

声優ブログの「ご報告」エントリー自動検出システムの検討

@MagnesiumRibbon

<http://voice-statistics.hatenablog.jp/>

Abstract

女性声優の結婚報告や事務所移籍の報告に代表される、いわゆる「ご報告」エントリーによる声優ファンへの精神的ダメージが問題になっている。そこで、ご報告エントリーを先読みして検出し、ユーザに警告を与えることで読んだあとのショックを和らげるようなシステムを構築できないかと考えた。本稿ではまずシステムで扱うご報告エントリーを定義し、またその中でも扱っている話題によって分類を試みた。次にそのようなエントリーに共通する特徴、あるいは通常のブログエントリーでは見られない特徴を列挙した。そして、自動検出システムの実現可能性について検討し、モデルを考案した。アメブロでオフィシャル化された全声優ブログを対象として実験をしたかったが、ピックアップした記事を手作業でアノテーションしたところで力尽きた。

Keywords: 声優統計, 声優ブログ, トピックモデル, 話題検出

1. はじめに

2012年に大きな話題となった声優の結婚ラッシュは2013年になってもとどまるところを知らず、また近いうちに結婚すると思われる声優のリストも長くなる一方である。[3]また、声優業界の再編の流れも止まらず、大物の移籍や新事務所の設立などのニュースが定期的に声ヲタたちの間をにぎわせている。このような状況下においても、声優が自身の重大なニュースをアナウンスする場が本人運営のブログにほぼ限られていることが問題となっている。すなわち、概ねファンサービスの場として機能している声優ブログにおいて進退に関わる記事が突如投稿されることで、覚悟無く記事を開いたファンに精神的なダメージを与えることが社会問題化している。このようなブログ記事のことを、よく記事タイトルに使われる言葉を取って「ご報告エントリー」と呼ぶことがある。

そこで、声優ブログが更新された段階で記事をチェックし、ご報告エントリーかどうかを判定し必要に応じて警告するシステムを開発することを考えた。もしこのようなシステムがあれば、事前に重大な報告が来ることを覚悟した上で記事を読むことができ¹、ダメージコントロールをすることが出来る。あるいは、「ご報告」と題するブログであっても実際には重大でない

記事（偽ご報告エントリー：仕事告知を報告と呼んでも不自然ではない等）にひるむ必要もなくなる。また、報告の種類について情報を提示することが出来れば、特定の話題についての報告にのみ注意を払っている場合にも有効に活用することが出来ると考えられる。

本稿ではまずご報告エントリーについてももう少し形式的な議論を行ったのち、ご報告エントリーに典型的な特徴について述べる。その上でご報告エントリー自動検出システムの設計を行い、具体的なモデルを与える。最後に実験を行うデータを収集し、アノテーションを行ってデータを分析する。

2. 既存研究

声優ブログと声優の結婚に関する先行研究としては、「ブログを用いた女性声優の結婚時期予測問題」[3]が挙げられる。しかし本稿では将来の予測については取り扱わない。あくまで、ブログにアップされた記事を人間より早く読む事でご報告エントリーかどうかを判別する問題に取り組む。また、先行研究では結婚が近づくにつれ声優ブログの内容が徐々に変化するという仮説を扱っているが、今回はあくまで結婚報告を含めたご報告エントリー記事そのものの特徴を扱う。

3. ご報告エントリー

ここでは本稿で取り上げる「ご報告エントリー」の定義を述べる。

Email address: magnesium.ribbon@gmail.com

(@MagnesiumRibbon)

¹もちろん記事を読む事を避けるという方法もあるが、ファンであれば望まない報告であるリスクがあっても記事を読まないという選択はないであろう。