A Comparison of Closed-Source and Open-Source Code Static Measures
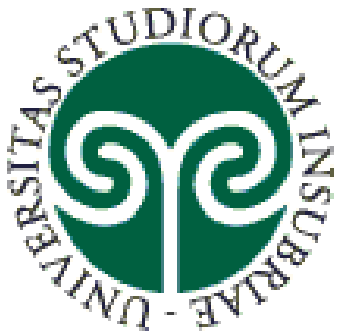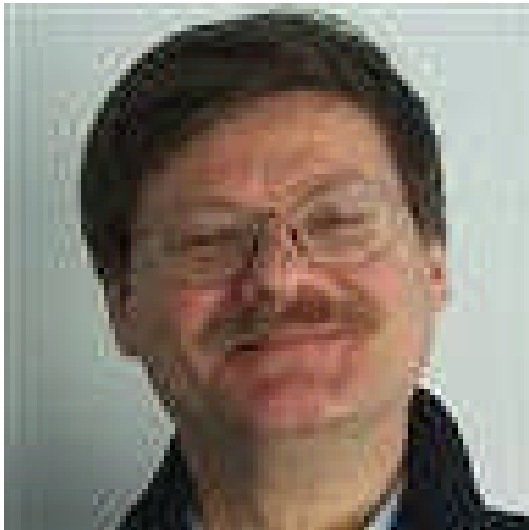
Luigi Lavazza

Università degli Studi dell'Insubria, Varese, Italy

Luigi Lavazza

Professional experience

- Professor of Computer Science at the University of Insubria at Varese, Italy.
- Scientific consultant in digital innovation projects at CEFRIEL – Politecnico di Milano.

Scientific Activity

- Research: Empirical software engineering, software metrics and software quality evaluation; project management and effort estimation; Software process modeling, measurement and improvement; Open Source Software.
- Several international research projects
- Reviewer of EU funded projects.
- Co-author of over 180 scientific articles.
- PC member of several international Software Engineering conferences
- Editor in chief of the IARIA International Journal On Advances in Software (2013-2018).
- IARIA fellow since 2011

Motivation and goals

- Situation:
  - Most empirical research in software engineering uses open-source project data
    - Because obtaining closed-source project data is very hard
- Question:
  - Do the results obtained by analysing OSS apply to CSS as well?
- Objective:
  - Let's check if measures from OSS and CSS are similar.
- The objective was addressed via an empirical study

The datasets

- Static measures of Java code were collected from
  - 3 industrial CSS projects
    - Ind1 and Ind2 are client and contract management systems from a large service company,
    - Ind3 is the back-end of a web application
  - 5 widely used OSS projects
    - Log4J, Jcaptcha, PdfBox, JasperReports, Hibernate
- Measures were collected via SourceMeter
  - https://www.sourcemeter.com/

The metrics

- Due to time and space limitations, this initial work considers only a set of the most widely used method-level measures.

Metric name	Abbreviation
Halstead Calculated Program Length	HCPL
Halstead Volume	HVOL
Maintainability Index (Original version)	MI
McCabe's Cyclomatic Complexity	McCC
Lines of Code	LOC

- Methods having McCC=1 (e.g., getters and setters) were excluded.

Statistic tests used

- Kruskal-Wallis test shows if the projects are all equivalent with respect to the considered metric

- When such equivalence does not hold, Wilcoxon rank sum tests were used to check difference among project pairs.

- When differences where found, Hedges g was used to evaluate the effect size.
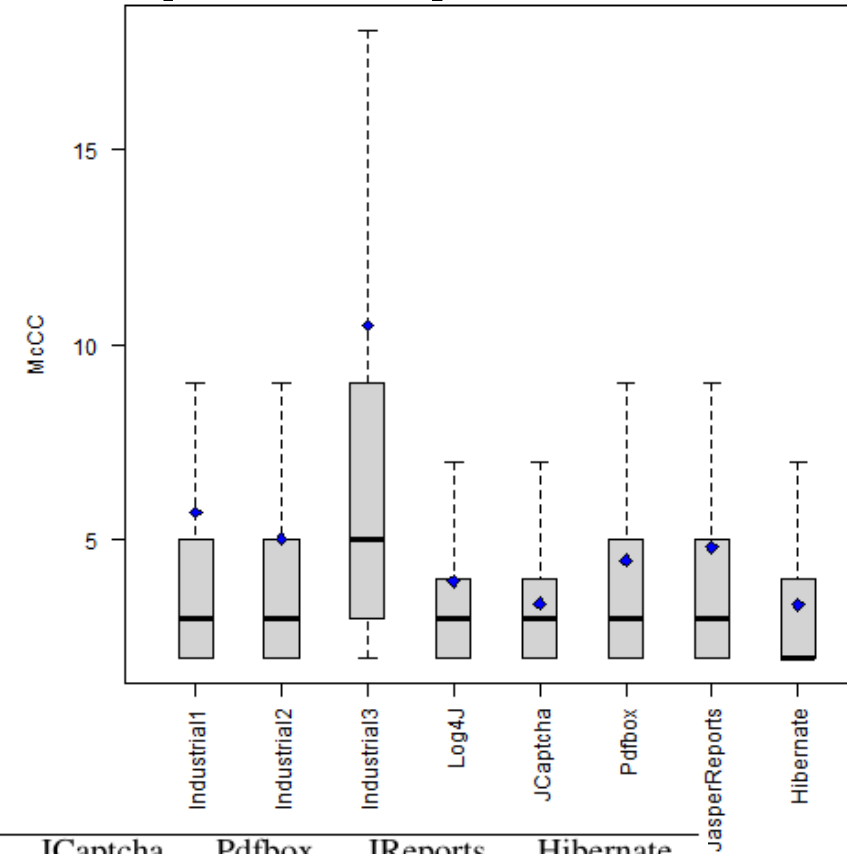
Results for LOC
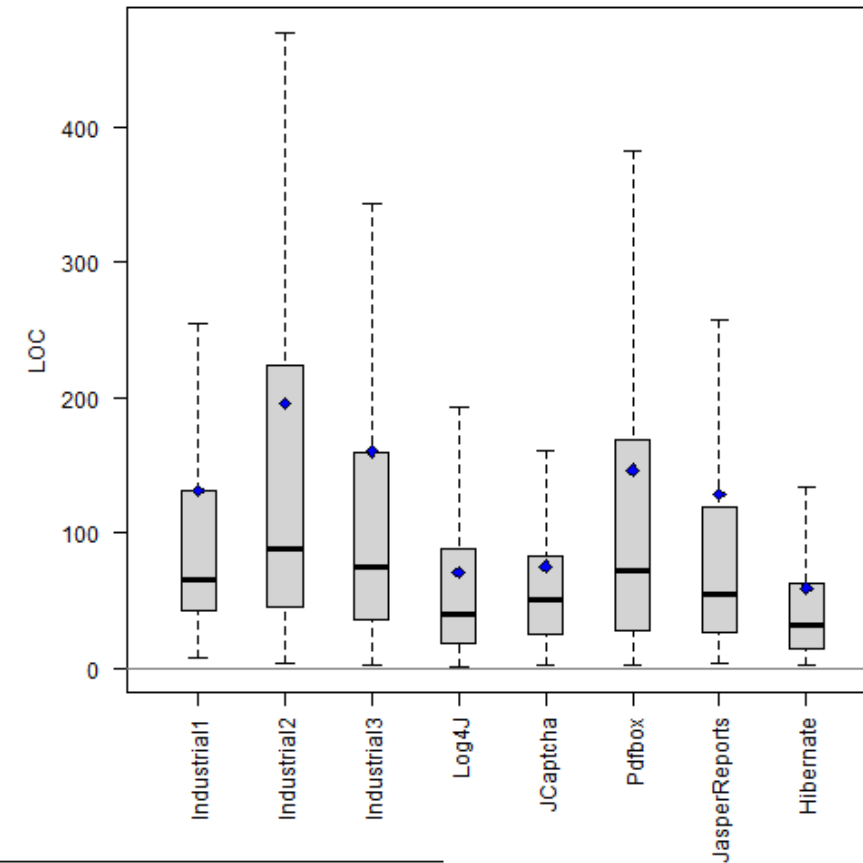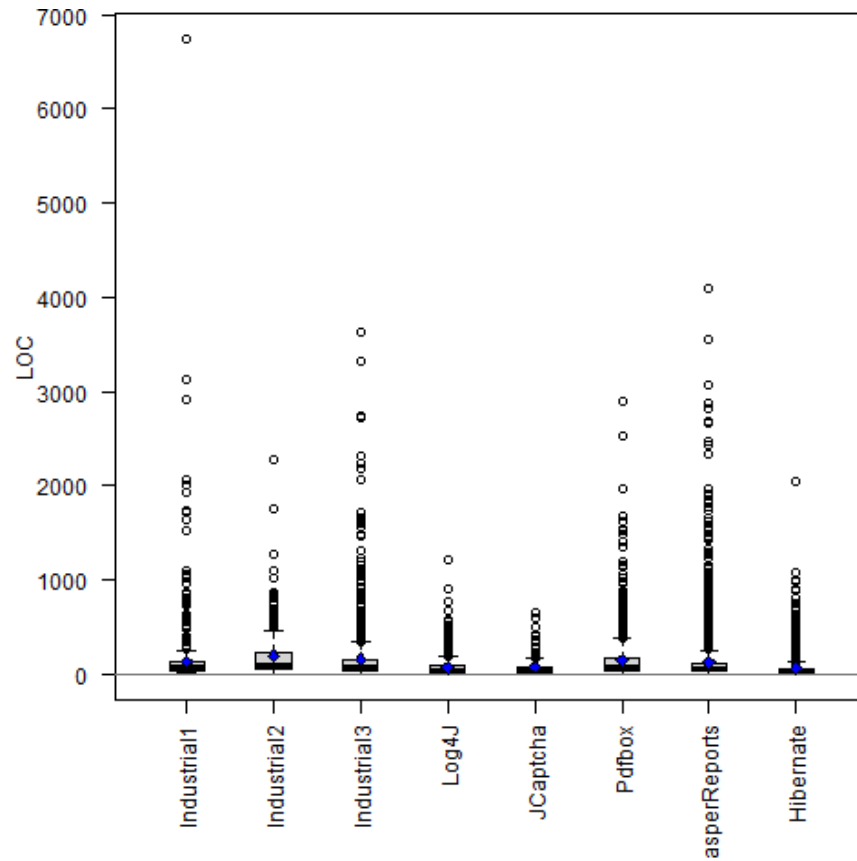
	Ind1	Ind2	Ind3	Log4J	JCaptcha	Pdfbox	JReports	Hibernate
Ind1	–	n	s	s	n	n	n	s
Ind2	n	–	s	s	s	=	n	s
Ind3	s	s	–	s	s	s	s	s
Log4J	s	s	s	–	n	s	s	n
JCaptcha	n	s	s	n	–	s	s	n
Pdfbox	n	=	s	s	s	–	n	s
JReports	n	n	s	s	s	n	–	s
Hibernate	s	s	s	n	n	s	s	–

Results for McCabe complexity
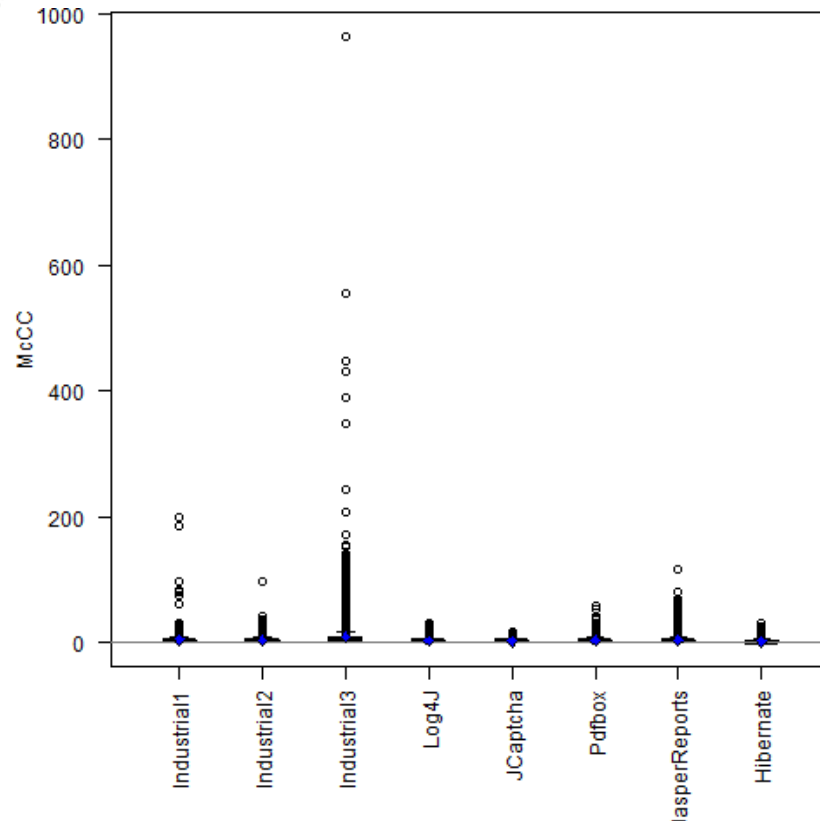
	Ind1	Ind2	Ind3	Log4J	JCaptcha	Pdfbox	JReports	Hibernate
Indl1	–	n	n	s	s	n	=	s
Ind2	n	–	s	s	s	=	n	s
Ind3	n	s	–	s	s	s	s	s
Log4J	s	s	s	–	n	n	n	s
JCaptcha	s	s	s	n	–	s	s	n
Pdfbox	n	=	s	n	s	–	n	s
JReports	=	n	s	n	s	n	–	s
Hibernate	s	s	s	s	n	s	s	–

Results for Halstead Volume
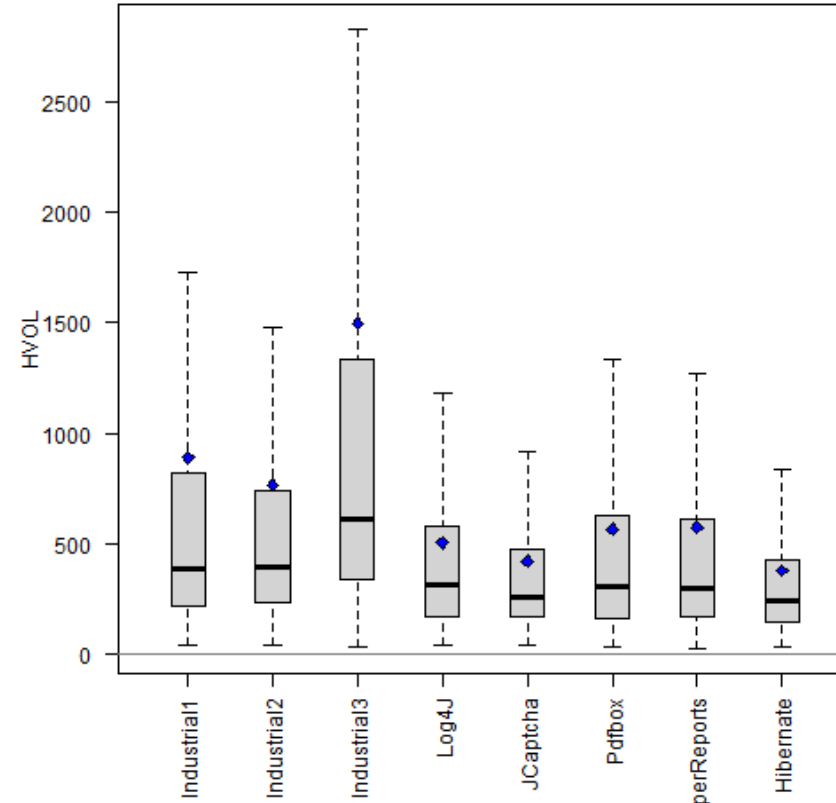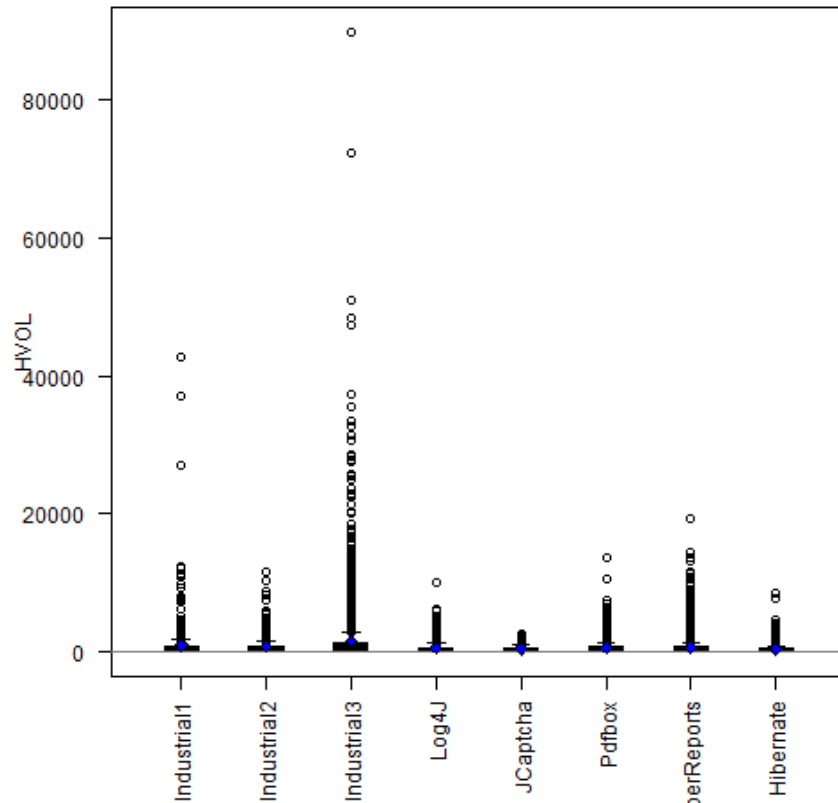
	Ind1	Ind2	Ind3	Log4J	JCaptcha	Pdfbox	JReports	Hibernate
Ind1	–	=	n	s	s	s	s	s
Ind2	=	–	s	s	s	s	s	m
Ind3	n	s	–	s	s	s	s	s
Log4J	s	s	s	–	n	n	=	s
JCaptcha	s	s	s	n	–	n	n	n
Pdfbox	s	s	s	n	n	–	n	s
JasperReports	s	s	s	=	n	n	–	s
Hibernate	s	m	s	s	n	s	s	–

Results for Halstead Computed Program Length
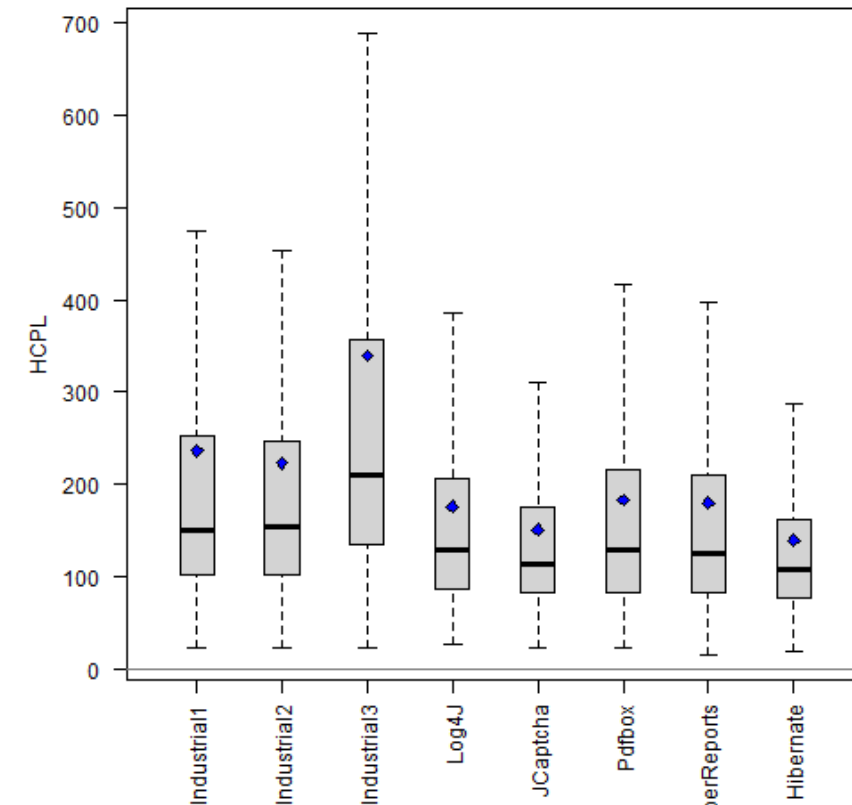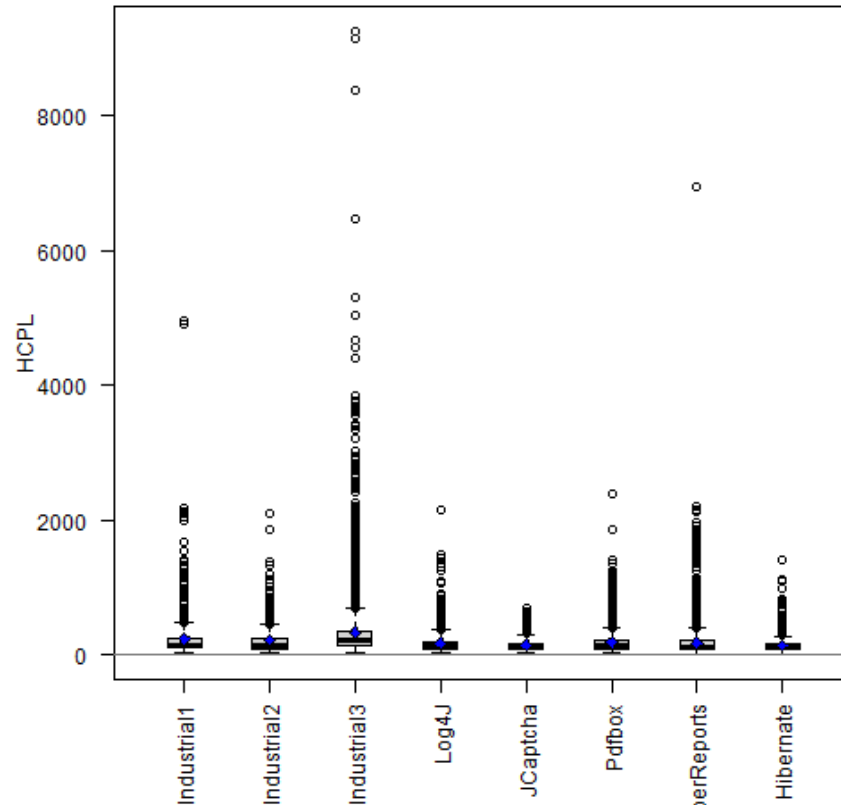
	Ind1	Ind2	Ind3	Log4J	JCaptcha	Pdfbox	JReports	Hibernate
Ind1	−	=	s	s	s	s	s	m
Ind2	=	−	s	s	s	s	s	m
Ind3	s	s	−	s	s	s	s	m
Log4J	s	s	s	−	n	=	n	s
JCaptcha	s	s	s	n	−	n	n	n
Pdfbox	s	s	s	=	n	−	n	s
JasperReports	s	s	s	n	n	n	−	s
Hibernate	m	m	m	s	n	s	s	−

Results for Maintainability Index
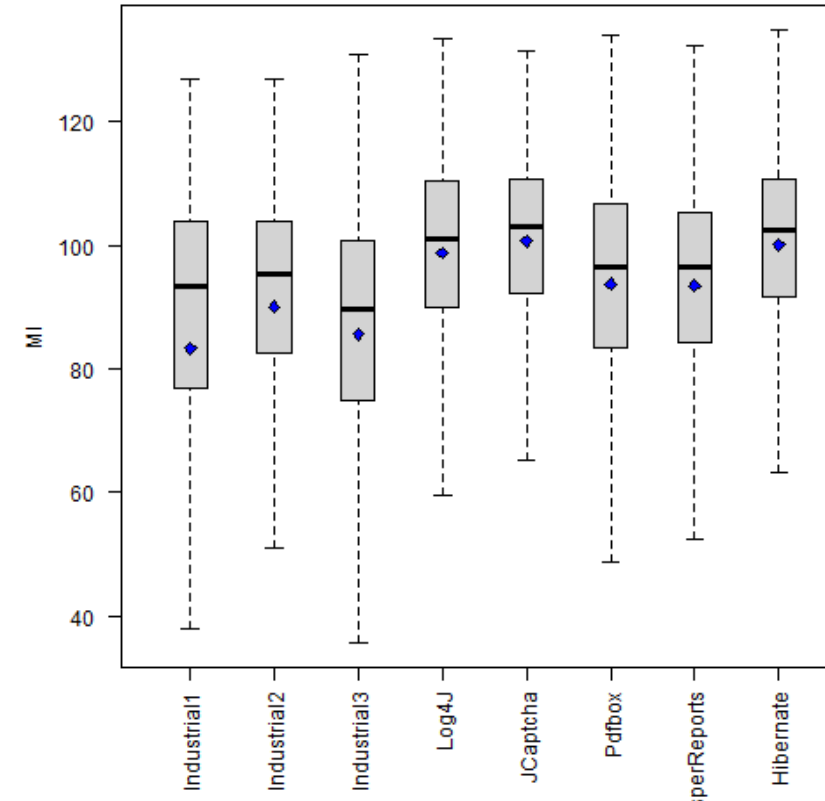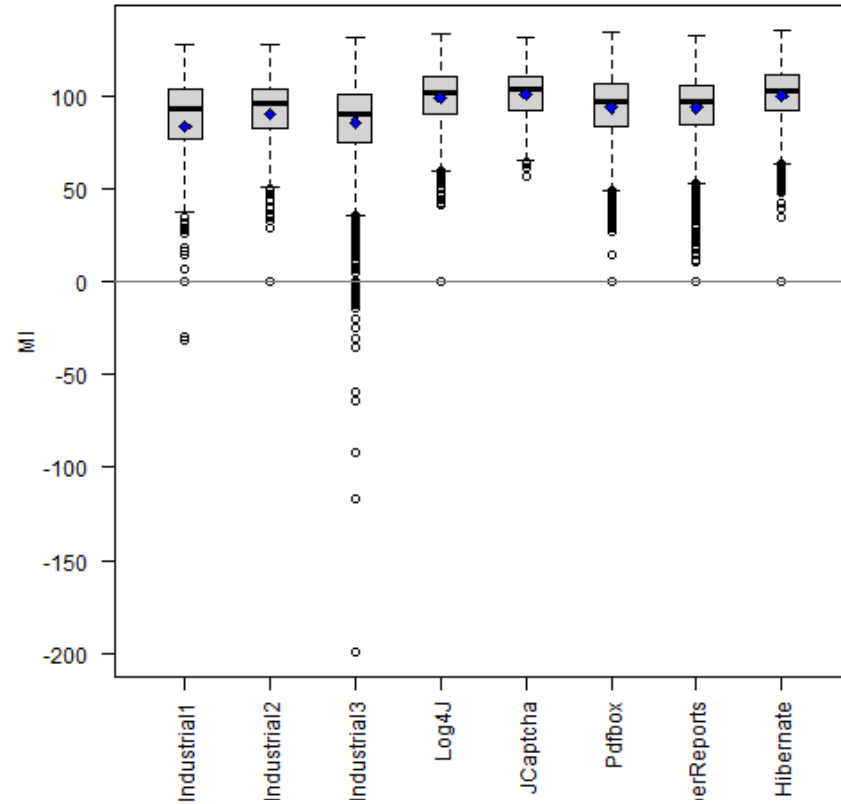
	Ind1	Ind2	Ind3	Log4J	JCaptcha	Pdfbox	JReports	Hibernate
Ind1	–	s	n	m	m	s	s	m
Ind2	s	–	n	s	m	n	n	m
Ind3	n	n	–	m	m	s	s	m
Log4J	m	s	m	–	n	s	s	n
JCaptcha	m	m	m	n	–	s	s	=
Pdfbox	s	n	s	s	s	–	n	s
JasperReports	s	n	s	s	s	n	–	s
Hibernate	m	m	m	n	=	s	s	–

Conclusions

- Our analysis shows that CSS projects appear sufficiently similar to OSS ones
- At least, the study did not highlight any difference that could disrupt the hypothesis that studies that analysed OSS data yield results that are applicable to CSS as well.
- However, there are a couple of important limitations:
  - Only method-level metrics were considered
  - Only a few projects were involved in the study