

学術的動画の評価傾向とその要因

林 正治 田中 克明 長谷 海平 高見澤 秀幸 松村 芳樹 中島 康 万代 勝信

(一橋大学 情報基盤センター)

本論文では学術的動画を提供するアーカイブの構築を目的とした動画の評価手法について述べる。動画の評価要因について質問紙調査と因子分析を実施した。その結果、「満足度因子」「学術的価値因子」「クオリティ因子」の3因子を抽出した。また、検証的因子分析を実施し、得られた3因子による学術的評価の因子モデルの妥当性を検証した。得られた因子得点を活用した動画アーカイブ構築の可能性を示した。

Analysis of Value Factors for Academic Moving Images

Masaharu HAYASHI / Katsuaki TANAKA / Kaihei HASE / Hideyuki TAKAMIZAWA /
Yosiki MATSUMURA / Yasushi NAKAJIMA / Katsunobu MANDAI

(Center for Information and Communication Technology, Hitotsubashi University)

In this paper, we describe an evaluation method of academic moving images for the purpose of construction of an academic moving image archive. We conducted a survey by asking college students and by using exploratory factor analysis. We have found three major factors which are “satisfaction”, “academic value” and “moving image quality”. The result of confirmatory factor analysis, the validity of the model of academic evaluation axis with the three major factors was inspected. Finally, we show visualization method of academic moving images archive using the academic evaluation axis.

1. はじめに

大量の動画を蓄積する動画アーカイブを活用する上で、利用者に有用と思われる動画を提示したり、発見を促したりする仕組みは重要である[1]。こうした仕組みはYouTubeやニコニコ動画などの動画共有サイトの機能として提供されているが、参照数が多い動画が推薦されやすくなるなど、少数の評価が反映されにくいという課題が存在する[2]。

現在、我々の研究グループでは、大学での講義、講演、イベント等の動画を蓄積する動画アーカイブの活用を検討している。こうした動画の多くは大学の研究業務に関連する性質上、学術的価値の高い動画と考えられる。しかしこれらは、決して娯楽性の高い動画ではないため、自ずと視聴回数が低くなる。我々はこうした特徴を持つ動画の活用を考えている。動画の学術的価値をふまえた上で利用者に動画を提供することを検討している。

本稿では我々が実施した先行研究[3]をもとに、学術的価値による評価に基づく動画アーカイブを構築するために、動画の学術的価値の評価要因を実験的に明らかにすることを目的とする。

2. 質問紙調査

大学における動画アーカイブの評価要因を因子分析およびテキスト分析によって抽出するこ

とを目的とした質問紙調査を実施した。我々が実施した先行研究[3]では動画の評価要因として、動画に対する満足度と動画の持つ学術的価値に関する因子が発見されている。しかしながら、先行研究では、動画の理解度が低いのに関わらず動画の学術的価値を高く評価する傾向が確認された。そこで本研究では、先行研究で用いた質問項目に加えて、動画の理解度を尋ねる項目(Q3~Q6)、動画の品質を尋ねる項目(Q22, Q23)、学術的価値を尋ねる項目(Q14~Q19)を追加した(表1)。

調査は2015年9月16日~18日に本研究に携わったことがない大学生19名(商学部生11名, 社会学部生4名, 経済学部生1名, 法学部生3名)に対して行い、動画アーカイブからランダムに抽出した127本の動画に関して得られた合計710件の回答を分析に用いた。得られた回答の記述統計量を表2に示す。先行研究[3]と同じく床効果が観測されたQ2を除く項目を分析に使用した。

表1 質問項目(色塗りは追加項目)

ID	項目	回答方式
Q1	動画の概要	記述式(100字以内)
Q2	講演者を知っていますか	知らない(1) - 知っている(10)
Q3	内容に関する背景知識を持っていますか	持っていない(1) - 持っている(10)

Q4	あなたの学問分野に関係のある内容でしたか	関係ない (1) - 関係ある (10)
Q5	興味がある内容でしたか	興味ない (1) - 興味ある (10)
Q6	興味がある/ない理由	記述式 (100字以内)
Q7	内容は理解できましたか	理解できない (1) - 理解できた (10)
Q8	内容に満足しましたか	満足しない (1) - 満足した (10)
Q9	満足した/しない理由	記述式 (100字以内)
Q10	また視聴したいと思いますか	視聴したくない (1) - 視聴したい (10)
Q11	視聴したい/したくない理由	記述式 (100字以内)
Q12	他の人に勧めたい内容でしたか	勧めたくない (1) - 勧めたい (10)
Q13	勧めたい/勧めたくない理由	記述式 (100字以内)
Q14	大学として保存する価値のある映像でしたか	価値ない (1) - 価値ある (10)
Q15	価値ある/無いと思う理由	記述式 (100字以内)
Q16	新たな知見は得られましたか	得られない (1) - 得られた (10)
Q17	研究に利用できる内容でしたか	利用できない (1) - 利用できる (10)
Q18	利用できる/できない理由	記述式 (100字以内)
Q19	内容に関する発表資料、関連論文等があれば参照しますか	参照しない (1) - 参照する (10)
Q20	学術的価値があると思いますか	価値ない (1) - 価値ある (10)
Q21	価値ある/無いと思う理由	記述式 (100字以内)
Q22	映像はみやすかったですか	みにくい (1) - みやすい (10)
Q23	音声は聞きやすかったですか	聞きにくい (1) - 聞きやすい (10)

3. 評価要因の分析と視覚化

3.1. 探索的因子分析

集計した回答 710 件分の回答を使って探索的因子分析を行った。分析には統計ソフトウェア R[5]に psych パッケージ[6]を導入し、fa 関数を用いた。各項目について Spearman の順位相関係数を求めた (表 3)。求めた順位相関係数に対して最尤法とプロマックス回転による因子分析を実施した。因子数はガットマン基準、スクリープロットから 3 因子と仮定した。分析の結果、得

られた因子負荷量について Q3, Q4 の共通性が 0.3 未満の数値を示したため、全体への影響度が低いものと判断し、Q3, Q4 を除いて、再度因子分析を行った。その結果、Q10, Q14, Q16 について、2 つ以上の因子に対して絶対値 0.3 以上の因子負荷量が得られた。これらは評価尺度を曖昧にする項目と考えられたため、Q10, Q14, Q16 を削除後、改めて因子分析を行った。最終的に得られた回転後の因子負荷量行列を表 4 に示す。

得られた因子負荷量について絶対値 0.1 以下の項目を除いた上で各因子における Cronbach の α 係数を求めた。その結果、すべての因子において α 係数が 0.7 以上を示し、十分な信頼性が示された (表 5)。累積寄与率は第 3 因子までで 65.5% であった。各因子の特徴を以下に示す。

第 1 因子は動画の内容に対する理解 (Q7)、満足 (Q8)、他者への推薦 (Q12) に関する項目から強い影響を受けている。また、内容に対する興味 (Q5) に関する項目から中程度の影響、内容についてのさらなる情報の要求 (Q19) についての項目から弱い影響を受けている。

第 2 因子は研究的価値 (Q17)、学術的価値 (Q20) に関する項目から強い影響を受けている。また、動画の内容についてのさらなる情報の要求 (Q19) に関する項目から中程度の影響を受けている。内容に対する興味 (Q5) に関する項目との弱い影響も確認できる。一方で、内容に対する理解 (Q7) に関する項目から弱い負の影響を受けている。

第 3 因子は映像、音声の質に関する項目 (Q22, Q23) からの強い影響を受けている。

表 2 評価回答の統計量 (N=710)

項目	平均	標準偏差	歪度	尖度
Q2	2.958	2.914	1.329	0.360
Q3	4.727	2.653	0.378	-0.905
Q4	3.945	2.686	0.606	-0.776
Q5	5.504	2.629	-0.243	-1.015
Q7	6.961	2.788	-0.761	-0.569
Q8	5.835	2.749	-0.280	-1.140
Q10	3.735	2.631	0.603	-0.950
Q12	5.790	2.765	-0.310	-1.087
Q14	6.401	2.603	-0.637	-0.637
Q16	5.038	2.961	-0.010	-1.376
Q17	3.721	2.822	0.603	-1.111
Q19	4.228	2.949	0.385	-1.263
Q20	4.310	3.036	0.269	-1.459
Q23	7.256	2.415	-0.741	-0.433
Q22	7.287	2.405	-0.778	-0.305

表 3 測定項目間の Spearman 順位相関係数

	Q3	Q4	Q5	Q7	Q8	Q10	Q12	Q14	Q16	Q17	Q19	Q20	Q22
Q4	0.360*												
Q5	0.262*	0.524*											
Q7	0.461*	0.153*	0.321*										
Q8	0.266*	0.301*	0.627*	0.551*									
Q10	0.055	0.334*	0.613*	0.118*	0.577*								
Q12	0.239*	0.209*	0.556*	0.382*	0.718*	0.577*							
Q14	0.076	0.199*	0.404*	0.168*	0.549*	0.510*	0.628*						
Q16	0.062	0.408*	0.641*	0.255*	0.644*	0.684*	0.569*	0.514*					
Q17	-0.119*	0.403*	0.440*	-0.094*	0.298*	0.584*	0.280*	0.327*	0.606*				
Q19	-0.004	0.345*	0.483*	0.014	0.329*	0.596*	0.354*	0.362*	0.573*	0.635*			
Q20	-0.169*	0.323*	0.376*	-0.164*	0.269*	0.564*	0.253*	0.398*	0.580*	0.793*	0.628*		
Q22	0.108*	0.110*	0.138*	0.194*	0.294*	0.113*	0.291*	0.266*	0.130*	-0.064	-0.052	-0.133*	
Q23	0.094	0.099	0.174*	0.267*	0.364*	0.135*	0.321*	0.294*	0.198*	-0.018	-0.066	-0.066	0.677*

*p<0.05

3.2. 検証的因子分析

探索因子分析で発見した 3 つの因子と各項目の関係を検証的因子分析で確認する。探索的因子分析の結果、相関が認められた各因子と項目との間に相関があることを仮定し、検証的因子分析を実施した。検証的因子分析には R[5] に lavaan パッケージ[7]を導入し、cfa 関数を用いた。その結果、図 1 のモデルの適合評価指標は、Goodness of Fit Index (GFI) の値が 0.97, Comparative Fit Index (CFI) の値が 0.975, Root Mean Square Error of Approximation (RMSEA) の値は 0.075 となった。一般的に、GFI と CFI の値は 0.9 以上でモデルが受容される。RMSEA の値は 0.1 以上でモデルのあてはまりが悪いとされ、0.05 以下でモデルのあてはまりが良いとされる。したがって、検証的因子分析の結果は必ずしも最適というわけではないが、許容範囲内であると判断できる。検証的因子分析の各項目の因子負荷量を表 6 に示す。また、その因子間相関行列を表 7 に示す。

各項目の因子負荷量には、探索的因子分析と比較して特に大きな変化は認められなかった。よって、探索的因子分析で発見した 3 つの因子と各項目の関係は妥当であると考えられる。各因子間の相関については、全般的に弱い相関が認められた(表 7)。しかしながら、第 1 因子と第 3 因子、第 2 因子の相関は認められなかった。

表 4 探索的因子分析の因子負荷量

項目	第 1 因子	第 2 因子	第 3 因子
Q5	0.601	0.280	-0.035
Q7	0.751	-0.382	-0.060
Q8	0.894	0.042	0.038
Q12	0.698	0.105	0.088
Q17	-0.028	0.902	0.049
Q19	0.179	0.648	-0.091
Q20	-0.060	0.912	-0.001
Q22	0.028	-0.025	0.768
Q23	0.045	0.028	0.854

表 5 探索的因子分析結果の概要

項目	第 1 因子	第 2 因子	第 3 因子
固有値	2.251	2.304	1.343
寄与率 (%)	0.250	0.256	0.149
累積寄与率 (%)	0.250	0.506	0.655
α 係数	0.806	0.782	0.764

3.3. 因子の命名

探索的因子分析および検証的因子分析の結果、3 つの因子を抽出し、各項目との関係の妥当性を検討した。次に 3 つの因子と各項目との関係を考察し、各因子に名称を付与する。

第 1 因子は動画の内容に対する理解 (Q7)、満足 (Q8)、他者への推薦 (Q12) に関する項目から強い影響を受けている。また、内容に対する興味 (Q5) に関する項目から中程度の影響、内容についてのさらなる情報の要求 (Q19) についての項目から弱い影響を受けている。第 1 因子は動画に対する純粋な満足度を表す因子と考えられる。以上の特徴から第 1 因子を「満足度因子」と命名する。「満足度因子」については先行研究[3]でも観測されており、先行研究の分析結果を支持する結果が得られた。

第 2 因子は研究的価値 (Q17)、学術的価値 (Q20) に関する項目から強い影響を受けている。また、動画の内容についてのさらなる情報の要求 (Q19) に関する項目から中程度の影響を受けている。内容に対する興味 (Q5) に関する項目との弱い影響も確認できる。一方で、内容に対する理解 (Q7) に関する項目から弱い負の影響を受けている。Q7 による負の弱い影響は動画内容の専門性の高さが要因として考えられる。以上の特徴から第 2 因子を「学術的価値因子」と命名する。この因子についても、同等の特徴を持つ因子が先行研究[3]で発見されている。ただし、先行研究[3]で得られた因子と比較して、理解度 (Q7) を示す項目の影響度は低く、動画の学術的価値について、より明確に示した因子と解釈できる。

第3因子は映像、音声の質に関する項目(Q22, Q23)からの強い影響を受けている。映像の見やすさ、音声の聞き取りやすさ、映像作品としての完成度を表す因子と捉えることができる。以上の特徴から第3因子を「クオリティ因子」と命名する。本因子は先行研究[3]では観測されなかった因子である。

表 6 検証的因子分析の因子負荷量

項目	第1因子	第2因子	第3因子
Q5	0.589	0.274	0.000
Q7	0.705	-0.376	0.000
Q8	0.933	0.000	0.000
Q12	0.751	0.057	0.000
Q17	0.000	0.888	0.000
Q19	0.148	0.668	0.000
Q20	0.000	0.888	0.000
Q23	0.000	0.000	0.875
Q22	0.000	0.000	0.774

表 7 因子間相関行列

	第1因子	第2因子
第2因子	0.336	
第3因子	-0.093	0.429

*p<0.05

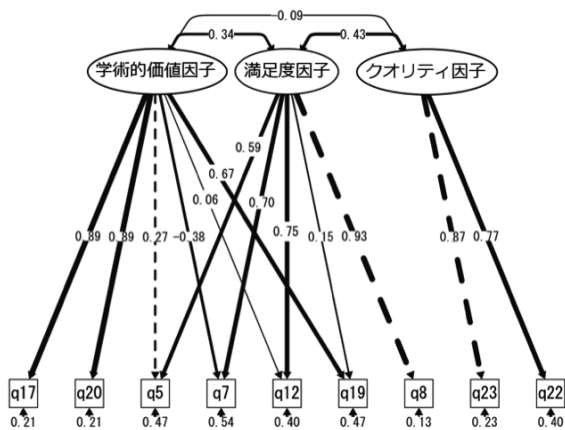


図 1 学術的動画評価の因子モデル

3.4. 因子得点

探索的因子分析で得られた因子負荷量をもとに回帰法による因子得点を求め、動画別の平均値を求めた。縦軸を「満足度因子」得点、横軸を「クオリティ因子」得点、点の大きさを「学術的価値因子」得点とした動画別の因子得点平均の分布を示す(図2)。

図2からは、「満足度因子」得点と「クオリティ因子」得点がともに低いが、「学術的価値因子」得点が高い動画が確認できる(図2左下)。また、

「満足度因子」得点および「クオリティ因子」得点が高く、かつ「学術的価値因子」得点が低く評価されている動画も確認でき、動画の満足度、映像・音声の質に関係なく学術的価値が評価されていることわかる。また、「クオリティ因子」得点については、全体的に高く評価されていることから、多くの動画は映像・音声の質に問題がないことが推測できる。「クオリティ因子」得点が高いと「学術的価値因子」得点が低くなる傾向がみられる。

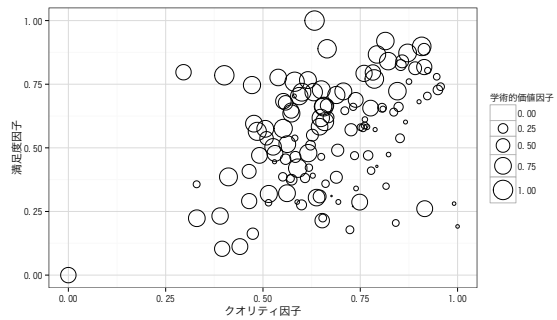


図 2 動画別の因子得点平均

3.5. テキスト分析

自由記述回答文を形態素解析、名詞と動詞を抽出、各回答文における出現回数により単語と出現回数からなる文書ベクトルを作成後、Probabilistic Latent Semantic Indexing (PLSI) [4]によりトピック抽出を行った。

資料概要(Q1)に記述されたテキストをPLSIにより5トピックに分類して得られた各トピックの特徴語上位10個を表8に示す。また、PLSIによる計算結果の一部として、各回答文の全回答文における生起確率 $p(d)$ が求められる。各映像資料について $p(d)$ の合計値を求め、求められた値の大きい順に映像資料への回答文の生起確率による順位付けを行った。この順位に基づき、自由記述による回答と、因子分析に用いた選択式回答との Spearman の順位相関係数を求めた(表9)。

表 8 PLSIによるトピックごとの特徴語

No	特徴語
1	社会, 挨拶, 閉会, 討論, 保障, 一橋大学, 日本, 説明, 内容, 政策
2	一橋大学, 説明, 紹介, キャンパス, オープン, 動画, 法学部, 大学, 教育, 商学部
3	問題, 経済, 復興, 必要, 貧困, 日本, 震災, 大震災, 東日本, 賃金
4	金融, 機器, 日本, 企業, 公園, 組織, 必要, 税, 説明, 財政
5	経済, 企業, 教育, 人材, 経営, TPP, 社会, 制度, グローバル, 大学

表 9 自由記述回答の順位相関係数

	Q1	Q6	Q9	Q11	Q13	Q15	Q18	Q21
Q5	0.250*	0.299*	0.250*	0.117	0.204*	0.099	0.250*	0.250*
Q7	0.065	0.136	0.065	0.055	0.090	0.094	0.065	0.065
Q8	0.265*	0.228*	0.265*	0.080	0.196*	0.109	0.265*	0.265*
Q12	0.191*	0.150	0.191*	0.060	0.226*	0.042	0.191*	0.191*
Q17	0.478*	0.377*	0.478*	0.268*	0.201*	0.182*	0.478*	0.478*
Q19	0.360*	0.296*	0.360*	0.173*	0.188*	0.134	0.360*	0.360*
Q20	0.464*	0.336*	0.464*	0.228*	0.214*	0.202*	0.464*	0.464*
Q22	-0.149	-0.094	-0.149	-0.044	-0.007	-0.098	-0.149	-0.149
Q23	-0.035	-0.033	-0.035	-0.008	0.064	-0.019	-0.035	-0.035

*p<0.05

4. 考察

4.1. 因子分析

本研究では質問紙調査の結果を探索的因子分析、検証的因子分析を行うことで、学術的動画の評価傾向を表す3つの因子「満足度因子」「学術的価値因子」「クオリティ因子」を明らかにした。

「満足度因子」「学術的価値因子」については先行研究[3]でも観測された因子であり、先行研究の結果を裏付けることができた。さらに、「クオリティ因子」の発見により、映像・音声に問題がある動画が存在し、そのような動画の学術的価値が高く評価される可能性があることがわかった。

検証的因子分析に用いた学術的動画評価の因子モデルについては、RMSEA値から改善の余地があることが示唆される。しかしながら、本研究が対象とする大学における動画アーカイブには多種多様な動画が記録されており、モデルの適合度を向上させるには限界があるものと考えられる。

4.2. 因子得点

因子得点からは「学術的価値因子」得点が「満足度因子」得点、「クオリティ因子」得点とは異なる次元で与えられていることが明らかになった。「満足度因子」得点、「クオリティ因子」得点が低い場合においても「学術的価値因子」得点が高く評価されることがある。また、逆に「満足度因子」得点、「クオリティ因子」得点が高得点であっても、「学術的価値因子」得点が低く評価されることがある。

とくに「クオリティ因子」得点が低く、かつ「学術的価値因子」得点が高く評価されている動画については、動画の撮り方や表現方法に問題がある可能性が高い。これは動画アーカイブを構築する上で撮影技術の向上も検討課題となることを示唆している。動画アーカイブの活用を考慮した場合、ただ撮り貯めるだけでは不足であり、映像の専門家とアーカイブする動画の作り方について

の検討が必要である。

4.3. テキスト分析結果からの考察

学術的価値(Q20)とその理由記述(Q21)の相関係数から、p(d)の順位と数値回答による順位に相関が存在することが明らかとなった(表9)。資料概要の自由記述(Q1)がQ17、Q20とも正の相関を持つことから、被験者が資料を理解しつつ学術的価値が高いと評価したことを知るためには、資料の概要を記述させることが有効であると言える。

PLSIによるトピックごとの特徴語(表8)からは、トピック1はシンポジウム・討論会に関する動画、トピック2は大学広報に関する動画、トピック3は経済復興に関する動画、トピック4は金融財政に関する動画、トピック5は経済・経営に関する動画を示すものと推測できる。これらのトピックは動画の発見を促す仕組みで利用できるものと考えられる。

表 10 YouTube 閲覧回数と各因子得点との相関係数

	満足度因子	学術的価値因子	クオリティ因子
閲覧回数	0.171	-0.404*	0.411*

*p<0.05

4.4. 動画共有サイトの閲覧数と因子得点の比較

実験で用いた動画について、閲覧回数が必ずしも動画の評価と一致しないことを確認するために、YouTubeにおける動画の閲覧回数と各因子得点との比較を行った。YouTubeにおける動画の閲覧回数は、実験実施日より以前の2015年9月13日にYouTube Data API[8]を用いて取得した。閲覧回数とSpearmanの順位相関係数を表10の通りである。

この結果から、YouTubeにおける動画の閲覧回数と「学術的価値因子」得点の間には中程度の負の相関があることが明らかとなった。「クオ

リティ因子」得点と閲覧回数との間には、中程度の正の相関があることが明らかとなった。また、「満足度因子」については有意な相関は観測できなかった。以上のことから、YouTube における動画の閲覧回数は学術的価値とは異なる評価軸にあることが示された。

5. 評価尺度を活用した動画アーカイブプロトタイプ

本研究で得られた動画の評価尺度を用いて動画アーカイブのプロトタイプシステムを構築した(図 3)。各評価尺度の影響度をサムネイル画像の位置、サイズに反映させた。サムネイル画像をクリックすると、対応する動画ページにアクセスすることができる。

プロトタイプシステムの構築により、本研究で発見した学術的動画評価尺度による動画提供の可能性を明らかにすることができた。しかしながら、実際のシステム化の際には、効果的な評価を実施させるためのユーザインタフェースの開発や動的な評価尺度の計算が必要になる。また、PLSI によって抽出されたトピックの利用など、解決しなければならない課題は多い。

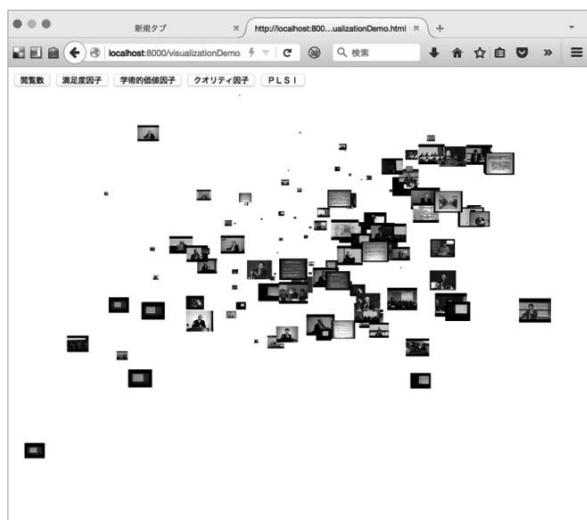


図 3 評価尺度を用いた動画の視覚化

6. おわりに

本稿では学術的動画の評価傾向について因子分析、テキスト分析の両面から分析し、その要因を明らかにした。また、得られた評価尺度を拡張した動画アーカイブの試験構築を行った。今後は本研究で得られた動画評価尺度の入力インタフェースの開発、学術的動画評価尺度およびトピックモデルを利用した動画アクセス手法についての研究を行う予定である。

謝辞

本研究は一橋大学研究プロジェクト「映像資料アーカイブ運用に関する汎用的フレームワークの構築」および科研費 15K00446, 24300310, 25560140 の一部助成を受けたものである。

参考文献

- 1) 住吉英樹：アーカイブを活用するコンテンツ検索・推薦システム, 映像情報メディア学会誌, Vol.63, No.11, pp.1504-1508, 2009
- 2) 平澤真大, 小川祐樹, 諏訪博彦, 太田敏澄, ニコニコ動画のログデータに基づくソーシャルノベルティのある動画の発見手法の提案, 情報処理学会論文誌, Vol.54, No.1, pp.214-222, 2013
- 3) 長谷海平, 田中克明, 林正治, 高見澤秀幸, 松村芳樹, 中島康, 万代勝信：映像資料アーカイブの汎用フレームワーク構築に向けて, 情報コミュニケーション学会第 12 回大会発表論文集, pp.148-151, 2015
- 4) Hofmann, T., “Probabilistic latent semantic indexing”, Proc. of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp.50-57, 1999.
- 5) The R Project for Statistical Computing (online), available from (<http://www.r-project.org/>) (accessed 2015-11-07)
- 6) Package ‘psych’ (online), available from (<https://cran.r-project.org/web/packages/psych/index.html>) (accessed 2015-11-07)
- 7) The lavaan Project (online), available from (<http://lavaan.ugent.be/>) (accessed 2015-11-07)
- 8) YouTube Data API V3 (online), available from (<https://developers.google.com/youtube/v3/>) (accessed 2015-11-07)