

平成22年 5月31日現在

研究種目： 基盤研究 (A)
 研究期間： 2006～2009
 課題番号： 18200020
 研究課題名 (和文) データ科学の新領域の開拓－文化財データ解析－
 研究課題名 (英文) Exploitation of the new domain of Data Science
 - Informatics of the cultural properties -
 研究代表者
 村上征勝 (MURAKAMI MASAKATSU)
 同志社大学・文化情報学部・教授
 研究者番号：00000216

研究成果の概要 (和文)：

本研究では、人文社会科学の中でも、情報化・データ化の研究が遅れていた文化財分野を対象とし、考古学、舞踊、古典文献、浮世絵等の具体的な研究課題の解明を通じて、文化財特有のデータの性質や基盤化、またそれをを用いた新しい文化研究の方法論および解析方法の開拓を試みた。データの基盤化に際して、STISを開発し、世界各地の文化遺産や言語情報などをこれに格納し、それらを束ねる実体として文化を復原、評価する研究手法の開発を試みた。

最終年度には国際シンポジウムを実施し、本課題の成果を公開した

研究成果の概要 (英文) : In this research task, the new domain of the data science about cultural properties was exploited. In particular about cultural properties, various researches were not progressing in the field until now. The distinctive character of cultural-properties data was investigated, that database was developed, and analysis of the cultural phenomenon was carried out by this project. STIS (Spatio-Temporal Information System) which unifies various information were developed and various phenomenon analysis modules were implemented. Also Our project held international symposium in the last fiscal year, advanced the share in question, and exhibited the result.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	15,000,000	4,500,000	19,500,000
2007年度	6,700,000	2,010,000	8,710,000
2008年度	9,100,000	2,730,000	11,830,000
2009年度	6,400,000	1,920,000	8,320,000
年度			
総計	37,200,000	11,160,000	48,360,000

研究分野： 総合領域

科研費の分科・細目： 情報学・統計科学

キーワード： 多変量解析・時空間データ

1. 研究開始当初の背景

(1) データ科学領域の背景

データ科学は、これまで、経済学や社会学、心理学などの人文社会科学分野において有用な分析手法の開拓と発展がなされた。また、医療分野、保険分野、行政分野なども新しい対象領域として、一定の成果を上げつつある。しかし、文化現象をあつかう芸術学・歴史学・人類学・文学などの分野では、一部の先行事例研究はあるものの、体系的な方法論や体制作りまで含めた、データ科学としての研究基盤はまだ成立していない状況であった。

(2) 文化財研究領域の背景

文化財研究領域では、文化遺産のドキュメンテーションとアーカイブ、マネジメントの目的から、これまで蓄積されてきた文化財情報、あるいはこれから蓄積される様々な文化現象に関わる情報を、どのように体系化し、分析していくかの方法論が模索されていた。特に分析科学や情報学とも関係する文化財科学分野では、“文化遺産情報学”や“時空間情報科学”のような、これまでの歴史学や芸術学とは一線を画した細分領域の必要性が指摘されてきた。

2. 研究目的

(1) 問題の整理と帰着点の検討

データ科学領域の研究手法を文化研究において適用する場合の問題を整理し、その問題を具体的なテストケースによって提示することが目的である。特に、文化財情報とはどのような性質をもっていて、それをどのように整理すれば分析できるデータとなるのか、という基本的な問題を解決し、その帰着点として、データから何らかの新しい文化現象を可視化して見せることが重要である。

(2) 超学領域での研究者の連携

対象となる文化財は、人類のあらゆる行動の客体的所産であり、時に有形・時に無形という様々な対象を扱うことになる。そのため、それまでの学領域を越えて研究者が連携することが必要であり、それらの研究者が同様の利便性を持って調査・研究データが利用できるように情報・データ基盤を構築することが不可欠である。

(3) 国際的視座での文化財のデータ化

文化財、言い換えれば人類の歴史や文化は、“現在”の国境や行政界、政治的領域などによって“単位”化できる対象ではない。人類

学で指摘されるような「地域知 (Local Knowledge) =文化現象」であることをデータ科学の手法によって明らかにするためには、東アジアだけでなく、遠く西アジアや中央アジアなど、少なくともユーラシア的規模での国際協力と情報共有を必要とする。

(4) 文化現象分析の手法の開発

データ科学領域の研究でも、一部の先行事例を除けば、文化を可視化するだけの計量的アプローチという実践事例は存在しない。そのため、「地域知」を再構築すること、少なくともその再構築に手続き的再現性を伴った分析手法を数多く必要とすることは明らかであり、それを達成することが、本研究課題の最も重要な目的である。

3. 研究の方法

(1) データ基盤の構築

研究の目的を達成するために最も重要な基礎的研究の方法として、まず、文化財データ・情報の基盤化を進めた。具体的には、大型の記憶媒体を伴った計算機にデータを格納することになるが、これについて、ISO 情報標準に準拠した、「文化財情報標準」を設計し、これにより、多様な文化財の情報化を実施した。また、これらの基盤化、データベースの構築には、XML を利用した時空間情報のリレーションを付加したデータベース STIS (Spatio-Temporal Information system) を設計・開発・実装し、多様な文化財情報のデータ基盤とした。

(2) 情報化・調査の実践的实施

文化財情報は、その文化現象を理解するために計測取得されるデータの塊であり、その第一次情報は、全て文化財調査のドキュメンテーション過程から発生する。そのため、既存に存在するデータのほぼ全ては、特定の目的のために取得された再現性のないデータであり、これを用いて分析手法を開発したり、適用研究を実践することは、避けなくてはならない。そのため、本課題では、できるだけ文化財情報の第一次情報を取得できるよう、実際に文化財調査を実施し、その成果をデータ基盤に繁栄できるように調査、計測を実施した。

(3) 国際的な問題意識と現状の共有

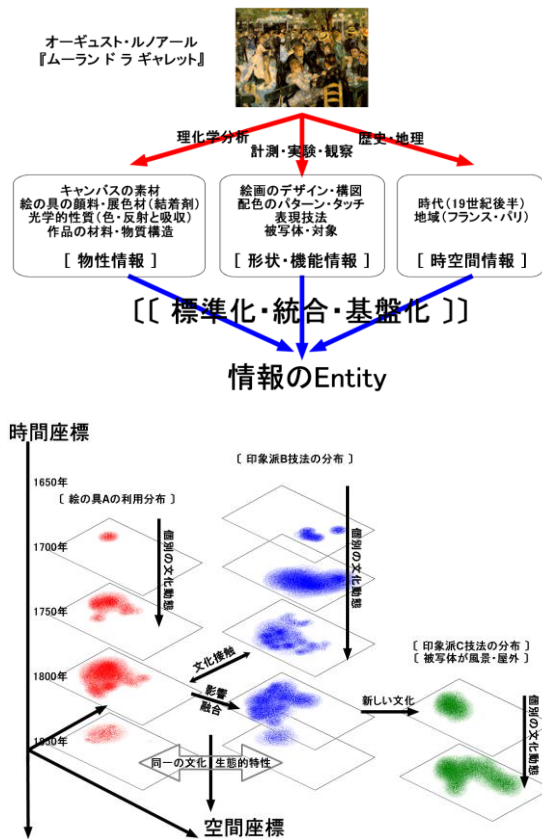
文化財ドキュメンテーションで発生する第一次情報の精度や確度は、データ科学領域で指摘される統計科学的な精度・確度の問題

だけでなく、調査者の調査目的や調査対象によって、その性質が代わることが明らかである。そのため、本研究課題では、国際機関と協力し、国際的な文化財情報のドキュメンテーション方法に関するマスタープランニングにイニシアチブを持って参画すると同時に、その精度や確度に対する主体的な評価機構としての役割も果たせるよう準備を進めた。そうすることで、本研究課題終了後も蓄積され続ける文化財データに関して、一定の基準によって、開発した分析手法が適用できるよう準備を進めた。

(4) STISによる文化現象の可視化

時間と空間の座標によって統合化された文化財情報は、必要に応じた情報相互の関連を様々な手法によって評価され、可視化されることとなる。

例えば、文化財から取得された様々なデー



タは、その文化財の属性データとして標準化されたメタデータを付されて、1つのデータのエンティティとしてデータベースに格納される。格納されたデータベースでは、まずそのエンティティの情報のオントロジー構造を解析し、データAとデータBが何で関連しているのかを評価する。その関連が評価された情報は、次の段階で時空間座標上に展開され、それら相互の時空間での関連(現象)として評価される。こうした方法により、①

文化財の属性データベースからオントロジー構造を評価し、②そのオントロジー構造が時空間でどのように展開・関連するかを解析する。本研究課題で開発したSTISはこれを可能とする1つの技術基盤であり、①と②の評価手法の開拓が、本研究課題の中核である。

4. 研究成果

(1) STISの開発

統合的な現象解析を可能とする方法論の開発と情報のオントロジーを評価するシステムを構築した。これにより、本研究課題では扱う事ができなかった、より多様な文化現象に関する、実践的に適用が可能となる。また、従来実践されてきた経済現象や社会現象に対しても、文化現象と同じような方法論が適用可能であり(そもそも、経済や社会という“類の概念”は文化的所産そのものである)、これまでデータ領域分野で実践されてきた、よりロバストな方法によって解析が可能となる。

(2) 文化財情報標準の提唱

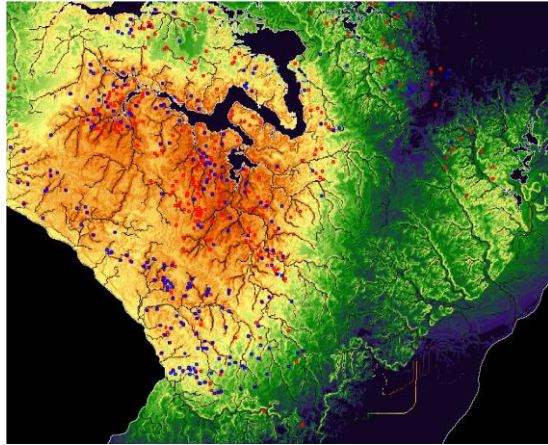
データ基盤を構築するために設計した文化財情報標準は、世界的にも例がない、新しい標準の成果である。特に、文化研究機関で進められている文化財のドキュメンテーション作業に関連して、本研究課題で設計した文化財情報標準の適用も進められている。また、実際の第一次情報取得のための調査・研究を通じて、中国や中央アジア諸国、カンボジアやインド、オマーン、トルコなどの東・東南・西・南アジア各地域の研究機関と連携し、実際の適用実験も進めた。

(3) 文化現象の可視化

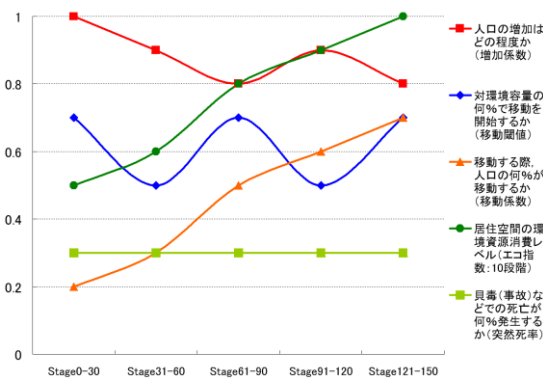
取得した文化財情報によって、これまで明らかにできていなかった文化現象の可視化を実施した。

日本の考古学遺跡に関しては、縄文時代の貝塚遺跡を対象として、その人類の行動生態が文化現象とどのような関連にあったかについて、実際にSTISにエージェントシミュレーションなどの新しい解析手法を適用した方法によって明らかにした。特に生態的行動のパラメータの時空系列での行動変遷が、文化の変容に果たした背景について明らかにした。また、この成果はNHKの科学番組サイエンスZEROでも最先端の考古学研究として取り上げられ、連携研究者の津村が、本研究を紹介した。

また、弥生時代遺跡に関しては、これまで定性的に理解されてきた集落存在の背景について、同様に空間的なパターンを見出し、そこから、定性的解釈であった“階層性”や“機能分化”の社会を否定する成果をえた。



文化の時空間分布

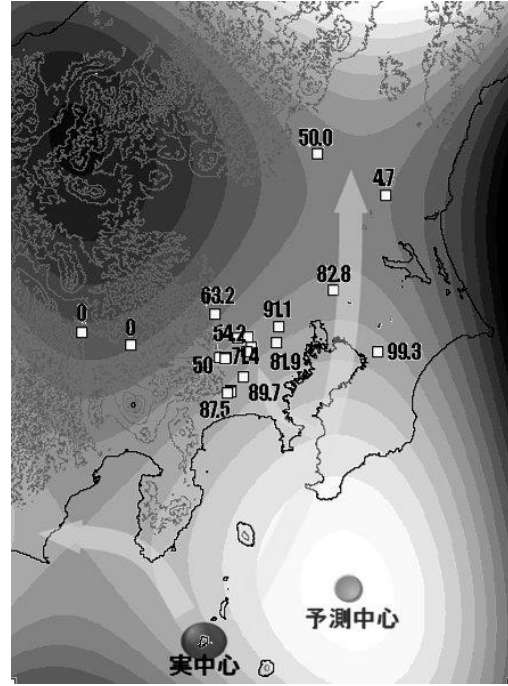


行動パラメータの推移

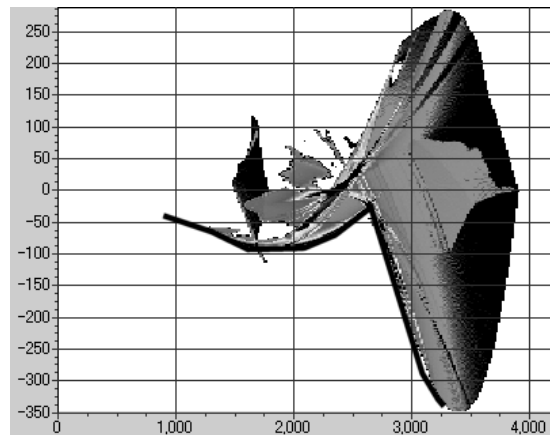
特に、集落環境の多様化と人類行動の多様化の関連について、それが社会構造によって起こるのではなく、時空間構造によって確率的に（ランダムに）発生した現象であることを明らかにした。

1年目に取得した電子顕微鏡を用いたデータ化とこれによる文化現象解析についても進めた。黒曜石のような特定の元素組成を持つ岩石が、日本の多くの遺跡で発見されている。本研究課題では、それらの遺跡資料から第一次情報として分析データを取得し、これから多変量解析によって原産地を同定し、その分布の傾向面分析から、どのような黒曜石の物流が存在したかを再現した。特に、物流に関して、傾向面と出土量の画像相関から、原産地からの搬出距離逓減モデルを作成し、日本の先史時代の経済体系のパターンと時系列動態を評価した。

また、無形文化遺産については、特に言語現象に着目し、中国・日本における方言分布に関する現象評価を実施した。言語はその他の文化財よりも、より人間の直接的な行動と密接に関連しやすく、生態的特性を多分に保持している。そのため、方言分布調査を民俗学的手法により実施し、第一次情報を取得すると共にこれを STIS に格納して文化現象の評価を実施した。結果として、方言の動態が



黒曜石の物流現象



黒曜石の搬出距離逓減モデル

メディアを通じた情報の交流密度と関連することを定量的に評価した。

(4) 文化現象のデータベース化と数量解析
前述の研究に加え、方法論の開発のため以下のような研究を進めた。

- ・『法華経』28巻の成立順序推定のための梵文大乘経典のデータベースの完成と数量分析
- ・東洲齋写楽を始めとする浮世絵の数量分析
- ・日本舞踊の動作の数量分析
- ・視覚的表象と言語の直列的統合
- ・香道の書の分析
- ・茶具の研究
- ・古代寺院の建立時期の推定
- ・『源氏物語』井原西鶴作品の文章の計量分析

(5) 国際シンポジウムの開催

2010年1月、本研究の締めくくりとして、ユーラシア各地の文化財情報研究者を招聘して、国際シンポジウムを開催した。参加者代表国は、中国・カンボジア・ウズベキスタン・トルコ・モンゴルなどの文化遺産情報のドキュメンテーションを実際に進めている各国であり、関連するフランス・イギリス・エジプト・オマーン・ロシア・イタリアなど各国からの情報も収集した。

この国際シンポジウムでは文化財情報・データの標準化の問題、第一次情報取得・計量の問題、文化現象を可視化する分析手法の問題などが積極的に議論された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 25 件)

- ① Hiroomi Tumura, Utilization of Laser Range Finder and Differential GPS for High-Resolution Topographic Measurement at Hacitugrul Tepe, Turkey. *Geoarchaeology: An International Journal*, Vol. 24, No. 2. 2009, pp. 176-190, 査読有
- ② 田口哲也, 非都市化の文法、比較文化研究, No. 89, 2009, pp. 1-3, 査読有
- ③ 高橋美都, 愛媛県今治市立村上水軍博物館寄託 能島村上家旧蔵「香道の書」、文化情報学 4 巻 2 号, 2009, pp. 13-20, 査読有
- ④ 金 明哲, 文章の執筆時期の推定-芥川龍之介の作品を例として-, 行動計量学, 第 36 巻 第 2 号, 2009, pp. 89-103, 査読有
- ⑤ Mamiko Sakata, Mieko Marumo, and Kozaburo Hachimura, "Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication", An Analysis of Motion Features of Different Characters in Nihon-Buyo Dance Using Motion Capture, 2008, pp. 273-278, 査読有
- ⑥ 津村宏臣, GISを用いた弥生集団論-文化財情報から可視化される文化の実体と実態- 設楽博巳、藤尾慎一郎、松木武彦編『集落から読む弥生社会』同成社, 2008, pp. 96-111, 査読有
- ⑦ 津村宏臣, オマーン ラス・ジプス貝塚 『考古学研究』 54-3, 2007, pp. 114-117, 査読有
- ⑧ 津村宏臣, 先史海洋民族の生態史を辿る-オマーン東海岸の詳細分布調査と環境調査-, 『考古学が語る古代オリエント』, 2007, pp. 144-150, 査読有
- ⑨ Woong Choi, Tadao Isaka, Mamiko Sakata, Hidemitsu Morioka, Seiya Tsuruta, Kohei Furukawa, and Kozaburo Hachimura, Quantitative Analysis of Leg Movement and EMG Activity in Expert Japanese Traditional Dancer, "Proceedings of 16th IEEE International Conference on Robot & Human Interactive Communication", 2007, pp. 882-887, 査読有
- ⑩ 村上征勝, 浦部治一郎「浮世絵における役者の顔の描画法に関する数量分析」, 統計数理 Vol. 55. No. 2, 2007, pp. 223-233, 査読有
- ⑪ 金明哲, 村上征勝: ランダムフォレスト法による文章の書き手の同定, 統計数理, Vol. 55. No. 2, 2007, pp. 255-268, 査読有

[学会発表] (計 61 件)

- ① 津村宏臣, 文化遺産情報の資源化と共有化-国際的需要と実状- 国際シンポジウムデータ科学の新領域の開拓-文化遺産情報のアーカイブと文化の分析- 同志社大学 2010. 01. 28
- ② 津村宏臣, 敦煌石窟壁画データアーカイブへの GIS の応用-写真データの一括管理と画像解析 奈良文化財研究所第 14 回遺跡 GIS 研究会 奈良文化財研究所 2009. 11. 20
- ③ 祖父江早苗, 村上征勝, 東洲齋写楽の役者絵と数量分析, アート・ドキュメンテーション学会第二回秋季研究発表会, 国立西洋美術館, 2009. 10. 17
- ④ Mingzhe Jin, Masakatsu Murtakami, Yuejun Zheng, 2009. 8. 26, Authorship Identification with Same Classifiers, International Institute of Applied Statistics, 中国青島、中国海洋大学国際学術交流センター
- ⑤ 建石 徹, 津村宏臣, 黒曜石原産地データの時空間パターン分析と人・モノ・情報の移動 日本文化財科学会第26回大会 名古屋大学 2009. 07. 11
- ⑥ 津村宏臣, 沈力, 馮良珍, 無形文化財としての言語(方言)の時空間分布調査データのGIS解析と語形成理論 - 中国・山西省、霍州内部方言拡散の数理分析 - 日本文化財科学会第26回大会 名古屋大学 2009. 07. 11
- ⑦ 津村宏臣, 石器 3 次元データからの形態パラメータの設定と定量解析手法の開拓 日本文化財科学会第26回大会 名古屋大学 2009. 07. 11
- ⑧ Tetsuya Taguchi Japanese Zen Tosses Water on a Century of Fire, South and Southeast Asian Association for the Study, Universitas Hindu Indonesia (Denpasar, Bali, Indonesia) 2009. 6. 4
- ⑨ 中谷友美, 村上征勝, 「日本古代寺院にお

ける金堂基壇の平面形状の変化」, 日本情報考古学会第 26 回大会, 共立女子大学, 2009. 5. 31

- ⑩ 金明哲. 文章の執筆時期は推定できるか?, 日本計量国語学会第 52 回大会、武庫川女子大学、2008. 9. 20
- ⑪ 沈力・津村宏臣、霍州内部方言拡散的数理解析、中国語言学会第 14 届學術年会 温州大学 2008. 8. 28
- ⑫ 津村宏臣・鎌倉快之・澤田砂織・寺村裕史、文化財総合情報システム STIS の開発と応用、日本文化財科学会第 25 回大会 鹿児島国際大学 2008. 5. 15
- ⑬ 早川裕一・津村宏臣・鎌倉快之・樋泉岳二、遺跡立地・分布に関する社会生態的数値シミュレーションアプローチ、日本文化財科学会第 25 回大会 鹿児島国際大学 2008. 5. 14
- ⑭ 津村宏臣、GIS とエージェントシミュレーションー先史時代貝塚の立地論・分布論・景観論と文化財情報科学ー 日韓の先史時代集落シンポジウム 立命館大学 2008. 2. 17
- ⑮ 金明哲、芥川龍之介の文体の変化ー助詞を中心にー、日本行動計量学会第 35 回大会、同志社大学、2007. 9. 5
- ⑯ 津村宏臣、一般調査データの GIS 化と遺跡立地の定量解析ー時系列動態の波動と社会生態特性ー、第 160 回アナトリア考古学研究会、中近東文化センター、2007. 2. 24
- ⑰ 津村宏臣・村上征勝・八村広三郎・矢野環・建石徹・阪田真己子、文化情報学のパースペクティブーデジタルアーカイブへの新地平ー、人文科学とコンピュータシンポジウム 2006、情報処理学会、同志社大学、2006. 12. 14
- ⑱ 津村宏臣・寺村裕史・村上征勝、文化財科学情報の総合的情報システムの構築、日本情報考古学会第 22 回大会、日本情報考古学会、早稲田大学、2006. 9. 23

[図書] (計 9 件)

- ① 伊藤瑞叡、村上征勝、佐野泰夫、伊藤康裕、野田悟史、「主要梵大乘經典語彙用例総索引」, 主要梵大乘經典語彙用例総索引刊行会, 2009, pp. 3793
- ② 鄭躍軍、統計的社会調査ー心を測る理論と方法、勉誠出版、2008、pp. 323
- ③ 富士昭雄、谷脇理史、西島孜哉、村上征勝、竹下義人、佐藤勝明、伴野英一、「新編西鶴全集第五卷一本文篇・索引篇上・下巻」, 勉誠出版, 2007, 上 pp. 1059, 下 pp. 727
- ④ 鄭躍軍、金明哲、村上征勝、「データサイエンス入門」, 勉誠出版, 2007, pp. 227

6. 研究組織

(1) 研究代表者

村上 征勝 (MURAKAMI MASAKATSU)
同志社大学・文化情報学部・教授
研究者番号: 00000216

(3) 連携研究者

岡部 篤行 (OKABE ATSUYUKI)
東京大学・工学研究科・教授
研究者番号: 10114050
篠原 徹 (SHINOHARA TORU)
国立歴史民俗博物館・研究部・名誉教授
研究者番号: 80068915
八村 広三郎 (HACHIMURA KOZABURO)
立命館大学・情報理工学部・教授
研究者番号: 70124229
矢野 環 (YANO TAMAKI)
同志社大学・文化情報学部・教授
研究者番号: 10111410
二宮 修治 (NINOMIYA SHUJI)
東京学芸大学・教育学部・教授
研究者番号: 30107718
金 明哲 (JIN MINGZHE)
同志社大学・文化情報学部・教授
研究者番号: 60275469
田口 哲也 (TAGUCHI TETSUYA)
同志社大学・文化情報学部・教授
研究者番号: 00145103
高橋 美都 (TAKAHASHI MITO)
同志社大学・文化情報学部・准教授
研究者番号: 70326172
岡田 健 (OKADA KEN)
東京文化財研究所・文化遺産国際協力センター・国際情報研究室長
研究者番号: 40194352
鄭 躍軍 (ZHENG YUEJUN)
同志社大学・文化情報学部・教授
研究者番号: 80280527
下嶋 篤 (SHIMOJIMA ATSUSHI)
同志社大学・文化情報学部・教授
研究者番号: 40303341
福田 智子 (FUKUDA TOMOKO)
同志社大学・文化情報学部・准教授
研究者番号: 50363388
宿久 洋 (YADOHISA HIROSHI)
同志社大学・文化情報学部・教授
研究者番号: 50244223
梅崎 昌裕 (UMEZAKI MASAHIRO)
東京大学・医学研究科・准教授
研究者番号: 30292725
杉尾 武志 (SUGIO TAKESHI)
同志社大学・文化情報学部・准教授
研究者番号: 60335205
津村 宏臣 (TSUMURA HIROOMI)
同志社大学・文化情報学部・准教授
研究者番号: 40376934
阪田 真己子 (SAKATA MAMIKO)
同志社大学・文化情報学部・准教授
研究者番号: 10352551