

[2017]九州大学情報統括本部年報 : 2017年度

<https://hdl.handle.net/2324/2203028>

出版情報 : 九州大学情報統括本部年報. 2017, pp.1-, 2018-10-01. Information Infrastructure Initiative, Kyushu University

バージョン :

権利関係 :

第1章 学術情報研究部門

1.1 スタッフ一覧

職名	氏名	研究キーワード
教授	廣川 佐千男	機械学習、検索エンジン、テキストマイニング、推論システム、証明論
准教授	鈴木 孝彦	演繹データベース、論理プログラミング、機能論理プログラミング、CADデータベース、認証、PKI、属性認証
准教授	伊東 栄典	Webマイニング、Webサービス、情報検索、情報統合、XML、分散システム、ネットワーク、協調システム、ソフトウェア工学
助教	中藤 哲也	学術情報分析、テキストマイニング、文字列照合、機械学習

1.2 研究事例紹介

「利用者投稿型コンテンツの多様性動向分析」

伊東栄典

以下は、科学研究費助成事業（課題番号：15K00451）についての成果報告に加筆したものである。

1.2.1 背景

近年CGM（Consumer Generated Media）やUGC（User Generated Content）と呼ばれるコンテンツが人気である。YouTube やニコニコ動画、小説家になろう（syosetu.com）、画像のPixivなどのCGMサイトには毎日多数の投稿があり、膨大なコンテンツが蓄積されている。殆どは他愛のないコンテンツであるけれども、中にはTVや出版社などの従来メディアで提供されていたコンテンツと同程度の品質をもつコンテンツも増えている。動画投稿サイトには楽曲も多数投稿されている。小説・音楽・静止画・動画は有機的に連携し、多様な電子コンテンツが爆発的に生産されている。

多くのCGMサイトが、コンテンツへの評価尺度に閲覧回数を用いている。他にも閲覧者からの評価値、お気に入り（ブックマーク）登録数を用いる場合もある。これらの評価尺度は概ねコンテンツの人気を表す。閲覧回数によるランキングは、CGM開始期には有効に働き、作者のコンテンツ作成動機や品質向上を促す効果があったと考えられている。

開始初期に多様な分野のコンテンツが存在していたCGMサイトであるが、2014年頃からコンテンツの多様性喪失（画一化）が指摘されていた。2014年時点で、ニコニコ動画の音楽カテゴリに投稿される楽曲では、上位100位のうち90%がボーカロイド音楽で、最近数年は物語に基づくボーカロイド曲が人気になっている。「小説家になろう」でも上位100位の小説はほぼ全部「ファンタジーな異世界に転生するもの」になっていた。

多様性喪失の原因として次の仮説を考えている。現実世界の書店や図書館と異なり、PCやタブレットの画面では多数のコンテンツを一覧できない。そのためコンテンツを試すには一つ一つ見るしかない。しかし自力で試すには時間が掛かり過ぎる。一方、既に自分が知る分野は背景知識を持つため新規コンテンツ

への理解が速い。そのため、利用者は既知の分野のコンテンツを見たがる傾向があると思われる。作者も、人気を得やすい分野のコンテンツを作りたがる傾向がある。これらの理由からコンテンツに偏りが出て、多様性が失われていると考えている。

1.2.2 研究目的

本研究では、利用者投稿型コンテンツの多様性動向の分析を行う。研究対象は次の3種類、「ニコニコ動画」に投稿される音楽（楽曲）、「小説家になろう」に投稿されるオンライン小説、そして学術論文である。

本申請では上記に示した三種のデータについて、二つの分析を試みた。1つ目は多様性喪失（画一化）の定量評価である。多様性を評価する指標を提案し、それをコンテンツ集合に適用することで、多様性の動向を分析した。2つ目は利用者のコンテンツ選択モデルの確立である。直感的に上位コンテンツは益々閲覧数が増え、下位のもの閲覧されにくいと想定できる。コンテンツの閲覧回数分布を統計的に調べる。また、その分布に合致する利用者のコンテンツ選択行動モデルの確立を試みた。

1.2.3 研究方法

コンテンツの画一化を測るためには、多様性の指標が必要である。そこで以下に挙げた3つの指標を用いることにした。これらをニコニコ動画の動画、「小説家になろう」の小説、学術論文について適用し、推移を見た。

- ・ 情報エントロピー : 単語あたりのエントロピーによる多様性評価
- ・ ジニ係数 : 単語の出現頻度による占有率
- ・ Cos 類似度の総和 :

次に利用者のコンテンツ選択モデルの確立を試みた。まず動画と小説について閲覧回数の分布を調査した。次に、回帰分析で、その分布となるモデル構築を試みた。クラスタリングの適用により、どの分野に偏るため多様性が減少するのか調査した。

1.2.4 研究成果

「ニコニコ動画」の楽曲、「小説家になろう」の小説について、タイトルおよびタグに出現する単語について、情報エントロピー、ジニ係数、Cos 類似度の総和、の3つの指標の値を算出した。その結果、単語あたりの情報エントロピーは値の変化が乏しく使えないことが分かった。経済格差を示すジニ係数を単語の頻度に適用した所、値の変化が見えないため、単語に基づく多様性分析には適していないことが分かった。

Cos 類似度の総和指標を用いて多様性の動向を測定した。月ごとの新規投稿コンテンツについて測定した結果、多様性の減少を定量的に確認できた。図1にニコニコ動画のCos 類似度総和、図2にCos 類似度総和の推移を示す。

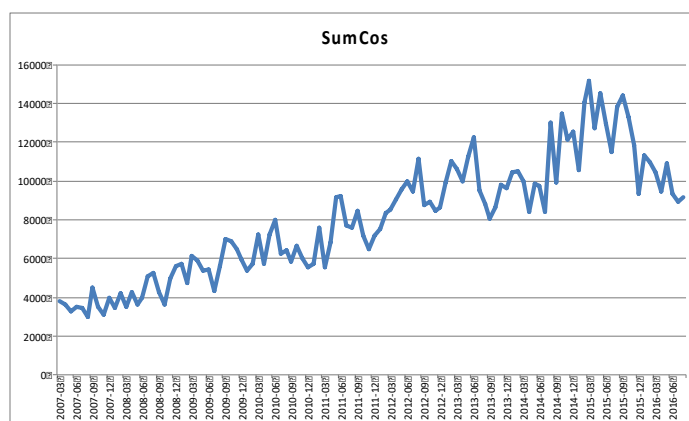


図1：ニコニコ動画 各月に投稿された動画のタグ単語のSumCos値

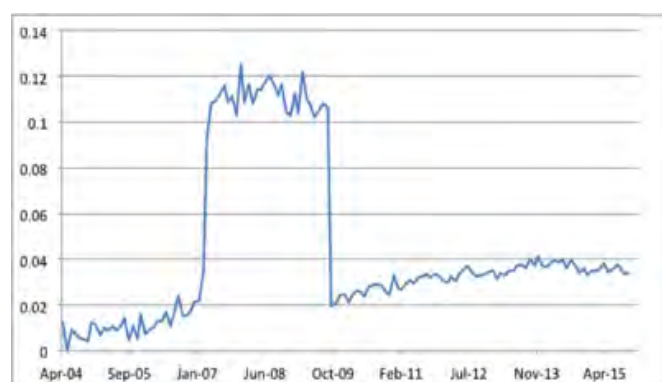


図2：小説家になろう 各月に投稿された小説のキーワード単語のSumCos値

学术论文についても同様の分析を適用した所、学术论文では多様性減少の傾向は見られなかった。LDA (Latent Dirichlet Allocation) による潜在トピック数の分析を学术论文集合に適用した。その結果、論文数が増えるとトピック数が増えるという、当たり前の結果が得られた。これらの結果から、学术论文では多様性がほぼ減少していないことが分かり、学术论文を分析対象から除外した。

動画と小説について閲覧回数の分布を調査した。統計ソフト R を用いた非線形回帰分析で分布を調べた。その結果、動画と小説の両方が閲覧回数の分布は対数正規分布であることが分かった。別件で調査した学内電子メールの容量分布も対数正規分布となり、人間社会における普遍的な分布である可能性が見えた。

閲覧回数が対数分布となるコンテンツ選択のモデルを構築するため、回帰分析を適用した。ニコニコ動画の閲覧回数を線形回帰分析および深層学習（6層のマルチパーセプトロン）で分析した。その結果、高い精度で閲覧回数を推定出来たものの、コンテンツの選択モデル構築には至らなかった。多様性が減少しているならば、どこかの分野に偏っているはずである。クラスタリングの適用により、偏りを分析した。ニコニコ動画の動画タイトル・説明文・タグからなる文書を Doc2Vec でベクトル化してクラスタリングを適用した。しかしながら、おおまかな傾向は分かるものの、詳細な偏り分析まで至らなかった。

1.2.5 主な発表論文等

〔雑誌論文〕(計 6件)

- [1] Eisuke Ito, Yuya Honda, Sachio Hirokawa: Empathy factor mining from reader comments of e-manga, Proc. of eKnow2018, IARIA, pp.107-112, Mar., 2018.(ISBN: 978-1-61208-620-0)
- [2] Eisuke Ito, Yuya Honda: Keyword diversity trend of consumer generated novels, Proc. of ICCESS2017., pp.140-147, 2017.
- [3] Kazuhisa Noguchi, Tomoya Iida, Eisuke Ito: An analysis of CGM contents pageview using SIR Model and GBM, Proceedings of ICCTD2017 (2017 8th International Conference on Computer Technologies and Development) pp.19-21, 2017.
- [4] Kyohei Kamihata, Eisuke Ito: A quantitative contents diversity analysis on a consumer generated media site, Proc. of AROB 21st 2016 (The Twenty-First International Symposium on Artificial Life and Robotics 2016), pp.436-440, 2016.
- [5] Zechen Wu, Eisuke Ito: Correlation Analysis between User's Emotional Comments and Popularity Measures Proc. of IIAI-AAI 2014/ESKM 2014, pp.280-283, September 3, 2014.
- [6] Xiao Lin, Eisuke Ito, Sachio Hirokawa: Chinese Tag analysis for foreign movie contents, Proc. of IEEE/ACIS ICIS2014, pp.163-166, June 4, 2014. Jun ZENG, Brendan FLANAGAN, Sachio HIROKAWA and Eisuke ITO: A Web Page Segmentation Approach Using Visual Semantics, IEICE TRANSACTIONS on Information and Systems, Vol.E97-D, No.2, pp.223-230, Feb.01 2014.

〔学会発表〕(計 9件)

- [7] 柴田知親, 伊東栄典: 回帰分析によるCGM動画再生回数推定, 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), 日本データベース学会, C5-2, 2012.
- [8] 佐嘉田悠樹, 伊東 栄典: クラスタリングによる利用者投稿動画の偏在性分析, 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), 日本データベース学会, G4-2, 2012.
- [9] 飯田委哉, 伊東栄典, 佐嘉田悠樹: クラスタリングによるオンライン小説の多様性動向分析, 情報処理学会 火の国情報シンポジウム 2018, B1-4, Mar.1, 2018.
- [10] 本田優也, 伊東栄典, 廣川佐千男: 読者コメント連動型漫画における人物の登場頻度分析, 人工知能学会 SIG-SWO-043-02, pp.02_1--02_6, Nov.25, 2017.
- [11] 佐嘉田悠樹, 伊東栄典: CGM 百科辞典を用いた利用者投稿動画クラスタリング, 平成29年度電気・情報関係学会九州支部連合大会, 11-2P-06, pp. 544-545. 2017.
- [12] 柴田知親, 伊東栄典: 回帰分析を用いた動画再生回数の推定, 平成29年度 電気・情報関係学会九州支部連合大会, 11-2A-08, pp.367-368. 2017.
- [13] 本田優也, 伊東栄典: 利用者投稿型小説サイトにおけるキーワードの多様性分析, 情報処理学会 火の国情報シンポジウム 2017, A2-4, Mar.1, 2017.
- [14] 上畑恭平, 伊東栄典: タイトルと説明文に着目した利用者投稿サイト動画の多様性分析, 情報処理学会 火の国情報シンポジウム 2016, 4A-3, 2016.
- [15] 上畑恭平, 伊東栄典: タグの類似度に着目した利用者投稿サイト動画の多様性分析, 信学技報, vol.115, no.381, AI2015-40, pp.83-88, 2015.

1.3 研究内容紹介

1.3.1 廣川佐千男

研究内容

研究のスローガンは、「データ*目的*技術」と体力です。文書群を「データ」として、具体的な「目的」に応じて、意味のある情報を抽出する「テキストマイニング技術」の研究を行っています。例えば、これまで、診療記録から長期入院患者の特徴抽出、営業日報から営業マンの売上推定、学生の授業コメントから成績予測、万葉集と古今集の自動識別、Web健康情報の信頼性評価、有価証券報告書から企業の比較、倒産情報から倒産理由抽出、特許情報から研究組織の構造推定などの研究を行い、国際会議やジャーナルあるいは特許として発表してきました。研究の手順としては、まず、対象の文書群についての専用検索エンジンを作り、分析観点を色々切り換え、数万回の検索を自動的に繰り返します。結果は統計的に集計したり、機械学習で分類したり、特徴語の関連を可視化したりします。計算機の力を借りて、人手では容易に実現できない、文書データの深い理解を目指しています。

所属学会名

人工知能学会, 情報処理学会, 電子情報通信学会

主な研究テーマ

- 手掛語と内容語の双対ブートストラップ・マイニング
- データマップ法と概念グラフによる次世代検索エンジンの研究開発
キーワード：検索エンジン, マトリックス検索, 概念グラフ, 大学発ベンチャー, 2006.08～.
- リンク情報と Web データの半構造性を融合した高品質コンテンツ・マイニング
キーワード：Web マイニング, データマイニング, WWW, XML, 文書検索, 検索エンジン, 特許検索, データベース, 2001.04～.
- 専門検索サイトの動的統合による次世代検索システムの研究開発
キーワード：検索エンジン, メタサーチ, 1993.04～.
- 構成的論理の推論構造の研究
キーワード：証明論, 型理論, ラムダ計算, 線形論理, 適切さの論理, 直観主義論理, 1997.01～.

研究プロジェクト

- ・ テキストマイニングによる医療プロセスのキーファクター抽出
2015.04～, 代表者: 廣川佐千男

研究業績

・ 原著論文

1. Eisuke Ito, Yuya Honda, Sachio Hirokawa, Empathy factor mining from reader comments of e-manga, Proc. of eKnow 2018, IARIA, pp.107-112, 2018
2. Yoshiki Mashima, Takuya Okada, Sachio Hirokawa, Kazuhiro Takeuch, Predicting Purpose of Program from Superficial Structure, Proc. ASEAN-AI2018, in press
3. Naoya Onimura, Takanori Yamashita, Naoki Nakashima, Hidehisa Soejima, Sachio Hirokawa, Machine Learning Support for Template Design of Clinical Notes, Proceedings of IEEE Conference on Biomedical and Health Informatics (BHI) 2018, 2018
4. N. Onimura, T. Yamashita, N. Nakashima, H. Soejima, and S.Hirokawa Interactive Visualization of Template Graph for Daily Clinical Notes, Proc. of the Eighth International Conference On Advances in Computing, Electronics and Electrical Technology - CEET 2018, pp.19-24, 2018
5. Kumiko Kanekawa, Tetsuya Nakatoh, Takahiko Suzuki, and Sachio Hirokawa: "Assessment of Doctoral Supervision of International Students," Proc. of International Conference New Perspectives in Science Education, pp.415-421, 2018
6. Takanori Yamashita, Naoki Nakashima, Sachio Hirokawa, Classification and feature extraction for text-based drug incident report 6th International Conference on Bioinformatics and Computational Biology, ICBCB 2018 Proceedings of 2018 6th International Conference on Bioinformatics and Computational Biology, ICBCB 2018, pp.145-149, 2018
7. Yao Lin, Kohei Yamaguchi, Tsunenori Mine, Sachio Hirokawa, Is SVM+FS better to satisfy decision by majority?, 3rd International Conference on Soft Computing and Data Mining, SCDM 2018 in Advances on Soft Computing and Data Mining - Proceedings of the 3rd International Conference on Soft Computing and Data Mining SCDM 2018, pp. 261-271, 2018
8. Toshiro Minami, Sachio Hirokawa, Yoko Ohura, Kiyota Hashimoto, A part-of-speech-based exploratory text mining of students' looking-back evaluation 11th International Symposium on Natural Language Processing, SNLP-2016 and 1st Workshop in Intelligent Informatics and Smart Technology, 2016, Advances in Natural Language Processing, Intelligent Informatics and Smart Technology - Selected Revised Papers from the 11th International Symposium on Natural Language Processing SNLP-2016 and the 1st Workshop in Intelligent Informatics and Smart Technology, pp.61-72, 2018
9. Takanori Yamashita, Naoya Onimura, Hidehisa Soejima, Naoki Nakashima, Sachio Hirokawa, Graph Clustering System for Text-Based Records in a Clinical Pathway, Studies in Health Technology and Informatics, Vol.245, pp.649-652, 2018
10. Tetsuya Nakatoh, Kenta Nagatani, Kumiko Kanekawa, Takahiko Suzuki, Sachio Hirokawa, Cluster analysis of scientific citation context, 19th International Conference on Information Integration and Web-Based Applications and Services, iiWAS2017, International Conference on Information Integration and Web-Based Applications and Services, iiWAS2017 - Proceedings, pp. 111-115, 2017

11. Yudai Tanabe, Koki Kagari, Yuki Kitanaka, Kazuhiro Takeuchi, Sachio Hirokawa, Finding key integer values in many features for learners' academic performance prediction, 9th International Conference on Education Technology and Computers, ICETC 2017 eedings of the 9th International Conference on Education Technology and Computers, ICETC2017, pp.167-171, 2017
12. Yuki Kitanaka, Kazuhiro Takeuchi, Sachio Hirokawa, Predicting learning result of learner in e-learning course with feature selection using SVM, 9th International Conference on Education Technology and Computers, ICETC 2017 eedings of the 9th International Conference on Education Technology and Computers, ICETC 2017, pp.122-125, 2017
13. Tetsuya Nakatoh, Sachio Hirokawa, Toshiro Minami, Takeshi Nanri, Miho Funamori, Attribute-based quality classification of academic papers, Artificial Life and Robotics, pp.1-6, 2017
14. A Multi-model SVR Approach to Estimating the CEFR Proficiency Level of Grammar Item Features Brendan Flanagan, Sachio Hirokawa, Emiko Kaneko, Emi Izumi, Hiroaki Ogata 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017 eedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017, pp. 521-526, 2017
15. Jun Zeng, Yinghua Li, Feng Li, Junhao Wen, Sachio Hirokawa, A Point-of-Interest Recommendation Method Using Location Similarity, 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017 eedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017, pp.436-440, 2017
16. Toru Sugihara, Soichiro Aihara, Sachio Hirokawa, Takashi Nara, An Analysis of Characteristics of Student-Athletes from Questionnaire by SVM, 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017 eedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017, pp.163-166, 2017
17. Kumiko Kanekawa, Tetsuya Nakatoh, Takahiko Suzuki, Sachio Hirokawa, Who is the Last Author of Your Paper? 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017 Proceedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017, pp.221-224, 2017
18. Sachio Hirokawa, Takahiko Suzuki, Tsunenori Mine, Machine learning is better than human to satisfy decision by majority, 16th IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017 eedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017, pp.694-701, 2017
19. Takahiko Suzuki, Koki Miyata, Sachio Hirokawa, Difficulty of words and their ambiguity estimated from the result of word sense disambiguation, 11th International Conference on Knowledge, Information and Creativity Support Systems, KICSS 2016 Proceedings -11th 2016 International Conference on Knowledge, Information and Creativity Support Systems, KICSS 2016, pp. 1-5, 2017
20. Yusuke Adachi, Naoya Onimura, Takanori Yamashita, Sachio Hirokawa, Classification of imbalanced documents by feature selection, 2017 International Conference on Compute and Data Analysis, ICCDA 2017 proceedings of 2017 International Conference on Compute and Data Analysis, ICCDA 2017 Part F130280, pp.228-232, 2017
21. Jun Zeng, Feng Li, Brendan Flanagan, Sachio Hirokawa, LTDE A layout tree based approach for deep page data extraction, IEICE Transactionson Information and Systems, Vol.E100D, No. 5, pp.1067-1078, 2017
22. Hiroaki Ogata, Misato Oi, Kousuke Mohri, Fumiya Okubo, Atsushi Shimada, Masanori Yamada, Jingyun Wang, Sachio Hirokawa, Learning analytics for E-book-based educational big data in higher education, Smart Sensors at the IoT Frontier, pp.327-350, 2017

研究資金

- **科学研究費補助金**

1. 2015年度～2017年度，基盤研究（B），代表，テキストマイニングによる医療プロセスのキーファクター抽出

教育活動

- **教育活動概要**

大学院システム情報科学府情報知能工学専攻の協力講座として広域分散データ特論・演習を担当し、インターネット検索エンジンなどの基礎となる技術を教えています。大学院統合新領域ライブラリーサイエンス専攻博士課程も担当しています。また、学部の講義では工学部電子情報工学科において、プログラミング論 I を担当しています。同学科の学部学生についての卒業研究と、大学院システム情報科学府情報知能工学専攻ならびに大学院統合新領域学府ライブラリー専攻の修士、博士課程の指導も行っています。

1.3.2 鈴木 孝彦

研究内容

- ・ 九州大学教務事務システムの作成と運用支援
- ・ 九州大学学務情報システムの計画支援
- ・ 日本語 WordNet およびテキストマイニング
- ・ 機械学習

所属学会名

情報処理学会

主な研究テーマ

- ・ 統計数値の不整合と関連テキストのテキストマイニング
キーワード：Benford の法則，属性選択，SVM，2017.08～2019.03.
- ・ 特徴語抽出による感性分析
キーワード：感性分析，特徴語，語のベクトル化，2015.05～2017.05.
- ・ 混合道路交通における低速車両の移動ロケーションデータの収集と分析方法の研究
キーワード：ロケーションデータ，Web マイニング，低速車両，2012.04～2017.04.

研究業績

- ・ 原著論文
 1. Kumiko Kanekawa, Takahiko Suzuki, Tetsuya Nakatoh, Sachio Hirokawa, Analysis of Scientific Citation Context, Proc AICS 2017, 2017.08.
 2. Kumiko Kanekawa, Tetsuya Nakatoh, Takahiko Suzuki, Sachio Hirokawa, Analyzing Researcher Stage with Last Authorship Ratio: Who is the last author of your paper?, Proc. of 6th International Congress on Advanced Applied Informatics (IIAI-AAI2017), ACT CITY Hamamatsu, Hamamatsu, Japan, 9-13 July 2017., 2017.07.
- ・ 学会発表
 1. 草刈祐子, 鈴木孝彦, 廣川佐千男, 有用なレビューの特徴抽出, 火の国シンポジウム, 2018.03.

2. 鈴木孝彦, 廣川佐千男, 統計数値における不整合部分検出の試み, 情報・統計科学シンポジウム, 2017.10.

教育活動

- 担当授業科目

1. 2017年度・後期, 情報処理概論
2. 2017年度・後期, プログラミング言語特論
3. 2017年度・後期, 情報処理概論

1.3.3 伊東 栄典

研究内容

- ・ 情報検索・情報統合・情報連携

Web上や電子データとして蓄積されている膨大なデータから、意味のある知識抽出や、抽出した知識を統合する研究を行っている。具体的には、Webデータからの情報抽出、利用者コメントからの知識発見、コメントやリンク構造を利用したコンテンツ推薦などを行っている。ソーシャルブックマークからの新規情報の発見や、「クチコミ」と呼ばれる情報サービスからの知識抽出について研究している。

- ・ 集合知を利用した高品質コンテンツ検索

情報検索および知識発見に関する研究の具体的な対象として、ネット上のコンテンツを対象とした研究を行なっている。近年、ユーザ投稿型のコンテンツサービスが普及しており、動画・写真・静止画・小説などが増大している。これらのサービスでは、投稿機能だけでなく、視聴者からのコメント・タグ付け・リンクなどのフィードバック機能もある。視聴者からのフィードバックは、Folksonomyや集合知と呼ばれるもので、これを活用したコンテンツ検索を行なっている。

- ・ 電子認証基盤および認証フェデレーション構築

特定メンバーへの情報提供サービスや、オンラインでの商取引および申請のように、利用者を正しく認証する電子認証が求められている。本研究では電子認証機構および認可機構の構築と、情報サービス連携のための認証連携機構の構築を目的としている。多様な組織が柔軟にサービスを提供しあうための、電子認証基盤の構築についての研究開発を行う。その実現のために、同じポリシー・標準規格での認証基盤となるフェデレーションについての研究開発を行う。電子認証基盤として、認証のためのデータベース構築、アプリケーションとなる情報サービスでの認証・認可機構の実現などが課題となる。組織内で使う、柔軟な利用者認証・認可システムの実現を目指す。

所属学会名

情報処理学会，電子情報通信学会

主な研究テーマ

- ・ 利用者投稿型コンテンツの多様性動向分析
キーワード：コンテンツ，利用者投稿型，多様性，エントロピー，コンテンツ選択モデル，2014.04～2018.03.

- ・ 大規模データ解析
キーワード：大規模データ，分散処理，統計解析，クラウド・コンピューティング，2012.06～.
- ・ 情報検索・情報統合・情報連携
キーワード：Information retrieval/extraction/integration/discovery, Data mining, Web mining, 2000.04～.

研究プロジェクト

- ・ 利用者投稿型コンテンツの多様性動向分析
2015.04～2018.03, 代表者：伊東栄典，基盤研究C
九州大学サービス開始から数年経過した CGM (Consumer Generated Media) では、コンテンツの多様性喪失(画一化)が進んでいる。画一化は文化の発展を阻害する。本研究では「ニコニコ動画」の楽曲、「小説家になろう」の小説、学術論文を対象に多様性の動向を分析する。閲覧回数上位のコンテンツを対象に、タグやキーワードの種類数を時系列で見ることで、多様性喪失(画一化)を定量評価する。また、閲覧回数の分布から多様性喪失の遠因となる利用者のコンテンツ選択モデルを確立する。利用者の選択モデルは、新たなランキング手法の提案や、外部刺激によるコンテンツの多様性確保に寄与できる。比較対象として学術論文も分析する。

研究業績

・ 原著論文

1. Eisuke Ito, Yuya Honda, Sachio Hirokawa, Empathy factor mining from reader comments of e-manga, Proc. of eKnow2018 (The Tenth International Conference on Information, Process, and Knowledge Management), 107-112, 2018.03
2. 佐嘉田 悠樹, 伊東 栄典, クラスタリングによる利用者投稿動画の偏在性分析, 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), G4-2, 2018.03
3. 柴田 知親, 伊東 栄典, 回帰分析によるCGM動画再生回数推定, 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), C5-2, 2018.03
4. 飯田 委哉, 伊東 栄典, 佐嘉田 悠樹, クラスタリングによるオンライン小説の多様性動向分析, 情報処理学会 火の国情報シンポジウム2018, B1-4, 2018.03
5. 本田 優也, 伊東 栄典, 廣川 佐千男, 読者コメント連動型漫画における人物の登場頻度分析, 人工知能学会, 02_1-02_6, 2017.11
6. Eisuke Ito, Yuya Honda, Keyword diversity trend of consumer generated novels, Proceedings of ICES2017 (The Third International Conference on Electronics and Software Science), 2017.08

- **学会発表**

1. 柴田 知親, 伊東 栄典, 廣川 佐千男, 回帰分析を用いた動画再生回数の推定, 平成29年度 電気・情報関係学会九州支部連合大会, 2017.09
2. 佐嘉田 悠樹, 伊東 栄典, 廣川 佐千男, CGM 百科辞典を用いた利用者投稿動画クラスタリング, 平成29年度 電気・情報関係学会九州支部連合大会, 2017.09

研究資金

2015年度～2017年度, 基盤研究(C), 代表, 利用者投稿型コンテンツの多様性動向分析

教育活動

- **担当授業科目**

1. 2016--2018年度: 人工知能 (工学部電気情報工学科・理学部情報理学コース)
2. 2011--2018年度: プログラミング演習1 (工学部電気情報工学科)
3. 2009--2017年度: 暗号と情報セキュリティ特論 (システム情報科学府)

大学運営

- **学内運営に関わる各種委員・役職等**

1. 2012.06～, 情報統括本部・認証基盤事業室
2. 2011.04～, 情報統括本部・全学基本メール事業室

1.3.4 中藤 哲也

研究内容

研究活動：

- ・ 研究動向分析、研究調査手法に関する研究

研究者が研究を行うに際して、関連する研究の調査は非常に重要なタスクです。ジャーナルインパクトファクタなどの指標を用いて読むべき論文誌を選択したり、引用数を用いて重要な論文を探し出したりする作業は、地道で手間の掛かる作業ですが非常に重要です。本研究では、より適切な論文誌、あるいは論文を発見するための新たな指標を提唱しております。また、研究分野の全体像を把握し、キーとなる研究の研究を発見するために、関連研究の全体像を可視化する手法の提案も行っております。

- ・ マイナートピックの発見・抽出手法に関する研究

テキストマイニングの研究においては、特徴語の頻度をベースとした情報抽出を行うことが一般的です。一方で頻度の低い語の中にも、興味深いトピックを構成している語が存在し、それらをうまく特定することが出来れば、有用な情報として利用可能です。本研究では、観光情報を対象として、マイナーであるが重要な情報をうまく特定し、そのような情報抽出に信頼性を付与する新たな手法を提案しております。

業務：

学務情報処理に関する業務をサポートしております。

また、九州大学情報統括本部の情報基盤に関する広報を担当しております。

所属学会名

情報処理学会, 人工知能学会, 言語処理学会, 日本データベース学会, 電気学会, 電子情報通信学会

主な研究テーマ

- ・ 研究動向分析、研究調査手法に関する研究

キーワード：研究動向分析, 計量書誌学, 2014.04～.

- ・ テキストマイニング

キーワード：テキストマイニング, データマイニング, ウェブマイニング, 情報抽出, 検索エンジン, 観光情報, 2010.04～.

- ・ 近似文字列照合アルゴリズムを用いたウェブマイニング手法の研究
キーワード：文字列照合アルゴリズム，テキストマイニング，2002.04～.

研究業績

・ 原著論文

1. Tetsuya Nakatoh, Sachio Hirokawa, Toshiro Minami, Takeshi Nanri, Miho Funamori, Attribute-based quality classification of academic papers, Journal "Artificial Life and Robotics", <https://doi.org/10.1007/s10015-017-0412-z>, Volume 23, Issue 2, pp.235-240, 2018. (in press)
2. Tetsuya Nakatoh, Toshiro Minami, Reducing Computational Effort for Plagiarism Detection with Approximate String Matching, Proc. of SCDM-2018, AICS volume 549 by Springer, pp.429-435, 2018.02
3. Yasuhiro Yamada, Yuusuke Himeno, Tetsuya Nakatoh, Weighting of Noun Phrases Based on Local Frequency of Nouns, Proc. of SCDM-2018, AICS volume 549 by Springer, pp.436-445, 2018.02.
4. Tetsuya Nakatoh, Kenta Nagatani, Kumiko Kanekawa, Takahiko Suzuki, Sachio Hirokawa, Cluster Analysis of Scientific Citation Context, Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2017), 2017.12.
5. Tetsuya Nakatoh, Kenta Nagatani, Toshiro Minami, Sachio Hirokawa, Takeshi Nanri, Miho Funamori, Analysis of the Quality of Academic Papers by the Words in Abstracts, HIMI 2017, Part II, LNCS 10274, Proc. of the 19th International Conference on Human-Computer Interaction (HCI International 2017), 2017., 10.1007/978-3-319-58524-6_34, pp.434-443, 2017.07.
6. Kensuke Baba, Tetsuya Nakatoh, Toshiro Minami, Plagiarism Detection Using Score Vectors Weighted by Distributed Representation of Words, HIMI 2017, Part II, LNCS 10274, Proc. of the 19th International Conference on Human-Computer Interaction (HCI International 2017), 2017., 10.1007/978-3-319-58524-6_34, pp.341-350, 2017.07.
7. Kumiko Kanekawa, Tetsuya Nakatoh, Takahiko Suzuki, Sachio Hirokawa, Analyzing Researcher Stage with Last Authorship Ratio: Who is the last author of your paper?, Proc. of 6th International Congress on Advanced Applied Informatics (IIAI-AAI2017), ACT CITY Hamamatsu, Hamamatsu, Japan, pp.9-13, 2017.07.
8. Tetsuya Nakatoh, Emi Ishita, Satoru Uchida, Toru OGA, Performance Comparison on Automated Generation of Coding Rules: A Case Study on ISO 26000, International Journal of Service and Knowledge Management, Vol. 1, No. 1, pp.19-30, 2017.06.

研究資金

・ 科学研究費補助金

1. 2016年度～2020年度，基盤研究(B)，分担，国連グローバル・コンパクトを中心としたグローバルCSRレジームの研究
2. 2015年度～2017年度，基盤研究(C)，代表，Webからのマイナートピック抽出における信頼性確保に関する研究
3. 2015年度～2017年度，基盤研究(B)，分担，テキストマイニングによる医療プロセスのキーファクター抽出

- 日本学術振興会（科学研究費補助金以外）
 1. 2017年度～2019年度，課題設定による先導的人文学・社会科学研究推進事業・グローバル展開プログラム，分担，グローバル社会における排他主義とデモクラシーに関する総合的研究
- 学内資金・基金
 1. 2017年度～2017年度，平成29年度QRプログラム，代表，学術論文調査のための論文内容の比較手法に関する研究
 2. 2015年度～2017年度，九州大学教育研究プログラム・研究拠点形成プロジェクト，分担，デジタル人文学における研究資源オープン化と研究プラットフォーム構築に向けた基盤的研究

教育活動

- 担当授業科目
 1. 2017年度・前期，情報処理概論
 2. 2017年度・後期，情報システム論

大学運営

- 学内運営に関わる各種委員・役職等
 1. 2017.06～ スポークスパーソン
 2. 2014.04～ 情報統括本部 広報事業室 室長
 3. 1994.05～ 入学試験機械処理委員会委員