ソフトウェア開発データの分析と応用

伊原 彰紀[1]　阿萬 裕久[2]

概要：近年，産学の双方において「ビッグデータ」が一つの重要なトピックになってきている．ソフトウェア開発も例外ではなく，ソースコードやその開発に関わる作業履歴，不具合情報といった多種多様で膨大なデータに対してマイニング技術や統計解析手法を適用し，品質の維持・向上に向けたさまざまな取り組みが研究・実践されてきている．昨今では，ソフトウェア工学と他分野（自然言語処理，知能情報処理，ゲーム理論，社会学，心理学等）との融合研究が増加しており，さらなる研究分野の発展が期待される．本セッションは，そのような開発データの分析方法並びに開発・管理への応用方法について産学双方の立場から情報の共有と議論を行い，より優れた手法の開発と実践に向けた取り組みを支援する場としたい．具体的には，(1) 開発データの共有，(2) 手法やツール情報の共有，(3) 研究（論文）情報の共有を目的とし，参加者がそれぞれ共有できる情報や共有したい情報について，情報提供や問題提起を行いながら，情報共有サイト（Wiki）の構築を行っていきたい．(1) については，公開データ（GitHub，PROMISE repository 等）へのアクセス手段や，共有可能なデータまたはリンク集の情報；(2) については，データの取得方法，クレンジング等解析のための整形方法，具体的な解析手法（機械学習，マイニング技術，統計手法等），可視化手法（グラフの種類や可視化ツール），データセットの使い方をまとめた情報；(3) については，論文の紹介やサーベイ情報，解析手法についてのチュートリアル情報等をそれぞれ想定している．これらの情報に興味のある皆様の，産学双方からの幅広い情報・話題提供を歓迎する．

1　奈良先端科学技術大学院大学
2　愛媛大学

© 2015 Information Processing Society of Japan