On Updating Library Dependencies in Open Source Software Projects

Raula Gaikovina Kula, Dr. Eng
Software Engineering Lab (Matsumoto ken)

Special International Track on
Software Analytics (2) (Invited Talk)

NARA INSTITUTE OF
SCIENCE AND TECHNOLOGY
- Outgrow your limits -

Get the Slides
https://goo.gl/6UcAHC

About Me

NARA INSTITUTE OF SCIENCE AND TECHNOLOGY
- Outgrow your limits -

大阪大学
OSAKA UNIVERSITY

Software Ecosystems

DATA VISUALIZATION

MSR

Software Maintenance Processes

Outline of my Talk

1. **Background:** Third-party Libraries (5 mins)
    - Rise of Library Ecosystems
2. **Motivation :** What about a Library Update?(10 mins)
    - Library Updates in Practice
    - Awareness Mechanisms
3. **Main Study:** A Study on Library Updates (10 mins)
    - Library Updates in Practice
    - Awareness Mechanisms
4. **Future and Challenges:** Library Ecosystems in the Future (5 mins)
    - Ongoing work & Software Analytics

Imagine a Developer in Need of a function, feature ...

Libraries to the rescue

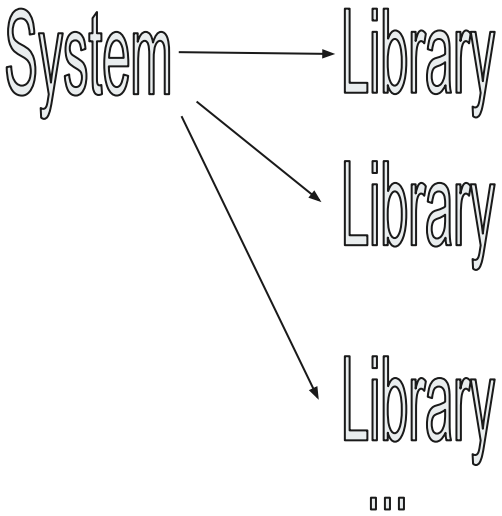
© https://www.youtube.com/watch?v=aIB5xohKUZI

1. Background - Third-Party Library Usage

Modern Landscape of Software Development is built on Third Party Libraries

- Suppose you want to write some unit tests in Java… just adopt the junit library

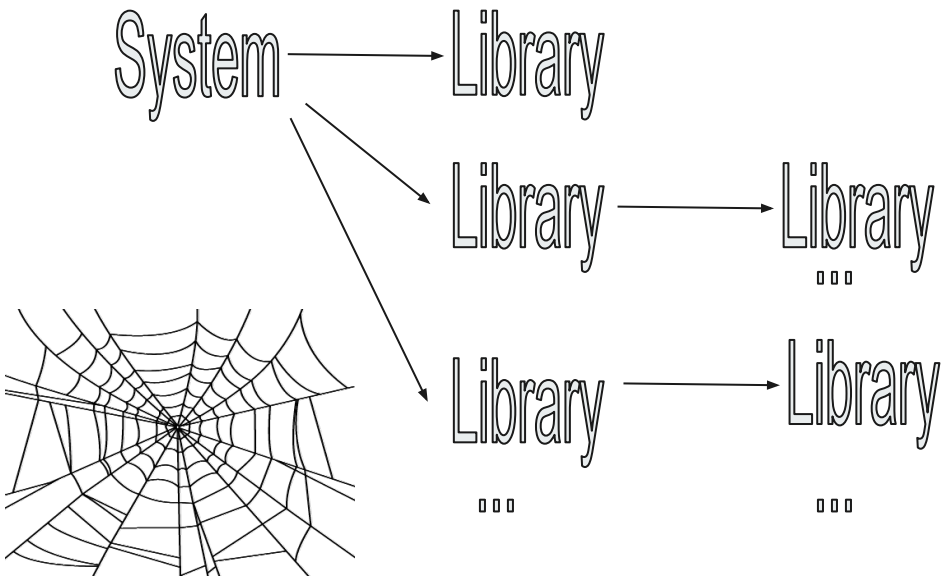
Library Dependencies



Biological Ecosystems

© https://www.tes.com/lessons/vfqedGrGvRj7dw/ecosystems

Results in Software (Library) Ecosystems



https://search.maven.org/

https://rubygems.org/

https://www.cpan.org/

https://cran.r-project.org/

475,000 building blocks

https://www.npmjs.com/

Disrupt the Ecosystem

© https://www.tes.com/lessons/vfqedGrGvRj7dw/ecosystems

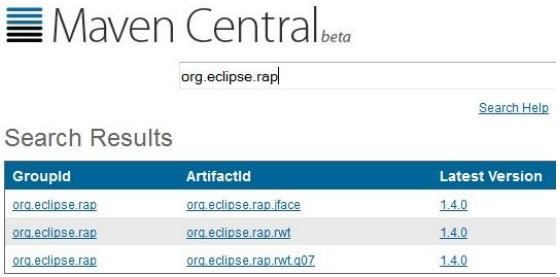
What is a Software Ecosystem?

Adapted from biological ecosystems:

Scypersrsky [1]: *"defined as a set of businesses functioning as a unit and interacting with a shared market for software and services, together with relationships among them."*

Lungu [2]: *"a collection of software systems, which are developed and co-evolve in the same environment"*

Stallman [3] : *"It is a mistake to describe the free software community, or any human community, as an "ecosystem", because that word implies the absence of (1) intention and (2) ethics"*



© https://www.scmagazineuk.com/researchers-developing-cyber-attack-predictive-model/article/540346/

2. Motivation

Updates are sometimes strong recommended,
especially with Security Vulnerabilities

CVE-2014-6271

CVE-2014-3566

CVE-2014-0160

Update Awareness: Security Advisories

Example:
https://cve.mitre.org/cgi-bin/cvename.cgi?name=cve-2014-0160

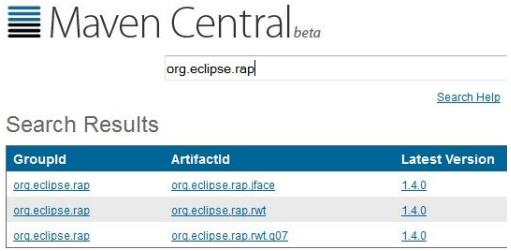
© https://www.singleblackmale.org/2010/02/17/spider-sense/

https://nvd.nist.gov/

Relentless mining of data ….

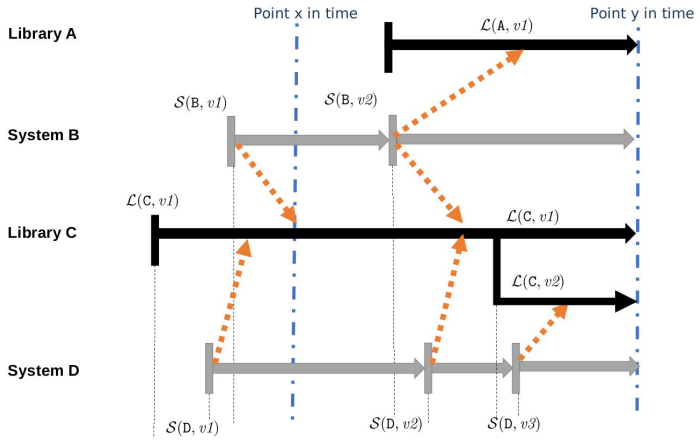
© https://www.ausimmbulletin.com/feature/innovation-in-mining/

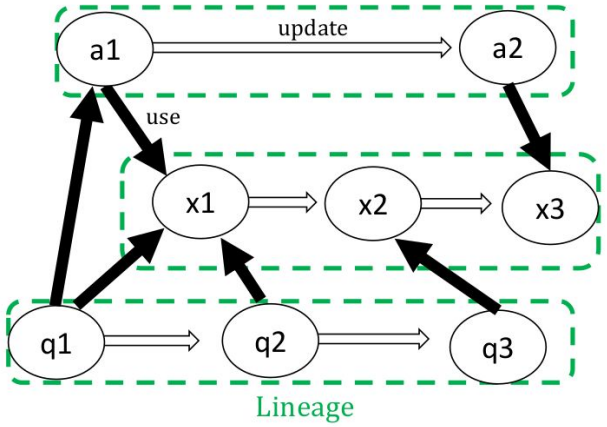
MSR Icon



https://search.maven.org/

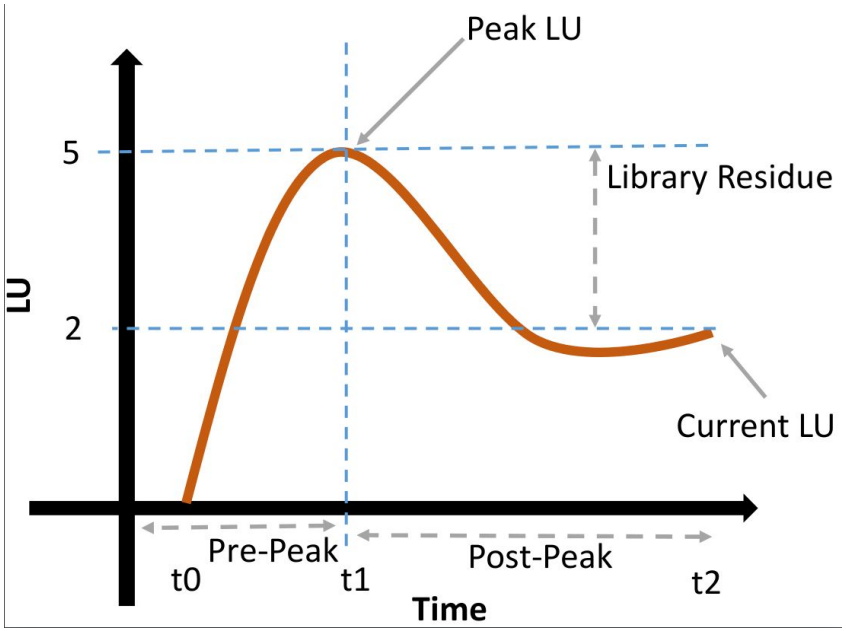
Systematic Modeling of System & Library Usage





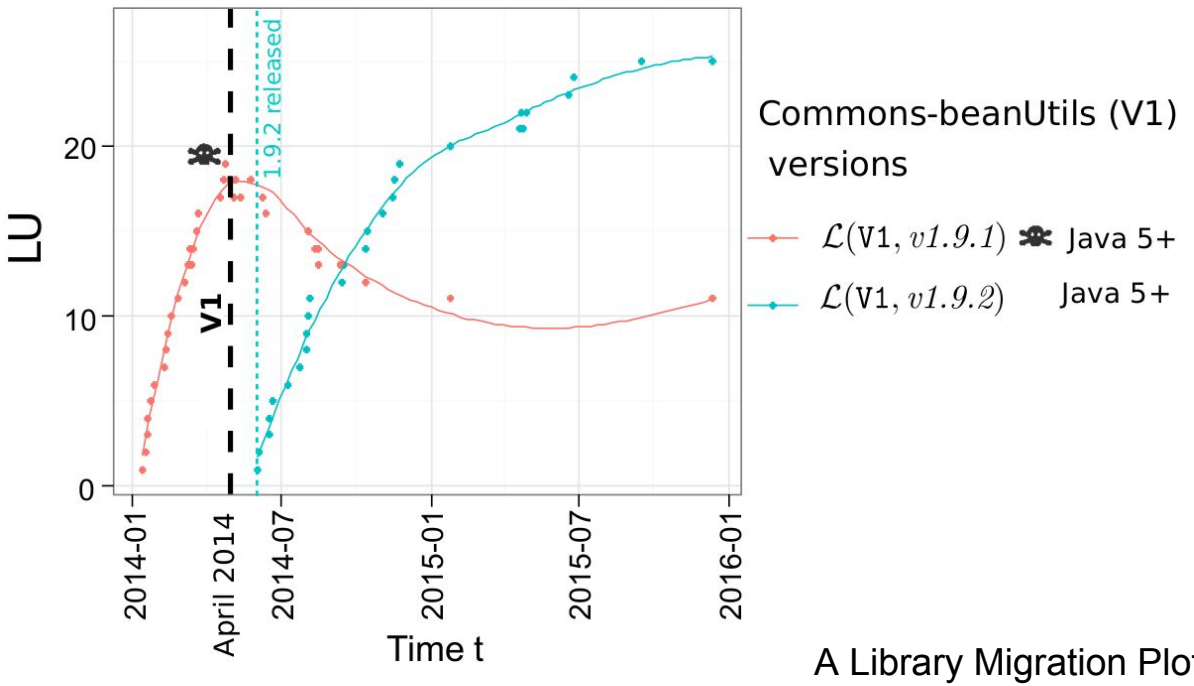
Library migration between systems and libraries. The orange arrow depicts dependency relations between them.

Library Tracking Model



Simple example of the LU-based metrics.
We show the Peak LU at time t1, current LU at time t2 and library residue (Peak LU / Current LU).

Visualizing Library Usage



A Library Migration Plot. In this example, the release of a related security advisory CVE-2014-0114 (black dashed line) that affects beanutils versions 1.9.1 (marked with crossbones).

3. An Empirical Study on Library Updates

Empirical Software Engineering
pp 1–34

Do developers update their library dependencies?

An empirical study on the impact of security advisories on library migration

Authors Authors and affiliations

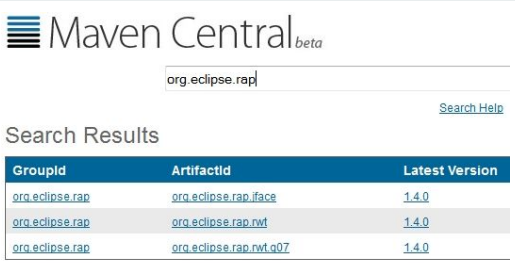
Raula Gaikovina Kula ✉, Daniel M. German, Ali Ouni, Takashi Ishio, Katsuro Inoue

Motivation: In 2014, Sonatype determined that over 6% of the download requests from the Maven Central repository included known vulnerabilities.

Related Work at the API level:

- Robbes et al. [4] - Smalltalk ecosystem
- Hora et al. [5] - Pharo
- Sawant et. al [6] - Java.
  - Bavota et al. [7] - Apache products.

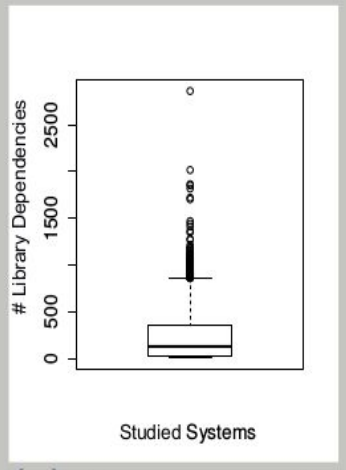
We study both library and system perspectives, as awareness and the nature (new release or security



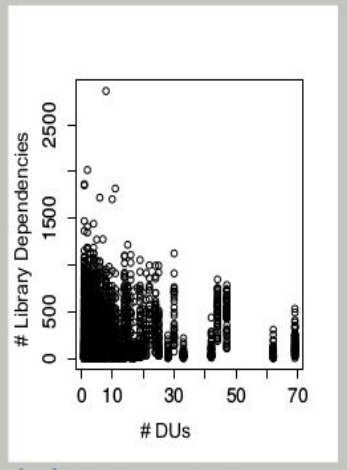
Library Migration in Practice

RQ1: To what extent are developers updating their library dependencies?

★ 4,659 GitHub Projects
★ 2,700 Java Maven libraries

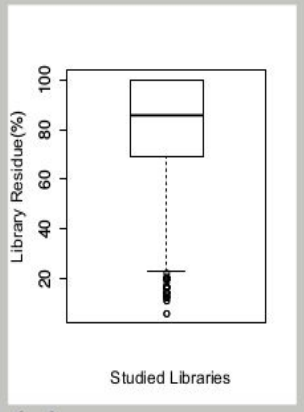


(a) Systems heavily rely on third-party dependencies

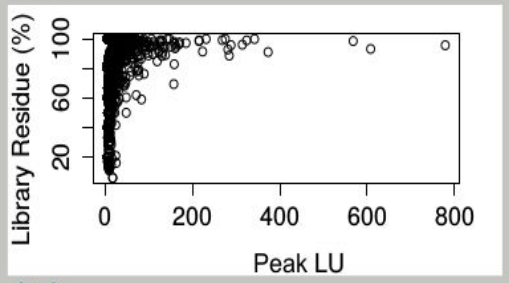


(b) There is little correlation between frequent updates and # dependencies

Figure: System Analysis



(a) Library migration is not common practice for older versions



(b) Systems are more likely to remain with older popular library versions

Figure: Library Analysis

Effectiveness of Awareness Mechanisms (1/2)

RQ2: To what extent are developers updating their library dependencies?

- ★ 3 new releases of popular libraries
- ★ 5 security vulnerabilities

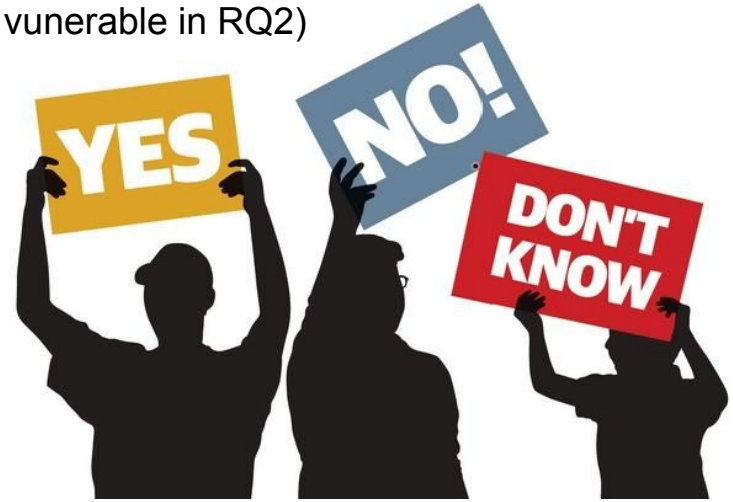
© https://www.singleblackmale.org/2010/02/17/spider-sense/

- ★ *"New release of a popular library (i) there exist patterns of consistent migration and patterns where an older popular library version is still preferred."*
- ★ *"For a security advisory disclosure we find cases of developer (ii) non responsiveness to security advisory disclosure, which is sometimes due to an incomplete patch or a latent security advisory."*

Effectiveness of Security Advisory (2/2)

RQ3: Why are developers non responsive to a security advisory?

- ★ 16 Projects contacted for feedback (detected as vunerable in RQ2)
- ★ Understand feedback

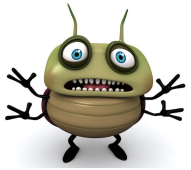


"69% of developers were unaware of their vulnerable dependencies and proceeded to immediately migrate to a safer dependency."

- ★ Developers evaluate based on project specific priorities.
- ★ Developers cite migration as a practice that requires extra migration effort and added responsibility.

The Complex Web of Updates & Library Ecosystems

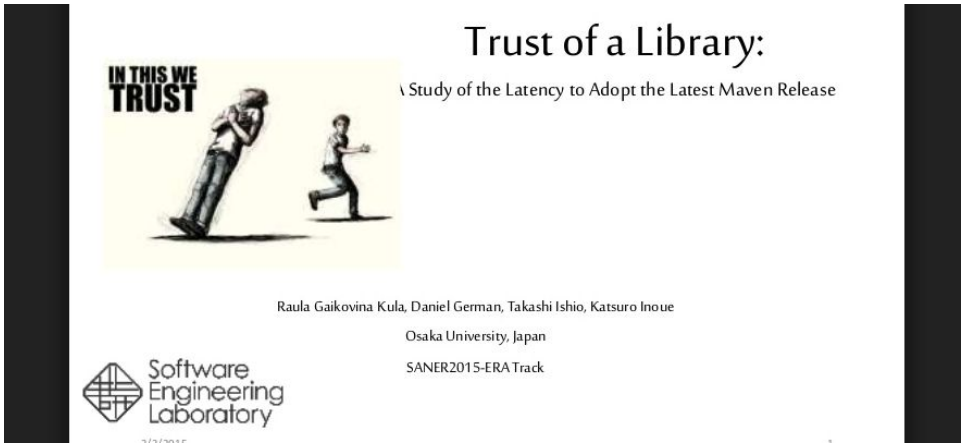
- We find that developers are wary of the latest versions

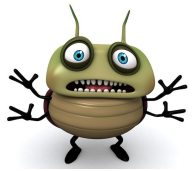


© https://www.boredpanda.com/old-superhero-paintings-eddie-liu/

The Complex Web of Updates & Library Ecosystems

- We find that developers are wary of the latest versions
- Leads to Dependency Hell with updates in the software ecosystem...





4. Future & Challenges

Perfect World: Real-time Updates, no breakages and the ecosystem is robust to handle any changes:

Reality: Work in Progress...

Challenges at the Software Ecosystem Level...

Library Recommendations and Visualizations (1\2)

Contents lists available at ScienceDirect

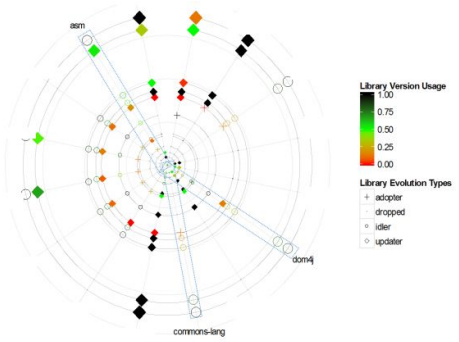
Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Search-based software library recommendation using multi-objective optimization

Ali Ouni[a,b,*], Raula Gaikovina Kula[a], Marouane Kessentini[c], Takashi Ishio[a], Daniel M. German[d], Katsuro Inoue[a]

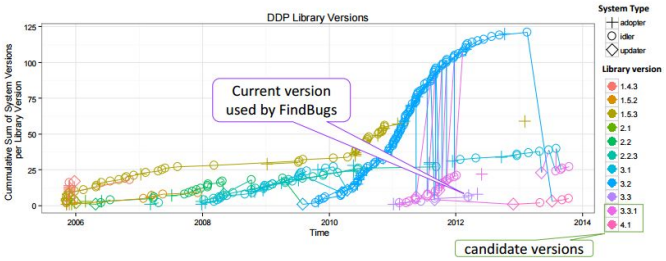
[a] Department of Computer Science, IST, Osaka University, Osaka, Japan
[b] Department of Computer Science and Software Engineering, UAE University, UAE
[c] Department of Computer and Information Science, University of Michigan, MI, USA
[d] Department of Computer Science, University of Victoria, Victoria, Canada



(a) SDP for FINDBUGS system

Visualizing the Evolution of Systems and their Library Dependencies

Raula Gaikovina Kula*, Coen De Roover*†, Daniel German*‡, Takashi Ishio*, Katsuro Inoue*
* Osaka University, Osaka, Japan † Vrije Universiteit Brussel, Brussels, Belgium
‡ University of Victoria, Canada
Email: {raula, coen, cderoove, ishio, inoue}@ist.osaka-u.ac.jp
dmg@uvic.ca



(b) LDP for ASM library

Library Recommendations and Visualizations (2\2)

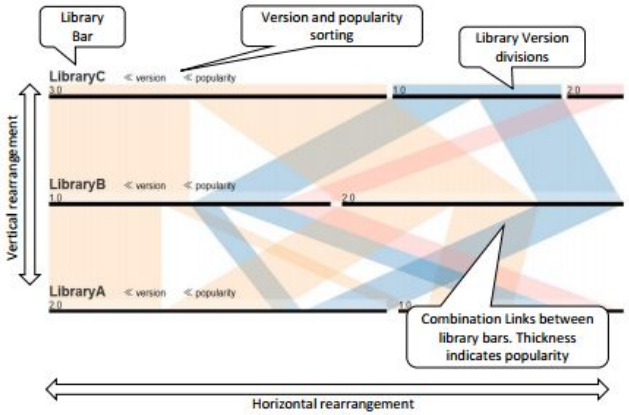
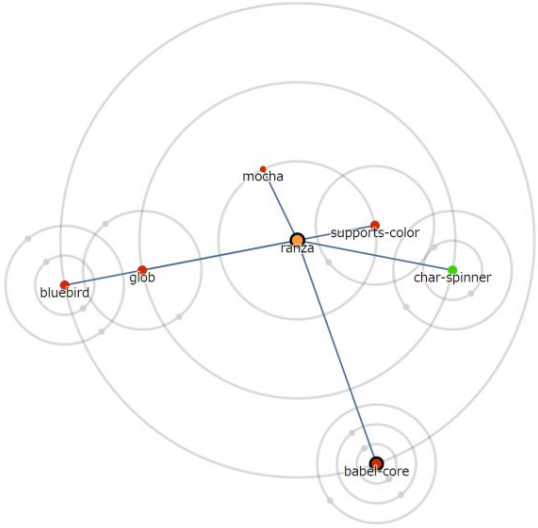


Fig. 2. VerXCombo - Parallel Sets Visualization

2015 IEEE 23rd International Conference on Program Comprehension

VerXCombo: An interactive data visualization of popular library version combinations

Yuki Yano, Raula Gaikovina Kula, Takashi Ishio, Katsuro Inoue
Osaka University, Japan
{y-yano, raula-k, ishio, inoue}@ist.osaka-u.ac.jp

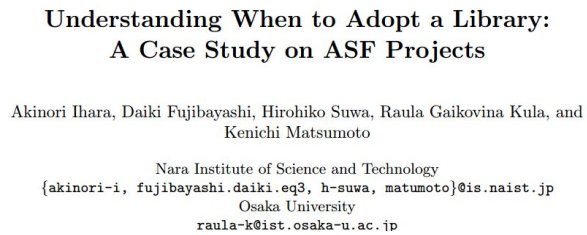


(a) *Ranza* overview, showing all 6 packages in use, 5 flagged as outdated (color and rotation) are candidates for an update.

SoL Mantra: Visualizing Update Opportunities Based on Library Coexistence

Boris Todorov*, Raula Gaikovina Kula†, Takashi Ishio†, Katsuro Inoue*
*Osaka University, Osaka, Japan
†Nara Institute of Science and Technology, Nara, Japan
Email: {boris-t, inoue}@ist.osaka-u.ac.jp, {raula-k, ishio}@is.naist.jp

Recent Ideas



Understanding When to Adopt a Library:
A Case Study on ASF Projects

Akinori Ihara, Daiki Fujibayashi, Hirohiko Suwa, Raula Gaikovina Kula, and Kenichi Matsumoto

Nara Institute of Science and Technology
{akinori-i, fujibayashi.daiki.eq3, h-suwa, matumoto}@is.naist.jp
Osaka University
raula-k@ist.osaka-u.ac.jp

- Studying Security Vulnerabilities in Other Ecosystems (npm)

https://snyk.io/

Recap and thank you...

- ★ **Background:** Third-party Libraries
  - Rise of Library Ecosystems
  - In Libraries we Trust
- ★ **Motivation :** What about a Library Update?
  - Library Updates in Practice
  - Awareness Mechanisms
- ★ **Main Study:** A Study on Library Updates
  - Library Updates in Practice
  - Awareness Mechanisms
- ★ **Future and Challenges:** Library Ecosystems in the Future
  - Ongoing work
  - Role of Software Analytics
  - Systematic Modelling