

MigrationAdvisor: Recommending Library Migrations from Large-Scale Open-Source Data

Hao He^{*^}, Yulin Xu^{*}, Xiao Cheng[^], Guangtai Liang[^], and Minghui Zhou^{*}

^{*}Department of Computer Science, Peking University

[^]Software Analysis Lab, Huawei Technologies Co., Ltd.

Presenter: Hao He (何昊), 1st year PhD candidate at Peking University

Advisor: Minghui Zhou (周明辉), Professor at Peking University

Problem Context – Fragility of Dependencies

- Dependency on a third-party library is **fragile** because both the project and the library are continuously and independently evolving

Table 2: Why open source projects fail?

Reasons	Group	Projects
Usurped by competitor	Environment	27 ■
Obsolete	Project	20 ■
Lack of time	Team	18 ■
Lack of interest	Team	18 ■
Outdated technologies	Project	14 ■
Low maintainability	Project	7
Conflicts among developers	Team	3
Legal problems	Environment	2
Acquisition	Environment	1

Sustainability failures
(Coelho and Valente, 2017)



Restrictive licenses

Scalability?

Performance?

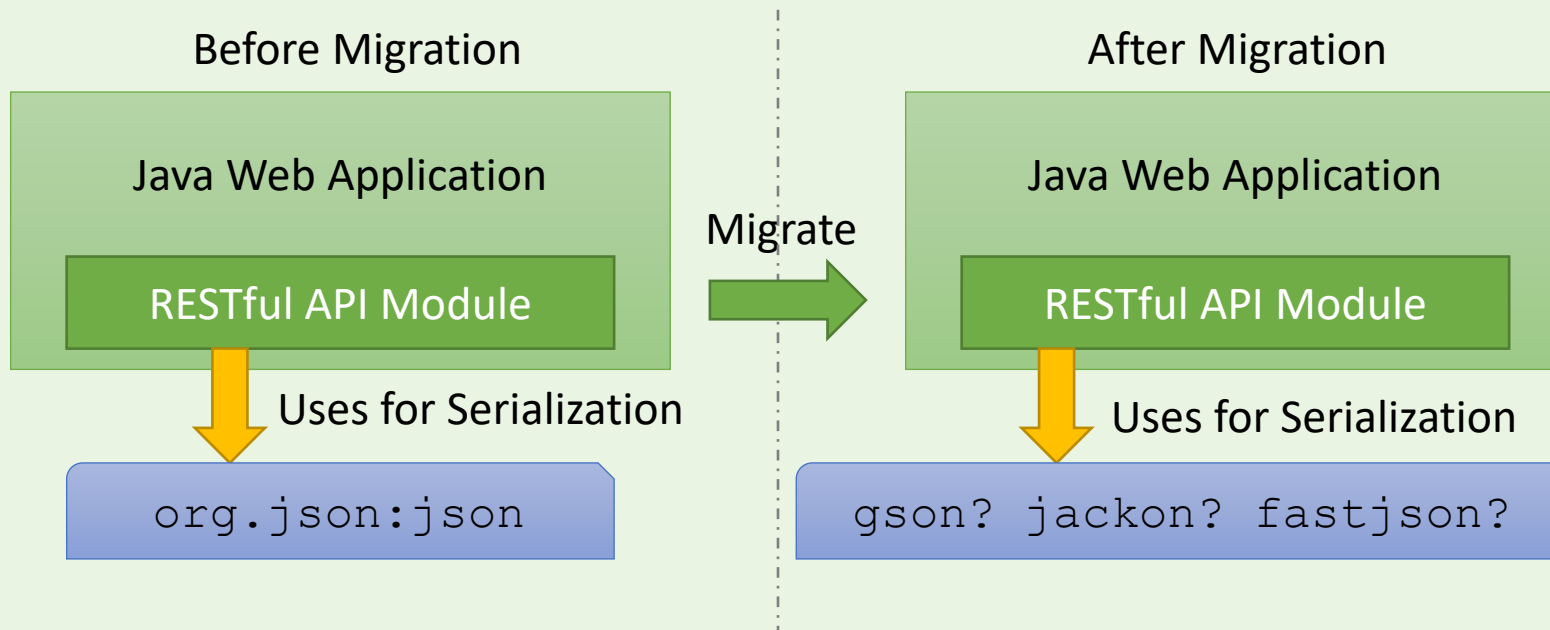
Security?

New Feature?

Unsatisfiable new requirements

Problem Context – Library Migration

- In these situations, a project has to migrate an **already in-use library** to **another similar or functionally equivalent library**



Problems for Maintainers

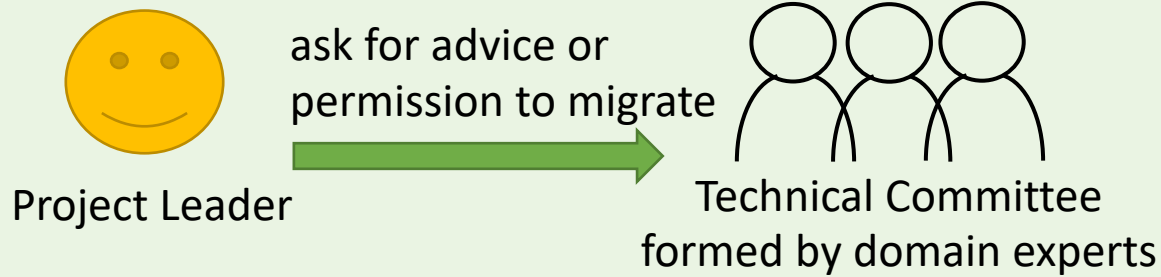
Can we migrate?

Shall we migrate?

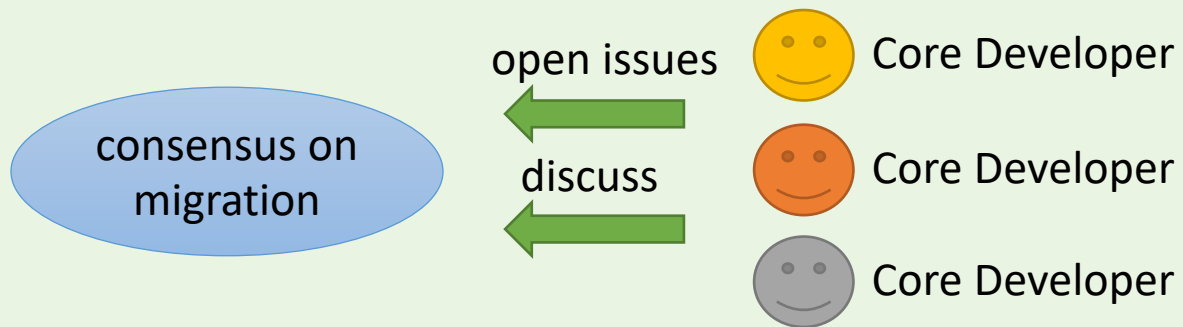
Which to migrate?

Problem Context – Current Situation

Industry projects (in a large IT company)







Typical open-source projects (Kabinna et al.)



Make decision from personal experience and opinion, but is the decision optimal?

1. Huang et al, ASE 2018
2. Chen et al, ASE 2016
3. Teyton et al, WCRE 2012
4. Alrubaye et al, ICSME 2019

Existing Tools and Services

Tool/Service	Disadvantage in the context of library migration
 Maven TM	Has library categories but not well maintained
 stackoverflow	Do not allow opinion-based discussions
 reddit Personal Blogs  alternativeTo <small>Crowdsourced software recommendations</small> DiffTech [1]	Community discussions/curated lists about competitor libraries are opinion-based, inherently controversial and may not be trust-worthy
SimilarTech [2]	No evidence on feasibility of a library migration
Teyton et al. [3] MigrationMiner [4]	Evidence-supported by mining existing migrations in GitHub, but use ineffective filtering-based algorithm and not scalable

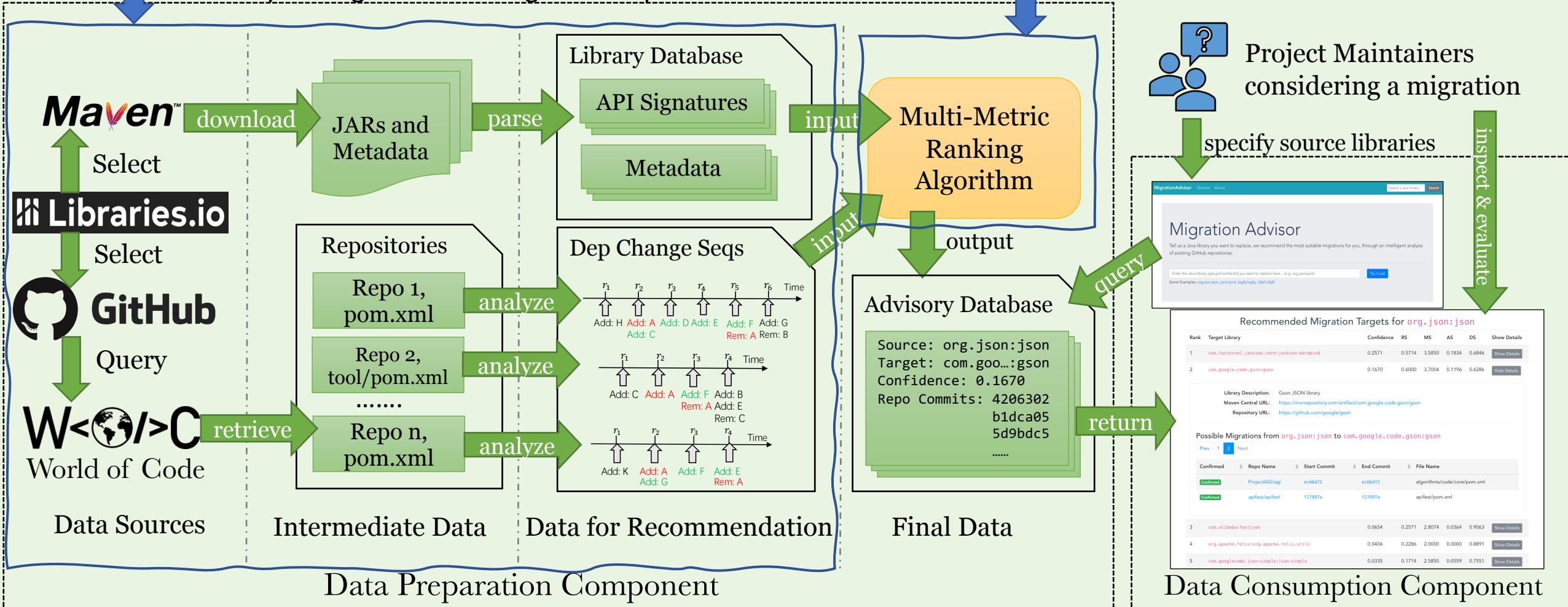
Goals of MigrationAdvisor

- **Input:** A library l_1 that a user specify
- **Output:** The list of libraries that the user can migrate from l_1 , and existing migrations for each returned library
- **Key Requirements**
 - **Evidence-Supported:** Generate results with existing migrations from GitHub repositories
 - **Accurate:** High precision and recall in the returned results
 - **Scalable:** Works on a large number of repositories and libraries
- **Our Core Contributions**
 - **Accurate:** a novel multi-metric ranking algorithm
 - **Scalable:** a carefully designed architecture and data processing workflow

MigrationAdvisor Workflow

Scalability through careful design and implementation

High accuracy through a multi-dimensional mining and analysis of libraries and repositories



MigrationAdvisor Data

Data	Count/Size	Time to Retrieve or Construct
Repositories	21,358	Several Minutes
Commits with Diffs	29,439,998	~1 Day
Parsed pom.xml Blobs	10,009,552	~1 Day
Dependency Change Sequences	147,220	Several Hours
Maven Artifacts (i.e. Java Libraries)	184,817	Several Minutes
Maven Artifact Versions	4,045,748	Several Hours
Different Classes in Maven	25,272,024	~3 Days
Non-zero API Count Pairs	4,934,677	~2 weeks
Migration Advisories	1,956,809	Several Hours
Confirmed Migration Advisories	14,334	~1 Week (manually)
Total Database Size (compressed dump)	~50GB	~3 Weeks

Tool Demonstration

- <http://migration-helper.net/>

MigrationAdvisor

Dataset

About

Search a Java library...

Search

Migration Advisor

Tell us a Java library you want to replace, we recommend the most suitable migrations for you, through an intelligent analysis of existing GitHub repositories.

Enter the Java library (groupId:artifactId) you want to replace here... (e.g. org.json:json)

Try it out!

Some Examples: [org.json:json](#), [junit:junit](#), [log4j:log4j](#), [c3p0:c3p0](#)

December 3, 2021

ICSE 2021 Demonstration

MigrationAdvisor

Dataset

About

Search a Java library...

Search

Basic Information of Source Library **org.json:json**

Library Description:

JSON is a light-weight, language independent, data interchange format. See <http://www.JSON.org/> The files in this package implement JSON encoders/decoders in Java. It also includes the capability to convert between JSON and XML, HTTP headers, Cookies, and CDL. This is a reference implementation. There is a large number of JSON packages in Java. Perhaps someday the Java community will standardize on one. Until then, choose carefully. The license includes this restriction: "The software shall be used for good, not evil." If your conscience cannot live with that, then choose a different package. The package compiles on Java 1.2 thru Java 1.4.

Maven Central URL:

<https://mvnrepository.com/artifact/org.json/json>

Repository URL:

<https://github.com/douglasrockford/JSON-java>

Recommended Migration Targets for **org.json:json**

Rank	Target Library	Confidence	RS	MS	AS	DS	Show Details
1	com.fasterxml.jackson.core:jackson-databind	0.2571	0.5714	3.5850	0.1834	0.6846	Show Details
2	com.google.code.gson:gson	0.1670	0.6000	3.7004	0.1196	0.6286	Hide Details

Library Description:

Gson JSON library

Maven Central URL:

<https://mvnrepository.com/artifact/com.google.code.gson/gson>

Repository URL:

<https://github.com/google/gson>

Possible Migrations from **org.json:json** to **com.google.code.gson:gson**

Prev

1

2

Next

Confirmed	Repo Name	Start Commit	End Commit	File Name
Confirmed	ProjectAGI/agi	ec66d72	ec66d72	algorithms/code/core/pom.xml
Confirmed	apifest/apifest	f27897a	f27897a	apifest/pom.xml

3	com.alibaba:fastjson	0.0654	0.2571	2.8074	0.0364	0.9063	Show Details
4	org.apache.felix:org.apache.felix.utils	0.0406	0.2286	2.0000	0.0000	0.8891	Show Details
5	com.googlecode.json-simple:json-simple	0.0335	0.1714	2.5850	0.0559	0.7551	Show Details
6	org.apache.johnzon:johnzon-core	0.0332	0.1429	2.3219	0.0000	1.0000	Show Details
7	com.fasterxml.jackson.core:jackson-core	0.0311	0.2571	2.0000	0.0078	0.6047	Show Details
8	org.springframework.boot:spring-boot-starter-json	0.0265	0.1143	2.3219	0.0000	1.0000	Show Details

Thank You!

MigrationAdvisor: Library Migration Recommendations
from Large-Scale Open Source Data

Tool Demo: <http://migration-helper.net/>

GitHub Repositories: <https://github.com/hehao98/MigrationHelper>
<https://github.com/hehao98/MigrationHelperFrontend>

Paper Preprint: <https://hehao98.github.io/files/2021-migration.pdf>

Contact: Hao He, heh@pku.edu.cn
Yulin Xu, kylinxyl@pku.edu.cn
Minghui Zhou, zhmh@pku.edu.cn

