

# Manutenção e Evolução de Software

## Proposta de projeto

Pedro Araujo Pires

### Objetivo

O objetivo deste projeto é a criação de um *dataset* com informações sobre a evolução de dez sistemas *open source*. Essas informações serão coletadas das seguintes fontes:

- Sistemas de controle de versão (Git e SVN)
- Sistemas de *issue tracking* (Jira e Bugzilla)
- Métricas dos códigos fonte.

### Metodologia

Para construir um modelo de evolução de cada sistema, serão necessárias as seguintes informações: (1) informações dos sistemas *issue tracking*; (2) informações sobre versões do código fonte, retiradas dos sistemas de controle de versão; e (3) informações sobre defeitos nas classes.

Primeiramente será construído um modelo de como o sistema foi modificado ao longo do tempo. Para isso será feito o parsing das informações do sistema de *issue tracking*, que consistem em quais arquivos foram modificados, quando essas modificações ocorreram, e uma mensagem do desenvolvedor que fez a alteração. Depois será construído um modelo do código fonte, de acordo com a linguagem FAMIX. Serão construídas representações de várias versões do código fonte, obtidas a cada duas semanas desde o início do sistema, totalizando 180 versões. Em seguida será feita uma *linkagem* entre os bugs reportados pelas ferramenta de *issue tracking* e os *commits* obtidos do sistema de controle de versão, para determinar quais classes foram afetadas por cada *bug*. Para fazer isso, será buscada nas mensagens dos *commits* o ID do bug no sistema de *issue tracking*, visto que é uma prática relativamente comum entre desenvolvedores inserir o ID do bug na mensagem do *commit*. Por último a plataforma Moose será utilizada para calcular 17 métricas (CK + 11) para cada versão dos sistemas, tornando os modelos mais completos.

### Sistemas

Os sistemas que serão analisados são:

- Eclipse JDT Core – ferramenta para gerar a AST de um programa Java.
- Eclipse PDE UI – plugin para o Eclipse para desenvolvimento de plugins.
- Lucene – ferramenta para busca textual,
- Equinox Framework – implementação do framework OSGi
- Geronimo – servidor para aplicativos Java/OSGi
- Hadoop MapReduce – biblioteca para computação distribuída
- TV-Browser – guia de TV online
- Hibernate – um *Object Relational Mapper*
- (Sistema a definir)
- (Sistema a definir)

## **Conexão com a literatura**

Este trabalho é baseado no artigo “An Extensive Comparison of Bug Prediction Approaches”, de Marco D'Ambros, Michele Lanza e Romain Robbes, onde os autores criaram um dataset com cinco sistemas (Eclipse JDT Core, Eclipse PDE UI, Equinox Framework, Lucene, Mylyn) de forma similar da descrita acima. O número de versões desses sistemas será estendido (de aproximadamente 100 para 180), e novos sistemas serão incluídos.