RI.
SE

JOHAN LINÅKER, EFI PAPATHEOCHAROUS, THOMAS OLSSON

How to characterize the health of an Open Source Software project?

A snowball literature review of an emerging practice

Open Source Software Health

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

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

Open Source Software Health

- Or from an ecosystem perspective*…
  - 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

* Adapted from e.g., Jansen (2014) and Ianisti & Levian (2004)

Photo by Markus Frieauff | https://unsplash.com/photos/IJ0KiXl4uys

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

Open Source Software 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

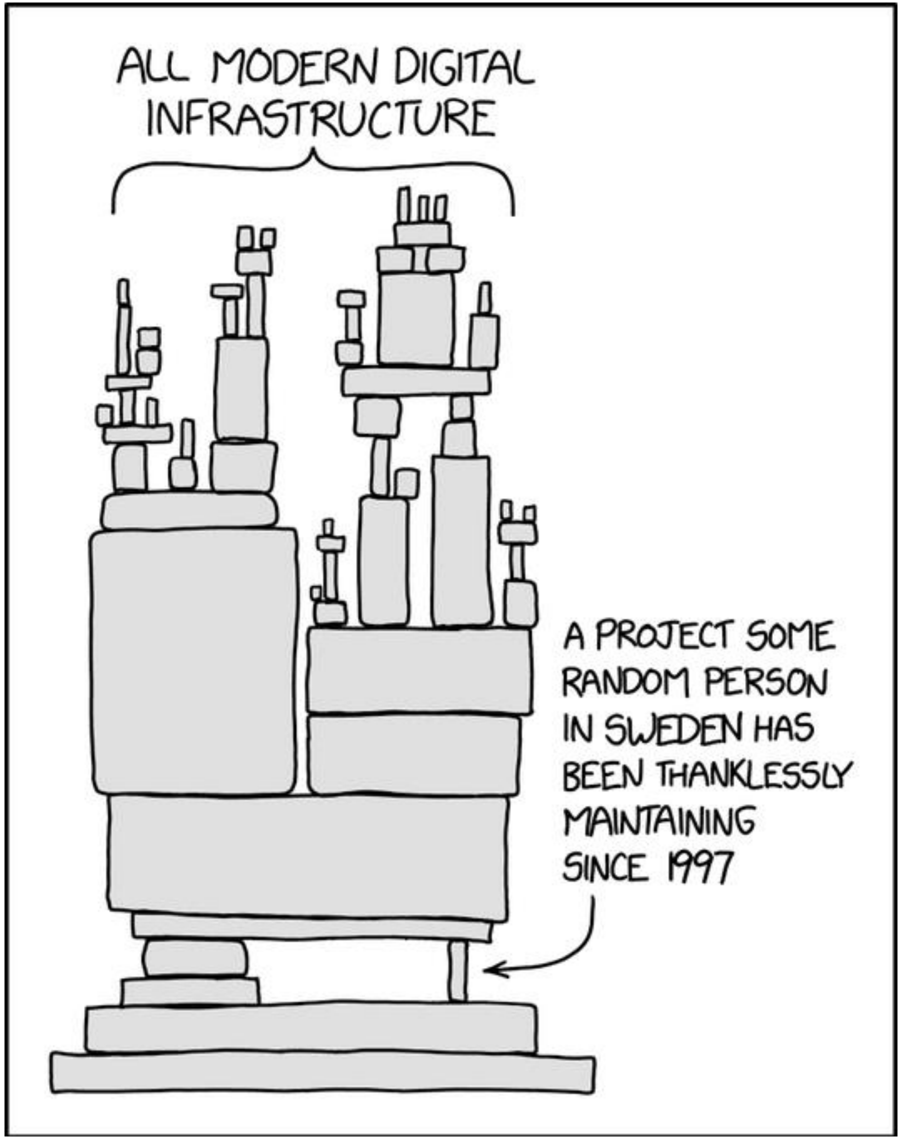
ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN SWEDEN HAS BEEN THANKLESSLY MAINTAINING SINCE 1997

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

Open Source Software 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

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

What characteristics affect the health of an OSS project?

- First part of a longer design science research project*
- Goals:
  - Enable health analysis at intake and acquisition of OSS, and ongoing consumption
  - Enable sourcing decisions and proactive health improving measures

* https://bit.ly/3AM5NR8

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

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

What can we find in literature?

- Snowball literature review
- From start-set of 8, two iterations back and forth, to 146 included primary studies
- Goal of reaching topic saturation, not paper saturation
  - Gain a comprehensive overview
  - Input for future design cycles

A-priori framework

- Level of abstraction*
  - Network-level
  - Project-level
- Socio-technical dimension**
  - Actors
  - Software
  - Orchestration

* Jansen (2014)
** Manikas and Hansen (2013)

Analysis process

- Structural coding per paper
- Axial coding per dimension
- Output:
  - Assessment framework
  - 107 characteristics
  - Divided over 15 themes

Actor-related characteristics

- Communication (4)
- Culture (6)
- Finance (2)
- Popularity (5)
- Stability (12)
- Technical activity (5)

Photo by Ryoji Iwata | https://unsplash.com/photos/IBaVuZsJJTo

Software-related characteristics

- Development process (7)
- Documentation (6)
- General characteristics (7)
- Licence (4)
- Scaffolding (10)
- Security (10)
- Technical quality (10)

```
            t.appeared = false;
            return;
        }

        //is the element inside the visible window?
        var a = w.scrollLeft();
        var b = w.scrollTop();
        var o = t.offset();
        var x = o.left;
        var y = o.top;

        var ax = settings.accX;
        var ay = settings.accY;
        var th = t.height();
        var wh = w.height();
        var tw = t.width();
        var ww = w.width();

        if (y + th + ay >= b &&
            y <= b + wh + ay &&
            x + tw + ax >= a &&
            x <= a + ww + ax) {

                //trigger the custom event
                if (!t.appeared) t.trigger('appear', settings.data);

            } else {

                //it scrolled out of view
                t.appeared = false;
            }
        };

        //create a modified fn with some additional logic
        var modifiedFn = function() {

            //mark the element as visible
            t.appeared = true;

            //is this supposed to happen only once?
            if (settings.one) {

                //remove the check
                w.unbind('scroll', check);
                      .inArray(check, $.fn.appear.checks);
                      checks.splice(i, 1);
```

Photo by Markus Spiske| https://unsplash.com/photos/1LLh8k2_YFk

Orchestration-related characteristics

- Orchestration (9)

Photo by Sivani B | https://unsplash.com/photos/bczrpU9n8f4

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

Application in practice – Intake

- Pre-trial on a large international software-producing organization
- Lightweight questionnaire/checklist for developers to cross off
- Questionnaire developed through iterations based on CHAOSS metrics
- Evaluation must be quick and easy (<15 min), potential to be automated
- Considered to raise awareness and decrease overall risk in the intake process
- Process owned by enterprise architects

Application in practice – 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

Limitations and threats to validity

- Not a systematic overview of literature
- Limited coverage of the network-level
- Did we really characterize the health of OSS projects?

RI.
SE

Future work

- Validate, prioritize and further characterize characteristics through an interview survey with a general and a case-specific sample (ongoing)
- Gather metrics and data sources from practitioners through observations, and contrast to what we've found in literature
- Investigate applicability and possibility to automate and quantify characteristics with industry partner (ongoing)

Photo by Tomasz Frankowski | https://unsplash.com/photos/kBUfvkbFIoE