

令和元年5月13日現在

機関番号：14401

研究種目：基盤研究(S)

研究期間：2013～2017

課題番号：25220003

研究課題名(和文)多様なソフトウェア資産の収集・分析・評価と効果的な利活用の研究

研究課題名(英文)Collecting, Analyzing, and Evaluating Software Assets for Effective Reuse

研究代表者

井上 克郎 (INOUE, Katsuro)

大阪大学・情報科学研究科・教授

研究者番号：20168438

交付決定額(研究期間全体)：(直接経費) 97,900,000円

研究成果の概要(和文)：ソフトウェアシステムの効率的な構築のために、過去のソフトウェア資産を効率的に利活用するための諸技術の開発、応用を行った。それらの成果として、ソフトウェア資産の利用関係をモデル化するSUGの開発、ライブラリの依存関係を視覚化する手法の構築などの他、高速にライブラリ等の再利用を検索する技術、リファクタリング性の評価技術、ソフトウェアライセンスの整合性検出技術など多数の研究開発を行った。

研究成果の学術的意義や社会的意義

ソフトウェアシステムは、現代社会に必要な要素で、その効率的な開発方法の探求は、強く求められている。本プロジェクトでは、過去のソフトウェア資産を有効に活用し、開発の効率を上げ、高い機能のソフトウェア・システム構築できるようにするための種々の技術開発を行った。特に、過去の資産を効率よく探し出す検索技術、多数の資産の間の関連を分析し評価する技術、個々の資産をいろいろな視点で評価する技術などを開発し、ソフトウェア資産の効率的な利活用の推進に貢献した。

研究成果の概要(英文)：For efficient development of software system, we have studied various methods and techniques on collecting, analyzing, and evaluating software assets. One of the important outcomes is developing a method of modeling dependency of software assets, named SUG (Software Universe Graph). Also, we have developed library dependency visualizers. In addition to these, we have studied fast library search techniques, refactorability evaluation method, software license inconsistency detection method, and many others.

研究分野：ソフトウェア工学

キーワード：ソフトウェア資産 再利用 依存関係グラフ プログラム検索 リファクタリング評価 依存関係可視化

## 様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

### 1. 研究開始当初の背景

近年、システムが必要とするソフトウェアの規模は巨大化しているが、実際にそのコードの全てを新規に開発することは稀である。既存のオープンソースソフトウェア (OSS) の一部や全部を流用したり、過去に開発した類似のソフトウェアの一部を改変して利用したりするケースが多い。このような開発形態により、開発の効率の向上を目指すとともに、過去の安定したソフトウェアを流用することによってソフトウェアの信頼性が高まることが期待されている。組織内外に存在するソフトウェア資産を多様なソフトウェア資産と称して自動収集し、統一的な評価のもと、再利用することができるようになると、今まで、手元にある限られたレポジトリを対象とした再利用より、飛躍的にデータ量が増加し、より深く、有意義な結果が得られることが期待された。

### 2. 研究の目的

このような背景のもとに、本プロジェクトでは、Software Analysis and Reuse Framework (SARF) と称する枠組みを考え、(1) ~ (3) の研究を遂行する。

- (1) 多様なソフトウェア資産の収集・分析方法の開発  
評価のため必要な情報を得るため、ソフトウェア資産を効率的に収集、分析する種々の技術を開発する。
- (2) ソフトウェア資産の価値の定量的な評価法の開発  
評価のためのソフトウェア資産の価値モデルを構築する。(1) で得た分析結果を利用し、個々のソフトウェア資産の有用性を示す指標を構築する。
- (3) 評価に基づいた利活用技術の開発  
価値モデルで得た定量的な評価に基づき、効果的にソフトウェア開発を支援するための手法とシステムの開発を行う。

### 3. 研究の方法

SARF プロジェクトでは、上記の各目的に対し、数人の研究代表者、分担者が中心となり研究を行った。また、よりプロジェクトを国際的なものとして強化・発展させるために、連携研究者を広く国内外から公募を行い、大阪大学の特任研究員 (特任助教) として雇用した。この結果、3 名の特任助教を採用することができ、SARF プロジェクトが一気に国際化し、活動が活発化された。また、国際的に著名な研究者に、研究協力者として参画してもらい、毎週 1 回程度の頻度で TV 会議で議論を行った。さらに、連携研究者や多数の学生が参加した。

### 4. 研究成果

これら (1) ~ (3) を統一的に進めるようにするための基盤として、資産間の関連をエコシステムとしてモデル化方法を開発し、それに基づいて収集、分析、評価が効果的に行えることを示した。(学会発表 16 Raula et. al., "Trusting a Library: A Study of the Latency to Adopt the Latest Maven Release", SANER 2015 / 学会発表 1 Raula et. Al., "A Generalized Model for Visualizing Library Popularity, Adoption, and Diffusion within a Software Ecosystem, SANER2018) このモデル化では、ソフトウェアエコシステムに対して、個々のソフトウェア資産の各版を頂点、その間の依存関係を有向辺とした Software Universe Graph (SUG) と呼ぶグラフを構築し、実際にどのようなソフトウェアが別のソフトウェアを利用しているか、そのような利用関係が時間とともにどのように変化するかを可視化できるようにした。このモデルを用いて、160 万あまりの Java プログラムを収集・分析を行い、被利用数 (popularity) や多様度 (variety、利用するシステム種類数) という評価基準を用いて、開発者が再利用する際の利用指針を提供した。図 1 はその解析の一例で、これは Java のバイトコード操作ソフトウェア ASM の進化と、その利用数・多様度の時間推移を示している。これによると、新しい版が公開された後も昔の版が長く利用されていることが分かるが、利用しているシステムの種類数を表す多様度を見るとその差が縮まりっており、新しいシステム開発者は、より新しい版を利用する傾向にあることが分かり、今後、新しいシステムでは、新しい版を利用しても良いことが分かる。

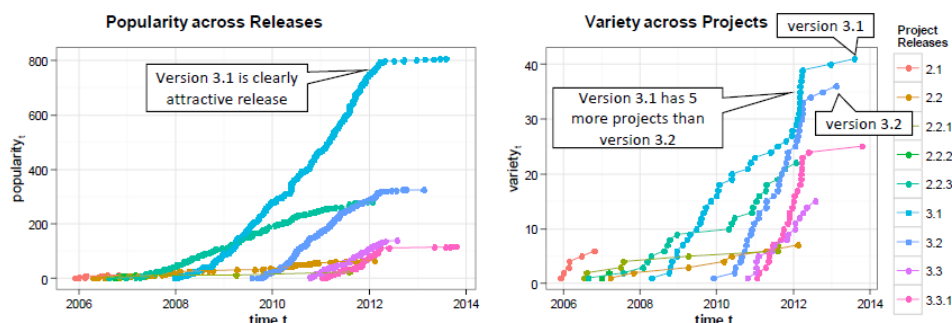


図 1 ASM の利用数、多様度の推移

また、図2は、Javaのパッケージ化のライブラリ commons-beansutils の利用の多様度の時間推移である。古い版 1.9.1 は、その脆弱性が発覚し警告レポートが2014年4月に出された後に、多くの利用者が新規版 1.9.2 に移行している。しかし、まだ、多くの利用者が残っており、危険なソフトウェアが今でも使われていることが分かり、また、新しい利用者は、古い版を利用すべきでないことが分かる。

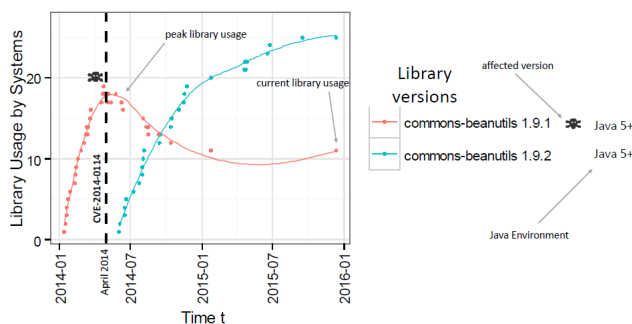


図2 Commons-beansutils の多様度の推移

さらに、SUG を用いて多数のライブラリの同時使用の様子を直感的に理解できるようにするための可視化システムの構築を行った。(学会発表 17 Yano, et. al. "VerXCombo: An Interactive Data Visualization of Popular Library Version Combinations", ICPC 2015.) 図3は、ライブラリ A の 1.0 版を使うシステムは、同時にライブラリ B の 1.0 を使っているのが約 7 割、2.0 を使っているのが約 3 割であることを示している。これにより、最も多くの利用がなされているのが、ライブラリ A が 2.0、B が 2.0、そして C が 3.0 の組合せであることが視覚的にわかる。

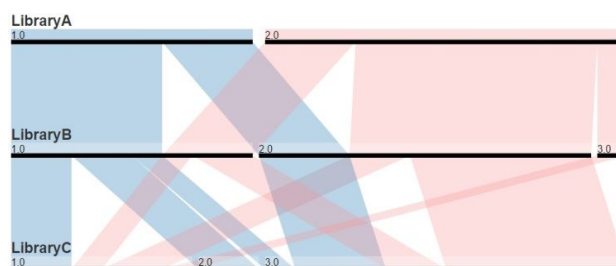


図3 ライブラリ A, B, C の同時利用割合の可視化

またライブラリの依存関係、更新状態などを直感的に視覚化する手法を考案し、システム構築を行った(図4)。(学会発表 7 Todorov, et. al., "SoLMAnta: Visualizing Updated Opportunities Based on Library Coexistence", VISSOFT, 2017.) このシステムでは、中心とするライブラリに依存するライブラリを軌道上に配置し、配色でその更新度合いを示し、現状のシステムが危険なライブラリに依存していないか、容易にわかるようにしている。

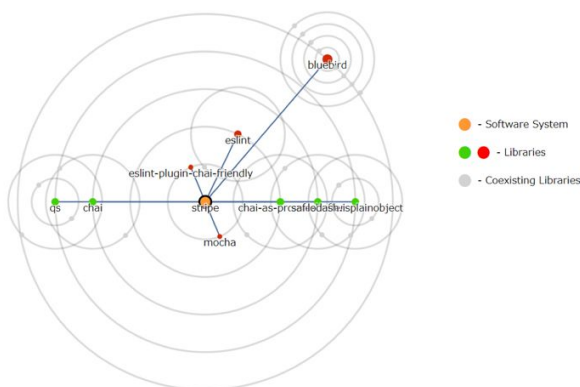


図4 ライブラリの依存と更新状態の表示

このようなソフトウェアエコシステム全体に対する取り組みの他、(1), (2), (3)のそれぞれの部分で活発な研究を行い、大きな成果を挙げている。

例えば、(1)に関しては、オープンソースソフトウェアを集め、その集合に対して、効果的にライセンスの違反の可能性のあるファイルを見つけるための分析手法を開発している。(学会発表 15 Wu, et. al., "A Method to Detect License Inconsistencies in Large-Scale Open Source Projects", MSR 2015.) また、ソフトウェアの再利用の状態を効率よく調べる手法の開発を行った。(学会発表 18 Kawamitsu, et. al., "Identifying Source Code Reuse across Repositories using LCS-based Source Code Similarity", SCAM 2014.) この手法では、ソースコードのファイル間の類似度を計算し、それを比較することによって、類似ファイルのペアを高速に発見する。また、類似ファイル集合を高速検索する技術の開発を行い、その評価を行った。(学会発表 2 Ishio, et. al., "Source File Set Search for Clone and Own Reuse Analysis", MSR 2017.) ここでは、b-bit ハッシュ技術を用いて、高速に多数のファイル群を比較する方法を開発した。

(2)に関する重要な成果としては、Web サービスソフトウェアが提供するインターフェース群の評価を行い、より良いモジュール化をするための手法を開発している。(雑誌論文 4 A. Ouni, et al., "Search-Based Web Service Antipatterns Detection", IEEE Trans on Services Computing, 2017.) この研究では、Web サービスが提供するインターフェースのアンチパターン(問題を引き起こす危険な構造)が解消され、良好になるモジュール構造を、多目的最適化法と呼ぶ探索的ソフトウェア工学手法を用いて実現した。また、複数の評価視点からコードの改良(リファクタリング)を評価する方法を提案し産業界への適用を議論した。(雑誌論文 6 Ouni, et. al., "Multi-criteria Code Refactoring Using Search-Based Software Engineering: An Industrial Case Study", ACM TOSEM 2017.) さらに再利用するソフトウ

エア断片(コードクローン)の有用性評価について、機械学習を用いて自動特定する手法を提案した。(雑誌論文 7 Yang, et al., "Classification Model for Code Clones Based on Machine Learning," Empirical SE, 2015.) また、大規模なりポジトリから類似機能を持つプロダクト(メソッド)を自動抽出する手法の有用性を評価した。(学会発表 19 Higo et al., "How Should We Measure Functional Sameness from Program Source Code? an Exploratory Study on Java Methods", FSE 2014.)

(3)に関する研究成果としては、ソフトウェアを再利用する際に、開発者により良いソフトウェアを提供するために、開発履歴を用いて、最適なりファクタリング手法を提示するシステムを開発した。(雑誌論文 8 A. Ouni, "Improving Multi-Objective Code-Smells Correction Using Development History", Journal of Systems and Software, 2015.) このシステムでは、Javaプログラムの不吉な匂い(Bad Smell)を検出し、探索的ソフトウェア工学手法を利用し、いろいろななりファクタリングの中で、不吉な匂いを小さくし、できるだけ過去と同様な変更で、意味の変更が起こらないものを効率よく探索し、提示する。

## 5. 主な発表論文等

[雑誌論文](計 43 件すべて査読あり)

1. Raula Gaikovina Kula, Ali Ouni, Daniel M. German, Katsuro Inoue: "An Empirical Study on the Impact of Refactoring Activities on Evolving Client-Used APIs", Journal of Information and Software Technology, Vol.93, Jan., pp.186-199, 2018.
2. Raula Gaikovina Kula, Daniel M. German, Ali Ouni, Takashi Ishio, Katsuro Inoue: "Do Developers Update Their Library Dependencies? An Empirical Study on the Impact of Security Advisories on Library Migration", Empirical Software Engineering Journal, Vol.23, Issue 1, pp.384-417, 2018.
3. Ali Ouni, Marouane Kessentini, Houari Sahraoui, Mel O Cinneide, Kalyanmoy Deb, Katsuro Inoue: "MORE: A Multi-objective Refactoring Recommendation Approach to Introduce Design Patterns and Fix Code Smells", Journal of Software: Evolution and Process, Vol.29, No.5, pp.e1843, 2017.
4. A. Ouni, M. Kessentini, K. Inoue and M. O. Cinneide: "Search-Based Web Service Antipatterns Detection," IEEE Transactions on Services Computing, vol. 10, no. 4, pp. 603-617, 2017.
5. Yuhao Wu, Yuki Manabe, Tetsuya Kanda, Daniel M. German, Katsuro Inoue: "Analysis of License Inconsistency in Large Collections of Open Source Projects", Empirical Software Engineering Journal (EMSE), Vol.22, Issue 3, pp. 1194-1222, 2017.
6. Ali Ouni, Marouane Kessentini, Houari Sahraoui, Katsuro Inoue, Kalyanmoy Deb: "Multi-criteria Code Refactoring Using Search-Based Software Engineering: An Industrial Case Study", ACM Transactions on Software Engineering and Methodology TOSEM, Vol.25, Issue 3, Article No.23, Aug. 2016.
7. Jiachen Yang, Keisuke Hotta, Yoshiki Higo, Hiroshi Igaki and Shinji Kusumoto: "Classification Model for Code Clones Based on Machine Learning," Empirical Software Engineering, Vol. 20, Issue 4, pp.1095-1125, 2015.
8. Ali Ouni, Marouane Kessentini, Houari Sahraoui, Katsuro Inoue, Mohamed S. Hamdi: "Improving Multi-Objective Code-Smells Correction Using Development History", Journal of Systems and Software (JSS), Elsevier, Vol.105-C, pp. 18-39, 2015.
9. Tetsuya Kanda, Takashi Ishio, Katsuro Inoue: "Approximating the Evolution History of Software from Source Code", IEICE Transactions on Information and Systems, Vol.E98-D, No.6, pp.1185-1193, 2015.
10. Eunjong Choi, Norihiro Yoshida, Katsuro Inoue: "An Investigation into the Characteristics of Merged Code Clones during Software Evolution", IEICE Transactions on Information and Systems, Vol.E97-D, No.5, pp.1244-1253, 2014.
11. 山中 裕樹, 崔 恩潯, 吉田 則裕, 井上 克郎: "情報検索技術に基づく高速な関数クローン検出", 情報処理学会論文誌, Vol.55, No.10, pp.2245-2255, 2014.
12. 秦野 智臣, 鹿島 悠, 石尾 隆, 井上 克郎: "Thin Slice のサイズに関する統計的評価", 情報処理学会論文誌, Vol.55, No.2, pp.971-980, 2014.
13. 佐々木唯, 肥後芳樹, 楠本真二: "プログラム文の並べ替えに基づくソースコードの可読性向上の試み," 情報処理学会論文誌, Vol. 55, No. 2, pp. 939-946, 2014.
14. 堀田圭佑, 肥後芳樹, 楠本真二: "CRD を用いたコードクローンの生存期間と修正回数に関する調査," 情報処理学会論文誌, Vol. 55, No. 2, pp. 947-958, 2014.
15. 堀田圭佑, 肥後芳樹, 楠本真二: "生成抑止, 分析効率化, 不具合検出を中心としたコードクローン管理支援技術に関する研究動向," コンピュータソフトウェア, Vol. 31, No. 1, pp. 14-29, 2014.

[学会発表](計 99 件)

1. Raula Gaikovina Kula, Coen De Roover, Daniel M. German, Takashi Ishio, Katsuro

- Inoue: "A Generalized Model for Visualizing Library Popularity, Adoption, and Diffusion within a Software Ecosystem", Proc. of 25th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER2018), pp.288-299, 2018.
2. T. Ishio, Y. Sakaguchi, K. Ito, K. Inoue: "Source File Set Search for Clone-and-Own Reuse Analysis", Proceedings of the 2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR2017), pp.257-268, 2017.
  3. Yoshiki Higo, Akio Ohtani, and Shinji Kusumoto, "Generating Simpler AST Edit Scripts by Considering Copy-And-Paste," In The 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE2017), pp. 532-542, 2017.
  4. Raula Gaikovina Kula, D. German, T. Ishio, A. Ouni, K. Inoue: "An Exploratory Study on Library Aging by Monitoring Client Usage in a Software Ecosystem", Software Analysis, Evolution and Reengineering SANER, pp. 407-411, 2017.
  5. Anfernee Goon, Yuhao Wu, Makoto Matsushita, Katsuro Inoue: "Evolution of Code Clone Ratios throughout Development History of Open-Source C and C++", 11th International Workshop on Software Clones, pp.47-53, 2017.
  6. Y. Sabi, Y. Higo, and S. Kusumoto, "Rearranging the Order of Program Statements for Code Clone Detection," In Proc. of the 11th IEEE International Workshop on Software Clones, pp. 15-21, 2017.
  7. Boris Todorov, Raula Gaikovina Kula, Takashi Ishio, Katsuro Inoue: "SoL MAntra: Visualizing Updated Opportunities Based on Library Coexistence", 2017 IEEE Working Conference on Software Visualization (VISSOFT), pp.129-133, 2017.
  8. Ali Ouni, Zouhour Salem, Katsuro Inoue, Makram Soui: "SIM: An Automated Approach to Improve Web Service Interface Modularization", IEEE International Conference on Web Services, ICWS2016-1012, 2016.
  9. Ali Ouni, Raula Gaikovina Kula, Katsuro Inoue: "Search-Based Peer Reviewers Recommendation in Modern Code Review", 32nd IEEE International Conference on Software Maintenance and Evolution (ICSME'16), pp.367-377, 2016.
  10. H. Nakajima, Y. Higo, H. Yokoyama, and S. Kusumoto, "Toward Developer-like Automated Program Repair –Modification Comparisons between GenProg and Developers–," In Proceedings of the 23rd Asia-Pacific Software Engineering Conference (APSEC2016), pp. 241-248, 2016.
  11. N. Ogura, J. Yang, K. Hotta, Y. Higo, and S. Kusumoto, "Hey! Are You Injecting Side Effect?: a Tool for Detecting Purity Changes in Java Methods," In Proc. of 24th IEEE International Conference on Program Comprehension (ICPC2016), pp.1-3, 2016.
  12. Takashi Ishio, Raula Gaikovina Kula, Tetsuya Kanda, Daniel M. German, Katsuro Inoue: "Software Ingredients: Detection of Third-party Component Reuse in Java Software Release", Proceedings of the 13th International Conference on Mining Software Repositories, pp.339-350, 2016
  13. Ali Ouni, Raula Gaikovina Kula, Marouane Kessentini, Katsuro Inoue, "Web Service Antipatterns Detection Using Genetic Programming", In Proc. of 24th ACM Genetic and Evolutionary Computation Conference (GECCO), pp.1351-1358, 2015.
  14. Ali Ouni, Marouane Kessentini, Houari Sahraoui, Mel Ó Cinnéide, Kalyanmoy Deb, Katsuro Inoue, "A Multi-Objective Refactoring Approach to Introduce Design Patterns and Fix Anti-Patterns", in Proc. North American Search Based Software Engineering Symposium (NasBASE), (Elsevier), pp.1-16, 2015.
  15. Yuhao Wu, Yuki Manabe, Tetsuya Kanda, Daniel M. German, Katsuro Inoue: "A Method to Detect License Inconsistencies in Large-Scale Open Source Projects", 12th Working Conference on Mining Software Repositories (MSR 2015), pp.324-333, 2015.
  16. Raula Gaikovina Kula, Daniel German, Takashi Ishio, Katsuro Inoue: "Trusting a Library: A Study of the Latency to Adopt the Latest Maven Release", 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2015), pp. 520 - 524, 2015.
  17. Y. Yano, Raula G. Kula, T. Ishio, K. Inoue: "VerXCombo: An interactive data visualization of popular library version combinations", International Conference on Program Comprehension (ICPC 2015), pp. 291-294, Florence, Italy, 2015.
  18. Naohiro Kawamitsu, Takashi Ishio, Tetsuya Kanda, Raula Gaikovina Kula, Coen De Roover, Katsuro Inoue: "Identifying Source Code Reuse across Repositories using LCS-based Source Code Similarity", International Working Conference on Source Code Analysis and Manipulation (SCAM 2014), pp.305-314, 2014.
  19. Yoshiki Higo and Shinji Kusumoto: " How Should We Measure Functional Sameness from Program Source Code? an Exploratory Study on Java Methods, " In In Proc. of the 22nd International Symposium on the Foundations of Software Engineering (FSE2014), pp. 294-305, 2014.



20. T. Ishihara, K. Hotta, Y. Higo, S. Kusumoto: "Reusing Reused Code," 20th Working Conference on Reverse Engineering (WCRE2013), pp. 457-461, 2013.

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕ホームページ <http://sel.ist.osaka-u.ac.jp/SARF/index.html>

## 6. 研究組織

### (1)研究分担者

研究分担者氏名：楠本 真二

ローマ字氏名：(KUSUMOTO, Shinji)

所属研究機関名：大阪大学

部局名：情報科学研究科

職名：教授

研究者番号(8桁)：30234438

研究分担者氏名：松下 誠

ローマ字氏名：(MATSUSHITA, Makoto)

所属研究機関名：大阪大学

部局名：情報科学研究科

職名：准教授

研究者番号(8桁)：60304028

研究分担者氏名：石尾 隆

ローマ字氏名：(ISHIO, Takashi)

所属研究機関名：奈良先端科学技術大学院大学

部局名：先端科学研究科

職名：准教授

研究者番号(8桁)：60452413

研究分担者氏名：岡野 浩三

ローマ字氏名：(OKANO, Kozo)

所属研究機関名：信州大学

部局名：工学部

職名：准教授

研究者番号(8桁)：70252632

研究分担者氏名：肥後 芳樹

ローマ字氏名：(HIGO, Yoshiki)

所属研究機関名：大阪大学

部局名：情報科学研究科

職名：准教授

研究者番号(8桁)：70452414

### (2)研究協力者

研究協力者氏名：ラウラ ガイコピナ クラ (連携研究者)

ローマ字氏名：(RAULA GAIKOVINA KULA)

研究協力者氏名：オウニ、アリ (連携研究者)

ローマ字氏名：(OUNI, Ali)

研究協力者氏名：デローバー、クーン (連携研究者)

ローマ字氏名：(DE ROOVER, Coen)

研究協力者氏名：吉田 則裕 (連携研究者)

ローマ字氏名：(YOSHIDA, Norihiro)

研究協力者氏名：真鍋 雄貴 (連携研究者)

ローマ字氏名：(MANABE, Yuki)

研究協力者氏名：チェ、ウンジョン (連携研究者)

ローマ字氏名：(CHOI, Eunjong)

研究協力者氏名：ガーマン、ダニエル

ローマ字氏名：(GERMAN, Daniel)

研究協力者氏名：ケセンティニ、マロウアン

ローマ字氏名：(KESSENTINI, Marouane)