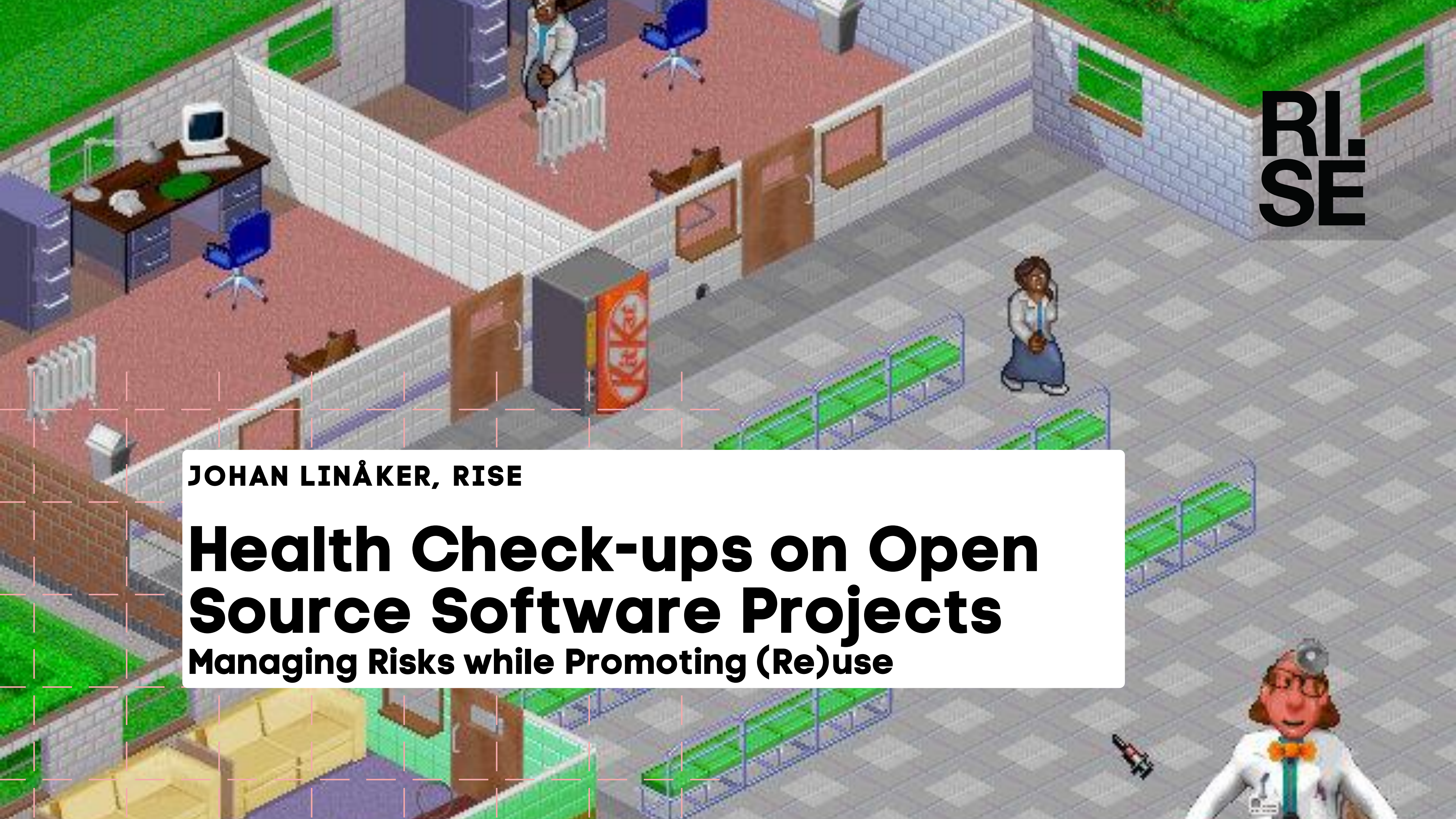


JOHAN LINÅKER, RISE

Health Check-ups on Open Source Software Projects

Managing Risks while Promoting (Re)use

Open Source Software Health

- An Open Source Software project's capability to stay viable and maintained over time without interruption or weakening

@johanlinaker

Photo by Jan Piatkowski | https://unsplash.com/photos/eopz9bkwROs

Open Source Software Health

- 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/IJ0KiXI4uys

Linus’ law

- ”Given enough eyeballs, all bugs are shallow”
- Requires that enough eyeballs actually reaches the codebase
- Free-riding, for both good and bad

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

The Tragedy of the commons

- Commonly exemplified through Hardin's open pastures (Hardin, 1968)
- May be considered as a Common Pool Resource (CPR)
- A resource system that is non-exclusive, and subtractable (Ostrom, 1990)

@johanlinaker

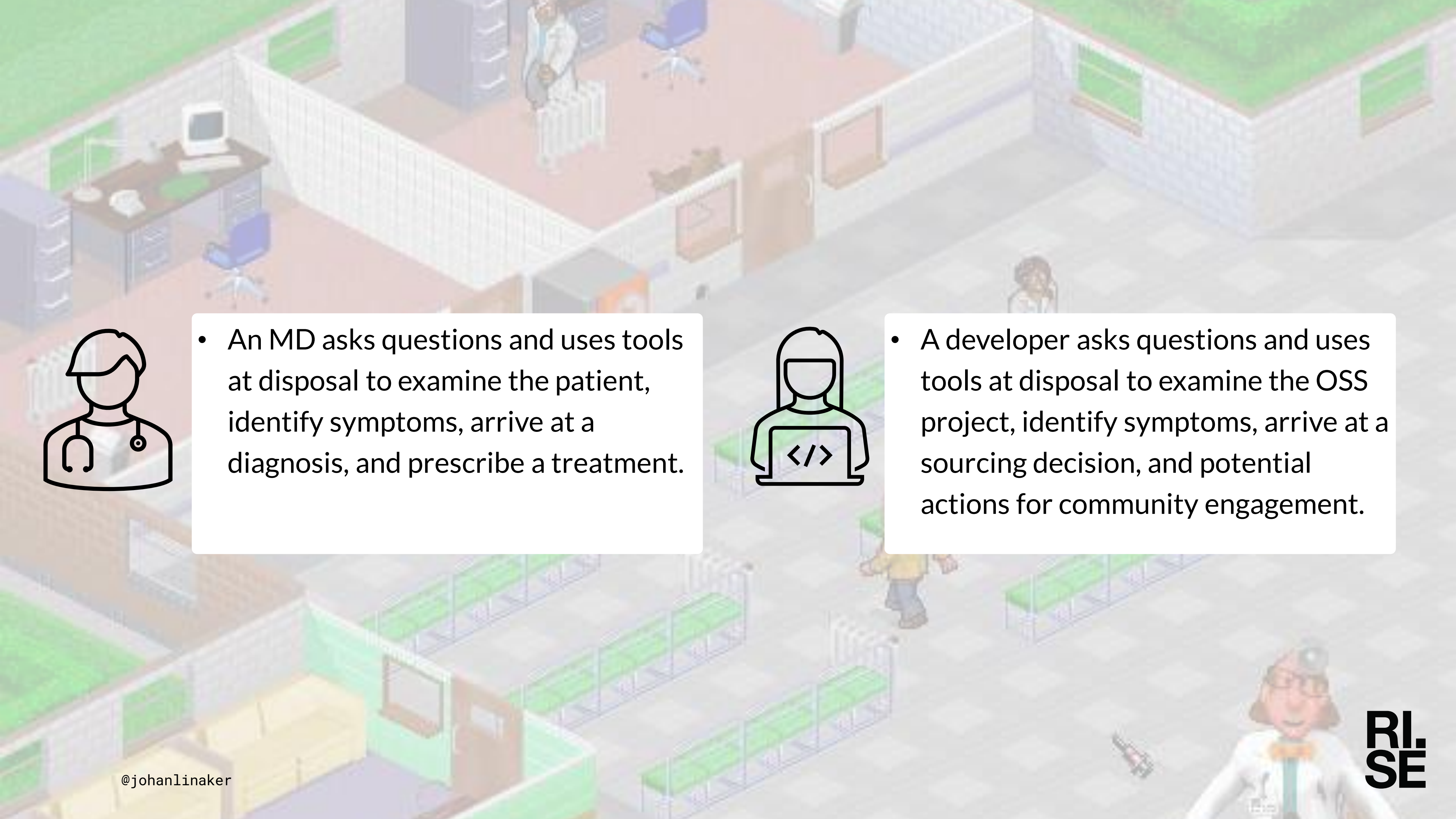
Photo by Martin Schmidli | https://unsplash.com/photos/qg8ihPPQkKo

Brain-time as a Common Pool Resource

- "Brain-time" and maintenance effort is subtractable
- Maintainers are humans, not robots
  - Burnout, changed family or working conditions
- Companies must adapt to stay competitive
  - Refactorization, new products, changed business model

@johanlinaker

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



- An MD asks questions and uses tools at disposal to examine the patient, identify symptoms, arrive at a diagnosis, and prescribe a treatment.

- A developer asks questions and uses tools at disposal to examine the OSS project, identify symptoms, arrive at a sourcing decision, and potential actions for community engagement.

@johanlinaker

Photo by Luke Chesser | https://unsplash.com/photos/JKUTrJ4vK00

Health and Security Management for OSS (HASMOSS)

- Two-year Vinnova-funded R&D-project
- Goals:
  - Enable health analysis at intake and acquisition of OSS, and ongoing consumption
  - Enable sourcing decisions and proactive health improving measures

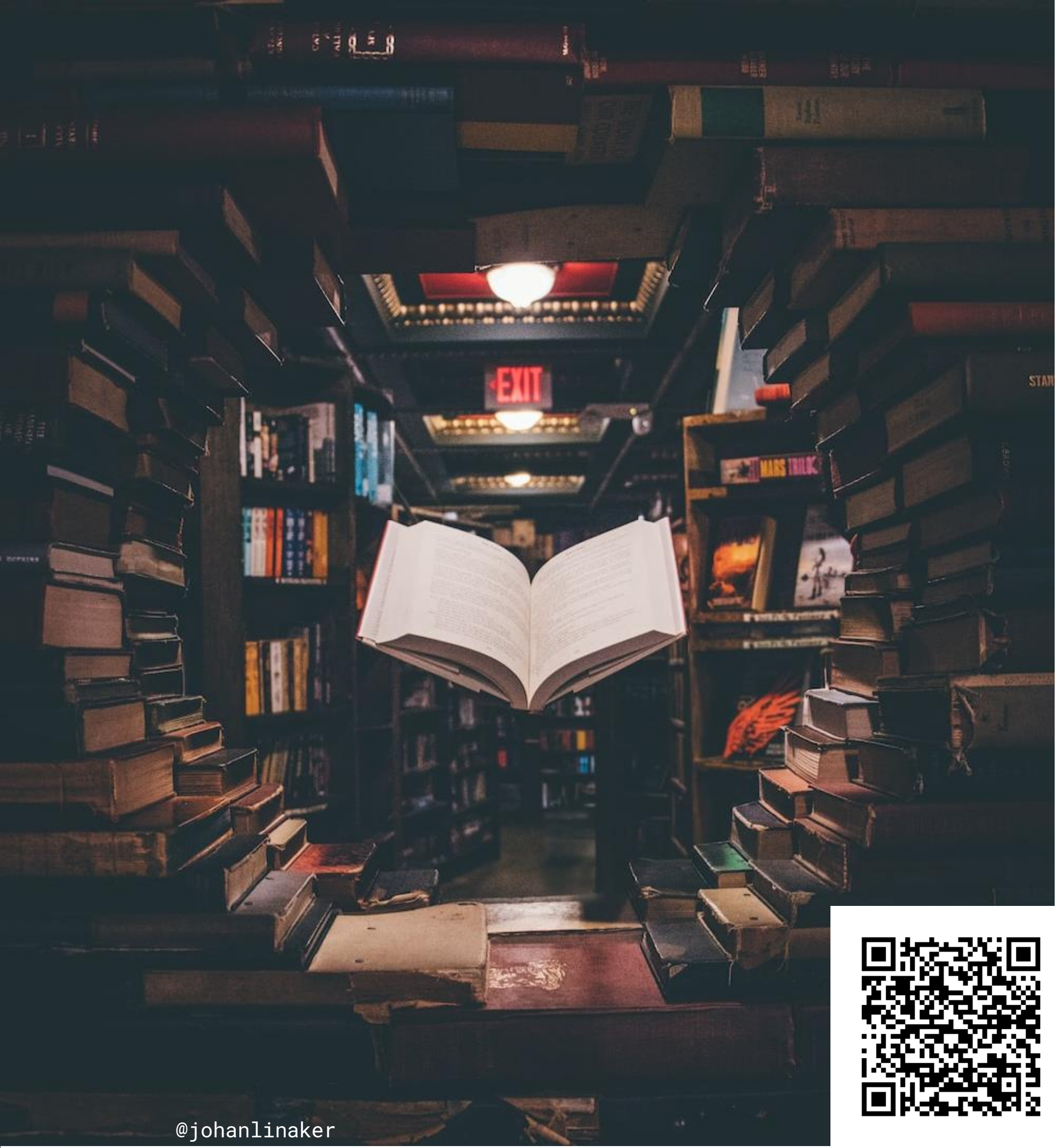


Photo by Jaredd Craig | https://unsplash.com/photos/HH4WBGNyltc

What can we find in literature?

- 146 studies
- 107 characteristics (+associated metrics
- Divided over 15 themes
- Supplementary material: https://doi.org/10.6084/m9.figshare.20137175
- Paper: https://www.ri.se/sites/default/files/2022-09/opensym2022-6%20%281%29.pdf

Photo by Austin Distel | https://unsplash.com/photos/smiling-man-reading-book-while-holding-mug-4r72LPFh4Ik

What does experts say?

- 17 interviews with industry and community experts
- 4 areas critical to classify projects, impacting what metrics to prioritize and how tough
- 21 areas of complementary metrics considering
  - Community productivity, and stability
  - Orchestration
  - Production process and outputs

Project Classifier

- Life-cycle stage
  - 1) inception, 2) growth, 3) stabilization, and 4) decline
- Project Complexity
  - scope, size, and technical complexity of the codebase
- Governance concentration
  - impact on the project’s openness to input and external influence on decisions and transparency of discussions
- Strategic Importance
  - importance of the OSS project from a business and technical perspective

Photo by Annie Spratt | https://unsplash.com/photos/open-book-page-vpFiAD-WWEs

Community Productivity

- Social activity
  - Activity from the community both in online channels, and physically offline.
- Responsiveness
  - Time to a response towards, e.g., discussions, pull requests, or issues
- External Visibility
  - Visibility to an audience beyond those actively engaged in the project.
- Development Activity
  - Including the many technical aspects and deliverables of the OSS project.
- Development Efficiency
  - effectiveness and ease in managing and moving the development forward

Photo by Andreas Klassen | https://unsplash.com/photos/man-holding-smartphone-looking-at-productivity-wall-decor-gZB-i-dA6ns

Community Stability

- Adoption
  - Usage and technical adoption of the OSS project as a dependency in downstream software projects and by end-users
- Organizational Diversity
  - diversity of organizations within an OSS community in terms of governance, contribution, and adoption of the underpinning project
- Demographical Diversity
  - Diversity on the individual level of the maintainer and contributors to an OSS project in gender, race, time zone, language, and cultural aspects
- Discussion Climate
  - In regard to sentiment, tone, and manner in answers, messages, and general communication within the OSS project, and how potential conflicts are managed.

Photo by Dan Hadar | https://unsplash.com/photos/brown-and-gray-stone-stack-on-beach-shore-during-daytime-HMG2ELxyos8

Community Stability

- Knowledge Concentration
  - Concentration or distribution of contributions and knowledge to specific individuals or groupings within an OSS project.
- Contributor Turnover
  - Attraction, retention, and attrition of maintainers and contributors to an OSS project
- Financial Sustainability
  - Financial situation of maintainers and contributors of OSS projects and whether it enables sustainable and dedicated time for maintenance of the projects.

Photo by Markus Spiske | https://unsplash.com/photos/gray-and-brown-rocks-f81VZFRDPCc

Orchestration

- Governance Structure
  - Explicitness, formality, and general recognition of the ecosystem’s governance structure and leadership
- Openness
  - To what extent the OSS project is welcoming to and accepting contributions and considering new ideas and general input and influence on the project’s development from existing and new contributors
- Licenses
  - License-related aspects and processes of managing and distributing the intellectual property maintained by the OSS project.

Photo by Christina @ wocintechchat.com | https://unsplash.com/photos/group-of-people-sitting-beside-rectangular-wooden-table-with-laptops-faEfWCdOKIg

Production process

- Development process
  - Presence and quality of development processes is seen by multiple interviewees as an important marker of a mature and sustainable OSS project
- Release Management
  - The release process should describe the governance and planning of how releases are made, and at what cadence
- Security Management
  - The implementation and management of proactive and reactive measures to prevent and address security concerns of the OSS project
- Scaffolding
  - The availability and quality of the development and communication infrastructure used in the OSS project

Production output

- Documentation
  - The presence and quality of documentation for the OSS project considering different stakeholders' perspectives, including developers and end-users
- Technical quality
  - The technical quality of the OSS and its source code, e.g., in terms of its architecture, source code and other quality attributes

Photo by Katja Vogt | https://unsplash.com/photos/two-white-ceramic-mug-CipURjPCXOo

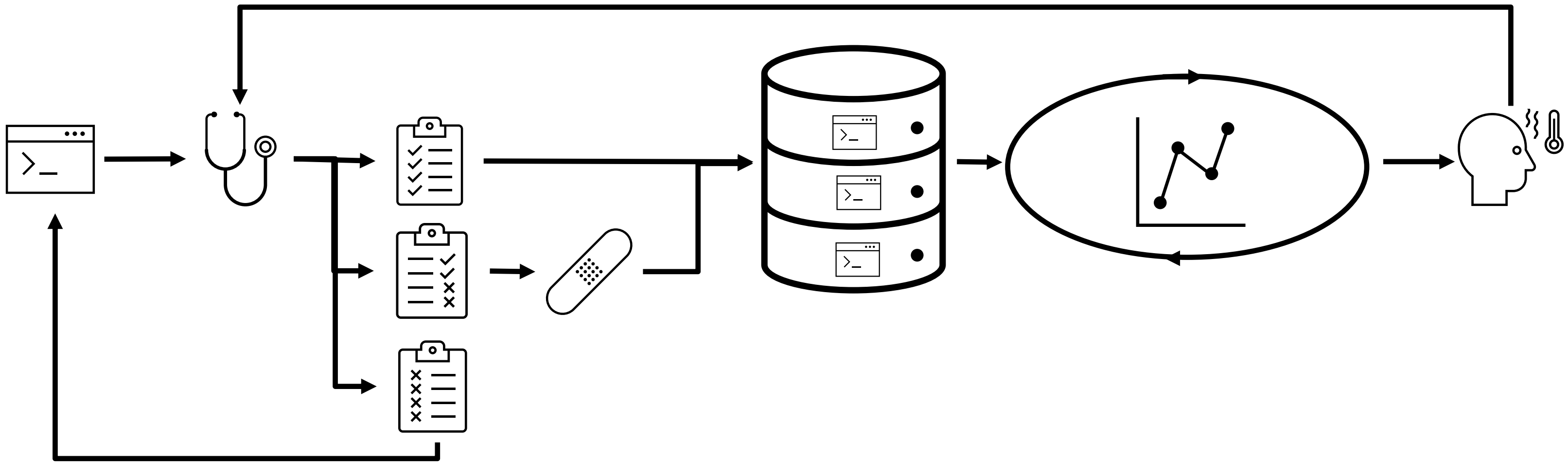
Extant work

- Community Health Analytics for Open Source Software (CHAOSS)
  - Framework with metrics for health analysis and assessments
- Open Software Security Foundation (OpenSSF)
  - Industry foundation focused on raising security of critical OSS
- SustainOSS
  - Community focused on sustainability and health topics

Going from theory to practice

- What:
  - Lower risk of OSS used and considered in the intake process
- How:
  - Set up an intake and screening process for new and existing OSS dependencies
  - Monitor health and make proactive decisions on sourcing options and community engagement
- Key requirements:
  - Decentralized, self-managed process
  - Enable but don't overburden developers
  - Enable follow-up and actionable insights

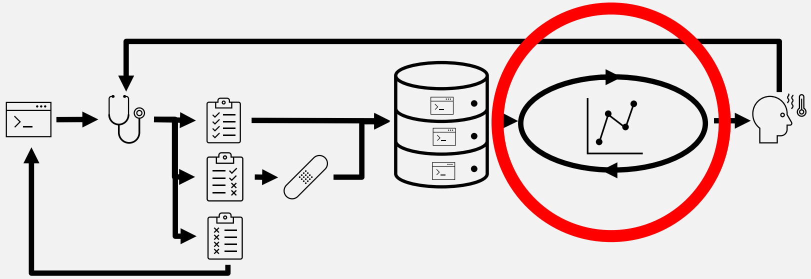
Semi-automating the health-check process



RI.
SE

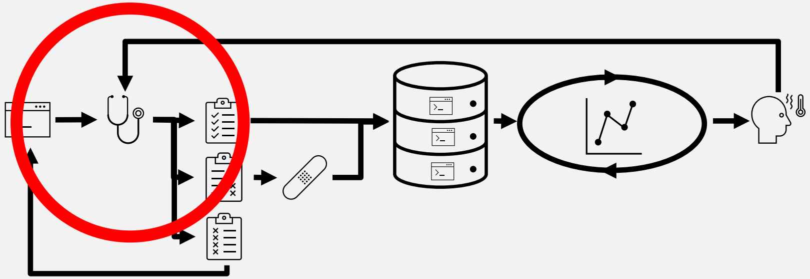
Quantitative screening

- Large amounts of dependencies commonly exist. Manual overview and inspection not applicable
- Tooling needed, intergated in CI/CD pipelines or partial-runs on regular occasions
- Runs high-level tests on dependencies tailored based on the type of ecosystem and dependencies
- Flags projects and directs attention where indicators together point towards a potential risk
- Manual inspections follow by developers or analysts
- Custom tooling and/or off-the shelf. See e.g., GrimorieLab and Debricked OSS Intelligence



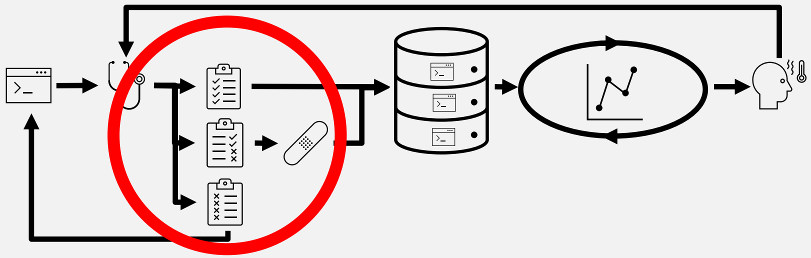
Manual inspections

- Analysis on single projects, either identified in screening, or as input to sourcing decision (intake process)
- Use of standardized checklist with automated tool support as needed
  - Trade-off between rigor and efficiency
  - Interview and map up main concerns from internal stakeholders
  - Consider types of projects used and need for tailoring
  - Needs simple answers (Yes/No) or clear categories (1-5, 6-10…)
- Lightweight documentation process, persisting and indexing analysis for future follow-up



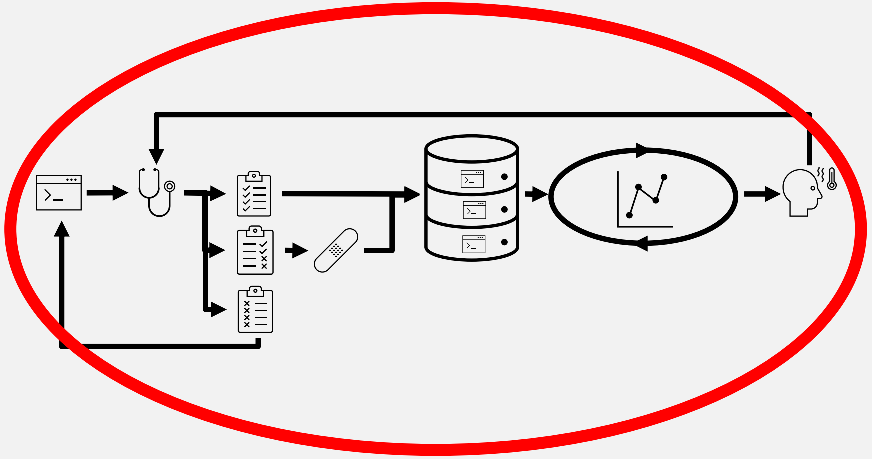
What to check for?

- Need to define the goals the analysis and the questions you want to answer
  - Main concerns and risks
  - types of OSS projects, in what domains, etc.
- Literature and practice have provided a knowledge base use together with existing initiatives, e.g., CHAOSS, OpenSSF
- Requires work up-front
- Evaluation at Scania
  - Focus group + user observations
  - Condensed into checklist of 14 health attributes



Training and follow-up needed

- Workshops for introducing checklists and analysis process
- Integrate as standard practice in development and Q&A
- Recurrent feedback session for presenting analysis of OSS projects
  - Encourage discussion, knowledge-sharing, and critical mindset
  - Contrast between types of projects, relevant questions to ask, and application/interpretation of metrics

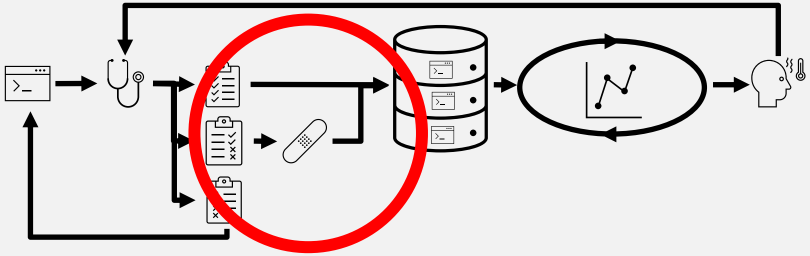


Sourcing and acquisition

- Pre-trial at large Swedish national agency
- Workshop format with internal stakeholders
- Goal was to evaluate health of to OSS e-archival solutions
- Questionnaire developed through iterations based on CHAOSS metrics
- Enable comparison between open and closed alternatives in an acquisition
- Evaluation needs to be thorough and detailed

Prescribing the necessary treatments

- Secure and enable the need human resources needed for a sustainable maintenance
- Originates either from the maintainers, or the community
- Requires investments and support of a human infrastructure in the projects



Human Infrastructrue in support of a sustainable maintenance

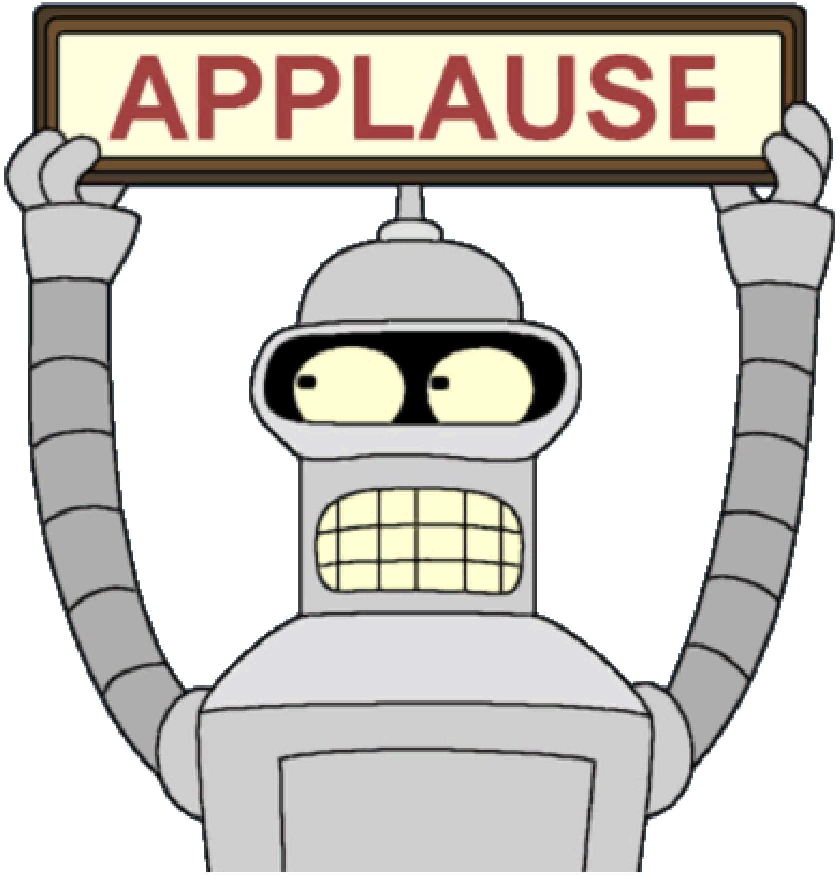
- Maintainer resources
  - Managing social expectations and peer-pressure
  - Balancing of workload with capacity
  - Finding time through funding
  - Work-life balance and prioritization

- Community resources
  - Embracing the episodic contributors
  - Mitigating toxicity
  - Promoting inclusiveness
  - Managing impact of project characteristics
  - Low-cost contributor support
  - Marketing and outreach
  - Distributing knowledge

Resource funding

- Full-time employment dedicated to projects
- Partially-dedicated employment
- Entrepreneurship, a common but risky endeavor
- Sponsorship, a diverse and limited source of income

Photo by Anne Nygård | https://unsplash.com/photos/brown-and-white-paper-bag-OtqaCE_SEMI



APPLAUSE