

# LDA を用いたレポート推薦機能を持つ e ポートフォリオシステム

## ePortfolio System with Reports Recommender Function Based on LDA

加藤 嘉浩<sup>\*1</sup>, 石井 隆稔<sup>\*2</sup>, 植野 真臣<sup>\*1</sup>  
Yoshihiro KATO<sup>\*1</sup>, Takatoshi ISHII<sup>\*2</sup>, Maomi UENO<sup>\*1</sup>

<sup>\*1</sup> 電気通信大学大学院情報システム学研究科

<sup>\*1</sup> Graduate School of Information Systems, University of Electro-Communications

<sup>\*2</sup> 首都大学東京

<sup>\*2</sup> Tokyo Metropolitan University

Email: y-kato@ai.is.uec.ac.jp

あらまし：本研究では、e ポートフォリオ内の学習者のレポートの意味・内容を分析し、他者の有用なレポートを推薦するシステムを開発した。具体的には、蓄積された学習者のレポートを文書分類手法であるトピックモデル (LDA) を用いて分析する。LDA は、文書中の単語頻度情報から、話題やカテゴリを意味するトピックを推定可能である。本システムの特徴は、①LDA を用いレポートの内容からトピックを推定可能である。②同一トピックで、表現や論述方法が異なる多様なレポートを推薦可能である。③推薦された他者のレポートを参考にし、学習者が自身のレポートを修正すると期待できる。

キーワード：e ポートフォリオ, 推薦システム, 学習コミュニティ, 認知的徒弟制, トピックモデル

### 1. はじめに

近年、学習理論の主流は、ヴィゴツキーに代表される社会的構成主義に移行しつつある。ヴィゴツキー理論は、その後、物理的なスキルや過程よりも「メタ認知」や「経験による学習」を重視した認知的徒弟制に引き継がれてきた。COLLINS らは、学校における具体的な教育法として、①モデリング (modeling)：学習者が観察・模倣できるように教師が課題遂行の成功例を見せる、②コーチング (coaching)：学習者が課題を模倣する間、教師が観察・励まし・ヒント・アドバイスなどのあらゆる学習促進を与える、③足場かけ (scaffolding)：コーチングの一部とも解釈されるが、学習者の課題実行を支援して成功に導くことであり、学習者の能力に適切に提供される。最終的には支援は減少し、フェーディングさせる、④詳述 (articulation)：学習者が自身の知識や思考を言語化し、外化する、⑤省察 (reflection)：学習者が自身の遂行と他者の遂行の差異を比較できるようにする、⑥探索 (exploration)：学習者自身が解決すべき問題を発見できるような段階に導く、を提案している<sup>(1)</sup>。

さらに、認知的徒弟制を実現するための有用な学習環境として、LAVE and WENGER は、多様な熟達レベルの参加者がスキルについて活発に話しあったり、従事したりするインフォーマルな「実践コミュニティ」を提唱している<sup>(2)</sup>。さらに、この考え方を学校教育にあてはめて、「学習コミュニティ」という教育アプローチが提案された<sup>(3)</sup>。学習コミュニティは、①多様な熟達レベルの参加者が参加し、それぞれはその貢献度によって評価され、成長のための支援を受ける、②知識やスキルを発達させるという共通目標を持つ、③「学び方」を学ぶ、④学習成果を共有できる、といった特徴を持つ。この概念に基づき、様々な CSCL システムが開発されてきたが、同

時に同一トピックを学習するメンバによって構成される学習コミュニティを支援するので、メンバの熟達レベルの多様性が小さく、他者から学び方や学習成果を学べる範囲は限定される。

一方、長期間にわたり、多様な学習者の学習成果物、成績、学習日記、学習履歴などを蓄積して公開・共有する e ポートフォリオが近年、普及しつつある。e ポートフォリオは、そもそも学習者個人のリフレクションを促進するためのツールとして普及してきたが、Web サーバ上に大量に長年蓄積されてきた多様な学習者データは、学習コミュニティの特徴①～④を潜在的に兼ね備え、学習者が他者から学ぶための有用なツールとなりえる。以上の文脈より、植野らは、個人の e ポートフォリオを構造化し、ハイパーリンクでつなぐことにより、多様なパスで有用な他者情報の発見を支援する e ポートフォリオ・システムを開発している<sup>(4)</sup>。また、e ポートフォリオが大量になってくると参照すべきものを発見困難になる問題を解決するために推薦システムを開発している。しかし、この推薦システムは表層的な統計データ (既学分野、学習頻度、分野毎の理解度など) を用いて推薦を行っており、学習成果物の内容は考慮していない。

そこで本研究では、e ポートフォリオ内に蓄えられた学習成果物 (例えば、レポート) の意味・内容的な分析を行い、それを学習へ有効活用することを目指す。具体的には、他者の有用なレポートを参考に本人のレポート作成支援を行うシステムを研究開発する。意味・内容的な分析には、文書分類手法であるトピックモデルを用いる。

トピックモデルとは、文書中の単語は文書の潜在的な意味 (トピック) に依存して出現すると仮定し、文書中に出現する単語の頻度からそのトピックを推定する手法である。トピックモデルの代表例として、

Latent Dirichlet Allocation (LDA)<sup>(5)</sup>があり、高精度かつ計算効率が良いことから、これを採用する。

LDA は、文書中に複数個のトピックが混在すると仮定し、それぞれのトピックがどれだけの割合で含まれているか（以降、トピック分布と呼ぶ）を出現単語により推定する手法である。LDA を用いた推薦システムでは、LDA におけるトピック分布を用いて分布間の疑似距離の算出を行い、これの近いものを推薦する。本研究においてもトピック分布間の疑似距離の算出を行う。

また、レポート特有の問題を考慮したレポートの推薦方法を提案する。学習者のレポートとの類似度が高いレポートを推薦する際、内容や表現などの差異が少ない場合、参考にする箇所が少なくレポートの推敲を促すことが難しいと考えられる。したがって、LDA により推定したトピック分布の類似度が高いレポート群を抽出し、その中から出現単語の類似度が低いレポートを推薦する。これにより、意味や内容が類似しつつ表現が異なった多様なレポートを推薦できると考える。

## 2. 本システム概要

### 2.1 学習データ

本システムで用いる学習者のレポートは、植野が長年開発してきた e ラーニング・ポートフォリオ”Samurai”<sup>(4)</sup>に蓄積されている、実際の大学院の講義で課されたものを使用した。

レポートの課題は、テーラーの科学的管理論、リエンジニアリングの手法、ベンチャー精神の育成についてなどである。

### 2.2 推薦メカニズム

他者のレポートを推薦しレポート作成支援を行う場合、当該学習者のレポートと推薦レポートが類似しすぎている場合、差異が小さいため参考にならないなどの問題が考えられる。

本システムでは、e ポートフォリオ内の学習者のレポートの内容を考慮して以下のように推薦することで、①、②の問題を解決し学習者にとって有用なレポートを推薦できると考える。

1. LDA を用いてレポート内容を分析する。
2. 学習者のレポートと類似したトピック情報を持つレポート群を抽出する。
3. その中で使用している単語が類似しておらず、評価の高いレポートを学習者に推薦する。

LDA を用い内容を分析することで、課題との関連性を制御し、使用単語の類似性を考慮することで表現方法などが多様なレポートを推薦可能である。

図 1 は、本システムのレポート推薦画面を示す。図 1 上部には、推薦レポートとその情報が示され、下部には当該学習者のレポート情報が表示される。表示される推薦レポートの情報は、課題名、書き出し、トピック分布である。当該学習者のレポート情報は、出現単語ランキング、各トピックの単語分布

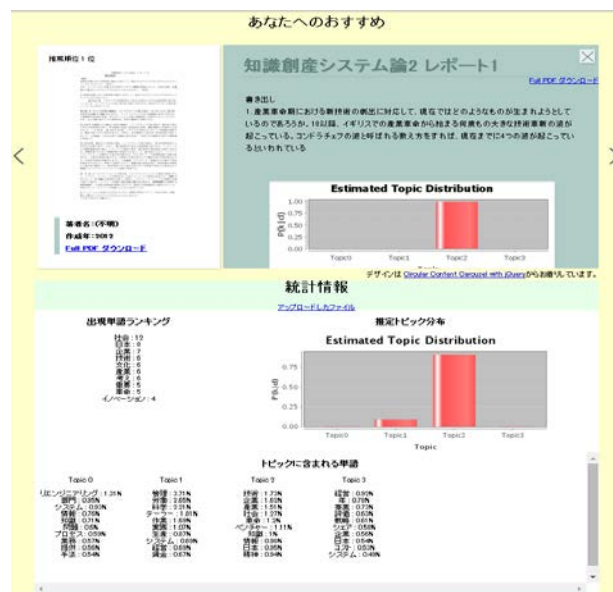


図 1 LDA を用いたレポート推薦システム

の上位語を表示している。詳細を表 1 に記載する。各トピックを代表する単語は、課題の内容を反映している。

表 1. 各トピックの単語分布の上位語

トピック 0	リエンジニアリング, 部門, システム, 情報, 知識, 問題, プロセス, 業務
トピック 1	管理, 労働, 科学, テーラー, 作業, 生産, 経営, 賃金
トピック 2	技術, 企業, 産業, 社会, 革命, ベンチャー, 知識, 日本, 精神
トピック 3	経営, 年, 事業, 評価, 戦略, シェア, 企業, 日本, コスト, システム

これらの情報と推薦されたレポートを参考にすることで、当該学習者に自信のレポートとの差異を学び、レポートの推敲を促す。

現在、4 年生大学の工学部の学生に本システムを利用してもらい、システム評価を行っている。

### 参考文献

- (1) Collins, A. and Brown, J.S. and Newman, S.E.: "Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics". In L.B. Resnick (Ed.), *Knowing, learning, and instruction: essays in honor of Robert Glaser* (pp. 453-494), Laurence Erlbaum Associates, NJ (1989)
- (2) Lave, J. and Wenger, E.: "Situated Learning: Legitimate Peripheral Participation", Cambridge University Press, CA, (1991)
- (3) Scardamalia, M. and Bereiter, C.: "Computer support for knowledge-building communities", *The Journal of the Learning Sciences*, 3, pp.265-283, (1994)
- (4) 植野真臣, 宇都雅輝: "他者からの学びを誘発する e ポートフォリオ", *日本教育工学会論文誌*, 35-3, pp.169-182, 2011
- (5) D.M. Blei, A.Y. Ng, and M.I. Jordan, *Latent Dirichlet Allocation*, *Journal of Machine Learning Research*, pp.993-1022. 2003.