

# KnowDIME: An Infrastructure based on Architecture-Driven Modernization for Improving Legacy System

Rafael S. Durelli<sup>1,3</sup>, Márcio E. Delamaro<sup>1</sup> and Valter V. de Camargo<sup>2</sup>

<sup>1</sup>Computer Systems Department University of São Paulo  
São Carlos, SP, Brazil.

<sup>2</sup>Computing Departament  
Federal University of São Carlos (UFSCAR)  
São Carlos, SP, Brazil.

<sup>3</sup>RMoD Team, INRIA, Lille, France

{rdurelli, delamaro}@icmc.usp.br<sup>1</sup>, valter@dc.ufscar.br<sup>2</sup>

***Abstract.***

## 1. Introduction

Model-driven software engineering (MDSE) is becoming the most promising paradigms in software engineering. MDSE promotes the use of models as key artifacts in all phases of the software development lifecycle. Out of the many approaches to model-driven engineering, the Architecture-Driven Modernization (ADM) which has been proposed by OMG (Object Management Group) and advocates conducting refactoring following the principles of MDA (Model-Driven Architecture) [?] has become a great candidate since it aims to promote (i) portability, (ii) interoperability and (iii) reusability. According to the OMG the most important artifact provided by ADM is the KDM metamodel, which is a vertical and horizontal standard metamodel that represents all elements of the existing information technology architectures.

On the other hand, refactoring has been known and highly used both industrially and academically. It is a form of transformation that was initially defined by Opdyke [ref] in the discipline of Object Oriented (OO) Programming as “a change made to the internal structure of the software while preserving its external behavior at the same level of abstraction”. Nowadays it is possible to identify several catalogs of refactoring for different languages. The most complete and influential was published by Fowler in [6] for refactoring of Java code. It is evident that refactorings are useful to improve the quality of source code, and thus, to increase its maintainability.

However, while software reengineers would greatly benefit from the possibility to assess different choices, in practice they mostly rely on experience or intuition because of the lack of approaches providing comparison between possible variations of a change. Therefore, software reengineers do not have the possibility to easily apply analyses on different source-code version branches of a system and compare them to pick up the most adequate changes. In this context, the motivations for moving from software refactoring to model refactoring are: (i) a model provides an abstract view of the system, hence, visualizations of the structural changes required are easier, (ii) problems uncovered at the design-level can be improved directly on the model and (iii) using refactoring in high

abstract level can allow the software engineer to explore alternate design paths in much cheaper than software refactoring.

To overcome the described problems, we devised a *plug-in* on the top of the Eclipse Platform named KDM-RE (**K**nowledge **D**iscovery **M**odel-**R**efactoring **E**nvironment). Firstly, this *plug-in* provides a model-driven refactoring based on KDM models. Secondly, it also supplies a multiple versions of a system at level models (KDM), enabling to the engineer to work interactively on multiple models and to explore alternate refactoring path. Finally, after the best alternate refactoring path is chosen by the engineer, a forward engineering is performed and the source code of the refactored target system is generated again. Notice that KDM-RE supports 16 refactorings that are heavily inspired by the refactorings given by Fowler [?]. These refactorings are split up into four groups. The first group named **Moving Features Between Objects** which consists of relatively simple refactorings such as moving and creating features. The second group called **Organizing Data** is a set of refactorings to be applied in order to make working with data easier. The third group is named **Dealing With Generalization** and it represents refactorings to mostly dealing with moving methods and fields around a hierarchy of inheritance. This paper is organized as followed: Section 2 provides information related to KDM-RE - Section 3 the architecture of KDM-RE is depicted - in Section 4 there are related works and in Section 5 we conclude the paper with some remarks and future directions.

## 2. KDM-RE

In this section the **K**nowledge **D**iscovery **M**odel-**R**efactoring **E**nvironment (KDM-RE) is presented. In Figure 1 we depict the overall process of our technique which was adapted based on the horseshoe modernization model. It is split into three parts, they are: (i) Reverse Engineering, (ii) Refactorings and (iii) Forward Engineering. Furthermore, our process is divided into six levels. Firstly, the user starts the process in the **Level-0** by choosing an eclipse project which contains the source-code to realize the refactorings. After that, into *Level-1* it is need to transform the source-code into a Platform-Specific Model (PSM). This PSM is an instance of the source-code which represents all abstraction of the source-code, in Figure ?? (a) illustrates how to get the instance of this PSM. More specifically, this model represents an Abstract Syntax Tree (AST) that represents the syntactic struct of Java source code.

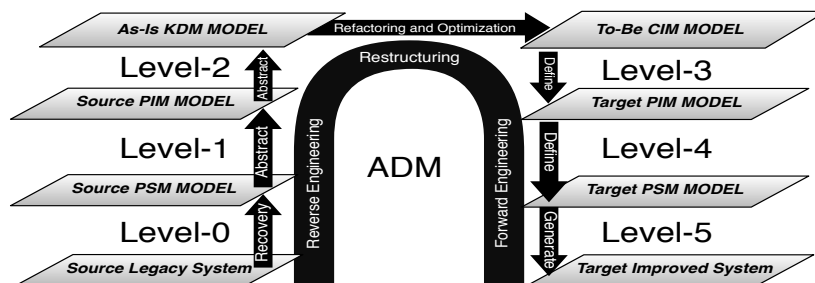


Figure 1. KDM-RE Process

As long as the user has triggered the earlier step, the next one is to transform the PSM to a Platform-Indented Model (PIM) based on the KDM. Therefore, in this step the tool performs a set of model-to-model to get an instance of the KDM which represents

the systems “AS-IS”. In Figure ?? (b) depicts how to get the instance of the KDM based on the earlier PSM.

### 3. Architecture

In Figure 2 is depicted the architecture of KDM-RE which is split in three layers. As shown in this figure, the first layer of our *plug-in* is the Core Framework. This layer represents that we devised the *plug-in* on the top of the Eclipse Platform. In this layer it is also possible to see that we used both Java and Groovy as programming language. Moreover, this layer contains Eclipse Plugins on which our tool is based on, such as MoDisco and EMF. We used MoDisco<sup>1</sup> that is an extensible framework to develop model-driven tools to support use-cases of existing software modernization and provides an Application Programming Interface - (API) to easily access the KDM model. Also, Eclipse Modeling Framework (EMF)<sup>2</sup> was used to load and navigate KDM models that were generated with MoDisco.

The second layer, the Tool Core, is where all refactorings provided by our *plug-in* were implemented. KDM-RE works intensively with KDM models, which are XML files. Therefore, we use Groovy to handle those types of files because of the simplicity of its syntax and fully integrated with Java. KDM-RE also provides a way to create multiple versions of the KDM file to allow the engineer to assess different refactorings in the same system. In order to optimize memory usage of multiple versions for large models and enabling to work interactively on multiple models our *plug-in* persists these models in a MongoDB. We chose MongoDB since it provides a high performance. After the engineer to choose a version refactored a forward engineering is carried out and the source code of the modernized target system is generated again. Finally, the top layer is the Graphical User Interface (GUI) that consists of a set of SWT windows with several options to perform the refactorings based on the KDM model.

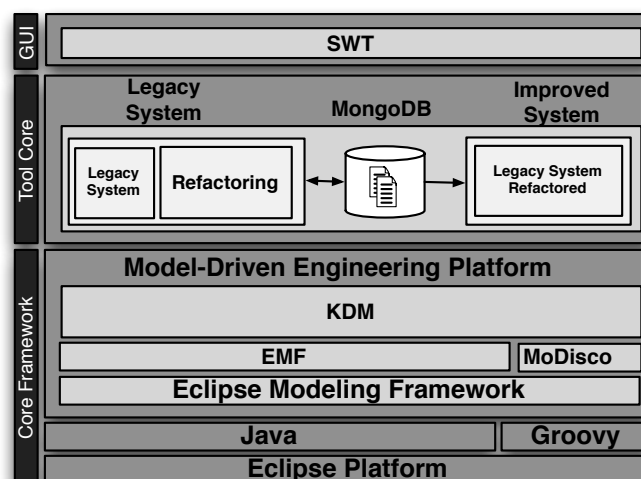


Figure 2. Architecture of KDM-RE

<sup>1</sup><http://www.eclipse.org/MoDisco/>

<sup>2</sup><http://www.eclipse.org/modeling/emf/>

## 4. Related Work

## 5. Concluding Remarks

In this paper is presented the KnowDIME to support the refactoring of legacy systems based on ADM, which uses the KDM standard. It follows the theory of the horseshoe modernization model, which is threefold: (i) **Reverser Engineering**, (ii) **Reestructuring** and (iii) **Forward Engineering**.

Firstly, all the artefacts of the legacy system must be transformed into PSMs by statically analyzing the legacy source code. Still in the first step, these PSMs are integrated into a KDM model, i.e., a PIM model, through a M2M transformation implemented using ATL. Secondly, the KnowDIME applies a set of model refactorings and model optimizations in this PIM. Afterwards, KnowDIME executes a set of M2M transformation taking as input the PIM and producing as output a model conforming to the KDM models into UML meta model. Finally, an improved system is obtained from this UML by means of a set of M2C transformation; additionally, the generated code can be complemented by the software engineer in accordance with the more detailed specifications of the application business rules, such as the implementation of specific behaviors and features not covered by the code generation.

We believe that KnowDIME makes a contribution to the challenges of Software Engineering which focuses on mechanisms to support the automation of software refactoring process. Long term future work involves conducting experiments to evaluate the level of maintenance provided by KnowDIME. It is worth highlighting that KnowDIME is open source and it can be downloaded at [www.dc.ufscar.br/~valter/crossfire](http://www.dc.ufscar.br/~valter/crossfire).

## 6. Acknowledgements

The authors would like to thank CNPq for Processes 241028/2012-4 and FAPESP for Process 2012/05168-4.