

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年6月6日現在

機関番号：31302

研究種目：挑戦的萌芽研究

研究期間：2010～2012

課題番号：22652070

研究課題名（和文） 計量的分析による歴史資料の研究

研究課題名（英文） Analysis of Historical Sources by the Text-Mining

研究代表者

楠 義彦 (KUSUNOKI YOSHIHIKO)

東北学院大学・文学部・教授

研究者番号：20234429

研究成果の概要（和文）：歴史学研究で用いる文書資料を計量的なテキストマイニングにより分析することは、歴史学研究者のみに可能な事柄である。これは文書資料を構成する語の時代的、地域的また文脈による多様性を計量的分析では捉えられず、通常の文書分析と並行して行わなければならないためである。解釈批判の実行は極めて高度な歴史学研究の能力と経験を必須の基礎としており、現状ではコンピュータで人の代わりをさせることはできない。仮にできたとしても、そのための準備作業に膨大な時間と労力がかかり、コスト高となってしまう。今後資料のデジタル化が進展や新たな方策を開発により、テキストマイニングの歴史学研究への利用可能性は高まるものと考えられる。

研究成果の概要（英文）：Could the text-mining be a tool of analysis of historical sources? Yes in some part, no in others. No one can analyze historical sources but historical researchers, so with the text mining. For the text-mining can not read between lines, meanings characteristic of age and region, and a great diversity of meaning of a word. We need an enormous amount of time, if possible, for preparations of the text mining, digitizations of historical sources, selections of keywords, and so on. Now we don't think that the text-mining is a useful tool for historical research.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	700,000	0	700,000
2011年度	500,000	150,000	650,000
2012年度	500,000	150,000	650,000
年度			
年度			
総計	1,700,000	300,000	2,000,000

研究分野：人文学

科研費の分科・細目：史学・西洋史

キーワード：魔女、歴史資料、テキストマイニング

1. 研究開始当初の背景

テキストマイニングにより文書を分析する研究は日本でも盛んに行われている。しかし、

それらは主としてマーケティングやアンケートの自由回答欄、ウェブ、文学作品の分析等のために行われており、歴史資料の分析に

はほとんど用いられていない。欧米では歴史資料を計量的に分析することが進められていたが、日本の西洋史研究者で取り組んだものは数えるほどしかおらず、テキストマイニングの可能性も限界もまったく明らかになっていない。

2. 研究の目的

以上のような状況を受け、実際に歴史学研究にテキストマイニングを応用する方法について検討することを目的とする。

3. 研究の方法

テキストマイニングによる分析結果がどの程度の信頼性と正確性を持つかの検討を行うために、研究代表者の専門領域の歴史資料を実際にテキストマイニングで分析し、現在の課題を明らかにする。

4. 研究成果

研究成果は(1)テキストマイニングを行うに当たっての手順(①～⑤)と現在の課題(⑥～⑧)と、(2)テキストマイニングの今後の利用可能性、(3)その他、の三つの部分からなる。

(1)

①テキストマイニングのソフトウェアとしては、コンコーダンスーとエクセルが最低限必要である。テキストマイニング用のソフトウェアは高価なものが多いが、フリーウェアで提供されるコンコーダンスーを用いれば、比較的安価に利用可能である。

②文書資料をデジタル化する。歴史的な文書には表記の揺れを統一し、モダンナイズ(現代の言語表記に修正すること)する範囲を決定する。OCRによる文書のデジタル化は時間短縮にはなるが、黒体活字による文書をOCRは解析不能なので、人力による手入力が必要になる。

③分析する際のキーワードを決定する。これはテーマに関する研究の経験から判断し、あまりキーワードが多くなりすぎないように気を付けたい。キーワードの数としては50程度が望ましいと考える。

④コンコーダンスーを用いて、単語の頻度情報、出現傾向、単語間の共起関係、共起語の文脈上の位置特定、文脈を抽出する。

⑤見かけ上の相関関係と真の相関関係を区別するために、できれば複数の語との共起関係を分析し、テーマを構成する語を明らかにしたいが、そのためにはテキストマイニングソフトウェアが必要である。

⑥文書作成者の経歴、社会的地位、文書作成の動機、文書の版型、欠落部分の推測はテキストマイニングでは不可能である。

⑦テキストマイニングを行う文書に対しても外的・内的両方の史料批判が必要であるが、

これは統計学者や情報工学者では行えず、歴史学研究者の主体的な参加が絶対に必要である。

⑧テキストマイニングはそれぞれの語が持つ意味や時代性などの多様性に鈍感であり、この意味でも歴史学研究者以外の者が行うことは不可能である。

(2)

①歴史資料のデジタル化が急速に進展しつつあるという現状を考慮すると、資料の収集、文書のデジタル化にかかる時間の短縮と用いる資料の種類が増加が予想される。先行研究が未使用の文書が研究テーマの検討に適しているかの判断がますます重要性を持つことになるだろう。その場合、文書の利用適性を判断するツールとしてテキストマイニングは有益である。キーワードの出現頻度や他の語との共起関係が強く出る場合には、そのテーマに関する研究に有効な文書であると判断できるだろう。

②人間が通常読める分量を超えた文書を参照できるという大きなメリットがある。

③歴史資料の再検討の契機としてテキストマイニングは有効である。現在行われている方法に何らかの工夫を凝らすことで、歴史学研究への利用可能性が拡大するという可能性は閉ざされていない。

(3)

テキストマイニングの歴史学での適応可能性を検討する過程で、以下の文章がデジタル化された。なお、「タグなし」の表記は、プレーンテキストのもの、「タグあり」はTree Taggerによる形態素解析後、現代英語の綴りと異なる単語を修正したものを示している。

①THOMAS ADY, A Candle in the Dark: shewing The Divine Cause of the distractions of the whole Nation of ENGLAND, and of the Christian WORLD, 1655. (タグなしとタグあり)

②IOHN COTTA, The Trial of Witch-craft, shewing the true and right methode of the Discouery: with a Confutation of erroneous ways, 1616. (タグなし)

③LAMBERT DANEAU, A Dialogue of Witches in foretime named Lottellers, and now commonly called Sorcerers, 1575. (タグなしとタグあり)

④JOHN DARELL, An Apologie, or defence of the possession of William Sommers, a yong man of the towne of Nottingham. Wherein this marke of God is cleared from the evil name of counterfayinge: and therevpon also it is shewed that in these dayes men may be possessed with Devils, and that being so, by prayer and fasting the vncleane spirit may be cast out, 1599. (タグなしとタグあり)

⑤GEORGE GIFFARD, A Dialogue concerning Witches and Witchcraftes. In which is laide open how craftely the Diuell deceiueth not onely the Witches but many other and so leadeth them awrie into many great errorrs, 1593. (タグなしとタグあり)

⑥G. GYFFORD, A Discourse of the subtile Practises of Deuilles by Witches and Sorcerers. By which men are and haue bin greatly deluded: the antiquitie of them: their diuers sorts and Names. With an Aunswer vnto diuers friuolous Reasons which some doe make to prooue that the Deuils did not make these Aperations in any badily shape, 1587. (タグなしとタグあり)

⑦HENRY HOLLAND, A Treatise against WITCHCRAFT: OR A Dialogue, wherein the greatest doubts concerning that sinne, briefly answered: a Sathanicall operation in the Witchcraft of all times is truly prooued: the moste precious preseruatiues against such euils are shewed: very needfull to be knowen of all men, but chiefly of the Masters and Fathers of families, that they may learn the best meanes to purge their houses of all vnlean spirits, and wisely to auoide the dreadfull impieties and greate daugers which come by such abhominations, 1590. (タグなしとタグあり)

⑧FRANCIS HUTCHINSON, AN Historical Essay concerning Witchcraft with observations upon matters of fact; tending to clear the Texts of the Sacred Scriptures, and confute the vulgar Errors about that Point. AND also two sermons: One in Proof of the Christian Religion; The other concerning the Good and Evil Angels, 1718. (タグなしとタグあり)

⑨KING JAMES I, Daemonologie, in forme of ane Dialogve, 1597. (タグなしとタグあり)

⑩WILLIAM PERKINS, A Discourse of the damned Art of Witchcraft; so farre forth as it is revealed in the Scriptures, and manifest by true experience, 1608. (タグなしとタグあり)

⑪REGINALD SCOT, The Discoverie of Witchcraft, 1584. (タグなしとタグあり)

⑫THOMAS COOPER, The Mystery of Witchcraft. Discouering, The Truth, Nature, Occasions, Growth and Power thereof. TOGETHER With the Detection and Punishment of the time. As also, The seuerall Stratagemes of Sathan, ensnaring the poore soule by this desperate practize of annoying the bodie: with the seuerall Vses

thereof to the Church of Christ, 1617. (タグなしとタグあり)

⑬JOHN STEARNE, A Confirmation and Discovery of Witch Craft, 1648. (タグあり)

⑭A True Relation of the Arraignment of eighteen Witches That were tried, convicted, and condemned, at a Sessions holden at St. Edmunds-bury in Suffolke, 1645. (タグあり)

⑮The Examination, Confession, Triall, and Execution, of Joane Williford, Joan Cariden, and Jane Hott: who were executed at Feversham in Kent, for being Witches, on Munday the 29th of September, 1645. (タグあり)

⑯H. F., A true and exact Relation Of the severall Informations, Examinations, and Confessions of the late Witches, arraigned and executed in the County of Essex, 1645. (タグあり)

⑰RICHARD BOULTON, The Possibility and Reality of Magick, Sorcery, and Witchcraft, demonstrated. Or, a Vindication of a Compleat History of Magick, Sorcery and Witchcraft in Answer to Dr. Hutchinson's *Historical Essay*. 1722. (タグあり)

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

①楠義彦、前近代歴史資料の計量的分析— Visitation Articles を事例として—、ヨーロッパ文化史研究、査読有、第14号、2013年、29—53頁

②楠義彦、教会関係文書を分析する、ヨーロッパ・グローバリゼーションと諸文化圏の変容 研究プロジェクト報告書、V、2012年、221—232頁

6. 研究組織

(1) 研究代表者

楠 義彦 (KUSUNOKI YOSHIHIKO)
東北学院大学・文学部・教授
研究者番号：20234429

(2) 研究分担者

なし ()

研究者番号：

(3) 連携研究者

なし ()

研究者番号：