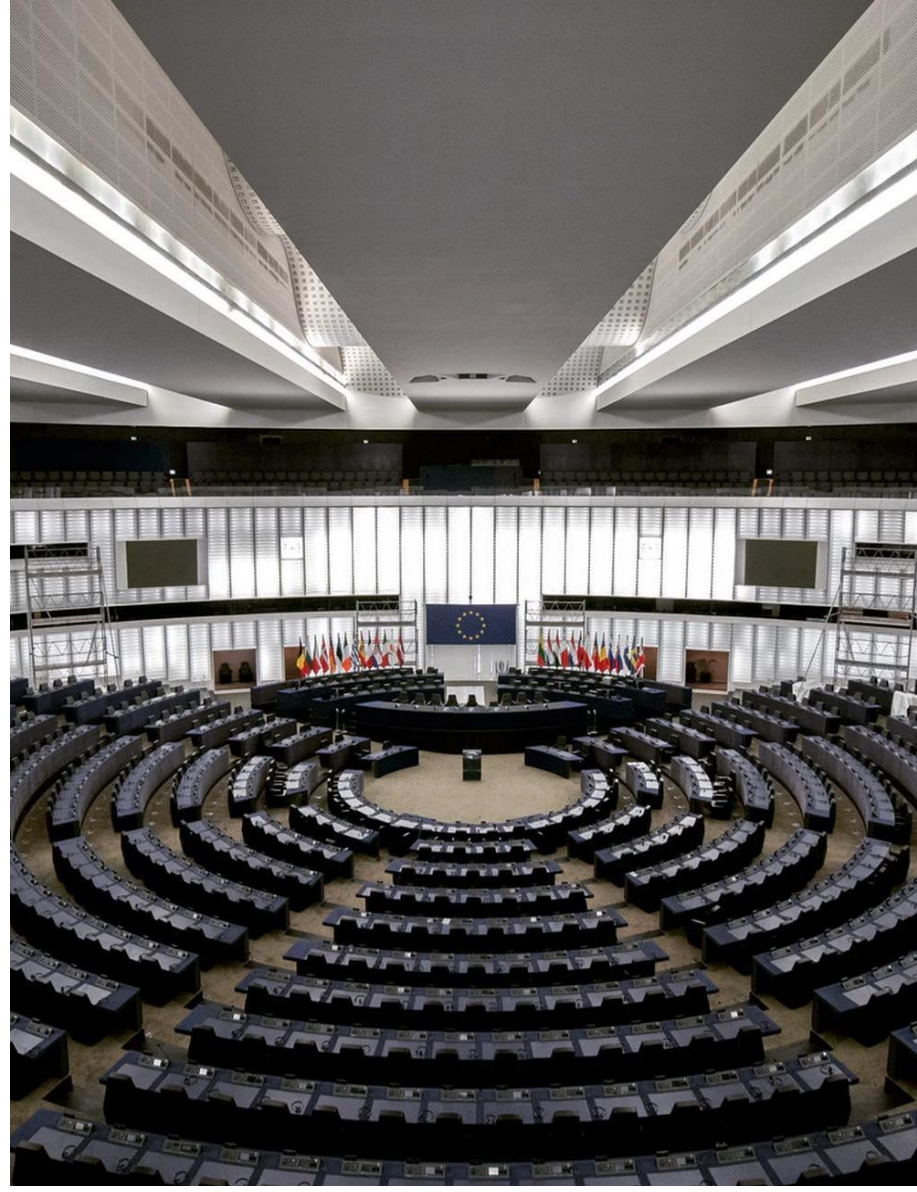


Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Community Structure and Governance

Leadership

Maintainership Maintainership Maintainership

Committers Committers Committers Committers
Committers Committers Committers Committers

Contributors Contributors Contributors Contributors Contributors
Contributors Contributors Contributors Contributors Contributors

Users Users Users Users Users Users Users Users Users Users
Users Users Users Users Users Users Users Users Users

@johanlinaker | https://linaker.se

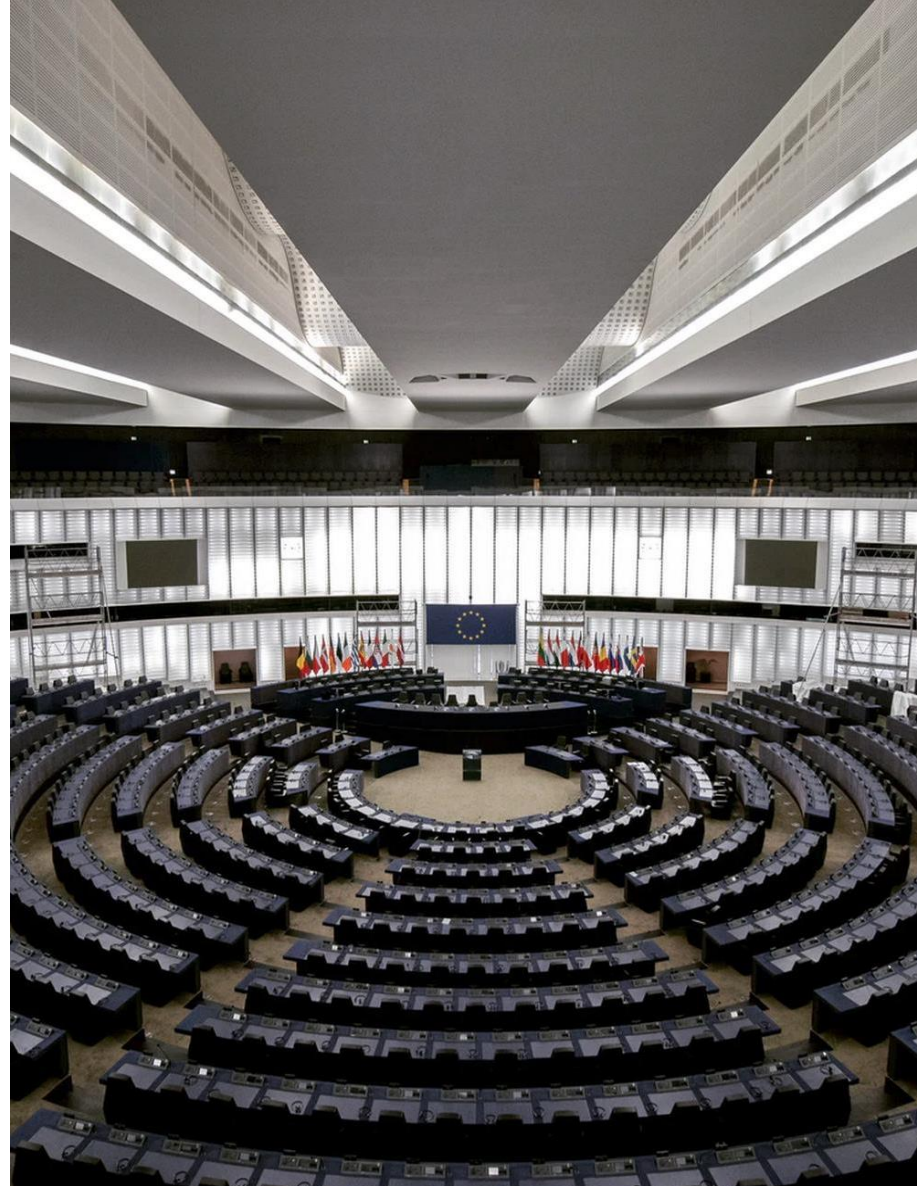


Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Governance structures

- Autocratic governance
  - Centralized steering where roles assignment and influence over development is decided top-down
  - Usually the actor(s) that founded the project
- Democratic governance
  - Decentralized steering where roles assignment and influence over development is decided collectively, and gained through active engagement and contributions
- Transitions and combinations common

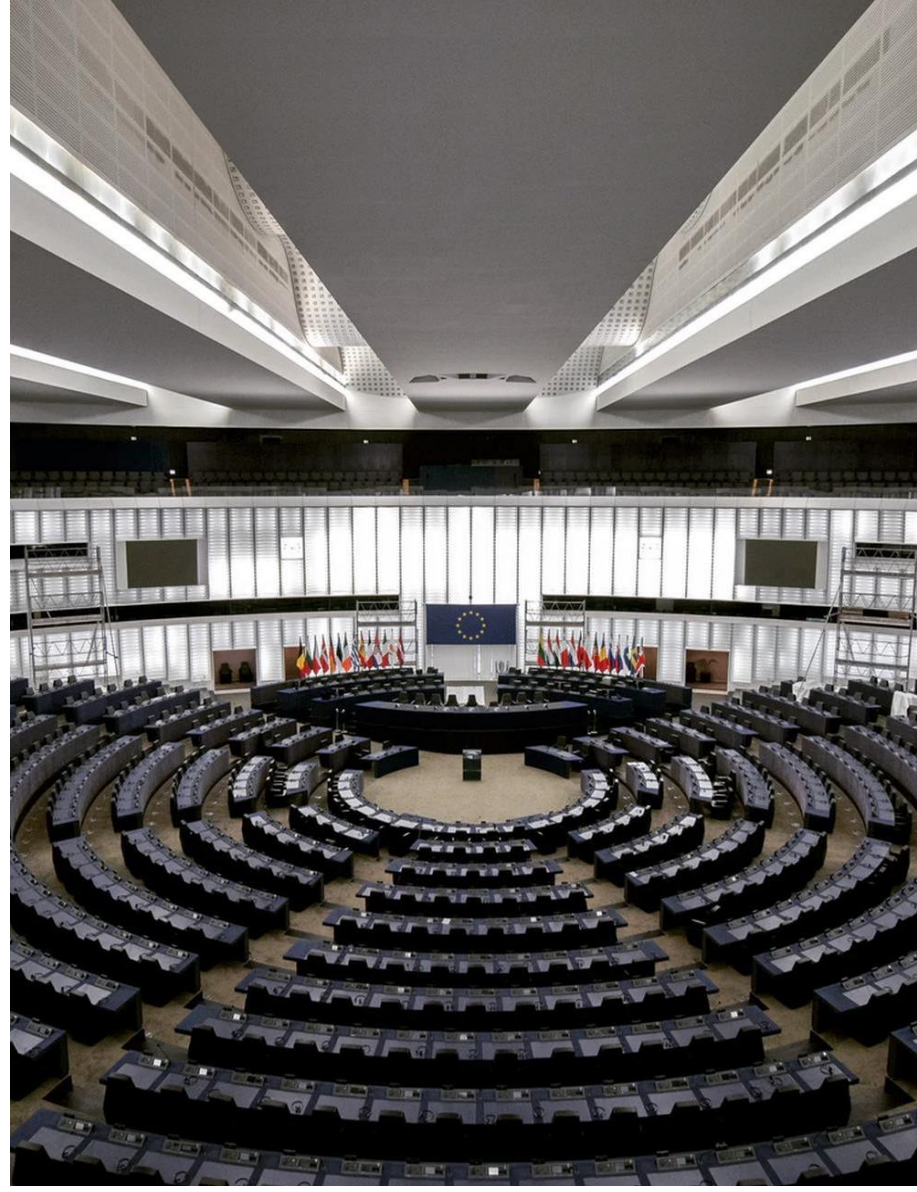


Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Governance structures

- Centralized governance
  - Formal steering and maintenance through a single or collective organization
  - Commonly pooled ownership of copyright
- Decentralized governance
  - Informal steering and maintenance through existing community
  - Distributed ownership of copyright
- Commonly transitions from decentralized to centralized structure

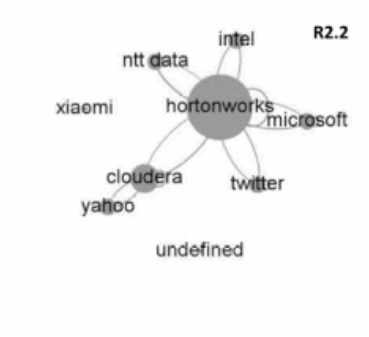
Photo by Leon | https://unsplash.com/photos/Oalh2MojUuk

Type of community

- Developer-driven community
  - Steering and maintenance typically performed by those who contribute to the development of the project
- User-driven community
  - Steering and maintenance typically performed by the end-users of the project.
  - Development performed through acquired resources

Photo by Vonecia Carswell | https://unsplash.com/photos/0aMMMUjiiEQ

Communities evolve constantly



RI.
SE

Relationship with community

- Symbiotic
  - Win-win for both firm and community
  - Contributing to influence projects according to internal agenda and improve health to mitigate security risks
- Commensalistic
  - Gain for firm, community indifferent
  - Use project and doing lighter contributions. Project in line with internal agenda and healthy with others already supporting it.
- Parasitic
  - Firm free-riding on community.
  - Using as is not giving anything back. Worst case expecting free work for nothing in return. Looked down on from communities.

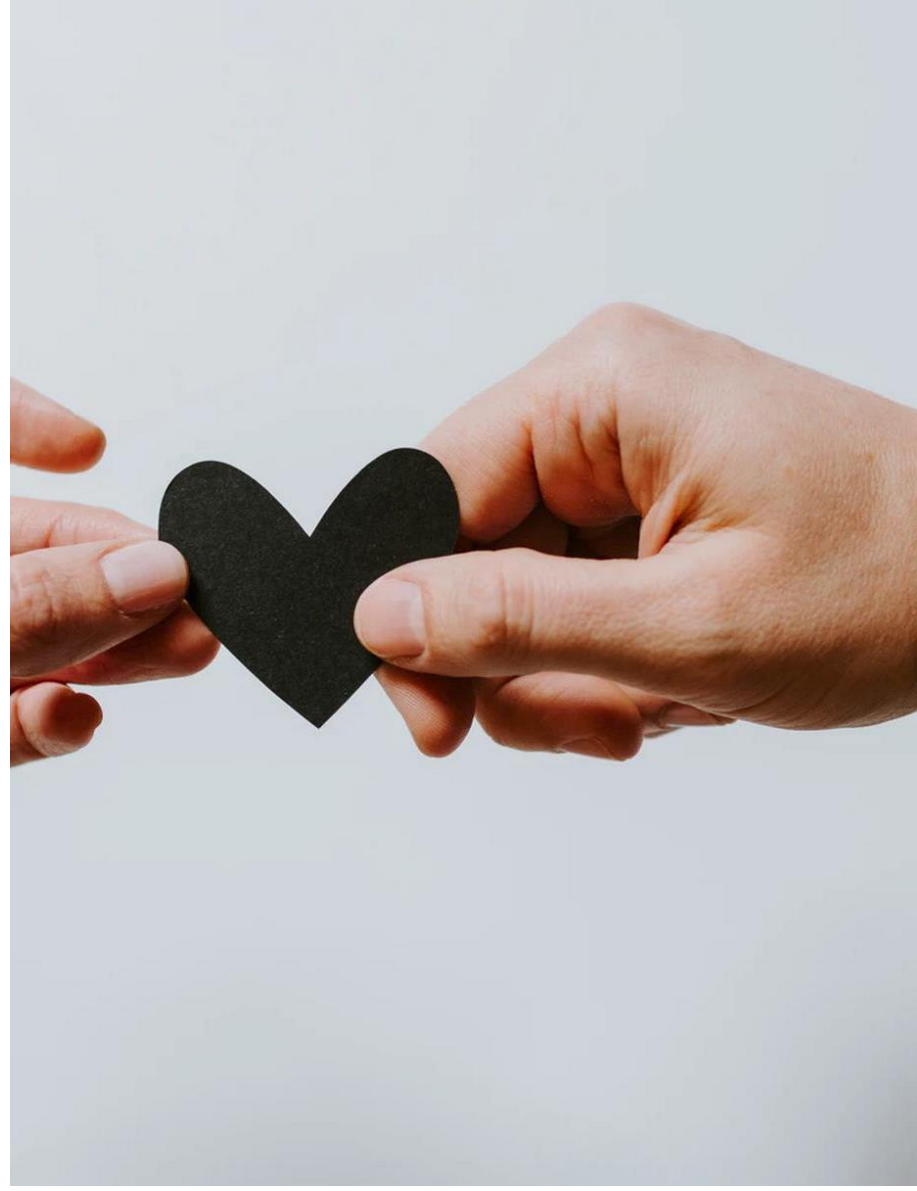
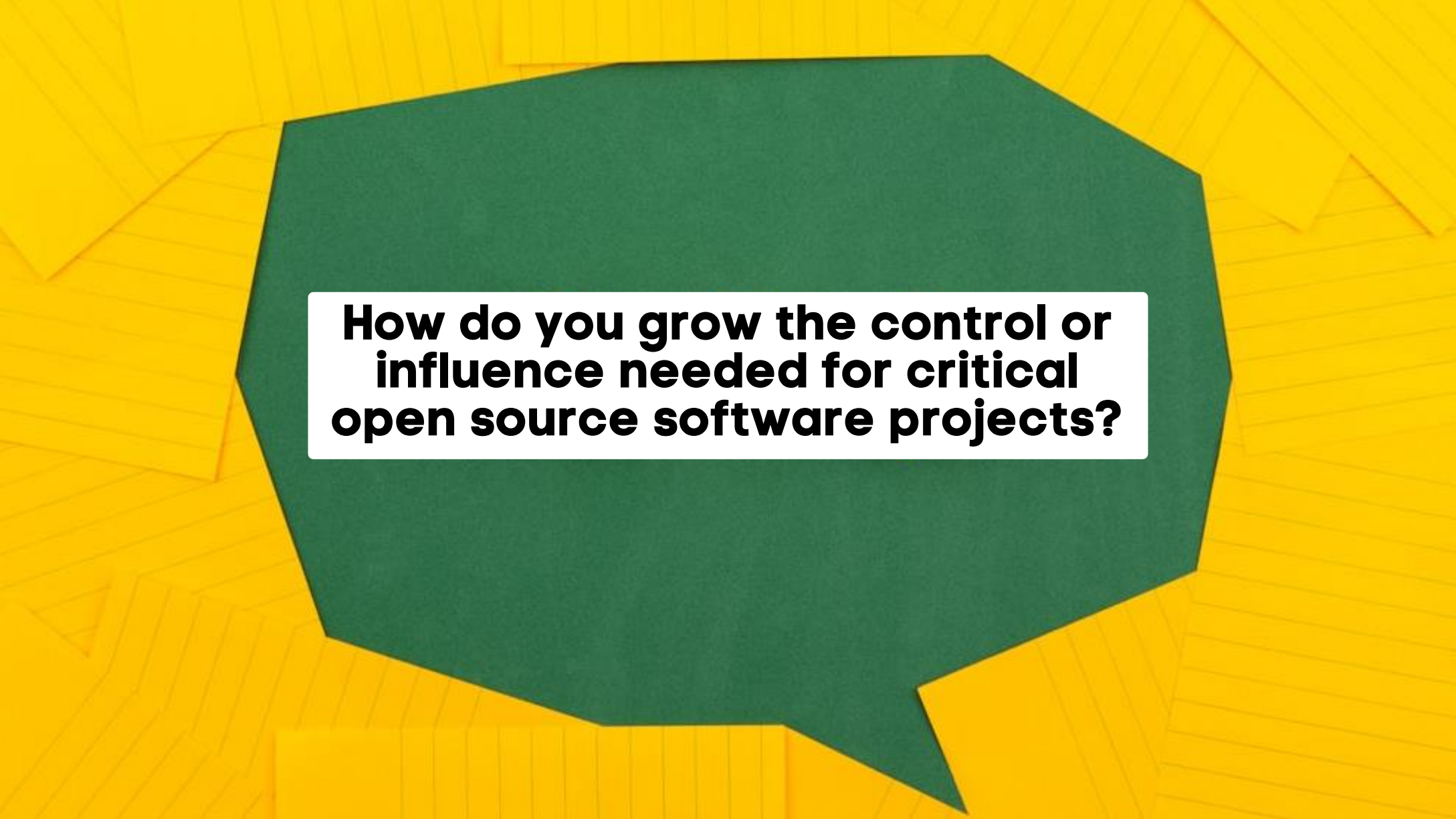


Photo by Kelly Sikkema | https://unsplash.com/photos/4le7k9XVYjE


How do you grow the control or influence needed for critical open source software projects?

OSS Project health

- The OSS project’s capability to stay maintained to a high quality, long-term without interruptions
  - Productivity: There is an active development of the project
  - Robustness: The development is open and spread out on several (independent) individuals
  - Openness: Users of the project can influence and contribute to the development of the project

Photo by Markus Frieauff | https://unsplash.com/photos/lJ0KiXl4uys

OSS and our Digital Infrastructure

- Open Source Software makes up a vitale building block in our digital infrastructure
- Needs maintenance as with physical infrastructure to stay secure and robust

Photo by Modestas Urbonas | https://unsplash.com/photos/vj_9l20fzj0

The Dualism of Quality

- Open Source Software is...
  - full of, or receptible to, vulnerabilties ready to be exploite
  - always more secure than proprietary alternatives

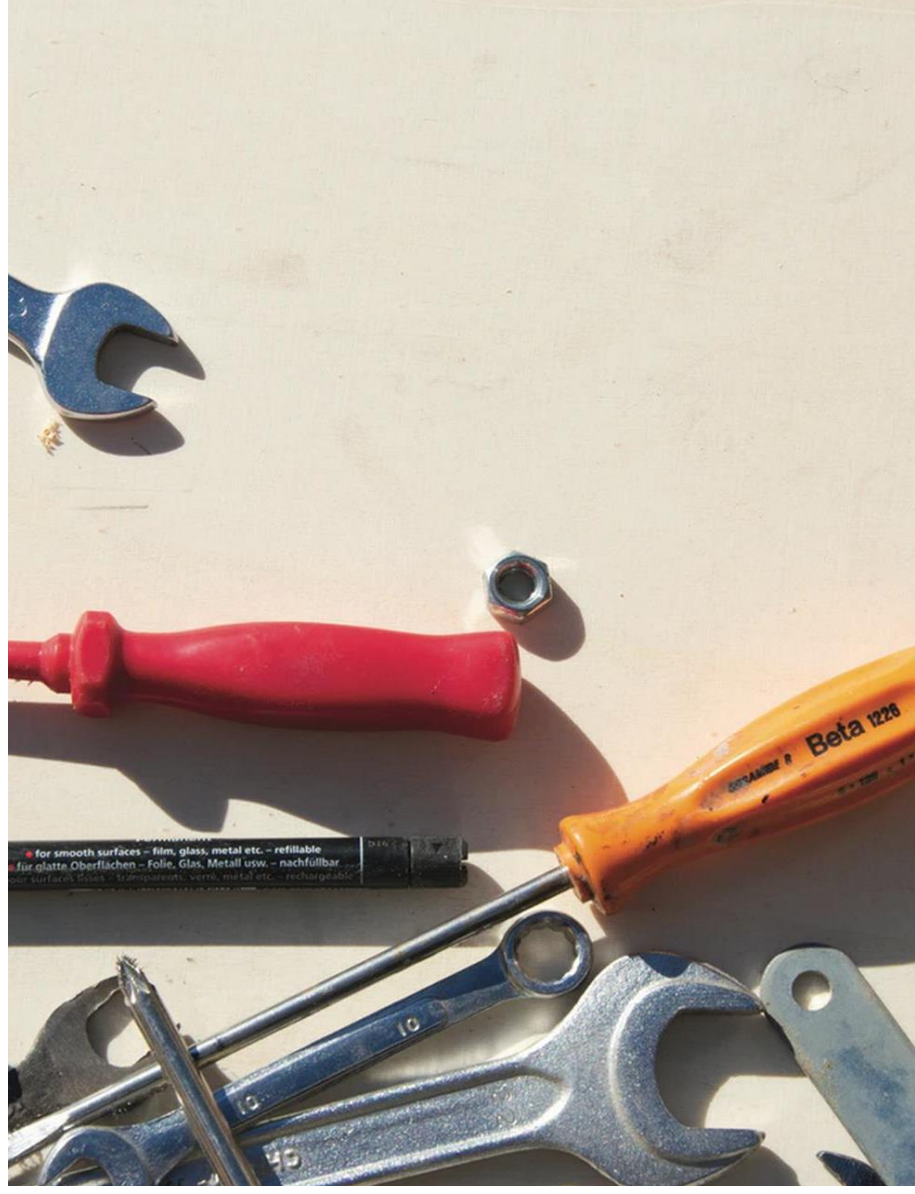


Photo by Elena Rouame | https://unsplash.com/photos/9JU2CKqtw0M

The ”Many-Eyes” effect

- Also known as Linus’ law →
  - ”Given enough eyeballs, all bugs are shallow”
  - Requires that enough eyeballs actually reaches the codebase

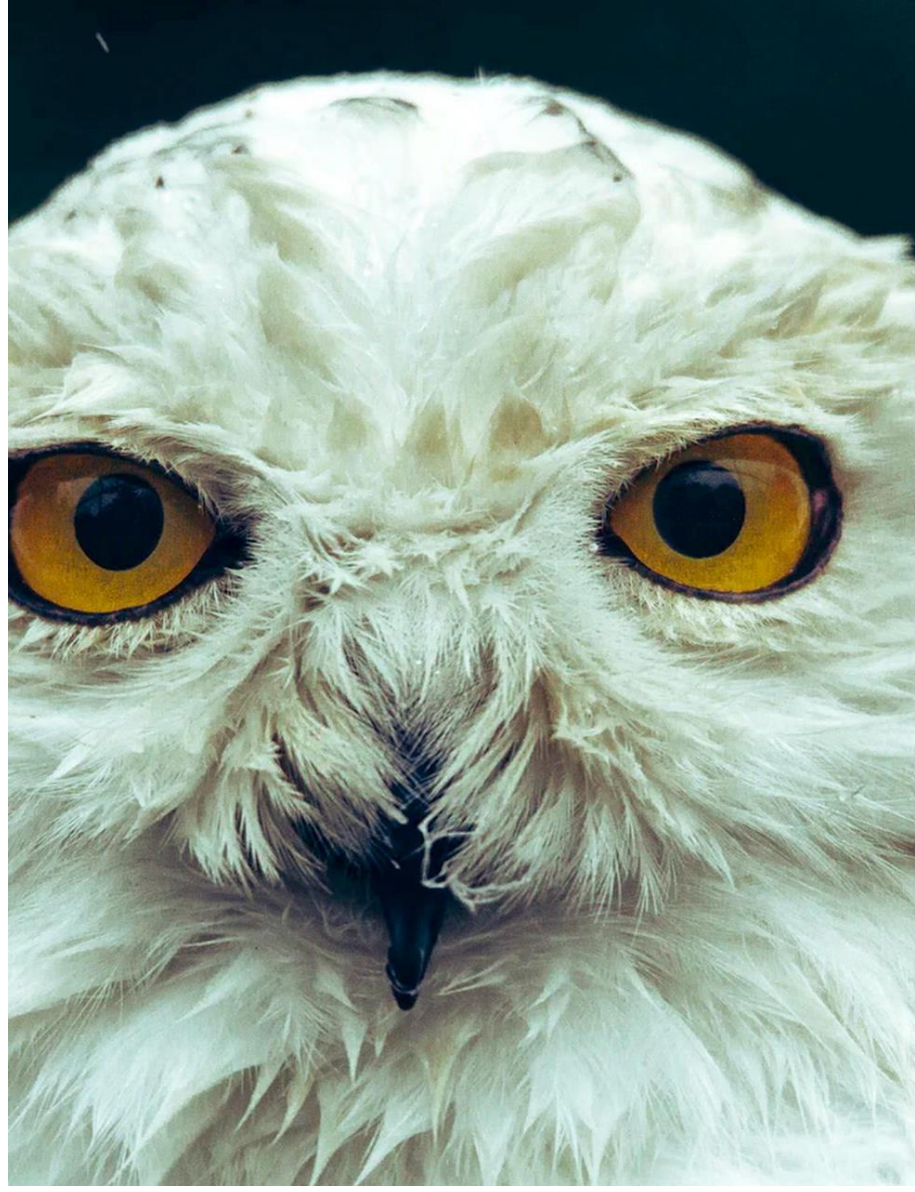


Photo by Frida Bredesen | https://unsplash.com/photos/c1fFv08N7qE

Development Resources are Depletable

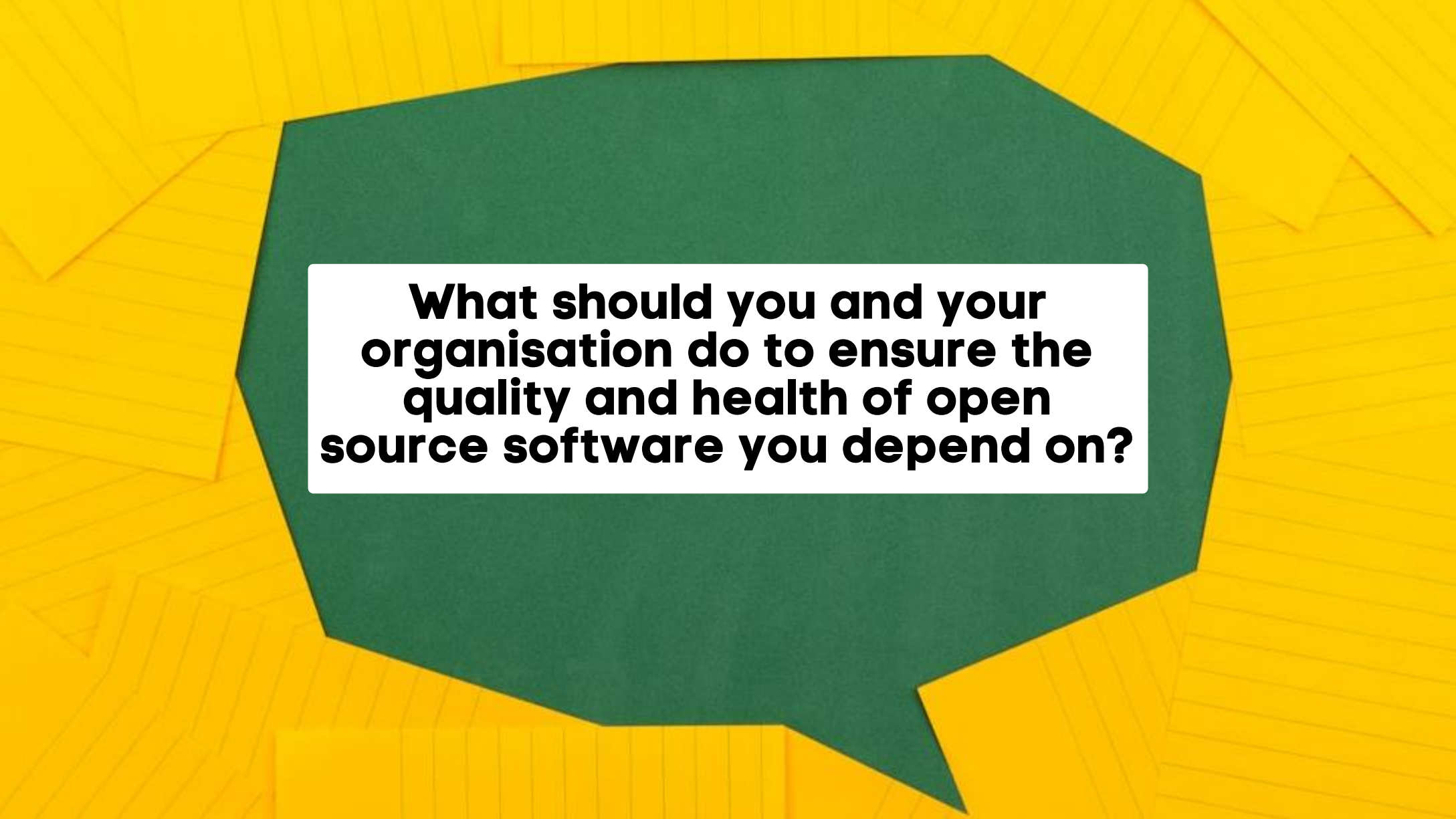
- Maintainers are humans, not robots
  - Burnout, changed family or working conditions
- Companies must adapt to stay competitive
  - Refactorization, new products, changed business model

Photo by NeONBRAND | https://unsplash.com/photos/KYxXMTpTzek

Who’s responsible for the SW quality?

- Maintainer(s)?
- Developer community?
- User community?
- Individuals vs. Companies vs. Government?

Photo by Scott Blake | https://unsplash.com/photos/x-ghf9LjrVg

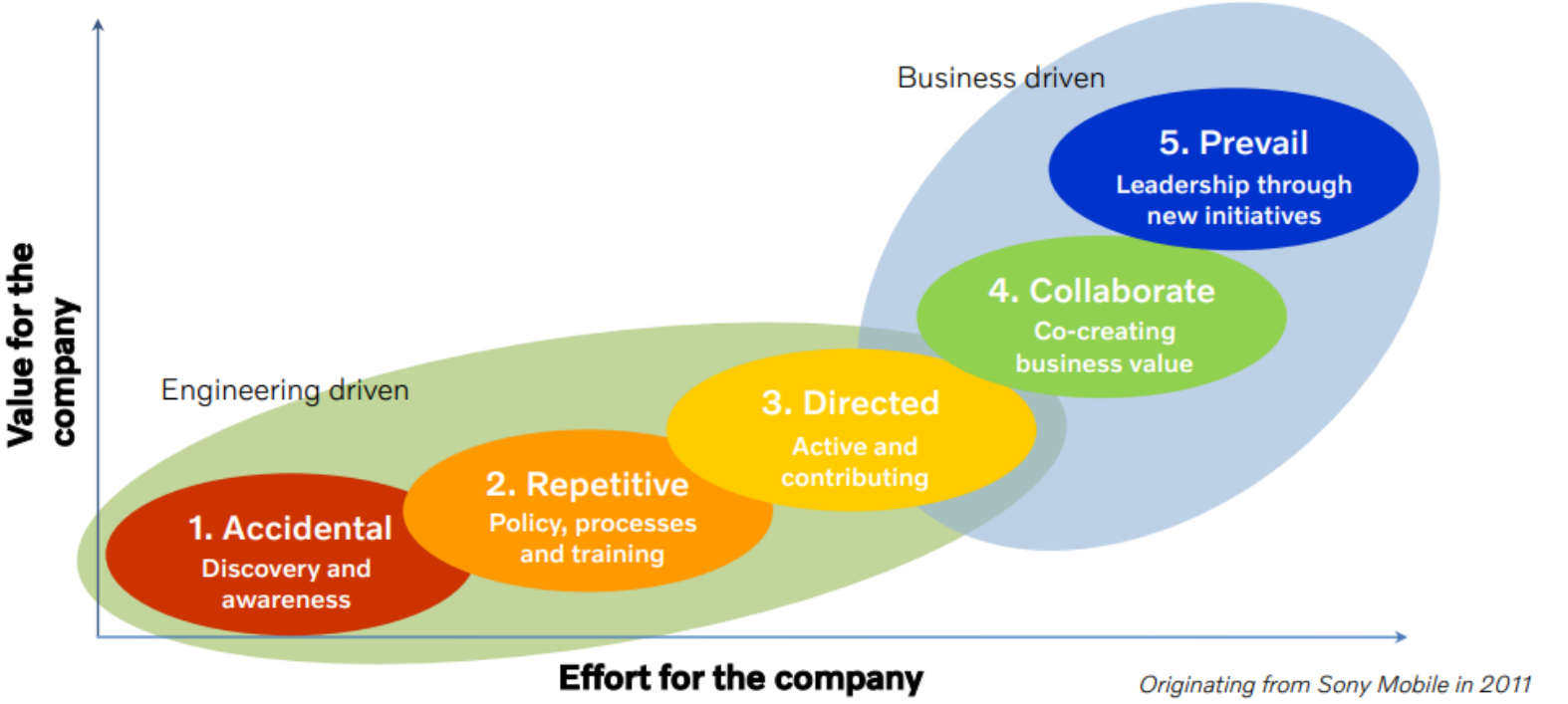

What should you and your organisation do to ensure the quality and health of open source software you depend on?

Importance of growing a healthy community

- Collectively grow and communicate a common vision for the project
- Be responsive and helpful in communication
- Grow an open, inclusive, and supportive culture
- Enable on-boarding and self-support through
  - detailed documentation,
  - standardized tooling
  - clearly defined development and governance processes

Photo by William White | https://unsplash.com/photos/people-raises-hands-TZCppMjaOHU

Maturing from consumption to leadership



Originating from Sony Mobile in 2011
/ Adapted by Carl-Eric Mols 2023

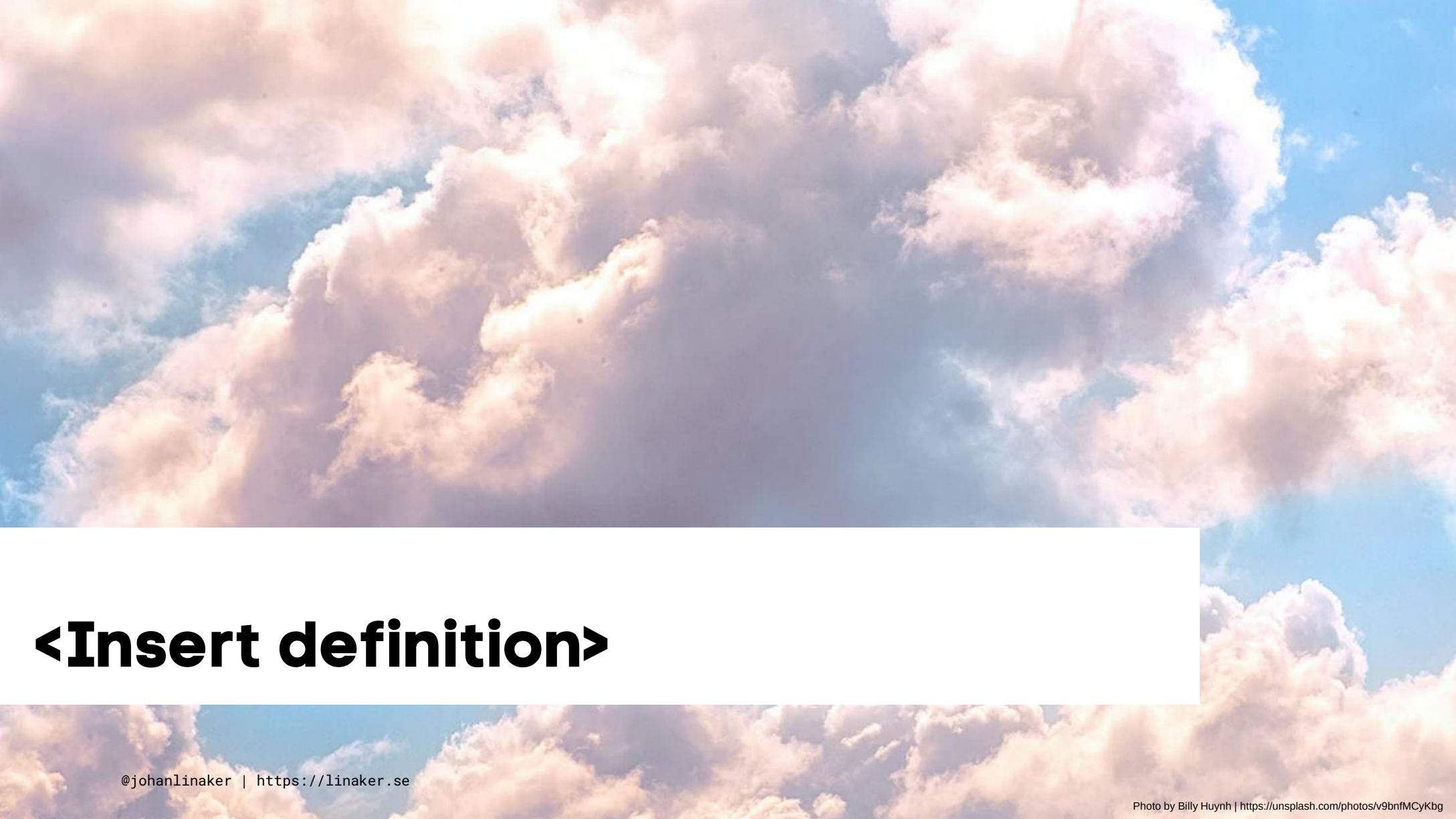
RI.SE

Open Source Program Offices (OSPOs)

- Center of competency and support
- Drives organizational readiness and maturity forward on open source
- Designs and executes an organization’s overarching open source strategy
- Provides voice of reason and objectivity on the benefits, risks, and costs of open source and how to balance between
- Supports use, development, and collaboration on open source

Photo by Quino Al | https://unsplash.com/photos/dhGFLj3rI0Q

So, what's Open Source AI?



<Insert definition>

Open Source AI (Systems)

- Definition being developed by the OSI
  - See: https://opensource.org/ai/open-source-ai-definition
- Open community effort working towards reaching consensus among key stakeholders
- Building on the four freedoms, and the AI systems definition by OECD

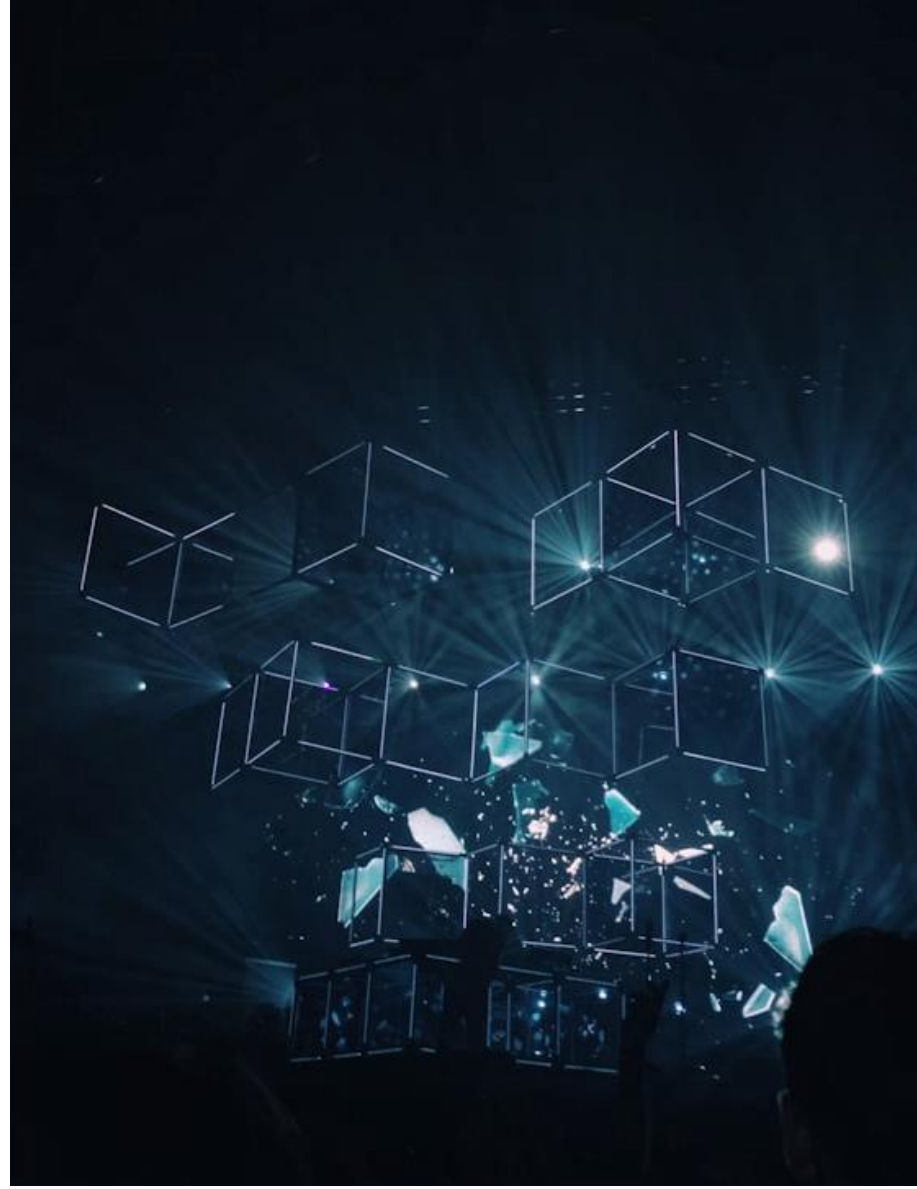


Photo by fabio | https://unsplash.com/photos/geometric-shape-digital-wallpaper-oyXis2kALVg

Open Source AI Systems

- An Open Source AI is an AI system made available under terms and in a way that grant the freedoms to:
  - **Use** the system for any purpose and without having to ask for permission.
  - **Study** how the system works and inspect its components.
  - **Modify** the system for any purpose, including to change its output.
  - **Share** the system for others to use with or without modifications, for any purpose.

Photo by Damir Kopezhanov | https://unsplash.com/photos/person-holding-green-and-silver-hand-tool-Z5tXjpTqwR8

Parts required for modifying an AI/ML system

- **Data Information**:
  - Sufficiently detailed information about the data used to train the system so that a skilled person can build a substantially equivalent system.
- **Code**:
  - The complete source code used to train and run the system. The Code shall represent the full specification of how the data was processed and filtered, and how the training was done.
- **Parameters**:
  - The model parameters, such as weights or other configuration settings.

RI.SE

Many models referred to as ”open source”

- But what is open? Are you able to
  - *Use the system for any purpose and without having to ask for permission?*
  - *Study how the system works and inspect its components?*
  - *Modify the system for any purpose, including to change its output?*
  - *Share the system for others to use with or without modifications, for any purpose?*
- Copyright ownership of data typically a main blocker

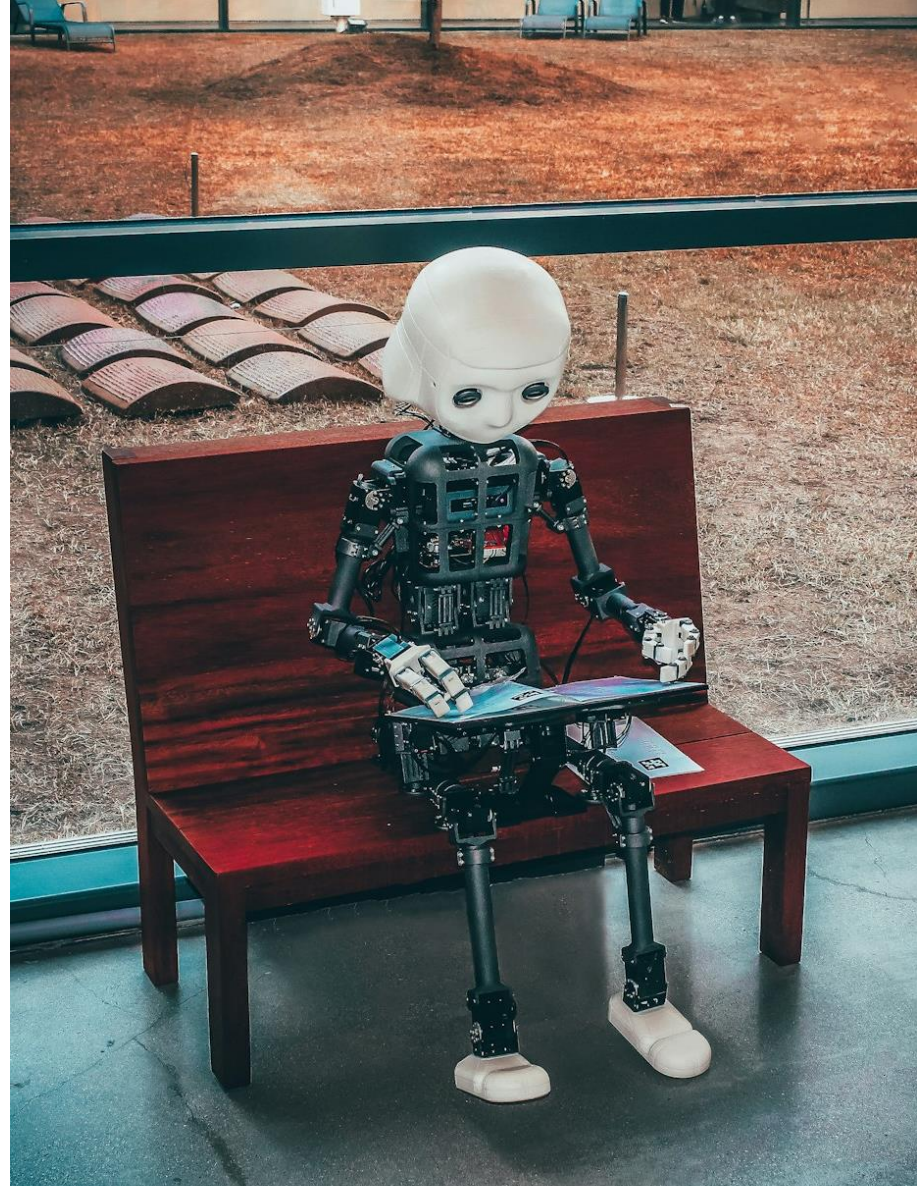
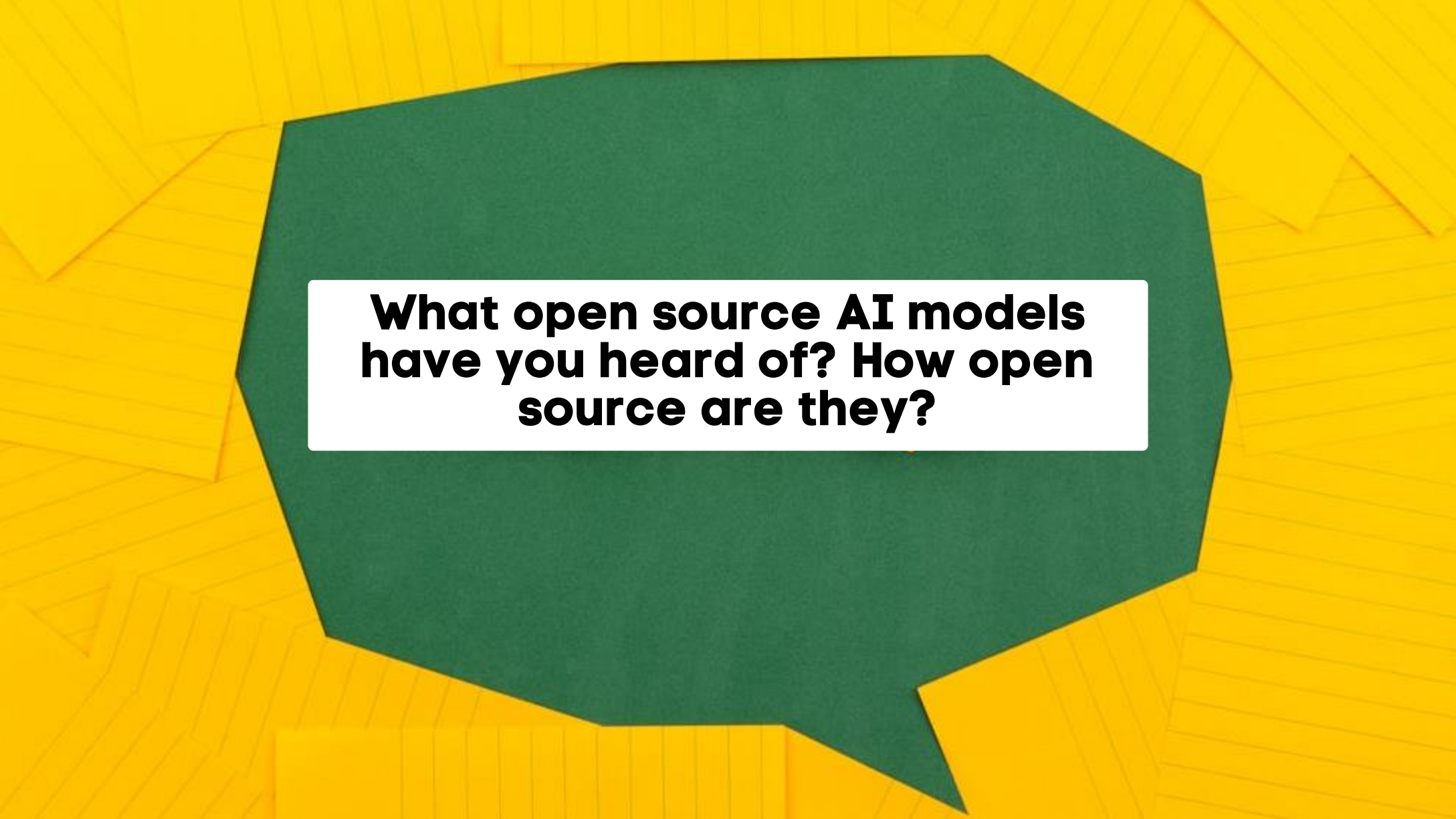


Photo by Andrea De Santis | https://unsplash.com/photos/black-and-white-robot-toy-on-red-wooden-table-zwd435-ewb4

“For [a machine learning system] to be open. I need to be able to question it,”
- Julia Ferraioli

https://aibusiness.com/ml/amazon-ml-expert-what-makes-machine-learning-truly-open-source

RI.
SE


What open source AI models have you heard of? How open source are they?

Incentives for going open source

- Very similar to open source software with certain nuances
- Social motivations
  - Democratizing access and inclusion in AI development
  - Knowledge sharing and community building
  - Expanding language and cultural representation
  - Transparency and explainability

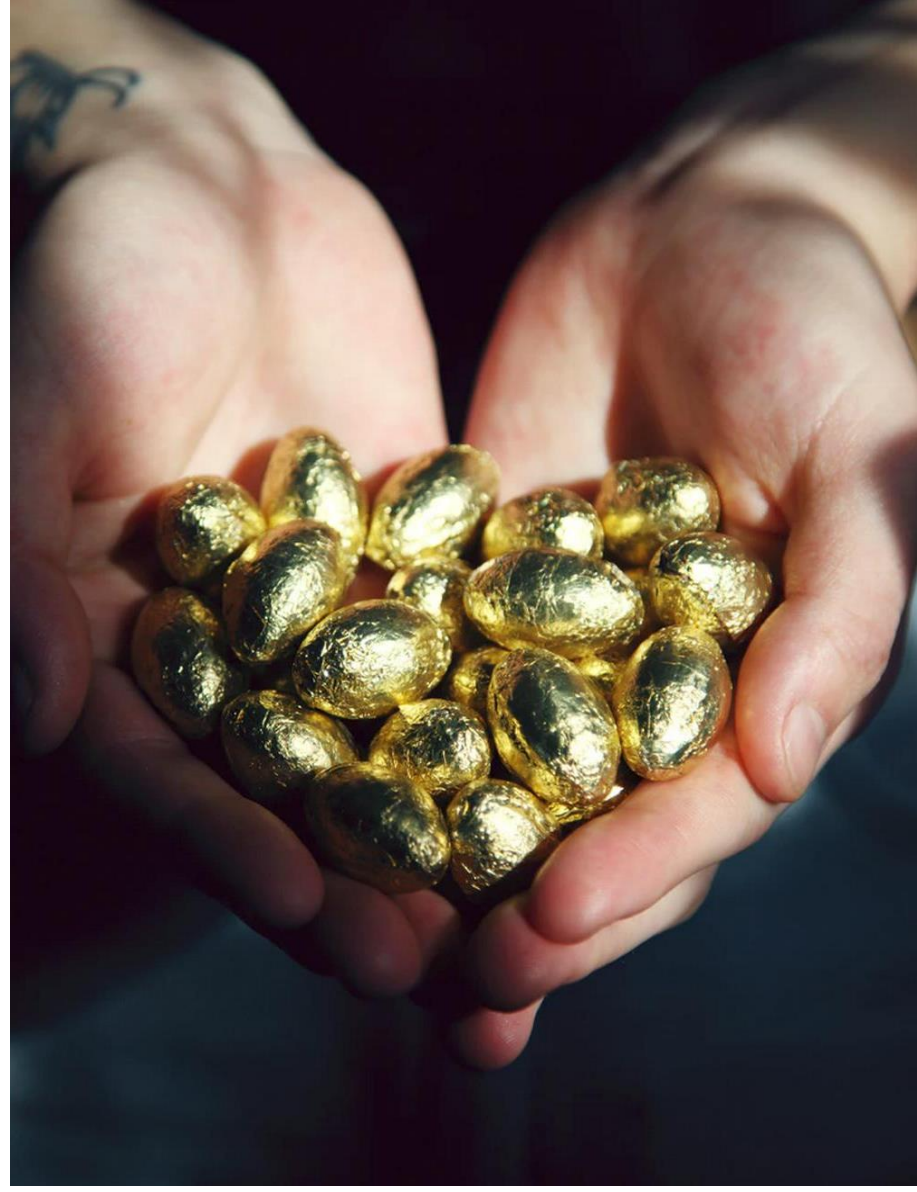


Photo by Annie Spratt | https://unsplash.com/photos/QckxruozjRg

Incentives for going open source

- Economic motivations
  - Ecosystem building
  - Resource efficiency
  - Market recognition
  - Dual business strategies
- Technological motivations
  - Promoting open science and reproducing models
  - Standardisation of tools and processes
  - Technical experimentation
  - Digital sovereignty

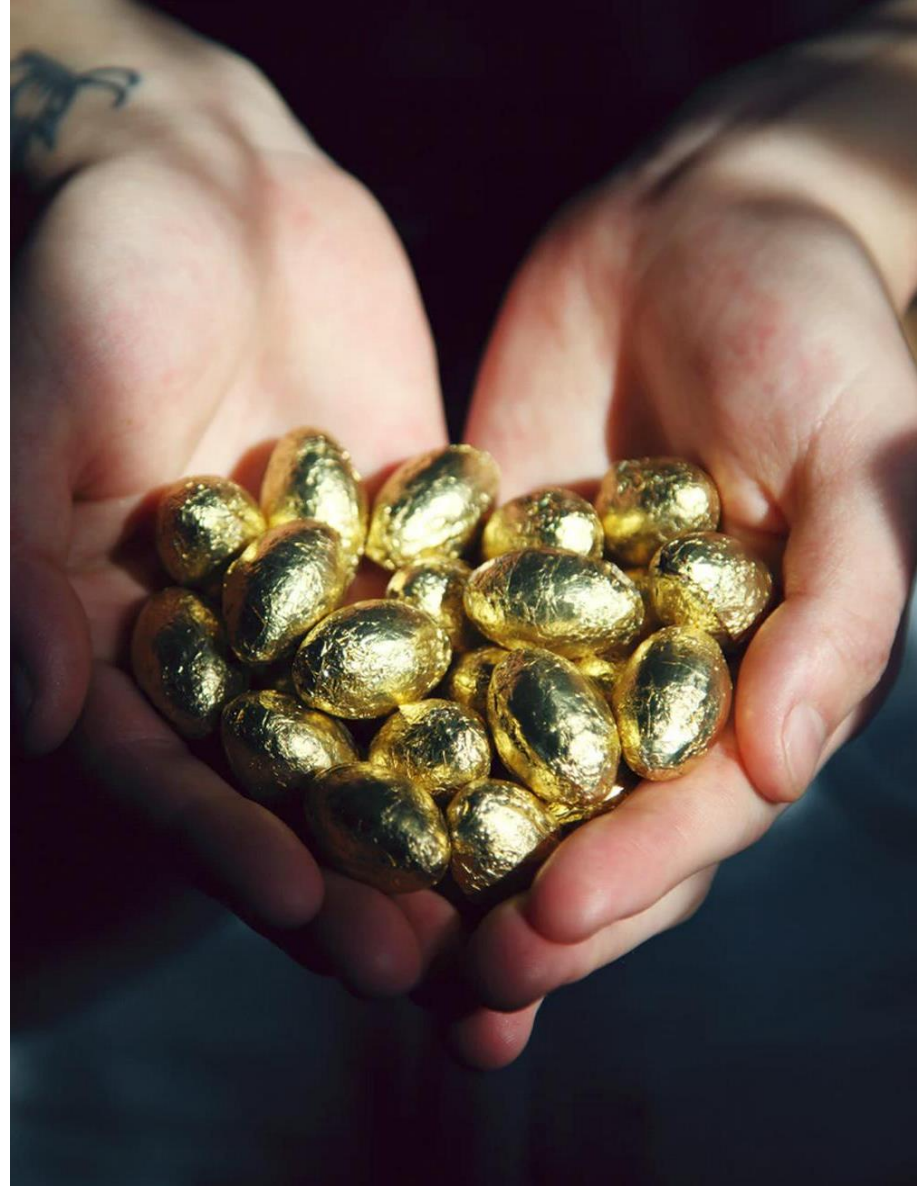


Photo by Annie Spratt | https://unsplash.com/photos/QckxruozjRg

Risks, costs and complexities

- Similar as for open source software but with additional complexities
- General:
  - Access to high quality and permissively licensed data limited
  - Access to computational resources limited and costly
  - Access to specialized expertise limited
  - Complex model development process including several disjunct steps and efforts
- Usually limited to resourceful, or venture-backed firms or research institutes

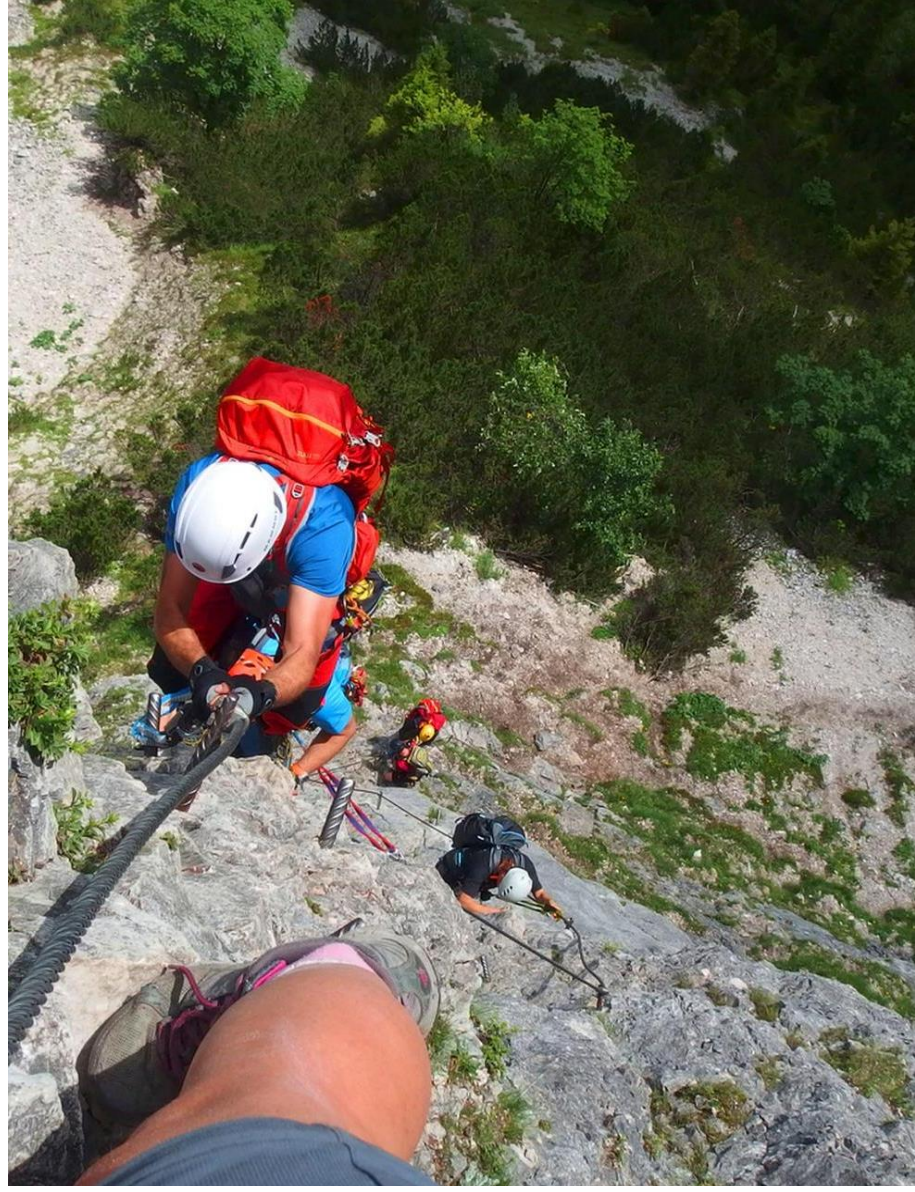


Photo by Maja Kochanowska | https://unsplash.com/photos/E-v6j5Y9xAs

Collaborative development varies

- Presence and form for collaboration may differ based on the component:
  - data (e.g., for training, validation, and testing),
  - source code (e.g., for training and inference),
  - model architecture (e.g., for design choices and hyperparameters), and
  - documentation (e.g., for training procedure and evaluation).

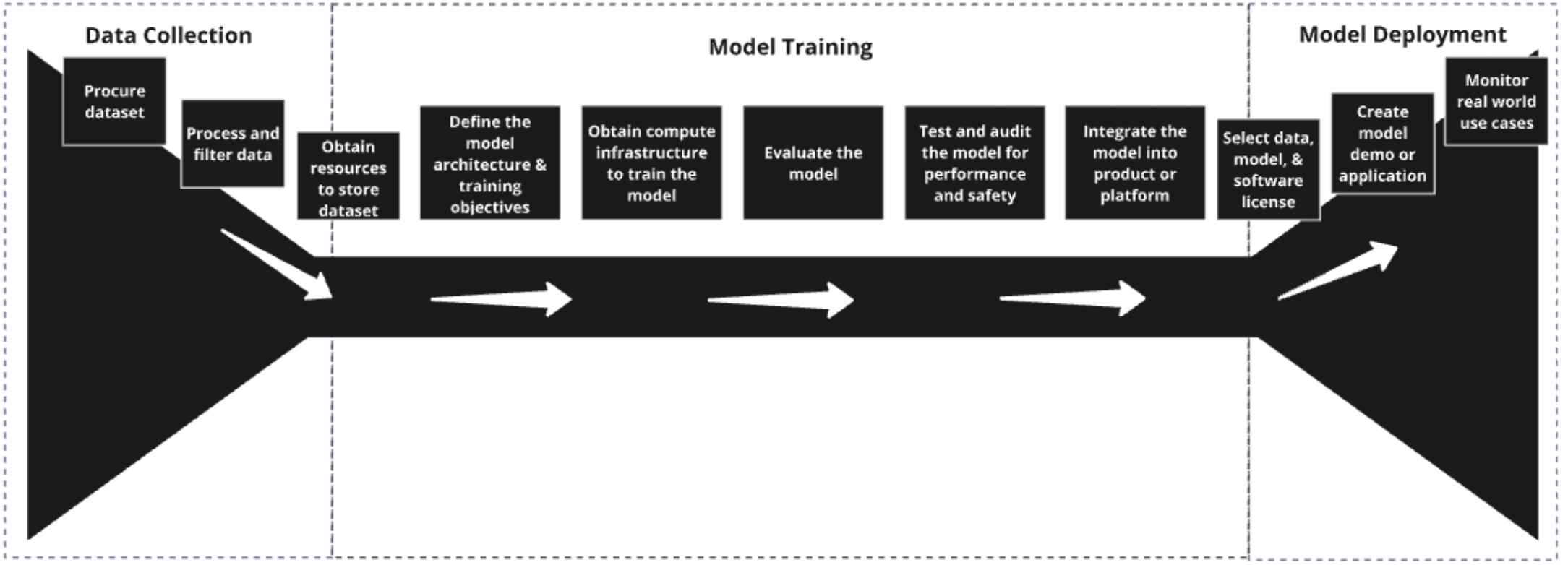
Photo by Desola Lanre-Ologun | https://unsplash.com/photos/woman-and-man-sitting-in-front-of-monitor-IgUR1iX0mqM

Single-vendor vs. community models

- A sliding scale without set definitions
  - Big tech: Llama (Meta)
  - Startups: Mixtral (Mistral)
  - Research Institutes: SEA-LION (AI Singapore), OLMo (AI2), Acquila (BAAI)
  - Community: Pythia (ElutherAI), BLOOM (BigScience Workshop), Bielik (Speakleash foundation)

Photo by Katrin Hauf | https://unsplash.com/photos/white-and-brown-weighing-scale-at-0-XyGfxCVbjCA

Model development pipeline



RI. SE

Data collection and curation

- Reuse of existing datasets
- Create and share internally developed training data
- Collecting data through community collaboration
- Inter-organisational partnerships for data collection
- Validating data provenance and licensing
- Data curation and mixture

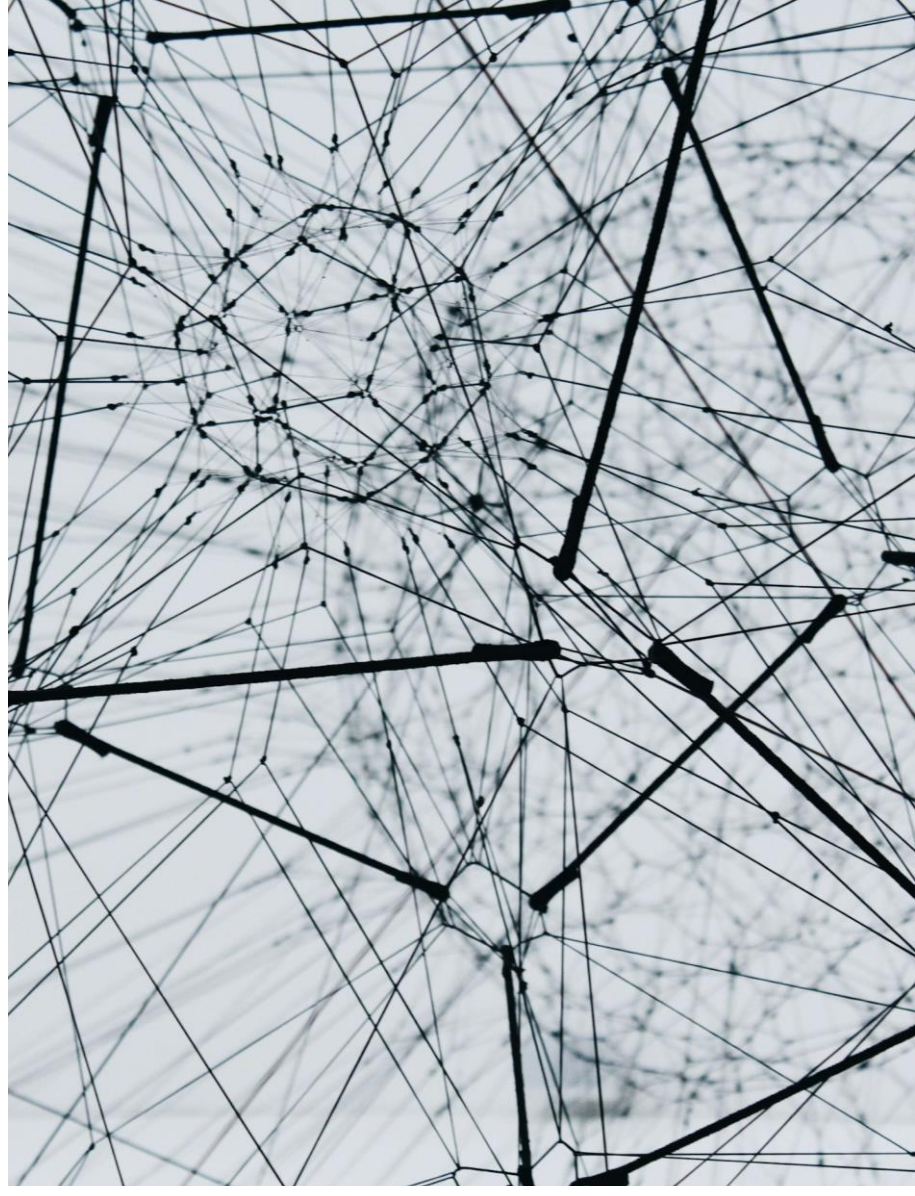


Photo by Alina Grubnyak | https://unsplash.com/photos/low-angle-photography-of-metal-structure-ZiQkhI7417A

Model training

- Open source frameworks for model training
- Compute donations and partnerships.
- Evaluation frameworks and tooling
- Evaluation datasets
- Model Cards, Technical Documentation, and Blog Posts

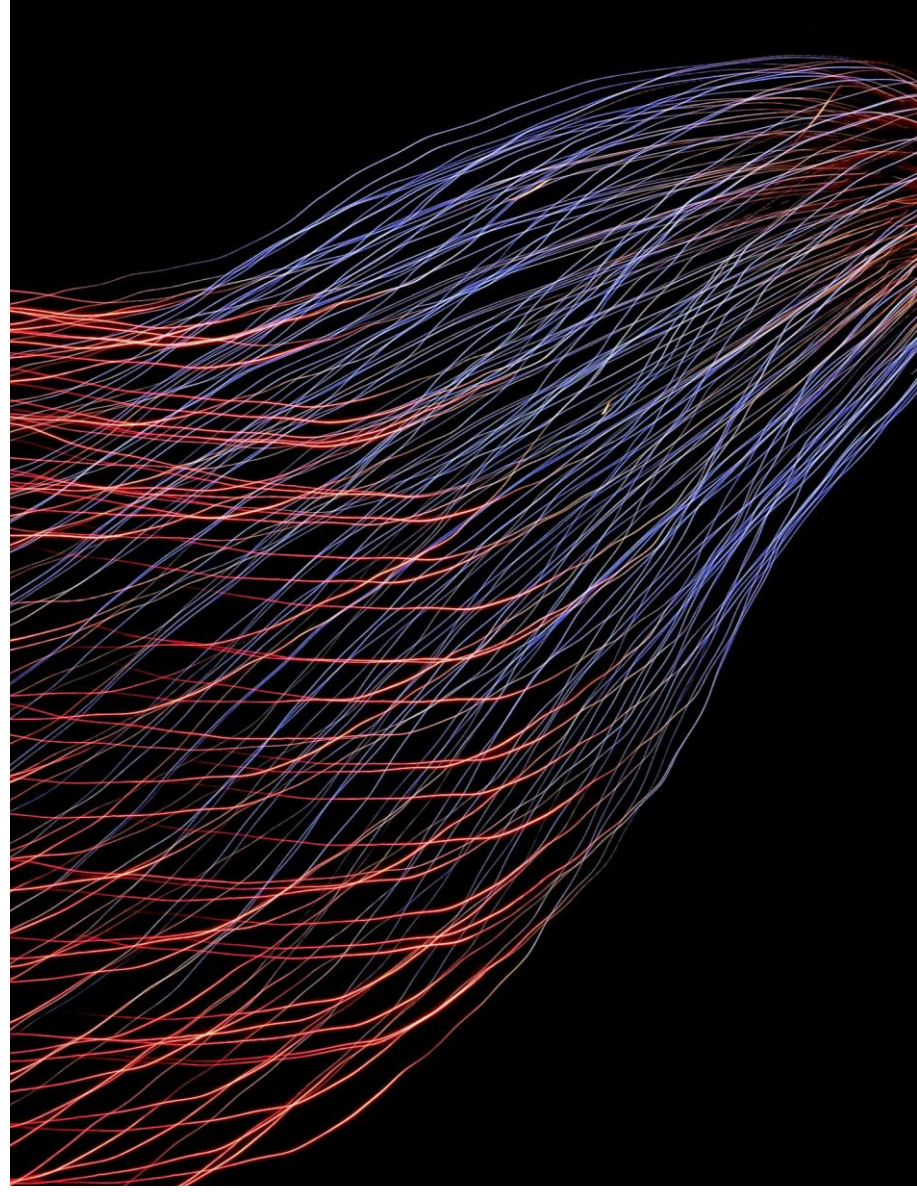


Photo by Hunter Harritt | https://unsplash.com/photos/red-and-blue-lights-from-tower-steel-wool-photography-Ype9sdOPdYc

Model deployment

- Benchmarking
- Derivative models
- Demos and applications
- Up-stream contributions
- Knowledge-sharing
- Quality assurance

Photo by Deng Xiang | https://unsplash.com/photos/graphical-user-interface--WXQm_NTK0U


What opportunities and challenges do you see for collaborating on AI model development?

Open Source <X> as a tool for gaining digital sovereignty

@johanlinaker | https://linaker.se

Reuters World Business Markets Sustainability More My News Subs

Dutch parliament calls for end to dependence on US software companies

By **Toby Sterling**

March 18, 2025 7:22 PM GMT+1 · Updated 11 hours ago

Europe's Digital Leaders Call for "Tech Sovereignty"

The anti-protectionist, pro-tech D9+ alliance switches sides as Europe debates targeting US digital services with tariffs and other measures.

European industry big win: Germany, France both support sovereign EU-based tech infrastructure

That was the easy part, now the real difficulties unfold.

Digital Sovereignty, Security, and Digital Innovation in Germany

Dirk Schrödter
Minister und Chef der Staatskanzlei des Landes Schleswig-Holstein

ARGUME

An expert's point of view on a current event.

Europe Must Avoid Becoming a Digital Colony

The EuroStack is the continent's last chance for technological sovereignty in the era of AI.

Digital sovereignty: less dependency, more European integration

The coalition partners are focusing strongly on digital sovereignty – an approach that could have far-reaching consequences for com

- **Building European integrated and resilient value chains**: "We are creating European integrated and resilient value chains for ke industries, from raw materials and chips to hardware and software." This is intended to reduce dependence on non-European manufacturers – both from China and the USA.
- **Strengthening open interfaces and standards**: "We ensure our digital sovereignty: we define open interfaces and open standarc across all levels and systematically promote open source [...]." This is a reaction to the increasing dependence on proprietary syste which are causing huge difficulties in the context of digital transformation and have had an enormous braking effect. With regard t source, two options are discussed in the document – either "define ambitious targets" or specifically "achieve an open source shar 50% by 2029" for the IT budget.
- **Exclusion of untrustworthy providers**: "We will exclude untrustworthy providers with legal certainty in future." This refers prima critical infrastructures and is likely a reference to the debate surrounding providers in the 5G network.

Framework agreement: MS-365 alternative OpenDesk to conquer the Bundeswehr

The IT system house of the BWI has concluded a framework agreement with Zendis for "sovereign communication and collaboration solutions" such as OpenDesk.

We are digitally cuffed by dependencies

- Software solutions siloes people in our organizations rather than connecting them
- Vendors and big-bang platforms define the needs, not the end-users
- Complete trust is required into how our data is managed, and decisions and operations are performed. If we want access to it, we need to pay up.
- License fees we pay are defined by the vendor, not through competitive procurements

Photo by gaspar zaldo | https://unsplash.com/photos/man-in-black-suit-jacket-wearing-silver-chain-bracelet-7Wg7NJy5JBE/

Dependencies few but strong

- We are heavily dependent on a few but very strong actors.
- The call to and need for breaking free of these dependencies is not something new.
- A worry about digital sovereignty of Denmark, or rather the lack of it.

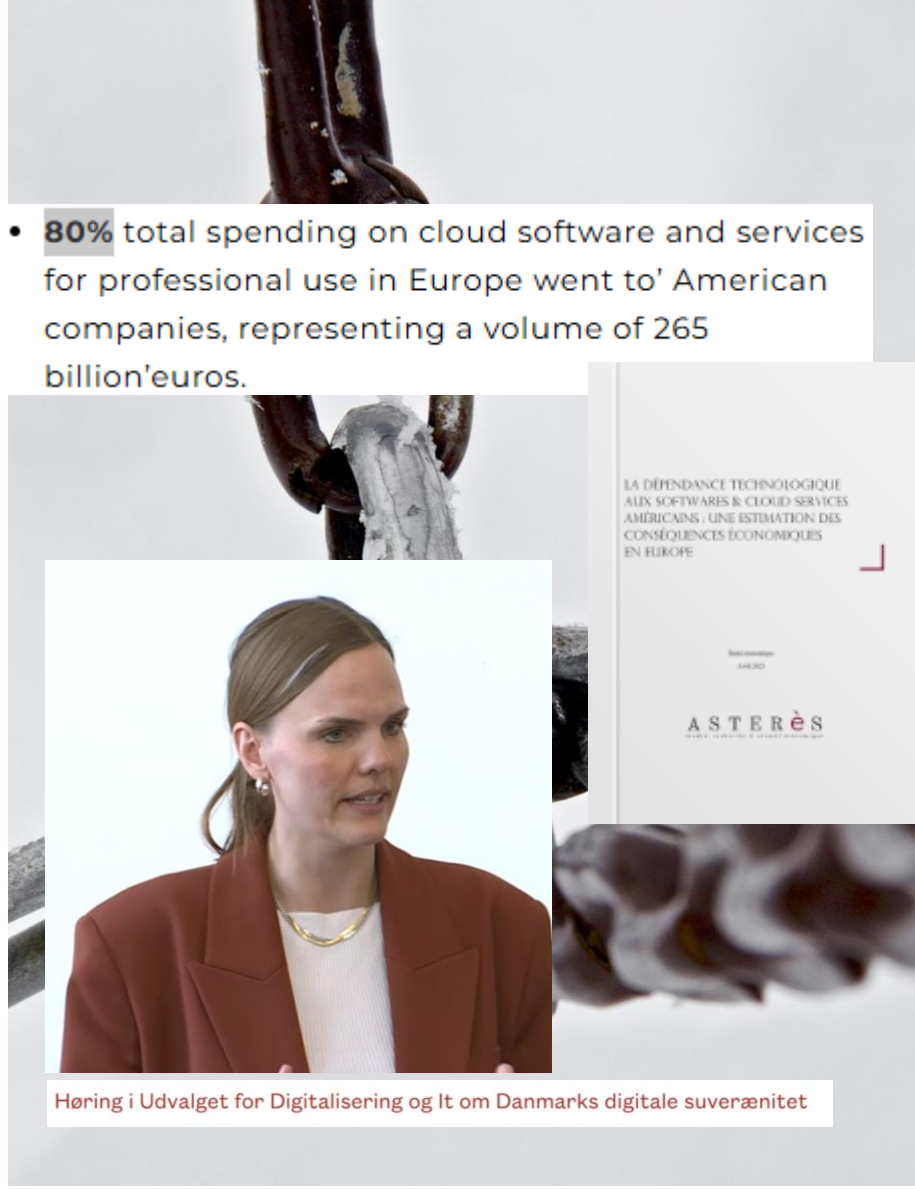


Photo by Bryson Hammer | https://unsplash.com/photos/selective-focus-photograph-of-gray-metal-chains-JZ8AHFr2aEg

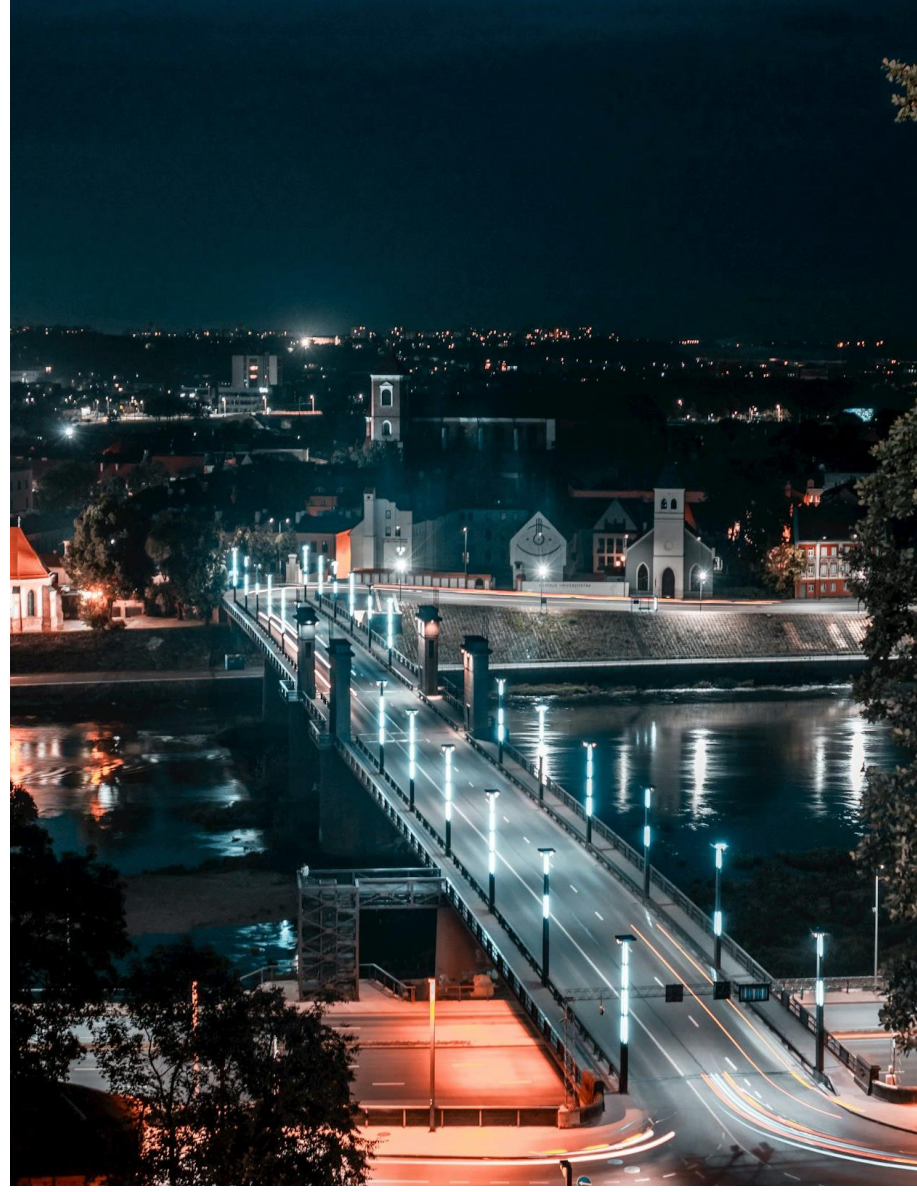
Sovereignty is about control, not isolation

- It’s about being able to maintain structural control of one’s digital sphere, based on one’s own values, norms and legislation
- Open strategic autonomy highlights how autonomy is gained by staying open to share and reuse, interoperability and collaboration
- Essentially nothing different from general vendor independence, but adding on a geopolitical dimension

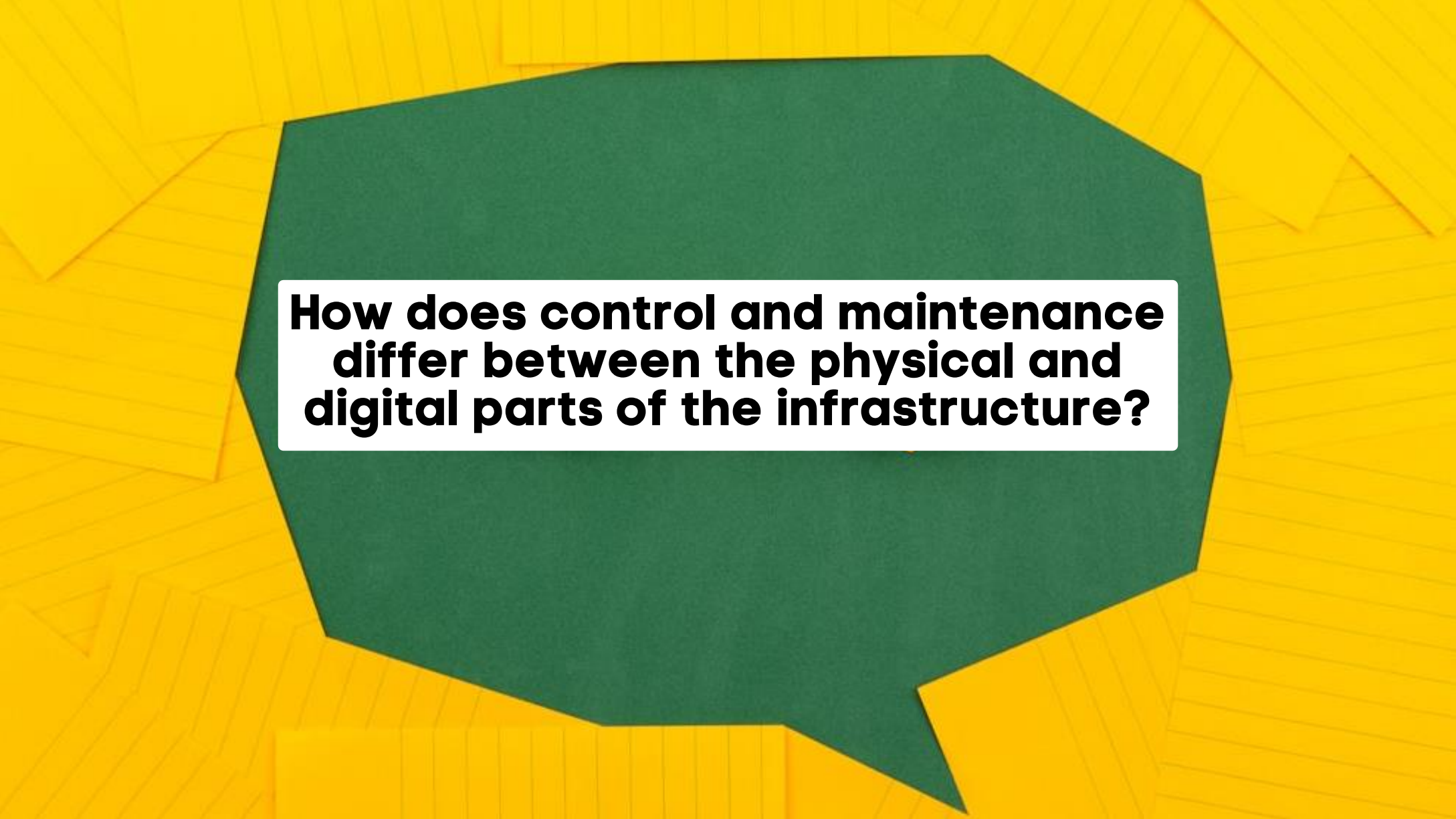
Photo by Adrià Crehuet Cano | https://unsplash.com/photos/children-playing-soccer-LIhB1_mAGhY

All parts of the infra-structure are equally critical

- Digital infrastructure is intertwined with the physical
- The amount of Analog Physical infrastructure is decreasing rapidly on behalf of the Digital Physical infrastructure
- The digital parts needs to be as robust, safe and secure as the rest of the infrastructure

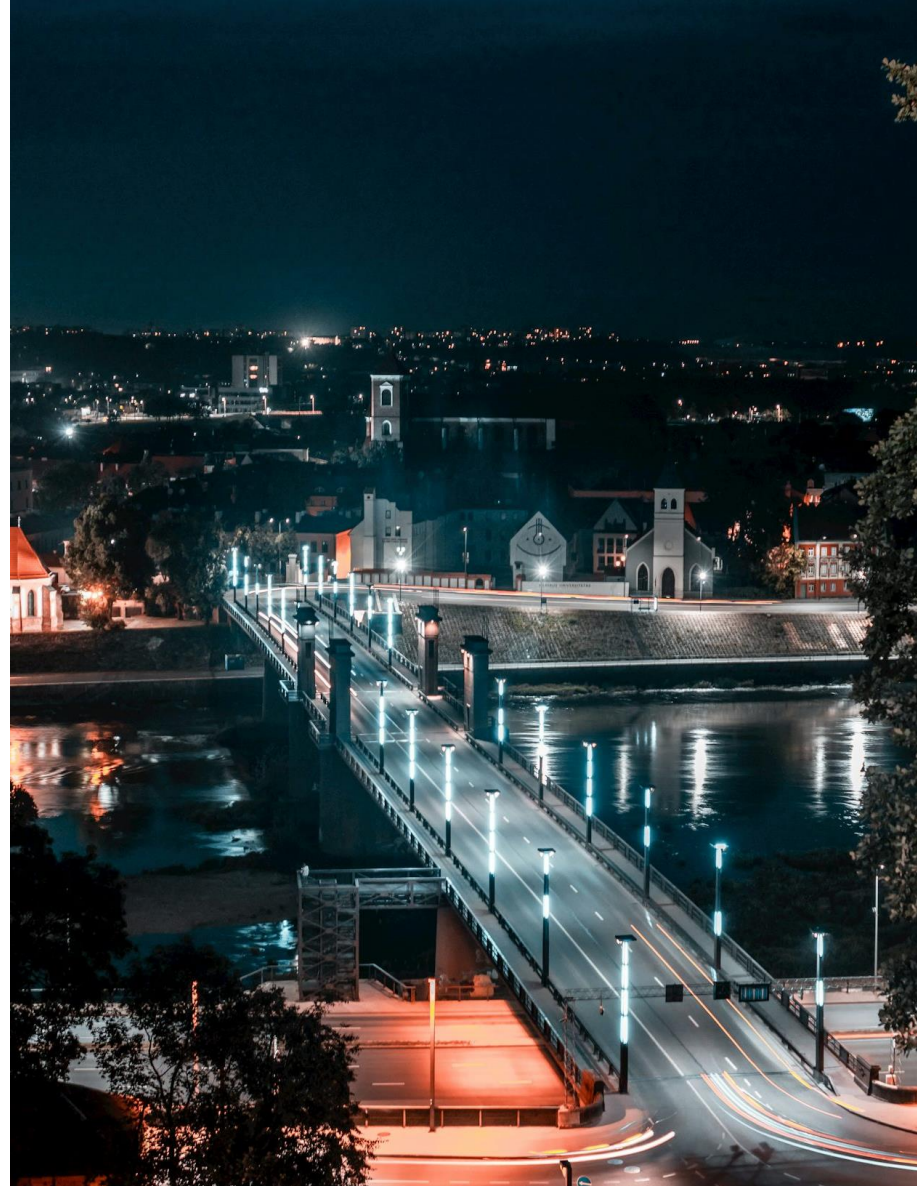


Tadas Petrokas | https://unsplash.com/photos/a-view-of-a-city-at-night-from-a-hill-dE45iUwsCOI

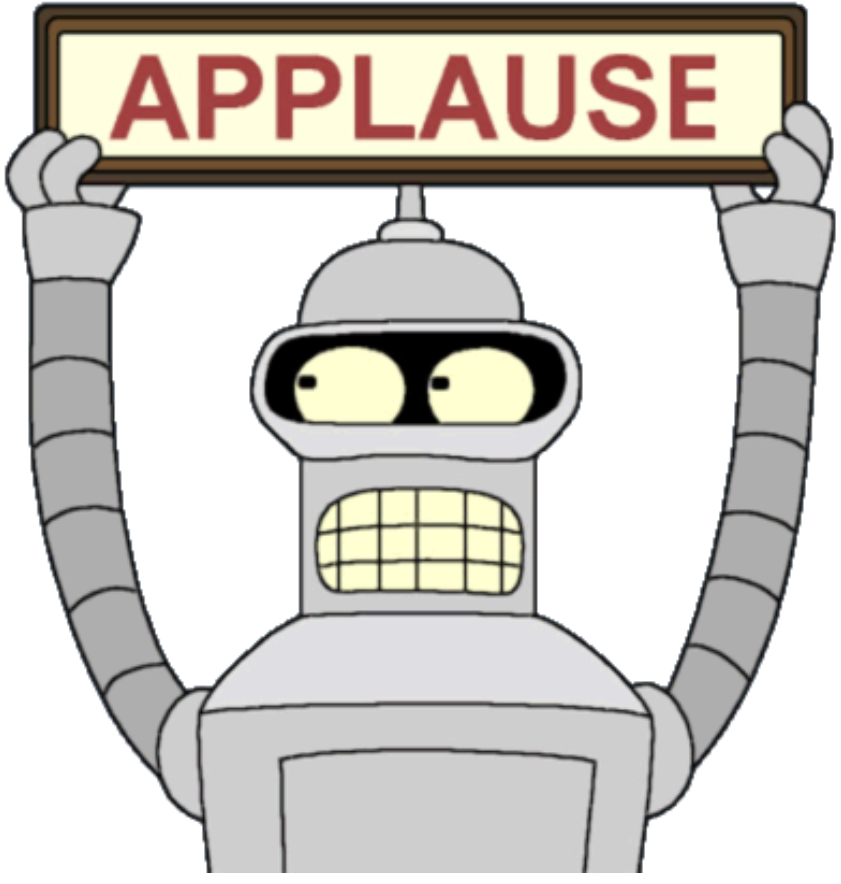

How does control and maintenance differ between the physical and digital parts of the infrastructure?

Control in open source comes through influence

- You need to invest and monitor in key open source technologies (software, AI, data…)
- Ensure your agenda and needs are satisfied, otherwise you need to invest more
- The one driving the development has the (implicit/explicit) control
- Capabilities for driving and ensuring maintenance of digital infrastructure critical for society’s robustness and resiliency



Tadas Petrokas | https://unsplash.com/photos/a-view-of-a-city-at-night-from-a-hill-dE45iUwsCOI



RI.
SE