

# Test Topics: TREC 2008 Legal Track, Relevance Feedback Task

This document provides summary information for each of the 12 assessed topics of the Relevance Feedback task of the TREC 2008 Legal Track. Most of the categories are the same as specified for the Ad Hoc task. The differences are as follows:

**Topic** : The 12 topics were selected from 2006 and 2007, whose numbers ranged from 6 to 101. There were 5 complaints in 2006 (labelled A, B, C, D and E) and 4 complaints in 2007 (labelled A, B, C and D).

**Sampling and Est. Rel.** : The number of pooled documents includes not just the residual output of the relevance feedback runs but also the special oldrel08 and oldnon08 runs (these latter 2 runs included at most 10 documents each). However, the oldrel08 and oldnon08 documents are not included in the other reported numbers (e.g., the reported number presented to the assessor does not include the up to 20 documents from the oldrel08 and oldnon08 runs).

**Est. Rel.** : The “Est. Rel.” is just the estimated number of *residual* relevant documents in the pool for the topic (i.e., the number of relevant documents after those judged in 2006 and 2007 are removed).

**Final Boolean Result Size  $B_r$**  : “ $B_r$ ” is the number of documents matching the final negotiated Boolean query after the documents judged in 2006 and 2007 are removed. “F1”, “Precision” and “Recall” are for the residual Boolean result set.

**Participant High F1@ $K_r$  and Median F1@ $K_r$**  : The median estimated F1 score is based on the median score of the 29 participant runs (this number does not include the 2 reference runs (refRF08B and randomRF08B)).

**5 Deepest Sampled Relevant Documents** : Only residual relevant documents are listed. The ranks following the run identifiers are residual ranks.

**Topic 14** (2006-A-9)

**Request Text:** All documents referring to the marketing, placement, or sale of lozenges to children.

**Initial Proposal by Defendant:** (marketing OR placement OR sale) AND "lozenges" AND "children"

**Rejoinder by Plaintiff:** lozenges AND (child! OR teen! OR juvenile! OR kid! OR adolescent)

**Final Negotiated Boolean Query:** lozenges AND (child! OR teen! OR juvenile! OR kid! OR adolescent)

**Sampling:** 526196 pooled, 386 assessed, 4 judged highly relevant, 13 other judged relevant, 363 judged non-relevant, 6 gray, “C”=2.36

**Est. Rel.:** 107.0 (including 32.6 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 429, **F1:** 32.9%, (**Precision:** 22.0%, **Recall:** 65.5%)  
**Participant High F1@K<sub>r</sub>:** 35.8% (UMKCTL08RF6), **Median F1@K<sub>r</sub>:** 0.7%  
**5 Deepest Sampled Highly Relevant Documents:** ztv07a00-16.9 (IowaSL08RF1A-40),  
lcz10d00-10.2 (otRF08fbF-24), hpz10c00-3.0 (otRF08F-7), isy24a00-2.5 (UMKCTL08RF1-6)  
**5 Deepest Sampled Other Relevant Documents:** cqf44e00-16.9 (otRF08fbF-40), fem62f00-  
14.8 (otRF08fbF-35), bbf46e00-14.4 (UMKCTL08RF6-34), wkq02c00-8.5 (UMKCTL08RF1-  
20), bkl93a00-6.8 (otRF08fbF-16)

**Topic 28** (2006-C-4)

**Request Text:** Any and all federal tax returns.

**Initial Proposal by Defendant:** "tax returns" AND (IRS OR "federal")

**Rejoinder by Plaintiff:** form AND ("internal revenue service" OR IRS OR "I.R.S.")

**Final Negotiated Boolean Query:** "form" AND ("internal revenue service" OR IRS OR  
"I.R.S. ")

**Sampling:** 509451 pooled, 880 assessed, 90 judged highly relevant, 29 other judged relevant,  
761 judged non-relevant, 0 gray, "C"=2.64

**Est. Rel.:** 408.2 (including 312.3 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 23426, **F1:** 2.5%, (**Precision:** 1.3%, **Recall:** 68.5%)

**Participant High F1@K<sub>r</sub>:** 28.0% (UMKCTL08RF2), **Median F1@K<sub>r</sub>:** 3.9%

**5 Deepest Sampled Highly Relevant Documents:** taj86d00-34.7 (otRF08fb-270),  
wec31c00-33.2 (UCBM25T10Th5-258), mfs40e00-14.5 (UCBM25T10Th5-113), kux43c00-10.6  
(SabL08rfbase-82), waj86d00-9.7 (IowaSL08RF3B-75)

**5 Deepest Sampled Other Relevant Documents:** mwc05f00-21.2 (UCPwrT5Th5-  
165), eth31c00-13.3 (SabL08rf1-103), bde07c00-9.7 (IowaSL08RF3A-75), mir76d00-8.9  
(IowaSL08RF3B-69), njs36c00-5.7 (SabL08rfbase-44)

**Topic 31** (2006-C-7)

**Request Text:** All documents discussing or referencing the unlawful business practices and  
unfair competition sections of the California Business and Professions Code.

**Initial Proposal by Defendant:** "California Business and Professions Code" AND  
("unlawful business practices" OR "unfair competition")

**Rejoinder by Plaintiff:** ("California Business and Professions Code" OR "Cal!  
Bus! & Prof!" OR "B&P Code") AND ((172! OR 167!) OR (unlawful w/3 (bus! OR  
practice!) OR "unfair compet!"))

**Final Negotiated Boolean Query:** ("California Business and Professions Code" OR  
"Cal! Bus! & Prof!" OR "B&P Code") AND ((172! OR 167!) OR (unlawful w/3 (bus!  
OR practice!) OR "unfair compet!"))

**Sampling:** 566516 pooled, 380 assessed, 37 judged highly relevant, 42 other judged relevant,  
301 judged non-relevant, 0 gray, "C"=3.41

**Est. Rel.:** 708.2 (including 169.4 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 258, **F1:** 36.6%, (**Precision:** 100.0%, **Recall:** 22.4%)

**Participant High F1@K<sub>r</sub>:** 49.4% (UMKCTL08RF3), **Median F1@K<sub>r</sub>:** 21.8%

**5 Deepest Sampled Highly Relevant Documents:** dvy52d00-49.2 (IowaSL08RF3A-  
168), nam60c00-18.5 (IowaSL08RF1B-63), yvs62c00-15.3 (UCBM25T10Th5-52), gtc72c00-

11.4 (UMKCTL08RF6-39), ius62d00-8.8 (UCBM25T10Th5-30)

**5 Deepest Sampled Other Relevant Documents:** cje05a00-274.3 (IowaSL08RF2C-947), ajh62d00-45.4 (otRF08fbF-155), ynt82c00-29.3 (UMKCTL08RF1-100), boo41a00-20.8 (otRF08F-71), txp87d00-19.4 (UCRFBM25BL-66)

**Topic 36** (2006-D-3)

**Request Text:** All documents expressly referencing one or more actual, planned, or contemplated meetings between the CEO of a tobacco company and a foreign governmental official or officials.

**Initial Proposal by Defendant:** "Chief Executive Officer" AND meeting! AND "foreign government official"

**Rejoinder by Plaintiff:** ("Chief Executive Officer" OR CEO OR "C.E.O.") AND (meeting! OR conversation! OR discussion! OR negotiat!) AND (foreign w/5 (official! OR ministr! OR delegat! OR representative!))

**Final Negotiated Boolean Query:** ("Chief Executive Officer" OR CEO OR "C.E.O.") AND (meeting! OR conversation! OR discussion! OR negotiat!) AND (foreign w/5 (official! OR ministr! OR delegat! OR representative!))

**Sampling:** 505148 pooled, 380 assessed, 0 judged highly relevant, 24 other judged relevant, 356 judged non-relevant, 0 gray, "C"=2.68

**Est. Rel.:** 716.5 (including 0.0 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 377, **F1:** 8.5%, (**Precision:** 12.1%, **Recall:** 6.5%)

**Participant High F1@K<sub>r</sub>:** 17.0% (UMKCTL08RF5), **Median F1@K<sub>r</sub>:** 0.5%

**5 Deepest Sampled Highly Relevant Documents:**

**5 Deepest Sampled Other Relevant Documents:** ojy43c00-433.5 (otRF08rvl-1189), kpn01e00-76.8 (SabL08rfbase-207), las81f00-51.7 (otRF08rvl-139), iip85a00-36.1 (UMKCTL08RF1-97), cur77c00-35.0 (UMKCTL08RF2-94)

**Topic 47** (2006-E-6)

**Request Text:** All documents linking ammonia to unusual odor in humans or animals.

**Initial Proposal by Defendant:** ammonia AND ("unusual odor" w/10 (humans OR animals))

**Rejoinder by Plaintiff:** (ammonia! OR NH3) AND odor AND (human! OR animal!)

**Final Negotiated Boolean Query:** (ammonia! OR NH3) AND odor AND (human! OR animal!)

**Sampling:** 431128 pooled, 382 assessed, 0 judged highly relevant, 77 other judged relevant, 295 judged non-relevant, 10 gray, "C"=2.76

**Est. Rel.:** 26981.9 (including 0.0 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 6419, **F1:** 12.4%, (**Precision:** 33.4%, **Recall:** 7.6%)

**Participant High F1@K<sub>r</sub>:** 33.6% (otRF08frw), **Median F1@K<sub>r</sub>:** 1.0%

**5 Deepest Sampled Highly Relevant Documents:**

**5 Deepest Sampled Other Relevant Documents:** hhs45a00-11084.2 (IowaSL08RFTr-68528), uly13a00-6129.7 (otRF08fv-24360), hya63a00-1872.0 (IowaSL08RFTr-5692), bfr80e00-1785.1 (SabL08rf1-5402), zyp75e00-958.3 (SabL08rf1-2774)

**Topic 60** (2007-A-9)

**Request Text:** Please produce any and all documents that discuss phosphate precipitation as a method of water purification.

**Initial Proposal by Defendant:** (phosphate w/3 precipitation) AND (water w/3 purification)

**Rejoinder by Plaintiff:** phosphat! AND (precip! OR septic OR method!) AND purif!

**Final Negotiated Boolean Query:** (phosphat! w/75 (precip! OR septic OR method!)) AND ((water! OR waste!) w/75 purif!)

**Sampling:** 472743 pooled, 580 assessed, 55 judged highly relevant, 222 other judged relevant, 299 judged non-relevant, 4 gray, "C"=2.38

**Est. Rel.:** 36821.0 (including 917.6 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 1387, **F1:** 6.9%, (**Precision:** 83.7%, **Recall:** 3.6%)

**Participant High F1@K<sub>r</sub>:** 35.1% (otRF08rvl), **Median F1@K<sub>r</sub>:** 3.9%

**5 Deepest Sampled Highly Relevant Documents:** jec59c00-327.3 (otRF08fv-1410), ewc34d00-158.7 (otRF08fbF-678), meo09e00-84.4 (IowaSL08RF1A-359), dxe51f00-62.6 (UMKCTL08RF2-266), feb69c00-43.3 (IowaSL08RF2A-184)

**5 Deepest Sampled Other Relevant Documents:** dsj31a00-6145.7 (UCRFPwrBL-37598), fxn91f00-4999.0 (UMKCTL08RF2-28245), fdg34d00-4408.6 (UMKCTL08RF1-23966), mfm00d00-3993.6 (otRF08fv-21147), yxd74e00-1602.5 (IowaSL08RF1A-7383)

**Topic 73** (2007-B-5)

**Request Text:** Any advertisements or draft advertisements that target women seen in the kitchen or cooking.

**Initial Proposal by Defendant:** advertisement AND "target! women"

**Rejoinder by Plaintiff:** ("ad campaign" OR advertis!) AND (woman OR women OR girl OR female) AND (kitchen OR cook!)

**Final Negotiated Boolean Query:** (("ad campaign" OR advertis!) w/25 (woman OR women OR girl OR female)) AND (kitchen OR cook!)

**Sampling:** 661375 pooled, 380 assessed, 24 judged highly relevant, 91 other judged relevant, 265 judged non-relevant, 0 gray, "C"=1.89

**Est. Rel.:** 101196.9 (including 12246.0 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 4003, **F1:** 3.0%, (**Precision:** 36.6%, **Recall:** 1.5%)

**Participant High F1@K<sub>r</sub>:** 50.5% (otRF08rvl), **Median F1@K<sub>r</sub>:** 1.1%

**5 Deepest Sampled Highly Relevant Documents:** qcs79d00-9295.9 (otRF08fv-32820), rzz34a00-2270.2 (UCBM25T10Th5-4839), ywh62e00-348.9 (SabL08rfbase-671), red42e00-98.0 (UMKCTL08RF6-186), voh81e00-71.7 (UMKCTL08RF6-136)

**5 Deepest Sampled Other Relevant Documents:** bdk51a00-14268.7 (otRF08rvl-94087), zfk11d00-14247.5 (otRF08rvl-93600), efp64e00-14124.5 (UCRFBM25BL-90849), pnt52f00-11512.2 (IowaSL08RFTr-51258), nye02d00-5183.4 (otRF08rvl-13221)

**Topic 79** (2007-C-1)

**Request Text:** All documents making a connection between the music and songs of Peter, Paul,

and Mary, Joan Baez, or Bob Dylan, and the sale of cigarettes.

**Initial Proposal by Defendant:** (music OR songs) AND (((peter w/2 paul) AND (paul w/2 mary)) OR "bob Dylan" OR "joan baez") AND sale

**Rejoinder by Plaintiff:** ((peter AND paul AND mary) OR dylan OR baez) AND (sale! OR sell! OR advertis! OR promot! OR market!)

**Final Negotiated Boolean Query:** (((peter w/3 paul) AND (paul w/3 mary)) OR (simon w/3 garfunkel) OR "Bob Dylan" OR "Joan Baez") AND (sale! OR sell! OR advertis! OR promot! OR market!)

**Sampling:** 574925 pooled, 380 assessed, 16 judged highly relevant, 134 other judged relevant, 230 judged non-relevant, 0 gray, "C"=2.67

**Est. Rel.:** 56162.3 (including 199.9 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 262, **F1:** 0.4%, (**Precision:** 34.2%, **Recall:** 0.2%)

**Participant High F1@K<sub>r</sub>:** 30.9% (otRF08fbF), **Median F1@K<sub>r</sub>:** 1.4%

**5 Deepest Sampled Highly Relevant Documents:** vxu12d00-133.1 (SabL08rf1-358), tpe68d00-36.3 (SabL08rf1-97), trk04f00-13.8 (otRF08fb-37), iyb02c00-3.0 (UMKCTL08RF3-8), dhi54d00-2.6 (IowaSL08RF3A-7)

**5 Deepest Sampled Other Relevant Documents:** upp12e00-10903.6 (randomRF08-64032), oij29d00-7877.8 (UMKCTL08RF2-34715), gzj27d00-7138.6 (UCRFBM25BL-29650), bri24d00-6985.9 (IowaSL08RF3A-28675), gns15d00-4061.6 (SabL08rf1-13613)

#### Topic 80 (2007-C-2)

**Request Text:** All documents making a connection between folk songs and music and the sale of cigarettes.

**Initial Proposal by Defendant:** "folk songs" AND (sale! OR sell! OR promot! OR advertis! OR market!)

**Rejoinder by Plaintiff:** folk AND (sale OR sell! OR promot! OR advertis! OR market!)

**Final Negotiated Boolean Query:** ("folk songs" OR "folk music" OR "folk artists") AND (sale! OR sell! OR promot! OR advertis! OR market!)

**Sampling:** 704996 pooled, 880 assessed, 13 judged highly relevant, 330 other judged relevant, 537 judged non-relevant, 0 gray, "C"=2.30

**Est. Rel.:** 46094.8 (including 5053.6 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 255, **F1:** 0.6%, (**Precision:** 60.4%, **Recall:** 0.3%)

**Participant High F1@K<sub>r</sub>:** 23.0% (otRF08frw), **Median F1@K<sub>r</sub>:** 3.3%

**5 Deepest Sampled Highly Relevant Documents:** jye65d00-5003.4 (UCPwrT5Th5-46696), tjg02f00-22.1 (IowaSL08RFTr-155), ezu51a00-6.9 (otRF08rvl-48), xor03d00-4.4 (otRF08fbF-31), iqz75d00-3.4 (IowaSL08RF2B-24)

**5 Deepest Sampled Other Relevant Documents:** nvr55e00-8134.5 (otRF08F-95951), luw03d00-5952.5 (UCRFBM25BL-59306), pry15a00-4095.5 (SabL08rf1-36040), kw148c00-2539.3 (UCRFBM25BL-20354), fzo68d00-2353.7 (UCRFBM25BL-18668)

#### Topic 83 (2007-C-5)

**Request Text:** All documents discussing using psychedelic colors to increase sales of cigarettes.

**Initial Proposal by Defendant:** "psychedelic color!" AND (increas! w/3 sales)

**Rejoinder by Plaintiff:** psyched\*lic AND (sale! OR sell! OR promot! OR advertis! OR market)

**Final Negotiated Boolean Query:** psyched\*lic AND color! AND (sale! OR sell! OR promot! OR advertis! OR market!)

**Sampling:** 634204 pooled, 380 assessed, 0 judged highly relevant, 6 other judged relevant, 372 judged non-relevant, 2 gray, "C"=1.85

**Est. Rel.:** 830.6 (including 0.0 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 232, **F1:** 0.6%, (**Precision:** 2.0%, **Recall:** 0.4%)

**Participant High F1@K<sub>r</sub>:** 5.0% (otRF08frw), **Median F1@K<sub>r</sub>:** 0.1%

**5 Deepest Sampled Highly Relevant Documents:**

**5 Deepest Sampled Other Relevant Documents:** hqt23a00-825.6 (otRF08frw-1597), cos98c00-1.0 (IowaSL08RF2A-5), lbw83e00-1.0 (IowaSL08RF3B-5), ayk24d00-1.0 (SabL08rf1-4), kza51d00-1.0 (SabL08rf1-3)

**Topic 85** (2007-C-7)

**Request Text:** All documents discussing or referencing generally accepted accounting principles in connection with the decision to record as sales products shipped to distributors on a sale-or-return basis, and the implementation thereof.

**Initial Proposal by Defendant:** ("gaap" OR "generally accepted accounting principle!") AND (revenue! OR records OR recording OR account!) AND (sale w/5 return)

**Rejoinder by Plaintiff:** ("gaap" OR "generally accepted accounting principle!" OR "fasb" OR "financial accounting standards board" OR "sab" OR "Staff accounting bulletin" OR "sas" OR (statement w/2 "auditing standards")) AND ((sale! OR allowance OR reserve! OR right! OR entitle! OR could) w/5 return!)

**Final Negotiated Boolean Query:** ("gaap" OR "generally accepted accounting principle!" OR "fasb" OR "financial accounting standards board" OR "sab" OR "Staff accounting bulletin" OR "sas" OR (statement w/2 "auditing standards")) AND (revenue! OR recording OR records OR account!) AND (sale! OR allowance OR reserve!) AND ((right! OR entitle! OR could) w/5 return!)

**Sampling:** 559944 pooled, 378 assessed, 4 judged highly relevant, 29 other judged relevant, 345 judged non-relevant, 0 gray, "C"=2.70

**Est. Rel.:** 746.7 (including 22.4 estimated highly relevant)

**Final Boolean Result Size (B<sub>r</sub>):** 1253, **F1:** 63.6%, (**Precision:** 47.8%, **Recall:** 95.2%)

**Participant High F1@K<sub>r</sub>:** 63.6% (otRF08fb), **Median F1@K<sub>r</sub>:** 2.9%

**5 Deepest Sampled Highly Relevant Documents:** fyu57c00-17.8 (otRF08fbF-48), dob40d00-2.6 (otRF08frw-7), lpo95c00-1.0 (IowaSL08RF1B-1), erj43a00-1.0 (IowaSL08RF2A-1)

**5 Deepest Sampled Other Relevant Documents:** nmp51d00-132.1 (otRF08fb-359), cka71d00-123.7 (otRF08frw-336), tcp05d00-121.8 (otRF08fb-331), hpe33f00-88.5 (IowaSL08RF3B-240), oit55f00-70.5 (otRF08fbF-191)

**Topic 89** (2007-D-1)

**Request Text:** Submit all documents listing monthly and/or annual sales for companies in the

property and casualty insurance business in the United States between 1980 and the present.

**Initial Proposal by Defendant:** (("monthly sales" OR "annual sales") AND ("property insurance" OR "casualty insurance")) AND (United States OR U.S.) AND (198! OR 199!)

**Rejoinder by Plaintiff:** (month! OR annual!) AND (sales OR sell! OR revenue) AND insurance AND (198! OR 199!)

**Final Negotiated Boolean Query:** (((month! OR annual!) w/15 (sales OR sell! OR revenue)) AND ((property OR casualty) AND insurance)) BUT NOT (England OR "Great Britain" OR U.K. OR UK)

**Sampling:** 528679 pooled, 378 assessed, 37 judged highly relevant, 86 other judged relevant, 215 judged non-relevant, 40 gray, "C"=3.09

**Est. Rel.:** 11660.8 (including 4803.8 estimated highly relevant)

**Final Boolean Result Size ( $B_r$ ):** 3559, **F1:** 2.9%, (**Precision:** 7.7%, **Recall:** 1.8%)

**Participant High F1@ $K_r$ :** 18.9% (SabL08rfbase), **Median F1@ $K_r$ :** 6.3%

**5 Deepest Sampled Highly Relevant Documents:** yza94a00-3776.1 (SabL08rfbase-14385), whr48d00-306.8 (SabL08rfbase-963), aej48d00-263.5 (UMKCTL08RF3-825), fvu48d00-210.7 (SabL08rfbase-658), elg67d00-73.5 (IowaSL08RFTr-228)

**5 Deepest Sampled Other Relevant Documents:** ayb70c00-5063.2 (UCRFPwrBL-20950), zrd71d00-1010.5 (IowaSL08RF2C-3289), ctv60f00-196.4 (randomRF08-613), ukp48d00-161.5 (SabL08rfbase-503), eft62d00-64.2 (UCPwrT10Th5-199)