

# 生きる悩みを訴えたブログのテキストマイニング

堀口裕太(和光大学)

## 1. 問題

本研究では、その対象としてウェブログ(以下、ブログと記す)を用いた。ブログは日常報告のような日記としての運用や、ニュース記事を引用してコメント付けを行うなど、様々な形で利用されている。近年ではメールアドレスさえ持っていれば登録することの出来る無料のブログサービスも普及しており、誰でも気軽に記事の投稿を開始することが可能である。また、従来紙を媒体として書かれてきた日記とは異なり、多くの場合この記事は第三者にも公開される。まったくの他人であっても閲覧が可能である状態で記事投稿を行い公開することは、単なる思考の文章化ではなく、自己表現の手段のひとつだと云えるだろう。そしてこのことは、広く他者に自分の言葉で思いや考えを発信することが出来るということでもある。

ブログやホームページなどのウェブサービスでは通常、定期的、或いは不定期的に記事が投稿され、最新記事が更新されていく。しかし、投稿を行ってきた個人や団体の都合によりこれがなされなくなる場合もある。投稿を行っていた人物が失踪したり、或いは亡くなったりしたことを理由に一切の更新がされなくなった状態にあるサイトは、ウェブ廃墟と呼ばれることがある。このウェブ廃墟に至るに際し、その理由が自殺であったとき、死の直前まで更新が続けられていたブログというものは当人の感情の変遷を辿る重要な材料になり得るだろう。

## 2. 目的

本研究では、故・二階堂奥歯氏の『八本脚の蝶』を対象に Text Mining Studio を用いた分析を行い、遺書のみから受け取る遺言だけでなく、連続性のあるブログ記事から逝去までの足跡をかえりみることの意義とその価値を確認したい。

## 3. 方法

**分析対象:** 『八本脚の蝶』の web サイト上の文字データを対象とした。URL: <http://homepage2.nifty.com/waterways/oquba/> のサイトからである。2001年06月13日に投稿された最初の記事から、2003年04月26日に投稿された最後の記事までの445記事を分析対象とした。著者は二階堂奥歯である。1977年に生まれ、大学卒業後は出版社で編集

者として働いていた。2003年4月26日に飛び降り自殺して亡くなった。なお、二階堂奥歯という名前はハンドルネームであり、本名は公開されていない。また、『八本脚の蝶』は2006年にポプラ社より書籍化され、2013年にはその第二版が出版されている。本文中には日常の報告のほか多くのブックレビューもみられ、読書家であった彼女の一面を垣間見ることが出来る。

**分析手順**：『八本脚の蝶』をテキストファイル化し、Microsoft Office Excel 2010にて日付と本文をタブ区切りした。そのデータをもとにText Mining Studioによって分析した。テキストマイニングを用いて、単語頻度解析や注目語分析などを行い、その後原文を参照するなどした。

**倫理的配慮**：本研究の対象は一般に公開されたブログであり、同様の内容で書籍化されていることから、著作権に配慮した。また、故人や遺族に対して悪意を以て分析に臨んだものでないことをここに明記する。

## 4. 結果と考察

### (1) 基本情報

表1 基本情報

| 項目        | 値     |
|-----------|-------|
| 総行数       | 445   |
| 平均行数(文字数) | 278.9 |
| 総文数       | 7313  |
| 平均文長(文字数) | 17    |
| 延べ単語数     | 45583 |
| 単語種別数     | 12702 |

タブ区切りデータを作成するにあたり、ひとつの記事を日付と本文に分けた際、本文がすべて一行となるようにした。そのため、表1に於ける総行数とは記事数と同意である。このとき、『八本脚の蝶』では一日に複数回更新された記事も散見されるため、記事数が更新日数と等しいわけではない。今回の分析結果から、総記事数が445(総行数)、平均行数が278.9(文字数)、平均文長が13.4(文字数)であったことがわかる。

表2 年ごとの記事数と更新日数

| 年(日)      | 記事数 | 更新日数 |
|-----------|-----|------|
| 2001(202) | 73  | 72   |
| 2002(365) | 239 | 161  |
| 2003(116) | 133 | 65   |

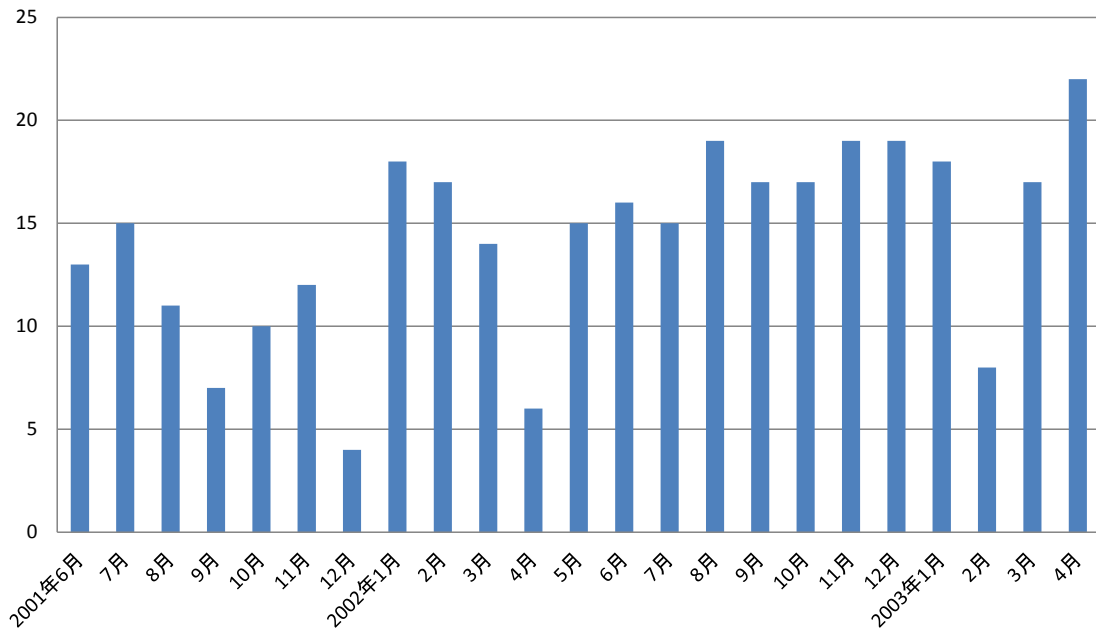


図1 記事数の月別推移

対象のブログは、開始されたのが2001年6月であり最後の記事が2003年4月のものであったので、約2年間の記録である。更新頻度(日数/更新日数)は、2001年は0.36、2002年が0.44、2003年が0.56であった。図1には月別の記事数の推移を表した。年間を通してみると更新頻度は増加傾向にあるものの大きな差はみられなかった。特に更新の少ない2001年12月には、忙しくてクマが出来たとの記述がある。なお、2003年の4月は更新頻度が最多であった。また、同月は同日の更新記事数も最多であるが、本やメールからの多くの引用がみられたほか、家族や恋人に宛てた遺言が投稿されていた。

(2) 単語頻度解析

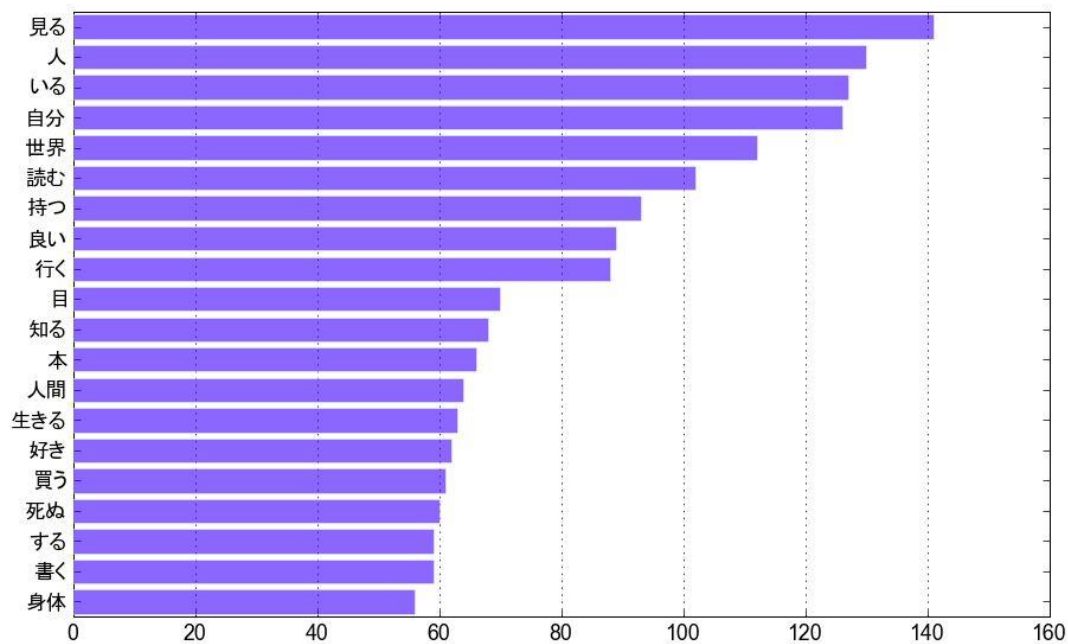


図2 『八本脚の蝶』の単語頻度

図2は、『八本脚の蝶』に於ける単語の使用頻度を表したものである。上位20位まで挙げられおり、最も頻度の高いものから順に、「見る」(141回)、「人」(130回)、「いる」(127回)、「自分」(126回)、「世界」(112回)、「読む」(102回)、「持つ」(93回)、「良い」(89回)、「行く」(88回)、「目」(70回)、「知る」(68回)、「本」(66回)、「人間」(64回)、「生きる」(63回)、「好き」(62回)、「買う」(61回)、「死ぬ」(60回)、「する」(59回)、「書く」(59回)、「身体」(56回)であった。

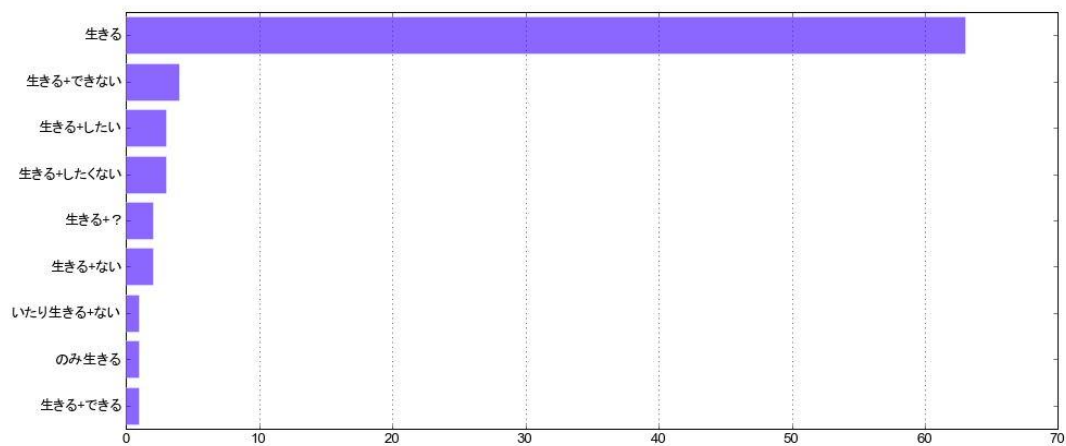


図3 単語頻度解析(「生きる」を含む)

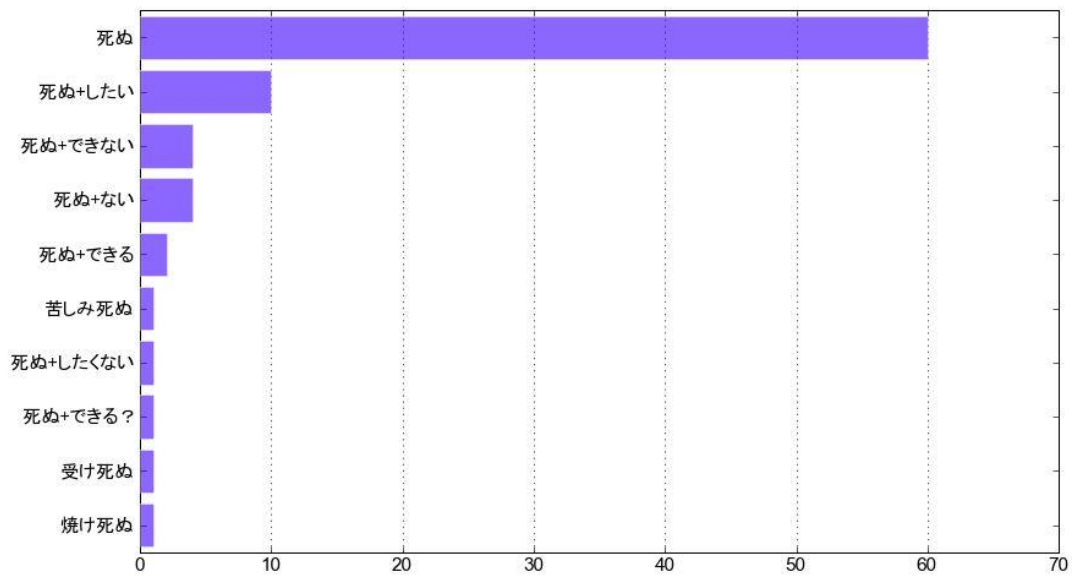


図 4 単語頻度解析(「死ぬ」を含む)

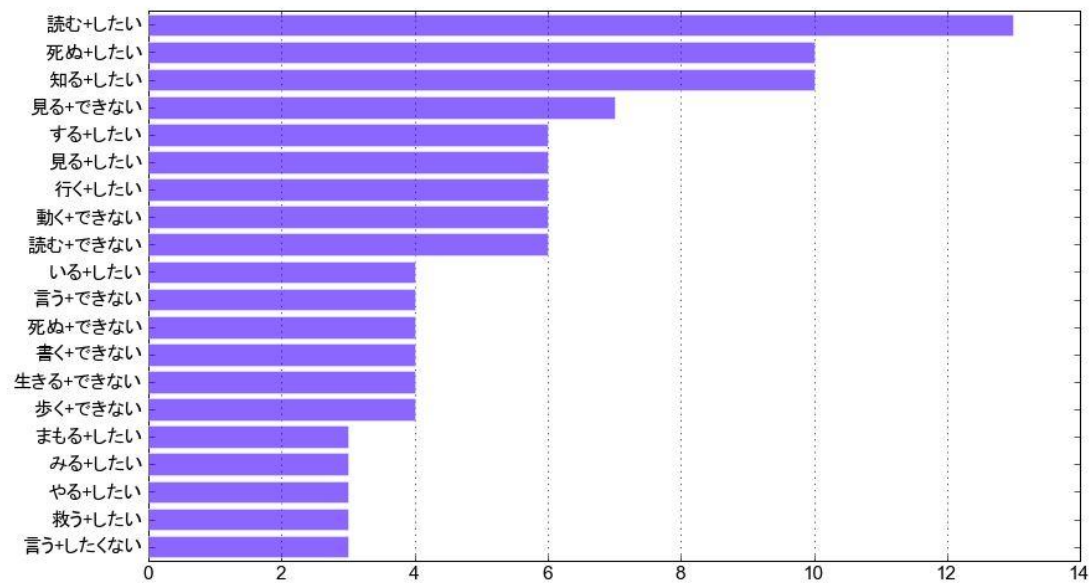


図 5 単語頻度解析(願望と不可能について)

『八本脚の蝶』の本文にて同程度の頻度での使用がみられた「生きる」「死ぬ」という単語について、それぞれ単語頻度分析を行い図 3、図 4 に表した。「生きる+できない」「生きる+したくない」というネガティブな語に、「生きる+したい」が並んでいる点が興味深い。一方の「死ぬ」を含む単語頻度分析の中で、「生きたい」に近い「死ぬ+したくない」という表現は殆どみられない。このことは、死というものが必ずしも避け得ないものでは

なく、自殺によって選択されるものであるからだろうと考えられる。「生きる+したい」と「生きる+できない」「生きる+したくない」、「死ぬ+したい」と「死ぬ+できない」「死ぬ+ない」の対立から、作者の生きづらさを窺い知ることが出来る。また、図 5 からは、そのほかの頻出単語の中にもこの対立がみられることがわかった。

### (3) 言葉のつながりからみる分析

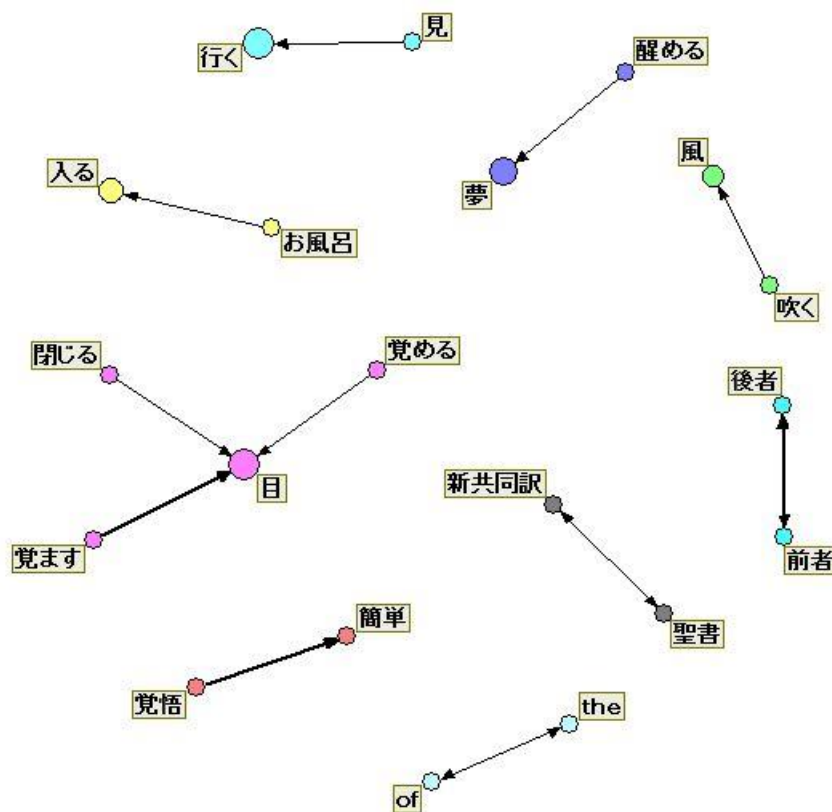


図 6 ことばネットワークによる解析

図 6 には、言葉のつながりの図示によって話題の傾向が表されている。ここで注目したいのは、「目」を中心としたつながりと、「夢」と「醒める」のつながりである。この「夢」という語は多くの場合夢見として用いられており、「夢から醒める」と「目が覚める」「目を覚ます」は同義であると云えるだろう。また、就寝の前に入浴をすると述べている記事も幾つかみられた。「お風呂に入り」「目を閉じる」ことは、作者にとって安寧の時間のひとつだったのではないかと推測する。このとき、夢の世界を離れる「目が覚める」ということは現実へ送還されることであり、生きる苦しみを実感せざるを得ない瞬間だったのではないだろうか。

## 5. まとめ

本研究では、主に単語の使用頻度やそのつながり・対立に焦点を合わせて分析を行った。ブログを構成しているものは文である。その一要素である単語を、ひとつひとつピックアップしながら文章を読むという作業はなかなかされないのではないかと。先に述べたとおり、対象のブログは書籍化されている。その帯には、「物語」を愛してやまなかった人が、あなたに読まれる物語になった」とある。私は、生きることが出来なくなり、ついに自殺してしまった女性の「物語」としてブログに触れた。それゆえに、文中の要素やそのつながりについて分析を行うテキストマイニングという形での研究に至ったのだらうと考える。

物語とは何なのか。本の中に描かれる物語には、フィクションと、ノンフィクションの二種類が存在する。一方は幻想で、もう一方は現実である。しかし、ノンフィクション作品であっても、読者にはそれが事実か否かを確かめ得ない場合もある。物語としての本は、多くの場合現実に在りながらほかの世界を知ることの出来るツールである。ノンフィクションの物語を読むということは、そこに描かれた人物の人生を垣間見るということであろう。すべての人の人生は、文章という媒体によって他者の手に取られ得る。ノンフィクションの物語として解釈される自身は、他者にとってはほかの世界の住人であり、「読まれる」という物語性を有した側面を持っていると云えるだろう。しかし、この物語性を主張したとき、大きな問題にぶつかる。世界には余りにもたくさんの物語が溢れている。先に述べたとおりであれば、物語は、存在する、或いは存在した人間と等しい数存在することになる。また、創作者によって幻想の物語は無数に生み出されてゆく。自分は自分の人生の主人公であるが、その物語が自分独特のものではなく、既に存在しており、誰かのパロディとして生きているのではないかという不安が生じたりするのである。「どんな真剣な行為もパロディであると知った」と、作者はブログの中で述べている。生きることそれ自体についてアイデンティティを見出すことが困難になるということは、大きな生きづらさにつながるだらうと考える。多くの書物を読み知識を得ることは、自身独特の思想や行動を形成するために活かすことの出来る方法であるかもしれないが、一方で、現実を陳腐なものに変え、飲み込むようなかたちで人を追い詰め得る要因になるのかもしれない。

## 6. 本研究の限界と今後の展望

本研究では約 2 年に渡って更新されたブログ記事を対象として分析を行ったが、その期間を通してなされていた具体的な感情の吐露について、十分なデータに表すことが出来なかった。全体を通してのひとつのブログの分析にとどまらず、その中で考察し得た時間という要素をより強く観点に組み込むことが改善点のひとつであるが、それを行うことが出来なかった本実験では以上の考察が限界であった。

自殺既遂者が遺した言葉は、家族や友人が故人の意向を知る重要な手掛かりになる。或る程度まとまった期間続けられたブログ等の記録からその推移を辿ることは、突然の消滅によって生じた空白を少しでも埋める助けになるかもしれない。このことは、自殺に至った経緯についての推測を、より確からしいものに近づけ、原因解明に役立てることが出来るだろう。意図的に遺したものであるか否かを問わず、物語となり得るかたちで自身の記録を記すことの意義とその価値が認められ、また、他者を物語として「読む」という理解のかたちが広く認知され受容・利用されることを望んでいる。

【参考・引用文献】

二階堂奥歯 八本脚の蝶 <http://homepage2.nifty.com/waterways/oquba/> (2013年10月2日取得)