

旧富岡製糸場に関する新聞記事の計量テキスト分析

群馬県立高崎工業高等学校 正会員 西尾 敏和
群馬県立高崎工業高等学校 学生会員 ○中澤 寛一

1. 研究の背景

1872（明治5）年、明治新政府により現在の群馬県富岡市に設置された模範工場は旧富岡製糸場（以下富岡製糸場と記す）である。1987（昭和62）年3月、戦中・戦後と長く製糸工場として活躍した富岡製糸場は操業を停止した。2005（平成17）年9月、地元富岡市に片倉工業株式会社から富岡製糸場の全ての建造物は寄贈された。2007（平成19）年1月、世界遺産暫定一覧表の追加物件「富岡製糸場と絹産業遺産群」に富岡製糸場は選定された。どのようにすれば富岡製糸場の歴史を客観的に把握することができるのかという問題を筆者らは発見した。

2. 研究の目的

2005（平成17）年9月から2007（平成19）年10月までの富岡市管理移行後25箇月間の新聞記事データを収集することにより、富岡製糸場の歴史を客観的に把握することが本研究の目的である。

3. 研究の方法

（1）「上毛新聞 Web データベース」による新聞記事データの取得

群馬県の県域地方新聞である上毛新聞の紙面に掲載した記事や写真を検索するために「上毛新聞 Web データベース」を活用する。Web 上で検索キーワード「富岡製糸場」と掲載日「2005/9～2007/10」を入力して富岡製糸場に関する新聞記事データを取得する。取得した新聞記事データを様々な機種のパソコンやソフトウェアで利用するために、文字データのみで構成されているテキストファイル形式で保存する。

（2）ソフトウェア「KH Coder」による新聞記事データの計量テキスト分析

新聞記事データは膨大な量であるため全体を見渡すことは難しい。新聞記事データの全体像を明確に把握することができれば細部を追うときにも解釈や分析がより容易になると考えている。グラフなどの形で新聞記事データの全体像をつかむことができる計量テキスト分析のソフトウェア「KH Coder」を活用する。「KH Coder」は大阪大学の樋口耕一氏が開発した。コーディングルールという客観的な基準を作成することにより新聞記事データを数値データに変換する。更に数値データを表やグラフに整理する。

①新聞記事データの要約

テキストファイルに保存した新聞記事データにホームページの表示によく用いられる HTML タグを付けて「KH Coder」へ入力する。新聞記事の構成を調べるため、記事に多く出現する150の言葉を表で確認する。

②新聞記事を表現するコードの作成

詳細に新聞記事の構成を調べるため、月ごとに特徴的な言葉をまとめる。次に数多くの特徴的な言葉をコンパクトにまとめた言葉を作成する。コンパクトにまとめた言葉をコード、コンパクトにまとめる作業をコーディングという。更に作成したコードが新聞記事をどのくらい表現しているのかを確認する。

③月ごとにコードの出現率を集計

月ごとにそれぞれのコードを含む文の数及び出現率を集計して月ごとの推移を表で確認する（表－1参照）。

キーワード 旧富岡製糸場, 新聞記事, 計量テキスト分析

連絡先 〒370-0046 群馬県高崎市江木町700 群馬県立高崎工業高等学校 土木科 TEL 027-323-5450

表－１ コードを含む文の数及び出現率

(この表から月ごとであるが客観的に富岡製糸場の歴史を把握することが可能である)

年月	コード「富岡製糸場」を含む文の数及び出現率	コード「県内地域」を含む文の数及び出現率	コード「取り組み・イベント」を含む文の数及び出現率	コード「世界遺産登録」を含む文の数及び出現率	コード「蚕糸絹業」を含む文の数及び出現率	コード「観光」を含む文の数及び出現率	コードを含む文の総数
2005/9	59 (32.24%)	46 (25.14%)	50 (27.32%)	37 (20.22%)	12 (6.56%)	7 (3.83%)	183
2005/10	132 (24.63%)	144 (26.87%)	110 (20.52%)	68 (12.69%)	63 (11.75%)	20 (3.73%)	536
2005/11	261 (26.26%)	169 (17.00%)	171 (17.20%)	136 (13.68%)	104 (10.46%)	59 (5.94%)	994
2005/12	124 (22.42%)	116 (20.98%)	119 (21.52%)	88 (15.91%)	84 (15.19%)	24 (4.34%)	553
2006/1	206 (32.80%)	117 (18.63%)	96 (15.29%)	95 (15.13%)	74 (11.78%)	28 (4.46%)	628
2006/2	103 (30.75%)	97 (28.96%)	88 (26.27%)	78 (23.28%)	36 (10.75%)	21 (6.27%)	335
2006/3	65 (32.34%)	65 (32.34%)	46 (22.89%)	40 (19.90%)	6 (2.99%)	32 (15.92%)	201
2006/4	115 (21.99%)	127 (24.28%)	66 (12.62%)	48 (9.18%)	9 (1.72%)	22 (4.21%)	523
2006/5	124 (26.27%)	98 (20.76%)	73 (15.47%)	58 (12.29%)	38 (8.05%)	18 (3.81%)	472
2006/6	81 (15.31%)	120 (22.68%)	82 (15.50%)	63 (11.91%)	73 (13.80%)	23 (4.35%)	529
2006/7	94 (26.93%)	92 (26.36%)	72 (20.63%)	49 (14.04%)	26 (7.45%)	28 (8.02%)	349
2006/8	104 (22.46%)	138 (29.81%)	91 (19.65%)	131 (28.29%)	58 (12.53%)	46 (9.94%)	463
2006/9	136 (27.36%)	101 (20.32%)	107 (21.53%)	105 (21.13%)	56 (11.27%)	31 (6.24%)	497
2006/10	164 (22.28%)	175 (23.78%)	137 (18.61%)	114 (15.49%)	70 (9.51%)	19 (2.58%)	736
2006/11	154 (18.10%)	204 (23.97%)	158 (18.57%)	177 (20.80%)	43 (5.05%)	42 (4.94%)	851
2006/12	103 (30.12%)	80 (23.39%)	71 (20.76%)	55 (16.08%)	48 (14.04%)	13 (3.80%)	342
2007/1	237 (21.99%)	182 (16.88%)	111 (10.30%)	222 (20.59%)	133 (12.34%)	53 (4.92%)	1078
2007/2	207 (24.70%)	207 (24.70%)	142 (16.95%)	259 (30.91%)	94 (11.22%)	50 (5.97%)	838
2007/3	179 (22.77%)	198 (25.19%)	130 (16.54%)	120 (15.27%)	102 (12.98%)	46 (5.85%)	786
2007/4	108 (15.56%)	190 (27.38%)	112 (16.14%)	87 (12.54%)	73 (10.52%)	47 (6.77%)	694
2007/5	130 (20.70%)	97 (15.45%)	102 (16.24%)	54 (8.60%)	64 (10.19%)	23 (3.66%)	628
2007/6	134 (20.65%)	96 (14.79%)	89 (13.71%)	82 (12.63%)	39 (6.01%)	36 (5.55%)	649
2007/7	156 (24.49%)	163 (25.59%)	120 (18.84%)	107 (16.80%)	73 (11.46%)	22 (3.45%)	637
2007/8	79 (11.08%)	168 (23.56%)	86 (12.06%)	81 (11.36%)	109 (15.29%)	43 (6.03%)	713
2007/9	124 (17.79%)	160 (22.96%)	76 (10.90%)	86 (12.34%)	77 (11.05%)	56 (8.03%)	697
2007/10	110 (26.00%)	93 (21.99%)	68 (16.08%)	66 (15.60%)	71 (16.78%)	12 (2.84%)	423
合計	3489 (22.75%)	3443 (22.45%)	2573 (16.78%)	2506 (16.34%)	1635 (10.66%)	821 (5.35%)	15335

4. 研究の結果

(1) 取得した 929 件の新聞記事データを要約したところ、「遺産」(2555 回)・「世界」(1744 回)・「富岡製糸場」(1577 回)・「製糸」(1420 回)・「富岡」(1397 回)・「登録」(909 回)といった富岡製糸場の世界遺産登録を表す言葉が多く出現していた。他には「観光」という語が 484 回出現していた。

(2) 月ごとに特徴的な言葉から「富岡製糸場」・「県内地域」・「取り組み・イベント」・「世界遺産登録」・「蚕糸絹業」・「観光」という 6 個のコードを作成した。新聞記事の 54.84% を 6 個のコードで表現できる。

(3) 例えばコードを含む文の数が一番多い 2007 (平成 19) 年 1 月の場合、「富岡製糸場」(21.99%) が「世界遺産登録」(20.59%) に関する暫定リストに追加された。「県内地域」(16.88%) では行政や地域住民等による「蚕糸絹業」(12.34%) に関する「取り組み・イベント」(10.30%) が実施された。今後「観光」(4.92%) に影響を与えるだろう。表－1 より月ごとであるが富岡製糸場の歴史を客観的に把握することが可能である。

5. おわりに

本発表では月ごとにコードの出現率を集計したが、より詳細に一日ごとや記事ごとに集計することも可能である。今後、本研究の成果を富岡市の歴史を活かした都市計画及びまちづくりに活用していきたい。本論文は本校工業科目「課題研究」で取り組んだ研究成果をまとめたものである。2009 (平成 21) 年 1 月、本校課題研究発表会においても研究発表を行った。

参考文献

- ・富岡製糸場の歴史については富岡市の「富岡製糸場 世界遺産推進ホームページ」を参考にした。
- ・樋口耕一 計量テキスト分析の方法と実践 大阪大学大学院人間科学研究科博士論文 2004 年