



GSLetterNeo Vol.95

2016年6月

テキストデータ群の重ね合わせによるヒストリと特徴のインタラクティブ表示(1)

松原 伸人、土屋 正人

matubara@sra.co.jp, m-tsuchi@sra.co.jp

◆はじめに

大量のデータをインタラクティブに操作する、Web ブラウザ上で動作するアプリケーションを開発しています。

「カタマリを見つけて辿れるヒストリデータブラウザ」(GSLetterNeo Vol.91, 92, 93, 94)は、歴史データや、開発ログといった大量のイベントを時間軸上に並べて表示することで、イベントが連続的に起きている時間帯や切れ目を見つけて辿れるようにする Web アプリケーションのプロトタイプです。

このプロトタイプでは、

- どのような実装をすれば大量のデータをインタラクティブに扱えるか？
- 実現したいインタラクションを、実装してみた結果としてインタラクティブ性能がどうなるか？

ということを試行錯誤して探っています。

◆プロトタイプの紹介

大量のテキストデータの、個々のテキストの特徴と時間関係をインタラクティブに見ていくビューアのプロトタイプを紹介します。



図 1 試作中のビューアの最初の画面



図 2 上にスクロールした画面



図 3 下にスクロールした画面

図 1 が試作中のビューアの画面で、使用したテキストデータは GSLetterNeo です。

GSLetterNeo の Vol.36 ~ Vol.94 のドキュメントごとに、

- 出版年月日、号、タイトル、著者(黒色)
- 各テキストから tf-idf¹で算出した特徴的な 10 語(青色)
- 各テキストから算出した使用数が少ない希少な 10 語(緑色)

を新しい順に縦にリスト表示しています。

画面上で上下にドラッグすると、レイヤーの面と平行に視界を動かします(図 1, 図 2, 図 3)。

このプロトタイプの場合、画面上部にあるテキストはサイズが大きく重なって表示され、画面下部のテキストはサイズがやや小さく重なりにくくなるように提示されています。

上下に視界を動かしながら、手前のテキストを読んだり、奥のテキストを読んだりします。テキストをクリックすると PDF ファイルを開きます。

◆ レイヤーと構成

ドキュメントのタイトルなどの情報を表示するレイヤー(図 4)、特徴語群を表示するレイヤー(図 5)、希少語群を表示するレイヤー(図 6)の 3 層を、画面奥に向かって重ねて配置することで、ドキュメントの時間関係と特徴がひとまとまりに見えるように表示しています。

図 4、図 5、図 6 では 3 つのレイヤーの仮想 3D 空間上での形とサイズを、Web Inspector でハイライト表示して示しています。



図 4 タイトルなどの情報を表示するレイヤー



図 5 特徴語を表示するレイヤー



図 6 希少語を表示するレイヤー

レイヤーを重ねて仮想空間に表示するビューアは、HTML と CSS と JavaScript で実装した web ブラウザ上で動く Web アプリケーションです。

各レイヤーの中身は、HTML の DIV などのエレメントで構成して CSS でレイアウトしています。

¹ 文書中の単語に関する重みの一種。

レイヤーの中身を作る際は、3つのレイヤーを仮想空間上に重ねて配置した時に、各ドキュメントのタイトル、特徴語群、希少語群がぴったり重なるように、各レイヤーのサイズおよびレイヤー内のオブジェクトの順番と大きさを揃えておきます。

仮想空間への各レイヤーの配置と視界の移動については、次号で紹介します。

夢を。

GSLetterNeo Vol. 95

2016年6月20日発行

発行者●株式会社 SRA 先端技術研究所

編集者●土屋正人

バックナンバを公開しています●<http://www.sra.co.jp/gsletter>

ご感想・お問い合わせはこちらへお願いします●gsneo@sra.co.jp



株式会社SRA

〒171-8513 東京都豊島区南池袋2-32-8

夢を。Yawaraka Innovation
やわらかいのべーしょん