

# De l'ombre à la lumière : plus de visibilité sur l'Eclipse

Boris Baldassari\*, Flavien Huynh\*, Philippe Preux\*\*

\* 76 Allées Jean Jaurès, Toulouse, France.

boris.baldassari@squoring.com – <http://www.squoring.com>

\*\* Université de Lille 3, LIFL (UMR CNRS) & INRIA, Villeneuve d'Ascq, France  
philippe.preux@inria.fr – <https://sequel.inria.fr>

**Résumé.** L'extraction de connaissances à partir de données issues du génie logiciel est un domaine qui s'est beaucoup développé ces dix dernières années, avec notamment la fouille de référentiels logiciels (Mining Software Repositories) et l'application de méthodes statistiques (partitionnement, détection d'outliers) à des thématiques du processus de développement logiciel. Cet article présente la démarche de fouille de données mise en œuvre dans le cadre de Polarsys, un groupe de travail de la fondation Eclipse, de la définition des exigences à la proposition d'un modèle de qualité dédié et à son implémentation sur un prototype. Les principaux concepts adoptés et les leçons tirées sont également passés en revue.

## 1 Introduction

L'évolution et le croisement des disciplines de la fouille de données et du génie logiciel ouvrent de nouveaux horizons pour la compréhension et l'amélioration du logiciel et des pratiques de développement. A l'aube de ce domaine émergent de nombreuses questions restent cependant d'actualité du point de vue de la conduite de programmes de mesure logicielle, notamment relevées par Fenton (1994) et Kaner et Bond (2004). Nous proposons dans cet article quelques pistes pour répondre à ces problématiques et mettre en place un processus de mesure fiable et efficace.

Il est utile pour la définition de la notion de qualité de s'appuyer sur des modèles ou standards reconnus : du point de vue de la *qualité produit*, la référence *de facto* semble être l'ISO 9126 et son futur successeur, la série 250xx SQuARE. La *maturité du processus de développement* est adressée par des initiatives largement reconnues telle que le CMMi ou l'ISO 15504. Afin de clarifier la démarche de mesure, l'approche *Goal-Question-Metric* proposée par Basili et al. et reprise par Westfal Westfall et Road (2005) permet une approche plus rigoureuse, qui préserve l'efficacité de l'analyse et le sens de ses résultats.

## 2 Topologie d'un projet de développement logiciel

Chaque référentiel d'information utilisé par le projet possède des informations exploitables. Il importe de lister les référentiels disponibles, puis d'identifier pour chacun d'eux les artefacts et mesures disponibles, et de leur donner un contexte sémantique.