

# テキストハイライト機能を持つ学習支援システムの開発

吉崎 弘一<sup>†</sup>中島 順美<sup>‡</sup><sup>†</sup>大分大学<sup>‡</sup>大分大学

## 1. 背景

Acrobat Reader 等の PDF ファイルビューアや Amazon Kindle 等の電子書籍リーダーの多くは、利用者がコンテンツ内で選択したテキストに、他と異なる背景色の「ハイライト」を設定する機能を持つ。一方、学習支援システムもその主要な機能として、デジタル教材を提示する機能を持つが、著者らが知る限り、従来の Web ベースの学習支援システムで、Web ブラウザ内で表示する教材のテキストに対して、ハイライトを設定する機能を標準実装したシステムは存在しない。なお、本研究で述べるテキストハイライト機能とは、デジタルコンテンツ内で利用者が選択した任意の文字列の見た目を変更し、操作以降に当該コンテンツを表示した際にも、他と異なる見た目を維持する機能を指す。

このような現状に対して本研究では、HTML/PDF ファイル形式の教材を対象に、テキストハイライトを設定する機能を、Web ブラウザ上で利用するオープンソースソフトウェアの学習支援システム LePo[1][2]に試験的に実装した。

本研究に関連する研究として、参考資料[3]が挙げられる。この研究では、HTML で構築された任意の Web ページの DOM (Document Object Model) を解析することで、ハイライトを挿入するブラウザ拡張機能を提案している。これに対して本研究では、学習支援システム内の HTML 及び PDF 教材をテキストハイライトの対象とし、特定の Web ブラウザを対象としていない点が異なる。

また、参考資料[4]では、教材 PDF ファイルに対してハイライト (マーカー) を付与する機能が実装されているが、選択した文字列を認識するアノテーションであるかは不明である。一方、本研究ではハイライト対象の文字列を認識してデータベースに保存し、ハイライト対象のファイルも PDF と HTML ファイルの 2 種としている。

## 2. 開発したテキストハイライト機能

本研究では、教授者がシステムにアップロードした HTML/PDF ファイルの教材を、ハイライトの対象としている。これらのファイルはシステ

ム内の iframe タグに表示され、利用者はオンラインで閲覧する。2 種のファイルに以下で述べる共通の DOM 解析を行い、ハイライトを設定した。

### 2.1. HTML ファイル教材へのハイライト

ブラウザで表示した HTML 教材から、利用者が選択した文字列をサーバに送信し、関連する情報と共にデータベースに保存する。その後、ブラウザ内で JavaScript を用いて、選択文字列に対するページ内検索を行い、文字列が見つかった場所に、スタイルシートで背景色を設定した独自タグを挿入する。その際に、選択した範囲に HTML タグが部分的に含まれていた場合でも、適切に独自タグを挿入する。ハイライトの表示箇所は、選択文字列に対する検索結果で指定するため、同一の文字列がページ内に複数存在する場合は、その全てにハイライトを表示する。

### 2.2. PDF ファイル教材へのハイライト

W3C が提唱する DOM は、HTML/XML ドキュメントのための API であるが、本研究では JavaScript ライブラリの PDF.js[5]を用いて、PDF ファイルを HTML に変換した後に、その DOM 解析を行った。

教授者がシステムにアップロードした PDF ファイルは、利用者が閲覧する際に、PDF.js を介して HTML ドキュメントに変換される。PDF.js では、PDF ファイルを、クライアントサイドの JavaScript で構文解析し、主に Canvas を用いたラスター画像として描画するレイヤーと、透明化した文字情報を保持するテキストレイヤーから構成された HTML ファイルとして表示する。後者の文字情報については、PDF.js が元々の PDF ファイルが持つ情報に基づき、ページ内の適切な場所に配置する。PDF ファイルのハイライト表示は、このテキストレイヤーから選択した文字列

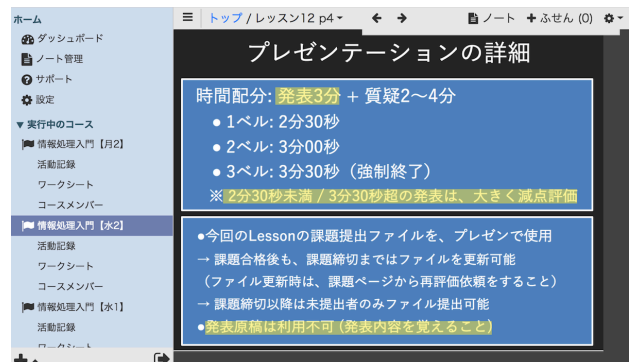


図 1 PDF 教材へのハイライト設定の一例

Development of Learning Management System with Text Highlight Function

<sup>†</sup> Koichi Yoshizaki, Oita University

<sup>‡</sup> Junmi Nakashima, Oita University

に対して、HTML 教材と同様の独自タグを挿入することで実現した(図 1)。なお、提案する手法でハイライトを設定できる PDF ファイルは、ページ内の文字列がフォントを用いて、そのまま可視化されている PDF ファイルであり、ページ内の情報を画像で保持し、文字列の情報を持たない PDF ファイルは設定できない。紙面をスキャナーで PDF ファイル化し、OCR によるテキストを透明レイヤーに保持した PDF ファイルは、作成方法に依存して、ハイライトを設定できる場合もある。

### 2.3. 利用者へのフィードバック

学習者が、HTML/PDF ファイルの教材に設定したテキストハイライトは、ページを表示した際に表示されるだけではなく、学習効果を高めるために、以下のフィードバックを行っている。

LePo では任意のコースを履修すると、そのコースの学習を記録する「レッスンノート」が、利用者ごとに作成される。コースは複数回のレッスンから構成されるが、レッスンで用いる教材ごとに、利用者が作成したハイライト等のアノテーションを自動的に集約表示し、アノテーションの散逸を防いでいる。また、教材に新たなハイライトを設定し、レッスンノートが更新されると、ログイン直後に表示するダッシュボードに、レッスンノートが更新されたコースを通知する仕組みも実装した。これにより、ハイライトを用いた学習の振り返りを促している。

## 3. 機能評価

### 3.1. ブラウザの対応状況

下記ブラウザに対して、テキストハイライト機能の動作を確認した。ブラウザのバージョンは、各 OS 上での最新安定版を用いた。現状では Windows と macOS では、対象の全ブラウザでハイライトの設定と表示が可能であり、iOS と Android では、ハイライトの表示のみ可能である。

- Windows 10: Internet Explorer 11 / Edge / Chrome / Firefox
- macOS 10.12: Chrome / Firefox / Safari
- iOS 11: Safari / Chrome
- Android 5: Chrome

### 3.2. PDF.js を用いた PDF ファイルの表示確認

PDF の表示に用いた PDF.js は、Firefox ブラウザの PDF ビューアとして標準実装され、広く使われている。その一方で、今回用いた PDF.js のバージョン 1.7.255 では、想定していなかった、いくつかの不具合が見られた。複数の環境で作成した PDF ファイルを PDF.js を用いて表示した際に、文中の空白文字がテキストレイヤーの文字列に保持されていたかどうかの確認結果を、表 1

表 1 様々な PDF ファイルと空白の表示

PDF 作成環境	PDF 作成操作	空白の保持
Windows 8.1 / PowerPoint 2013	名前を付けて保存 > PDF	保持しない
Windows10 / PowerPoint 2016	名前を付けて保存 > PDF	保持しない
Windows10 / PowerPoint 2016	印刷 > Microsoft Print to PDF	保持する
macOS 10.12 / Keynote 7	印刷 > PDF として保存	保持する
macOS 10.12 / PowerPoint 2016	印刷 > PDF として保存	保持しない

に示した。いくつかの環境で作成した PDF ファイルでは、PDF ファイル自体には空白を保持するものの、変換後の HTML ファイルのテキストレイヤーには空白を含まない分かち書きで文が記述されていたため、空白を含む文字列へのハイライトを、適切に設定できなかった。

## 4. まとめ

本研究では、学習支援システム LePo にアップロードした HTML/PDF 教材に、テキストハイライトを設定する機能を実装した。一部の環境で作成した PDF ファイルは、適切なハイライトが設定できない状況ではあるが、この点に注意すれば、実用的なテキストハイライト機能を実装することができた。今後は、開発したテキストハイライト機能の学習効果を評価していく予定である。

### 謝辞

本研究は JSPS 科研費基盤研究(C) (17K00490) の助成を受けたものである。

### 参考文献

- [1] LePo, <http://lepo.info/>
- [2] "LePo: An Open-Source Learning Management System with Text Annotation and Content Curation Functions", Koichi Yoshizaki, Hiroshi Hotta, The IAFOR International Conference on Technology in the Classroom - Hawaii 2017 Official Conference Proceedings, ISSN: 2432-1222, pp. 21-28, 2017
- [3] 三浦菜々, 池部実, 吉崎弘一, 吉田和幸, 学習支援システムと連携した情報収集のためのブラウザ拡張機能の開発, 第 69 回電気・情報関係学会九州支部連合大会, 314-314, 2016 年 9 月
- [4] BookRoll, <http://www.let.media.kyoto-u.ac.jp/wp-content/uploads/2017/10/48111e174f9ce44654e5e4a073896e1f.pdf>
- [5] PDF.js, <https://mozilla.github.io/pdf.js>