ON THE USE OF DATABASE FRAMEWORKS IN JAVA PROJECTS

A LARGE-SCALE HISTORICAL EMPIRICAL ANALYSIS

Mathieu Goeminne
mathieu.goeminne@umons.ac.be

Tom Mens
tom.mens@umons.ac.be

Alexandre Decan
alexandre.decan@umons.ac.be

Software Engineering Lab
GENERAL CONTEXT

- FNRS interuniversity research collaboration with the University of Namur.
- Study of software systems that use a database.
- Understanding the dynamics, issues and evolution of such systems.
- Need to take into account their specificities:
  - How is their connection managed?
  - How do the source code and the database co-evolve?
  - How can we assess and increase the quality of such systems?
DATABASE FRAMEWORKS?

Source code

- .java
- .jsp
- .php

Frameworks

- ORM
- Connector
- Session manager

Database

DBMS
ON THIS PRESENTATION

- Focus on database frameworks.

- Empirical study for answering 3 research questions:
  - **Q1:** Which combinations of DB frameworks are being used simultaneously?
  - **Q2:** How long do DB frameworks ‘survive’ in the projects that use them?
  - **Q3:** Does the introduction of a DB framework influence the survivability of another one?
CASE STUDY

- 20 Java DB frameworks.
- Top 5 retained: JDBC, Spring, JPA, Vaadin, Hibernate.
- 13,307 Java projects from the Github Java Corpus (MSR10): 3,707 of them use at least one retained framework.
SURVIVAL ANALYSIS: KAPLAN-MEIER ESTIMATOR

- Estimates the **survival function** of a population

- Takes into account **right censoring** (the event may occur after the last observation)

- **Mantel-Haenszel test** for determining if 2 survival functions differ.
Q1 - Which combinations of DB frameworks are being used simultaneously?

- In most cases (98%), different frameworks used in a project are used in combination.
- JDBC is used as the single DB framework in 56% of all projects.
- Most projects (83%) using Hibernate also use JDBC. Hibernate is almost never (3%) used alone.
- JDBC used as support for the other FW, providing unique services?

The table below shows the number of projects using different combinations of frameworks.

<table>
<thead>
<tr>
<th># simultaneous fw.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>total # frameworks used</td>
<td>2,443</td>
<td>22</td>
<td>776</td>
<td>328</td>
<td>104</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

# projects using X fw vs Y fw simultaneously

<table>
<thead>
<tr>
<th></th>
<th>Spring</th>
<th>JPA</th>
<th>Vaadin</th>
<th>Hibernate</th>
</tr>
</thead>
<tbody>
<tr>
<td>JDBC</td>
<td>645</td>
<td>565</td>
<td>143</td>
<td>192</td>
</tr>
<tr>
<td>Spring</td>
<td>558</td>
<td>76</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>JPA</td>
<td>98</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaadin</td>
<td></td>
<td>22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# projects using X and Y simultaneously
Q2 - How long do database frameworks ‘survive’ in the projects?

- If a framework is introduced, it remains in >45% of all the projects.

- Different trends. After 30 days:
  - Hibernate disappears from 11.7% of projects
  - Spring disappears from 3.7% of projects

- Partial order: JPA and Spring have higher survival rates than JDBC and Hibernate.

- No clear relation between the order and the popularity of the considered frameworks.
Q3 - Does the introduction of a database framework influence the survivability of another one?

- Visually, the introduction of some frameworks seems to have a favorable influence.

- After Bonferroni correction, no significant difference in the survival functions… at project level.

- In the future: file level analyse.
CONCLUSIONS

• **Important coexistence**

  • Especially between JDBC and the other considered DB frameworks.

• **JDBC is still often present in projects** despite the use frameworks providing more advanced, higher level services.

• In our study, there is **no evidence for competition** between frameworks.

  • An introduced framework has **an ≳45% of chance that it will remain**.

  • The disappearance of a framework doesn’t seem to be related to the introduction of an other one.
FUTURE WORK

• **Finer granularity:** from projects to packages, files, methods, …

• Analysis of (un)**frequently used framework services.**

• Looking for evolutionary patterns:
  
  • New frameworks introduced for their unique services?
  
  • Simple SQL queries are replaced by framework services, while complex queries remain as it?

• Extension to frameworks for **less structured databases**, NoSQL, etc.