Appendix A   Overview of Ontology Relations from Different Data Sources

O: Object property, **D:** Datatype property

Property	Type	Domain	Range	Functional	Inverse	Origin
rescon:prefLabelConcept	D	rescon:Concept	String	---	---	Thesaurus: descriptor name
rescon:altLabelConcept	D	rescon:Concept	String	---	---	Thesaurus: use for
skos:narrower	O	rescon:Concept	rescon:Concept	---	skos:narrower	Thesaurus: broader terms
skos:broader	O	rescon:Concept	rescon:Concept	---	skos:broader	Thesaurus: narrower terms
skos:related	O	rescon:Concept	rescon:Concept	symmetric	skos:related	Thesaurus: related terms
rescon:synonymous	O	rescon:Concept	rescon:Concept	symmetric	rescon:synonymous	Thesaurus: synonyms
rescon:component	O	rescon:CompoundConcept	rescon:Concept	---	rescon:belongsToCompoundConcept	Thesaurus: compounds
rescon:belongsToCompoundConcept	O	rescon:Concept	rescon:CompoundConcept	---	rescon:component	Thesaurus: compounds

Table 75: Overview of ontology relations which are derived from the thesaurus

Property	Type	Domain	Range	Functional	Inverse	Origin
rescon:prefLabelOrganization	D	rescon:Organization	String	---	---	Organization db: organization name
rescon:altLabelOrganization	D	rescon:Organization	String	---	---	parsing of organization name
rescon:altLabelShortOrganization	D	rescon:Organization	String	---	---	Organization db: short name
rescon:organizationHasInterest	O	rescon:Organization	rescon:Concept	---	rescon:isInterestOfOrganization	Organization db: keywords
rescon:organizationHasPart	O	rescon:Organization	rescon:Organization	---	rescon:partOfOrganization	Organization db: hierarchical relations between organizations

Table 76: Overview of ontology relations which are derived from the organization database

Property	Type	Domain	Range	Functional	Inverse	Origin
rescon:prefLabelPerson	D	rescon:Person	String	---	---	Person db: first and second name
rescon:altLabelPerson	D	rescon:Person	String	---	---	Person db: second name; combination of second name, first name
rescon:hasResearchInterest	O	rescon:Person	rescon:Concept	---	rescon:isResearchInterestOfPersonFromPersonDb	Person db: extraction from the field research interests
rescon:hasWorkInterest	O	rescon:Person	rescon:Concept	---	rescon:isWorkInterestOfPersonFromPersonDb	Person db: extraction from the field work interests
rescon:personHasEmail	D	rescon:Person	String	---	---	Person db: email address
rescon:personHasAffiliation	O	rescon:Person	rescon:Organization	---	rescon:organizationEmploysPerson	Link between organization and person databases

Table 77: Overview of ontology relations which are derived from the person database

Property	Type	Domain	Range	Functional	Inverse	Origin
rescon:prefLabelProject	D	rescon:Project	String	---	---	Project db: title
rescon:altLabelProject	D	rescon:Project	String	---	---	Project db: first part of title, delimited by a colon
rescon:altLabelShortProject	D	rescon:Project	String	---	---	Project db: short title
rescon: projectHasSubject	O	rescon:Project	rescon:Concept	---	rescon:isSubjectOfProject	Project db: extraction from the keyword field
rescon:projectHasMember	O	rescon:Project	rescon:Person	---	rescon:personWorksAtProject	Project db: extraction from the field organizational structure
rescon:projectIsCarriedOutByOrganization	O	rescon:Project	rescon:Organization	---	rescon:organizationCarriesOutProject	Project db: extraction from the field organizational structure
rescon:partOfProject	O	rescon:Project	rescon:Project	---	rescon:projectHasPart	Project db: is part of
rescon:hasSuccessorProject	O	rescon:Project	rescon:Project	---	rescon:hasPredecessorProject	Project db: is predecessor of

Table 78: Overview of ontology relations which are derived from the project database

Appendix B   Distributions of Evaluation Measure Scores for Baseline and Expanded Queries

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
$F_{\beta=1}$-measure			
$F_{\beta=5}$-measure			
Relative precision			

Table 79: Distributions of evaluation measure scores for the expansion with alternative labels

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
$F_{\beta=1}$-measure			
$F_{\beta=5}$-measure			
Relative precision			

Table 80: Distributions of evaluation measure scores for the expansion with term components

274

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
$F_{\beta=1}$-measure			
$F_{\beta=5}$-measure			
Relative precision			

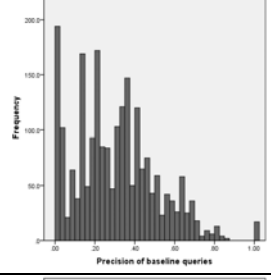**Table 81:** Distributions of evaluation measure scores for the expansion with broader terms
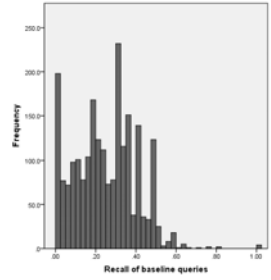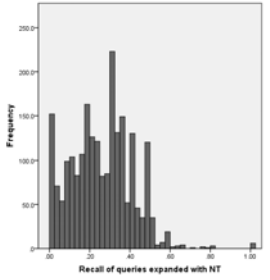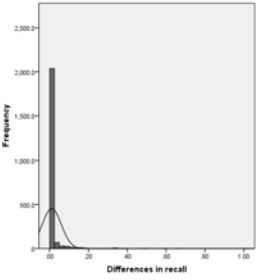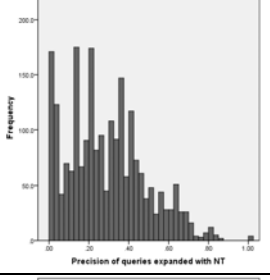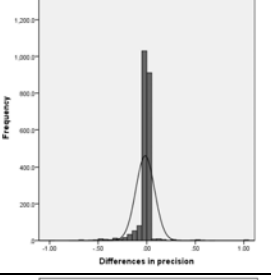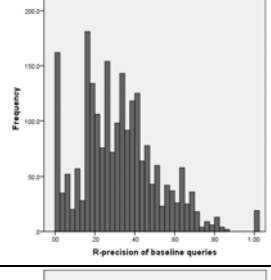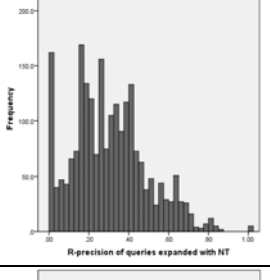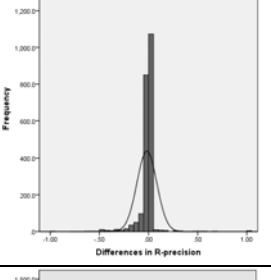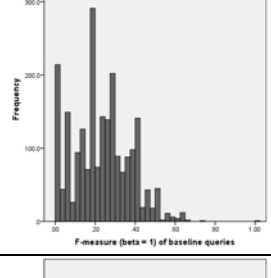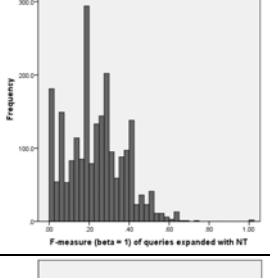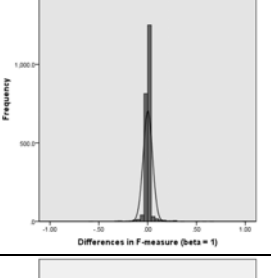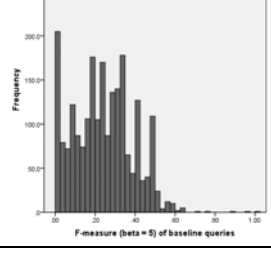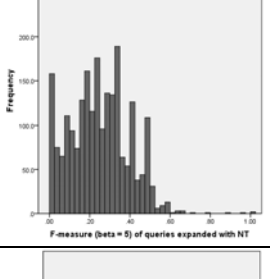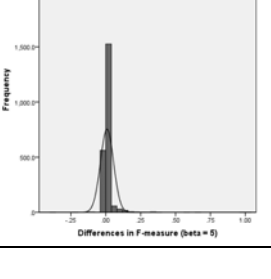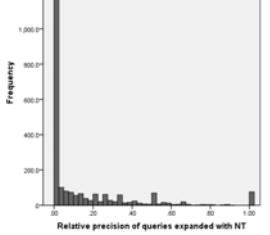
275

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
F$_{\beta=1}$-measure			
F$_{\beta=5}$-measure			
Relative precision			

Table 82: Distributions of evaluation measure scores for the expansion with narrower terms

276

	Baseline queries	**Expanded queries**	**Differences**
Recall			
Precision			
R-precision			
F$_{\beta=1}$-measure			
F$_{\beta=5}$-measure			
Relative precision			

Table 83: Distributions of evaluation measure scores for the expansion with related terms
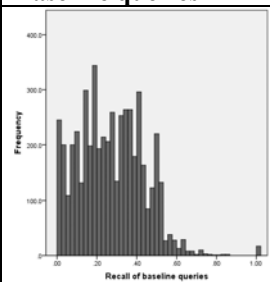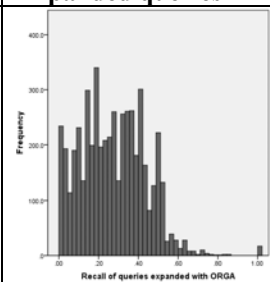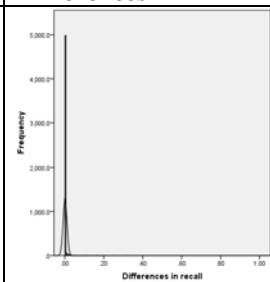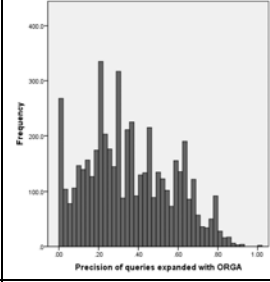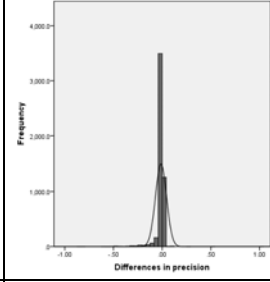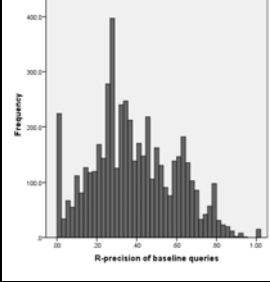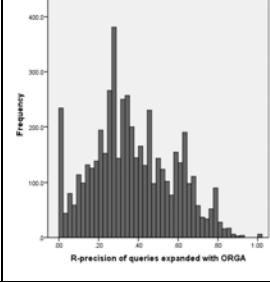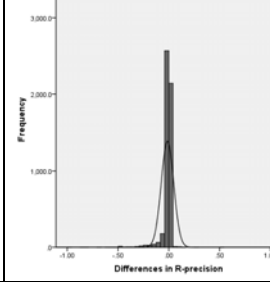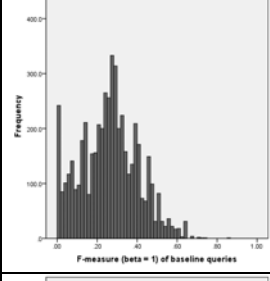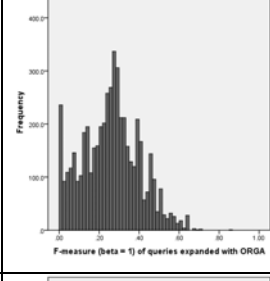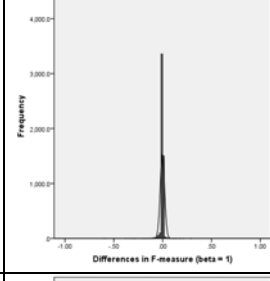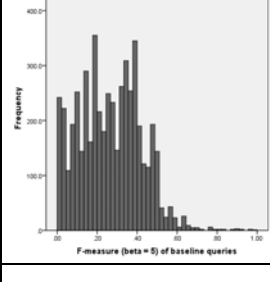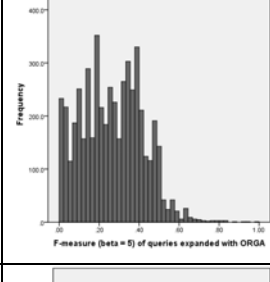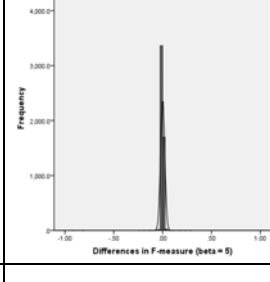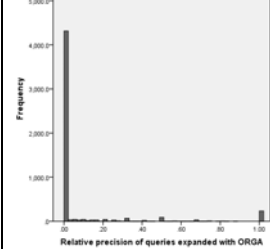
	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
F$_{\beta=1}$-measure			
F$_{\beta=5}$-measure			
Relative precision			

Table 84: Distributions of evaluation measure scores for the expansion with organization names
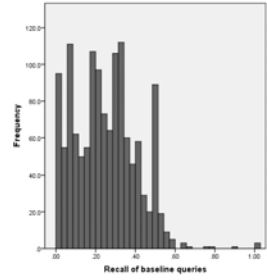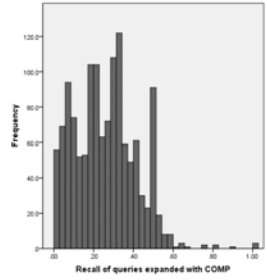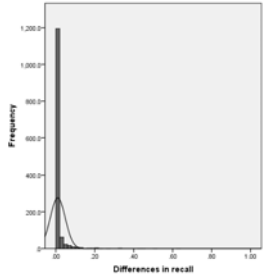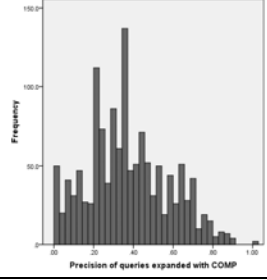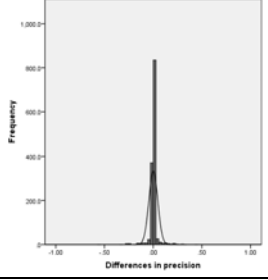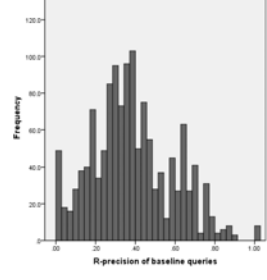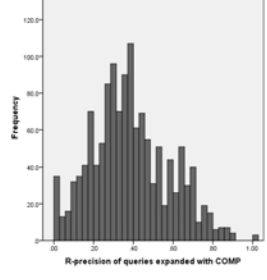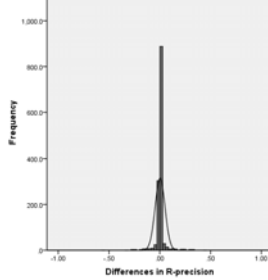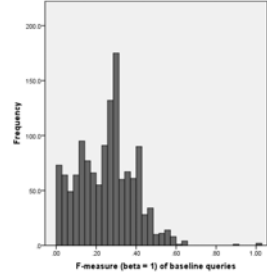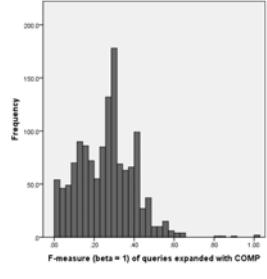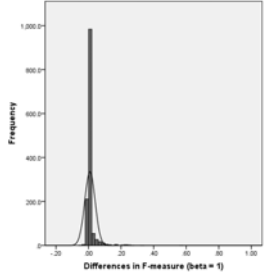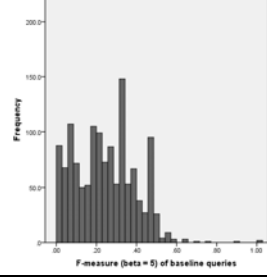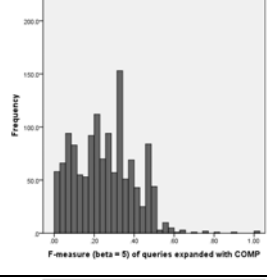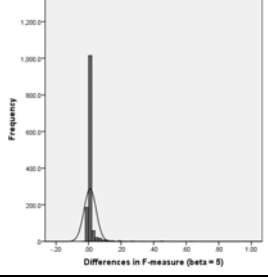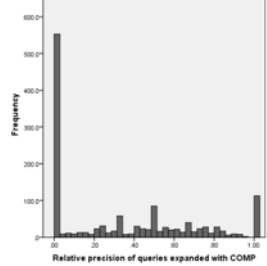
	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
$F_{\beta=1}$-measure			
$F_{\beta=5}$-measure			
Relative precision			

Table 85: Distributions of evaluation measure scores for the expansion with compounds

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
$F_{\beta=1}$-measure			
$F_{\beta=5}$-measure			
Relative precision			

Table 86: Distributions of evaluation measure scores for the expansion with synonyms
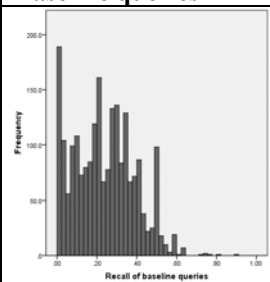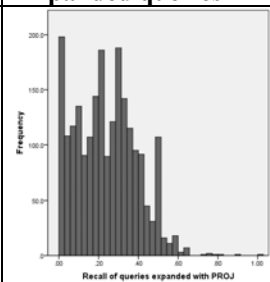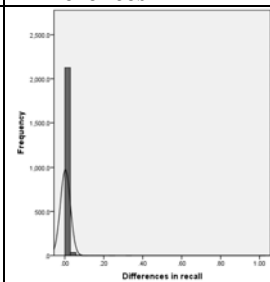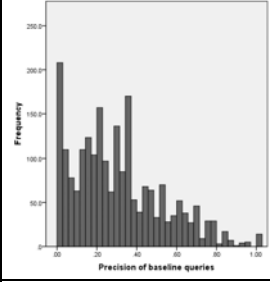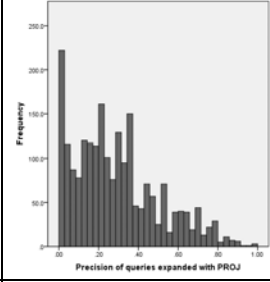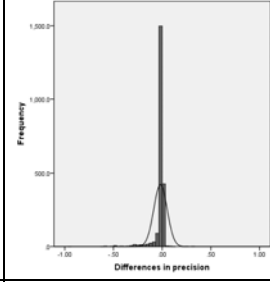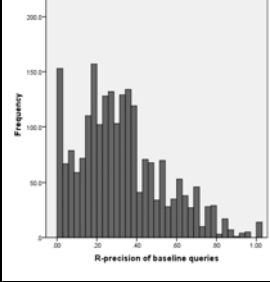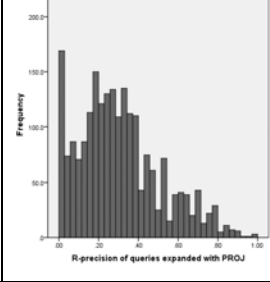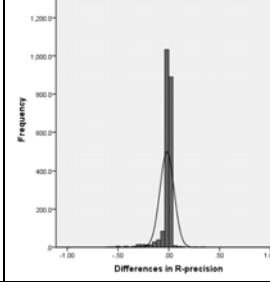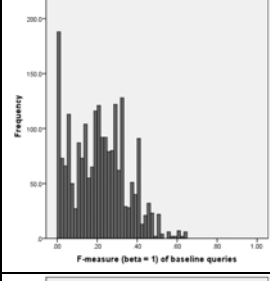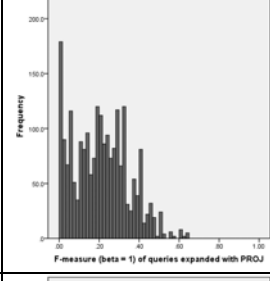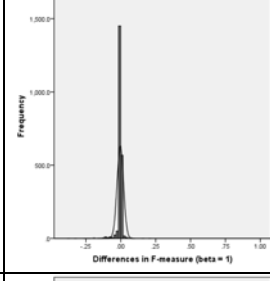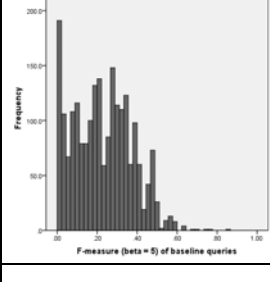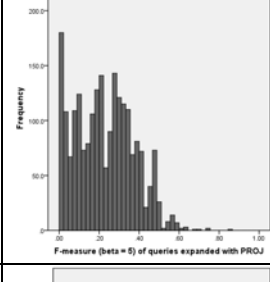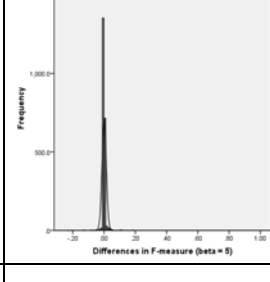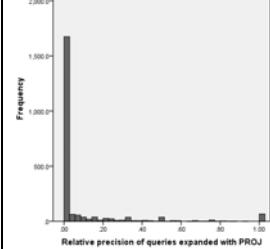
280

	Baseline queries	Expanded queries	Differences
Recall			
Precision			
R-precision			
F$_{\beta=1}$-measure			
F$_{\beta=5}$-measure			
Relative precision			

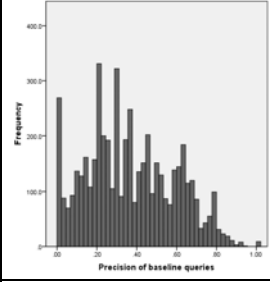**Table 87:** Distributions of evaluation measure scores for the expansion with project names

Appendix C   WS-R Test Results for the Comparison of Baseline and Expanded Queries

MR: Mean rank; **p**: two-tailed probability value

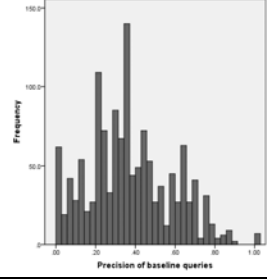|	Recall	Precision	R-precision	$F_{\beta=1}$	$F_{\beta=5}$
Baseline versus ALT	z = -48.610 p ≤ 0.001 (***) Ties: N = 1,338 $MR_{ALT<Baseline}$ = 0.00 $MR_{ALT>Baseline}$ = 1575.50	z = -0.982 p = 0.326 (ns) Ties: N = 731 $MR_{ALT<Baseline}$ = 1733.95 $MR_{ALT>Baseline}$ = 2043.75	z = -7.452 p ≤ 0.001 (***) Ties: N = 1,076 $MR_{ALT<Baseline}$ = 1514.63 $MR_{ALT>Baseline}$ = 1881.73	z = -38.225 p ≤ 0.001 (***) Ties: N = 684 $MR_{ALT<Baseline}$ = 1130.89 $MR_{ALT>Baseline}$ = 2145.13	z = -47.581 p ≤ 0.001 (***) Ties: N = 676 $MR_{ALT<Baseline}$ = 575.32 $MR_{ALT>Baseline}$ = 2203.84
Baseline versus TC	z = -18.645 p ≤ 0.001 (***) Ties: N = 202 $MR_{TC<Baseline}$ = 0.00 $MR_{TC>Baseline}$ = 232.00	z = -4.665 p ≤ 0.001 (***) Ties: N = 132 $MR_{TC<Baseline}$ = 273.36 $MR_{TC>Baseline}$ = 257.37	z = -4.134 p ≤ 0.001 (***) Ties: N = 153 $MR_{TC<Baseline}$ = 258.99 $MR_{TC>Baseline}$ = 252.77	z = -15.747 p ≤ 0.001 (***) Ties: N = 127 $MR_{TC<Baseline}$ = 158.50 $MR_{TC>Baseline}$ = 294.53	z = -18.767 p ≤ 0.001 (***) Ties: N = 127 $MR_{TC<Baseline}$ = 63.96 $MR_{TC>Baseline}$ = 302.79
Baseline versus BT	z = -33.063 p ≤ 0.001 (***) Ties: N = 880 $MR_{BT<Baseline}$ = 0.00 $MR_{BT>Baseline}$ = 729.00	z = -34.301 p ≤ 0.001 (***) Ties: N = 447 $MR_{BT<Baseline}$ = 1121.09 $MR_{BT>Baseline}$ = 216.84	z = -30.594 p ≤ 0.001 (***) Ties: N = 967 $MR_{BT<Baseline}$ = 709.16 $MR_{BT>Baseline}$ = 282.59	z = -29.508 p ≤ 0.001 (***) Ties: N = 448 $MR_{BT<Baseline}$ = 1152.09 $MR_{BT>Baseline}$ = 380.50	z = -3.515 p ≤ 0.001 (***) Ties: N = 447 $MR_{BT<Baseline}$ = 1089.07 $MR_{BT>Baseline}$ = 815.81
Baseline versus NT	z = -28.467 p ≤ 0.001 (***) Ties: N = 1,140 $MR_{NT<Baseline}$ = 0.00 $MR_{NT>Baseline}$ = 540.40	z = -21.316 p ≤ 0.001 (***) Ties: N = 599 $MR_{NT<Baseline}$ = 847.54 $MR_{NT>Baseline}$ = 709.50	z = -19.391 p ≤ 0.001 (***) Ties: N = 703 $MR_{NT<Baseline}$ = 823.44 $MR_{NT>Baseline}$ = 588.46	z = -0.606 p = 0.545 (ns) Ties: N = 584 $MR_{NT<Baseline}$ = 740.11 $MR_{NT>Baseline}$ = 911.80	z = -22.802 p ≤ 0.001 (***) Ties: N = 584 $MR_{NT<Baseline}$ = 409.40 $MR_{NT>Baseline}$ = 1037.84
Baseline versus RT	z = -20.356 p ≤ 0.001 (***) Ties: N = 467 $MR_{RT<Baseline}$ = 0.00 $MR_{RT>Baseline}$ = 276.50	z = -21.456 p ≤ 0.001 (***) Ties: N = 153 $MR_{RT<Baseline}$ = 462.15 $MR_{RT>Baseline}$ = 251.90	z = -19.753 p ≤ 0.001 (***) Ties: N = 293 $MR_{RT<Baseline}$ = 385.45 $MR_{RT>Baseline}$ = 215.89	z = -9.414 p ≤ 0.001 (***) Ties: N = 152 $MR_{RT<Baseline}$ = 470.02 $MR_{RT>Baseline}$ = 372.12	z = -9.074 p ≤ 0.001 (***) Ties: N = 151 $MR_{RT<Baseline}$ = 349.22 $MR_{RT>Baseline}$ = 491.57
Baseline versus ORGA	z = -25.683 p ≤ 0.001 (***) Ties: N = 4,252 $MR_{ORGA<Baseline}$ = 0.00 $MR_{ORGA>Baseline}$ = 440.00	z = -48.095 p ≤ 0.001 (***) Ties: N = 891 $MR_{ORGA<Baseline}$ = 2161.21 $MR_{ORGA>Baseline}$ =1714.04	z = -41.775 p ≤ 0.001 (***) Ties: N = 1,744 $MR_{ORGA<Baseline}$ = 1768.52 $MR_{ORGA>Baseline}$ = 1165.13	z = -35.749 p ≤ 0.001 (***) Ties: N = 890 $MR_{ORGA<Baseline}$ = 2053.74 $MR_{ORGA>Baseline}$ = 2483.99	z = -19.691 p ≤ 0.001 (***) Ties: N = 889 $MR_{ORGA<Baseline}$ = 1789.11 $MR_{ORGA>Baseline}$ = 3449.88
Baseline versus COMP	z = -24.210 p ≤ 0.001 (***) Ties: N = 551 $MR_{COMP<Baseline}$ = 0.00 $MR_{COMP>Baseline}$ = 391.00	z = -3.701 p ≤ 0.001 (***) Ties: N = 367 $MR_{COMP<Baseline}$ = 468.53 $MR_{COMP>Baseline}$ = 494.58	z = -6.232 p ≤ 0.001 (***) Ties: N = 368 $MR_{COMP<Baseline}$ = 486.85 $MR_{COMP>Baseline}$ = 479.83	z = -20.190 p ≤ 0.001 (***) Ties: N = 364 $MR_{COMP<Baseline}$ = 272.35 $MR_{COMP>Baseline}$ = 545.44	z = -24.391 p ≤ 0.001 (***) Ties: N = 364 $MR_{COMP<Baseline}$ = 119.14 $MR_{COMP>Baseline}$ = 571.98
Baseline versus SYN	z = -33.300 p ≤ 0.001 (***) Ties: N = 295 $MR_{SYN<Baseline}$ = 0.00	z = -4.750 p ≤ 0.001 (***) Ties: N = 140 $MR_{SYN<Baseline}$ = 719.80	z = -7.849 p ≤ 0.001 (***) Ties: N = 227 $MR_{SYN<Baseline}$ = 671.69	z = -30.225 p ≤ 0.001 (***) Ties: N = 121 $MR_{SYN<Baseline}$ = 361.68	z = -33.760 p ≤ 0.001 (***) Ties: N = 121 $MR_{SYN<Baseline}$ = 154.78

	$MR_{SYN>Baseline} = 739.50$	$MR_{SYN>Baseline} = 910.58$	$MR_{SYN>Baseline} = 854.50$	$MR_{SYN>Baseline} = 916.11$	$MR_{SYN>Baseline} = 909.15$
Baseline versus PROJ	$z = -20.973$ $p \leq 0.001$ (***) Ties: N = 1,588 $MR_{PROJ<Baseline} = 0.00$ $MR_{PROJ>Baseline} = 293.50$	$z = -31.369$ $p \leq 0.001$ (***) Ties: N = 240 $MR_{PROJ<Baseline} = 983.88$ $MR_{PROJ>Baseline} = 825.45$	$z = -26.309$ $p \leq 0.001$ (***) Ties: N = 688 $MR_{PROJ<Baseline} = 781.09$ $MR_{PROJ>Baseline} = 526.32$	$z = -21.036$ $p \leq 0.001$ (***) Ties: N = 239 $MR_{PROJ<Baseline} = 923.52$ $MR_{PROJ>Baseline} = 1161.93$	$z = -4.454$ $p \leq 0.001$ (***) Ties: N = 239 $MR_{PROJ<Baseline} = 753.63$ $MR_{PROJ>Baseline} = 1511.97$

Table 88: WS-R Test results for the comparison of baseline and expanded queries by different evaluation measures

Appendix D U-Test Results for the Comparisons of the Expansion Relations' Effects

z: z-value of the U-Test, **p:** Two-tailed probability value, **MR:** Mean rank, **B:** Baseline system, **AUT:** Automatic system, **INT**: Interactive system

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	$z = -1.792$ $p = 0.073$ (ns) $MR_{ALT} = 2591.10$ $MR_{TC} = 2481.82$	$z = -3.140$ $p \leq 0.01$ (**) $MR_{ALT} = 3359.89$ $MR_{BT} = 3515.00$	$z = -31.321$ $p \leq 0.001$ (***) $MR_{ALT} = 3862.07$ $MR_{NT} = 2328.38$	$z = -15.292$ $p \leq 0.001$ (***) $MR_{ALT} = 2907.34$ $MR_{RT} = 2078.65$	$z = -60.115$ $p \leq 0.001$ (***) $MR_{ALT} = 6441.58$ $MR_{ORGA} = 3382.89$	$z = -28.853$ $p \leq 0.001$ (***) $MR_{ALT} = 3191.77$ $MR_{COMP} = 1962.81$	$z = -9.868$ $p \leq 0.001$ (***) $MR_{ALT} = 2990.58$ $MR_{SYN} = 3486.45$	$z = -40.909$ $p \leq 0.001$ (***) $MR_{ALT} = 3973.36$ $MR_{PROJ} = 2006.44$
TC		$z = -4.769$ $p \leq 0.001$ (***) $MR_{TC} = 1363.42$ $MR_{BT} = 1540.79$	$z = -19.240$ $p \leq 0.001$ (***) $MR_{TC} = 1959.99$ $MR_{NT} = 1288.14$	$z = -10.744$ $p \leq 0.001$ (***) $MR_{TC} = 995.07$ $MR_{RT} = 742.93$	$z = -36.065$ $p \leq 0.001$ (***) $MR_{TC} = 4525.79$ $MR_{ORGA} = 2687.60$	$z = -16.334$ $p \leq 0.001$ (***) $MR_{TC} = 1289.19$ $MR_{COMP} = 854.12$	$z = -8.488$ $p \leq 0.001$ (***) $MR_{TC} = 1022.75$ $MR_{SYN} = 1293.30$	$z = -26.729$ $p \leq 0.001$ (***) $MR_{TC} = 2063.59$ $MR_{PROJ} = 1223.14$
BT			$z = -24.425$ $p \leq 0.001$ (***) $MR_{BT} = 2722.29$ $MR_{NT} = 1812.35$	$z = -14.055$ $p \leq 0.001$ (***) $MR_{BT} = 1828.63$ $MR_{RT} = 1334.19$	$z = -46.793$ $p \leq 0.001$ (***) $MR_{BT} = 5156.16$ $MR_{ORGA} = 3086.98$	$z = -18.870$ $p \leq 0.001$ (***) $MR_{BT} = 2076.62$ $MR_{COMP} = 1411.07$	$z = -4.114$ $p \leq 0.001$ (***) $MR_{BT} = 1989.95$ $MR_{SYN} = 2141.90$	$z = -32.501$ $p \leq 0.001$ (***) $MR_{BT} = 2811.85$ $MR_{PROJ} = 1658.47$
NT				$z = -8.630$ $p \leq 0.001$ (***) $MR_{NT} = 1529.98$ $MR_{RT} = 1816.11$	$z = -30.466$ $p \leq 0.001$ (***) $MR_{NT} = 4567.97$ $MR_{ORGA} = 3290.08$	$z = -4.228$ $p \leq 0.001$ (***) $MR_{NT} = 1723.27$ $MR_{COMP} = 1865.22$	$z = -36.494$ $p \leq 0.001$ (***) $MR_{NT} = 1415.98$ $MR_{SYN} = 2724.50$	$z = -16.670$ $p \leq 0.001$ (***) $MR_{NT} = 2472.85$ $MR_{PROJ} = 1916.33$
RT					$z = -29.925$ $p \leq 0.001$ (***) $MR_{RT} = 4201.17$ $MR_{ORGA} = 2851.95$	$z = -5.121$ $p \leq 0.001$ (***) $MR_{RT} = 1254.57$ $MR_{COMP} = 1115.90$	$z = -21.768$ $p \leq 0.001$ (***) $MR_{RT} = 962.92$ $MR_{SYN} = 1645.69$	$z = -19.662$ $p \leq 0.001$ (***) $MR_{RT} = 1998.26$ $MR_{PROJ} = 1408.92$
ORGA						$z = -32.725$ $p \leq 0.001$ (***) $MR_{ORGA} = 2934.87$ $MR_{COMP} = 4376.59$	$z = -58.896$ $p \leq 0.001$ (***) $MR_{ORGA} = 2750.60$ $MR_{SYN} = 5483.79$	$z = -10.036$ $p \leq 0.001$ (***) $MR_{ORGA} = 3540.27$ $MR_{PROJ} = 3919.06$
COMP							$z = -30.693$ $p \leq 0.001$ (***) $MR_{COMP} = 989.08$ $MR_{SYN} = 1976.65$	$z = -19.718$ $p \leq 0.001$ (***) $MR_{COMP} = 2132.09$ $MR_{PROJ} = 1521.54$
SYN								$z = -43.585$ $p \leq 0.001$ (***) $MR_{SYN} = 2800.52$ $MR_{PROJ} = 1299.93$

Table 89: U-Test results for the comparison by differences in recall

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	z = -5.471 p ≤ 0.001 (***) MR_{ALT} = 2620.54 MR_{TC} = 2283.14	z = -32.622 p ≤ 0.001 (***) MR_{ALT} = 3994.05 MR_{BT} = 2297.15	z = -9.249 p ≤ 0.001 (***) MR_{ALT} = 3491.16 MR_{NT} = 3078.23	z = -19.710 p ≤ 0.001 (***) MR_{ALT} = 2954.80 MR_{RT} = 1869.63	z = -13.882 p ≤ 0.001 (***) MR_{ALT} = 5229.25 MR_{ORGA} = 4443.29	z = -3.431 p ≤ 0.001 (***) MR_{ALT} = 2869.47 MR_{COMP} = 3048.75	z = -2.612 p ≤ 0.05 (*) MR_{ALT} = 3093.55 MR_{SYN} = 3225.81	z = -12.067 p ≤ 0.001 (***) MR_{ALT} = 3529.10 MR_{PROJ} = 2923.58
TC		z = -14.682 p ≤ 0.001 (***) MR_{TC} = 1935.35 MR_{BT} = 1378.05	z = -0.950 p = 0.342 (ns) MR_{TC} = 1416.29 MR_{NT} = 1451.00	z = -7.186 p ≤ 0.001 (***) MR_{TC} = 947.65 MR_{RT} = 773.88	z = -0.644 p = 0.520 (ns) MR_{TC} = 2937.69 MR_{ORGA} = 2893.42	z = -6.537 p ≤ 0.001 (***) MR_{TC} = 880.56 MR_{COMP} = 1058.13	z = -6.417 p ≤ 0.001 (***) MR_{TC} = 1070.22 MR_{SYN} = 1275.49	z = -1.360 p = 0.174 (ns) MR_{TC} = 1457.78 MR_{PROJ} = 1408.44
BT			z = -23.040 p ≤ 0.001 (***) MR_{BT} = 1843.92 MR_{NT} = 2737.01	z = -9.413 p ≤ 0.001 (***) MR_{BT} = 1574.84 MR_{RT} = 1916.24	z = -23.855 p ≤ 0.001 (***) MR_{BT} = 2855.20 MR_{ORGA} = 4134.99	z = -28.468 p ≤ 0.001 (***) MR_{BT} = 1461.20 MR_{COMP} = 2490.84	z = -29.321 p ≤ 0.001 (***) MR_{BT} = 1583.49 MR_{SYN} = 2677.66	z = -19.079 p ≤ 0.001 (***) MR_{BT} = 1899.80 MR_{PROJ} = 2638.90
NT				z = -17.706 p ≤ 0.001 (***) MR_{NT} = 1815.93 MR_{RT} = 1193.16	z = -8.244 p ≤ 0.001 (***) MR_{NT} = 3984.97 MR_{ORGA} = 3542.32	z = -17.362 p ≤ 0.001 (***) MR_{NT} = 1547.35 MR_{COMP} = 2158.42	z = -7.882 p ≤ 0.001 (***) MR_{NT} = 1868.89 MR_{SYN} = 2157.41	z = -8.125 p ≤ 0.001 (***) MR_{NT} = 2350.85 MR_{PROJ} = 2040.90
RT					z = -18.410 p ≤ 0.001 (***) MR_{RT} = 2142.49 MR_{ORGA} = 3260.79	z = -24.819 p p ≤ 0.001 (***) MR_{RT} = 780.90 MR_{COMP} = 1478.26	z = -17.536 p ≤ 0.001 (***) MR_{RT} = 1044.02 MR_{SYN} = 1599.08	z = -15.406 p ≤ 0.001 (***) MR_{RT} = 1230.22 MR_{PROJ} = 1768.92
ORGA						z = -27.558 p ≤ 0.001 (***) MR_{ORGA} = 2907.32 MR_{COMP} = 4482.69	z = -9.816 p p ≤ 0.001 (***) MR_{ORGA} = 3314.31 MR_{SYN} = 3852.41	z = -1.508 p = 0.132 (ns) MR_{ORGA} = 3677.17 MR_{PROJ} = 3595.96
COMP							z = -0.698 p = 0.485 (ns) MR_{COMP} = 1540.08 MR_{SYN} = 1562.71	z = -24.874 p ≤ 0.001 (***) MR_{COMP} = 2295.35 MR_{PROJ} = 1421.51
SYN								z = -9.351 p ≤ 0.001 (***) MR_{SYN} = 2161.73 MR_{PROJ} = 1820.90

Table 90: U-Test results for the comparison by differences in precision

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	z = -6.851 p ≤ 0.001 (***) MR_{ALT} = 2630.28 MR_{TC} = 2209.82	z = -32.700 p ≤ 0.001 (***) MR_{ALT} = 3967.22 MR_{BT} = 2346.23	z = -12.458 p ≤ 0.001 (***) MR_{ALT} = 3558.74 MR_{NT} = 2938.76	z = -19.894 p ≤ 0.001 (***) MR_{ALT} = 2954.53 MR_{RT} = 1865.81	z = -18.988 p ≤ 0.001 (***) MR_{ALT} = 5376.04 MR_{ORGA} = 4312.20	z = -0.971 p = 0.332 (ns) MR_{ALT} = 2897.94 MR_{COMP} = 2948.44	z = -2.809 p ≤ 0.01 (**) MR_{ALT} = 3089.85 MR_{SYN} = 3231.60	z = -15.360 p ≤ 0.001 (***) MR_{ALT} = 3579.49 MR_{PROJ} = 2814.94
TC		z = -12.575 p ≤ 0.001 (***) MR_{TC} = 1864.60 MR_{BT} = 1398.18	z = -0.969 p = 0.332 (ns) MR_{TC} = 1415.90 MR_{NT} = 1451.12	z = -6.292 p ≤ 0.001 (***) MR_{TC} = 933.93 MR_{RT} = 782.83	z = -0.156 p = 0.876 (ns) MR_{TC} = 2907.34 MR_{ORGA} = 2896.79	z = -7.055 p ≤ 0.001 (***) MR_{TC} = 871.31 MR_{COMP} = 1062.75	z = -7.524 p ≤ 0.001 (***) MR_{TC} = 1044.70 MR_{SYN} = 1285.06	z = -0.515 p = 0.607 (ns) MR_{TC} = 1433.63 MR_{PROJ} = 1415.18
BT			z = -21.028 p ≤ 0.001 (***) MR_{BT} = 1889.52 MR_{NT} = 2689.01	z = -6.861 p ≤ 0.001 (***) MR_{BT} = 1604.76 MR_{RT} = 1847.62	z = -20.596 p ≤ 0.001 (***) MR_{BT} = 2991.19 MR_{ORGA} = 4072.39	z = -30.203 p ≤ 0.001 (***) MR_{BT} = 1445.96 MR_{COMP} = 2517.58	z = -27.760 p ≤ 0.001 (***) MR_{BT} = 1613.49 MR_{SYN} = 2638.11	z = -17.064 p ≤ 0.001 (***) MR_{BT} = 1944.46 MR_{PROJ} = 2590.01
NT				z = -14.022 p ≤ 0.001 (***) MR_{NT} = 1773.81 MR_{RT} = 1284.91	z = -4.925 p ≤ 0.001 (***) MR_{NT} = 3857.34 MR_{ORGA} = 3596.81	z = -17.682 p ≤ 0.001 (***) MR_{NT} = 1544.08 MR_{COMP} = 2163.86	z = -11.612 p ≤ 0.001 (***) MR_{NT} = 1808.88 MR_{SYN} = 2232.55	z = -4.647 p ≤ 0.001 (***) MR_{NT} = 2283.56 MR_{PROJ} = 2108.56
RT					z = -13.970 p ≤ 0.001 (***) MR_{RT} = 2378.40 MR_{ORGA} = 3213.37	z = -23.254 p ≤ 0.001 (***) MR_{RT} = 807.95 MR_{COMP} = 1457.57	z = -17.867 p ≤ 0.001 (***) MR_{RT} = 1038.10 MR_{SYN} = 1602.48	z = -11.841 p ≤ 0.001 (***) MR_{RT} = 1318.57 MR_{PROJ} = 1726.83
ORGA						z = -25.875 p ≤ 0.001 (***) MR_{ORGA} = 2930.90 MR_{COMP} = 4389.21	z = -14.407 p ≤ 0.001 (***) MR_{ORGA} = 3251.22 MR_{SYN} = 4032.93	z = -0.862 p = 0.389 (ns) MR_{ORGA} = 3665.59 MR_{PROJ} = 3619.92
COMP							z = -3.293 p ≤ 0.001 (***) MR_{COMP} = 1492.09 MR_{SYN} = 1598.76	z = -22.561 p ≤ 0.001 (***) MR_{COMP} = 2238.77 MR_{PROJ} = 1455.24
SYN								z = -12.909 p ≤ 0.001 (***) MR_{SYN} = 2231.07 MR_{PROJ} = 1763.34

Table 91: U-Test results for the comparison by differences in R-precision

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	z = -5.484 p ≤ 0.001 (***) MR_{ALT} = 2620.15 MR_{TC} = 2285.77	z = -33.848 p ≤ 0.001 (***) MR_{ALT} = 3964.67 MR_{BT} = 2353.57	z = -20.567 p ≤ 0.001 (***) MR_{ALT} = 3687.73 MR_{NT} = 2680.84	z = -17.426 p ≤ 0.001 (***) MR_{ALT} = 2928.70 MR_{RT} = 1984.57	z = -49.494 p ≤ 0.001 (***) MR_{ALT} = 6153.12 MR_{ORGA} = 3635.19	z = -1.699 p = 0.089 (ns) MR_{ALT} = 2890.48 MR_{COMP} = 2977.96	z = -11.618 p ≤ 0.001 (***) MR_{ALT} = 2965.71 MR_{SYN} = 3549.40	z = -33.384 p ≤ 0.001 (***) MR_{ALT} = 3855.21 MR_{PROJ} = 2250.35
TC		z = -18.638 p ≤ 0.001 (***) MR_{TC} = 2041.12 MR_{BT} = 1347.95	z = -9.264 p ≤ 0.001 (***) MR_{TC} = 1691.91 MR_{NT} = 1368.44	z = -10.169 p ≤ 0.001 (***) MR_{TC} = 986.91 MR_{RT} = 748.26	z = -28.199 p ≤ 0.001 (***) MR_{TC} = 4170.76 MR_{ORGA} = 2733.61	z = -5.429 p ≤ 0.001 (***) MR_{TC} = 902.58 MR_{COMP} = 1047.14	z = -13.597 p ≤ 0.001 (***) MR_{TC} = 904.36 MR_{SYN} = 1337.70	z = -20.104 p ≤ 0.001 (***) MR_{TC} = 1904.06 MR_{PROJ} = 1271.93
BT			z = -7.518 p ≤ 0.001 (***) MR_{BT} = 2142.55 MR_{NT} = 2422.64	z = -7.325 p ≤ 0.001 (***) MR_{BT} = 1600.26 MR_{RT} = 1857.95	z = -30.210 p ≤ 0.001 (***) MR_{BT} = 4652.32 MR_{ORGA} = 3316.46	z = -16.527 p ≤ 0.001 (***) MR_{BT} = 1623.38 MR_{COMP} = 2206.30	z = -38.718 p ≤ 0.001 (***) MR_{BT} = 1438.65 MR_{SYN} = 2868.57	z = -13.989 p ≤ 0.001 (***) MR_{BT} = 2495.24 MR_{PROJ} = 1998.82
NT				z = -0.899 p = 0.369 (ns) MR_{NT} = 1629.38 MR_{RT} = 1599.57	z = -25.782 p ≤ 0.001 (***) MR_{NT} = 4430.80 MR_{ORGA} = 3349.43	z = -13.837 p ≤ 0.001 (***) MR_{NT} = 1602.33 MR_{COMP} = 2066.78	z = -27.241 p ≤ 0.001 (***) MR_{NT} = 1563.33 MR_{SYN} = 2540.01	z = -14.738 p ≤ 0.001 (***) MR_{NT} = 2440.93 MR_{PROJ} = 1948.92
RT					z = -22.345 p ≤ 0.001 (***) MR_{RT} = 3915.96 MR_{ORGA} = 2908.59	z = -12.055 p ≤ 0.001 (***) MR_{RT} = 991.07 MR_{COMP} = 1317.48	z = -24.468 p ≤ 0.001 (***) MR_{RT} = 909.17 MR_{SYN} = 1676.58	z = -12.898 p ≤ 0.001 (***) MR_{RT} = 1860.21 MR_{PROJ} = 1473.63
ORGA						z = -31.982 p ≤ 0.001 (***) MR_{ORGA} = 2941.65 MR_{COMP} = 4350.45	z = -49.351 p ≤ 0.001 (***) MR_{ORGA} = 2864.40 MR_{SYN} = 5154.45	z = -8.295 p ≤ 0.001 (***) MR_{ORGA} = 3559.84 MR_{PROJ} = 3872.88
COMP							z = -5.044 p ≤ 0.001 (***) MR_{COMP} = 1460.34 MR_{SYN} = 1622.61	z = -22.669 p ≤ 0.001 (***) MR_{COMP} = 2132.09 MR_{PROJ} = 1486.85
SYN								z = -36.775 p ≤ 0.001 (***) MR_{SYN} = 2671.34 MR_{PROJ} = 1405.29

Table 92: U-Test results for the comparison by relative precision

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	$z = -1.706$ $p = 0.088$ (ns) $MR_{ALT} = 2563.42$ $MR_{TC} = 2668.65$	$z = -44.420$ $p \leq 0.001$ (***) $MR_{ALT} = 4175.75$ $MR_{BT} = 1948.21$	$z = -29.275$ $p \leq 0.001$ (***) $MR_{ALT} = 3839.69$ $MR_{NT} = 2373.63$	$z = -25.906$ $p \leq 0.001$ (***) $MR_{ALT} = 3018.00$ $MR_{RT} = 1591.27$	$z = -48.845$ $p \leq 0.001$ (***) $MR_{ALT} = 6285.50$ $MR_{ORGA} = 3519.40$	$z = -15.792$ $p \leq 0.001$ (***) $MR_{ALT} = 3099.43$ $MR_{COMP} = 2273.94$	$z = -12.637$ $p \leq 0.001$ (***) $MR_{ALT} = 2949.75$ $MR_{SYN} = 3589.79$	$z = -37.872$ $p \leq 0.001$ (***) $MR_{ALT} = 3951.78$ $MR_{PROJ} = 2050.99$
TC		$z = -27.163$ $p \leq 0.001$ (***) $MR_{TC} = 2304.19$ $MR_{BT} = 1273.09$	$z = -19.886$ $p \leq 0.001$ (***) $MR_{TC} = 2002.21$ $MR_{NT} = 1275.49$	$z = -19.780$ $p \leq 0.001$ (***) $MR_{TC} = 1131.98$ $MR_{RT} = 653.59$	$z = -27.131$ $p \leq 0.001$ (***) $MR_{TC} = 4550.44$ $MR_{ORGA} = 2684.40$	$z = -13.336$ $p \leq 0.001$ (***) $MR_{TC} = 1240.72$ $MR_{COMP} = 878.32$	$z = -7.179$ $p \leq 0.001$ (***) $MR_{TC} = 1052.48$ $MR_{SYN} = 1282.14$	$z = -25.090$ $p \leq 0.001$ (***) $MR_{TC} = 2117.14$ $MR_{PROJ} = 1206.75$
BT			$z = -23.308$ $p \leq 0.001$ (***) $MR_{BT} = 1838.83$ $MR_{NT} = 2742.37$	$z = -12.756$ $p \leq 0.001$ (***) $MR_{BT} = 1538.02$ $MR_{RT} = 2000.68$	$z = -18.129$ p $p \leq 0.001$ (***) $MR_{BT} = 3066.25$ $MR_{ORGA} = 4038.87$	$z = -27.834$ $p \leq 0.001$ (***) $MR_{BT} = 1469.52$ $MR_{COMP} = 2376.24$	$z = -42.202$ $p \leq 0.001$ (***) $MR_{BT} = 1376.03$ $MR_{SYN} = 2951.11$	$z = -15.337$ $p \leq 0.001$ (***) $MR_{BT} = 1969.67$ $MR_{PROJ} = 2563.80$
NT				$z = -9.197$ $p \leq 0.001$ (***) $MR_{NT} = 1721.79$ $MR_{RT} = 1398.25$	$z = -18.954$ $p \leq 0.001$ (***) $MR_{NT} = 4386.35$ $MR_{ORGA} = 3368.65$	$z = -16.204$ $p \leq 0.001$ (***) $MR_{NT} = 1562.58$ $MR_{COMP} = 2133.03$	$z = -33.419$ $p \leq 0.001$ (***) $MR_{NT} = 1453.65$ $MR_{SYN} = 2677.34$	$z = -15.027$ $p \leq 0.001$ (***) $MR_{NT} = 2481.15$ $MR_{PROJ} = 1907.85$
RT					$z = -3.278$ $p \leq 0.001$ (***) $MR_{RT} = 2909.38$ $MR_{ORGA} = 3108.49$	$z = -16.370$ $p \leq 0.001$ (***) $MR_{RT} = 915.37$ $MR_{COMP} = 1375.39$	$z = -29.204$ $p \leq 0.001$ (***) $MR_{RT} = 809.08$ $MR_{SYN} = 1734.11$	$z = -3.051$ $p \leq 0.01$ (**) $MR_{RT} = 1524.37$ $MR_{PROJ} = 1631.05$
ORGA						$z = -35.891$ $p \leq 0.001$ (***) $MR_{ORGA} = 2809.13$ $MR_{COMP} = 4860.93$	$z = -43.833$ $p \leq 0.001$ (***) $MR_{ORGA} = 2835.39$ $MR_{SYN} = 5238.39$	$z = -1.165$ $p = 0.244$ (ns) $MR_{ORGA} = 3634.32$ $MR_{PROJ} = 3697.09$
COMP							$z = -24.918$ $p \leq 0.001$ (***) $MR_{COMP} = 1091.36$ $MR_{SYN} = 1899.82$	$z = -30.599$ $p \leq 0.001$ (***) $MR_{COMP} = 2420.09$ $MR_{PROJ} = 1345.08$
SYN								$z = -37.573$ $p \leq 0.001$ (***) $MR_{SYN} = 2728.38$ $MR_{PROJ} = 1358.77$

Table 93: U-Test results for the comparison by differences in the $F_{\beta=1}$-measure

	TC	BT	NT	RT	ORGA	COMP	SYN	PROJ
ALT	z = -0.442 p = 0.659 (ns) MR_{ALT} = 2580.52 MR_{TC} = 2553.26	z = -27.616 p ≤ 0.001 (***) MR_{ALT} = 3887.23 MR_{BT} = 2502.29	z = -29.333 p ≤ 0.001 (***) MR_{ALT} = 3840.69 MR_{NT} = 2371.61	z = -19.391 p ≤ 0.001 (***) MR_{ALT} = 2951.62 MR_{RT} = 1883.61	z = -55.187 p ≤ 0.001 (***) MR_{ALT} = 6477.12 MR_{ORGA} = 3351.80	z = -20.921 p ≤ 0.001 (***) MR_{ALT} = 3160.80 MR_{COMP} = 2067.14	z = -10.459 p ≤ 0.001 (***) MR_{ALT} = 2980.99 MR_{SYN} = 3510.73	z = -41.228 p ≤ 0.001 (***) MR_{ALT} = 4006.76 MR_{PROJ} = 1937.49
TC		z = -16.386 p ≤ 0.001 (***) MR_{TC} = 1985.74 MR_{BT} = 1363.71	z = -19.063 p ≤ 0.001 (***) MR_{TC} = 1979.09 MR_{NT} = 1282.42	z = -14.619 p ≤ 0.001 (***) MR_{TC} = 1056.46 MR_{RT} = 702.87	z = -29.548 p ≤ 0.001 (***) MR_{TC} = 4697.60 MR_{ORGA} = 2665.33	z = -15.566 p ≤ 0.001 (***) MR_{TC} = 1281.14 MR_{COMP} = 858.14	z = -7.839 p ≤ 0.001 (***) MR_{TC} = 1037.11 MR_{SYN} = 1287.91	z = -26.304 p ≤ 0.001 (***) MR_{TC} = 2150.88 MR_{PROJ} = 1196.43
BT			z = -3.664 p ≤ 0.001 (***) MR_{BT} = 2209.80 MR_{NT} = 2351.85	z = -3.905 p ≤ 0.001 (***) MR_{BT} = 1635.49 MR_{RT} = 1777.14	z = -9.482 p ≤ 0.001 (***) MR_{BT} = 4084.01 MR_{ORGA} = 3575.31	z = -6.042 p ≤ 0.001 (***) MR_{BT} = 1755.67 MR_{COMP} = 1974.19	z = -30.201 p ≤ 0.001 (***) MR_{BT} = 1569.25 MR_{SYN} = 2696.43	z = -6.324 p ≤ 0.001 (***) MR_{BT} = 2374.06 MR_{PROJ} = 2129.09
NT				z = -1.434 p = 0.152 (ns) MR_{NT} = 1604.13 MR_{RT} = 1654.57	z = -30.408 p ≤ 0.001 (***) MR_{NT} = 4815.63 MR_{ORGA} = 3182.92	z = -6.990 p ≤ 0.001 (***) MR_{NT} = 1684.22 MR_{COMP} = 1930.29	z = -34.762 p ≤ 0.001 (***) MR_{NT} = 1431.80 MR_{SYN} = 2704.70	z = -21.518 p ≤ 0.001 (***) MR_{NT} = 2603.67 MR_{PROJ} = 1782.74
RT					z = -14.334 p ≤ 0.001 (***) MR_{RT} = 3801.95 MR_{ORGA} = 2931.23	z = -2.065 p ≤ 0.05 (*) MR_{RT} = 1143.13 MR_{COMP} = 1201.15	z = -24.731 p ≤ 0.001 (***) MR_{RT} = 899.04 MR_{SYN} = 1682.41	z = -11.330 p ≤ 0.001 (***) MR_{RT} = 1866.74 MR_{PROJ} = 1470.57
ORGA						z = -34.962 p ≤ 0.001 (***) MR_{ORGA} = 2820.07 MR_{COMP} = 4818.80	z = -48.831 p ≤ 0.001 (***) MR_{ORGA} = 2765.01 MR_{SYN} = 5442.06	z = -4.281 p ≤ 0.001 (***) MR_{ORGA} = 3584.37 MR_{PROJ} = 3814.98
COMRP							z = -28.872 p p ≤ 0.001 (***) MR_{COMP} = 1018.11 MR_{SYN} = 1954.85	z = -26.853 p ≤ 0.001 (***) MR_{COMP} = 2338.49 MR_{PROJ} = 1395.08
SYN								z = -41.055 p ≤ 0.001 (***) MR_{SYN} = 2798.28 MR_{PROJ} = 1301.76

Table 94: U-Test results for the comparison by differences in the $F_{\beta=5}$-measure

Appendix E   Category-Specific R-Precision Values

■ B: low; ■ B: middle; ▨ B: high; ▦ B: very high
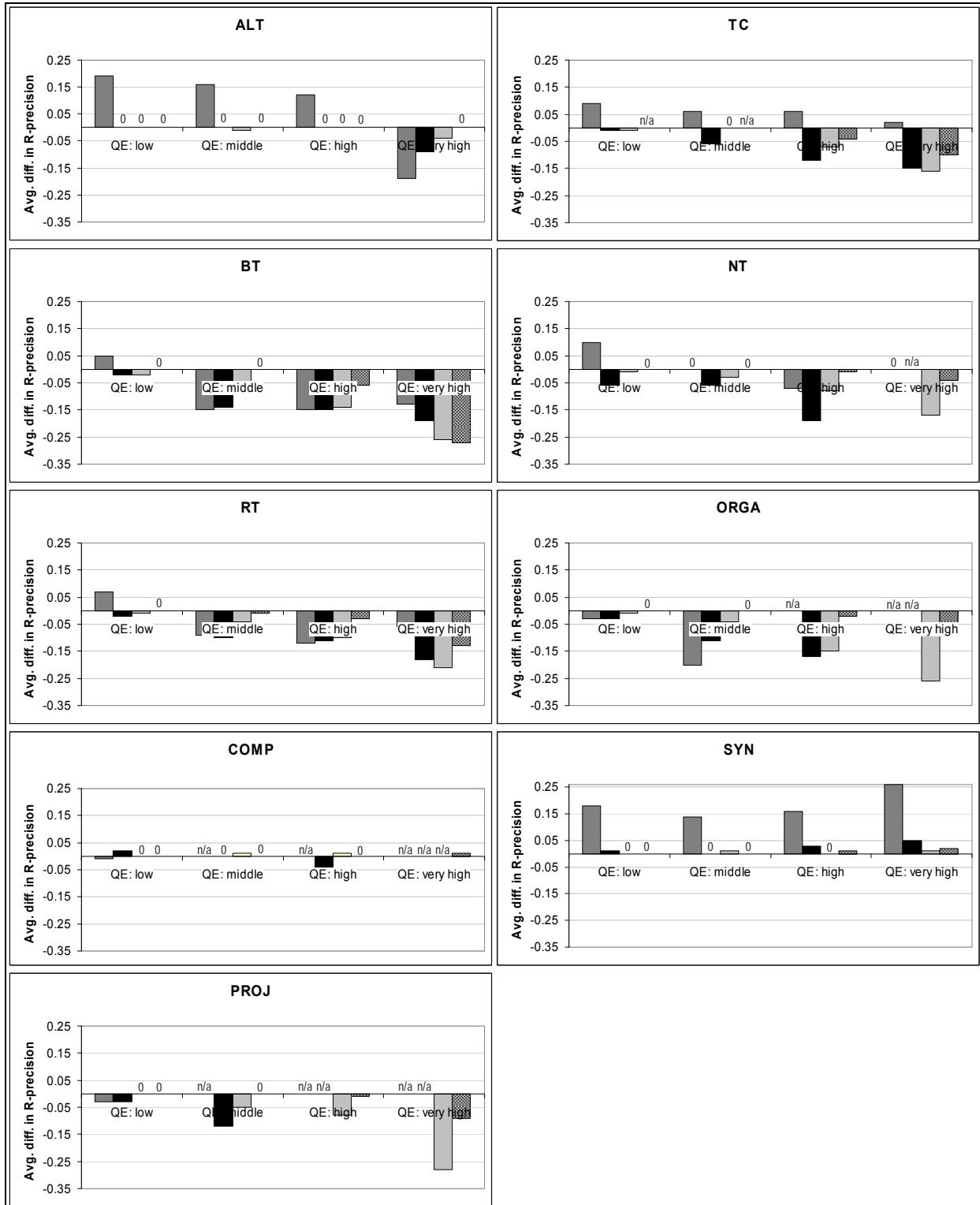
Table 95: Frequency category-specific R-precision values of different expansion relations

Appendix F   Category-Specific $F_{\beta=5}$-Measure Values

☐B: low; ■B: middle; ☐B: high; ▨B: very high
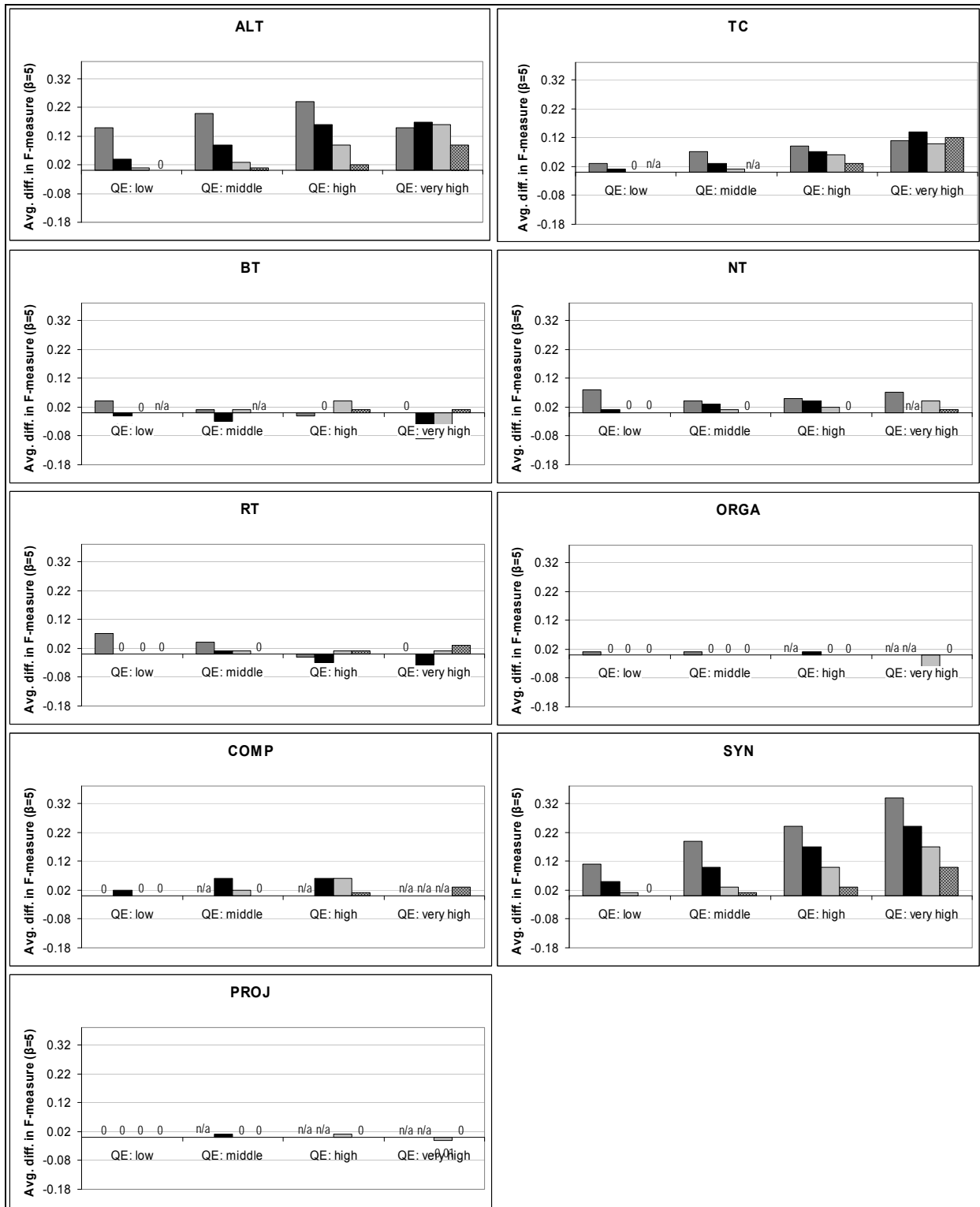
Table 96: Frequency category-specific $F_{\beta=5}$-measure values of different expansion relations

Appendix G   Expansion Effects for Project Queries

Baseline queries	Expanded queries	Differences
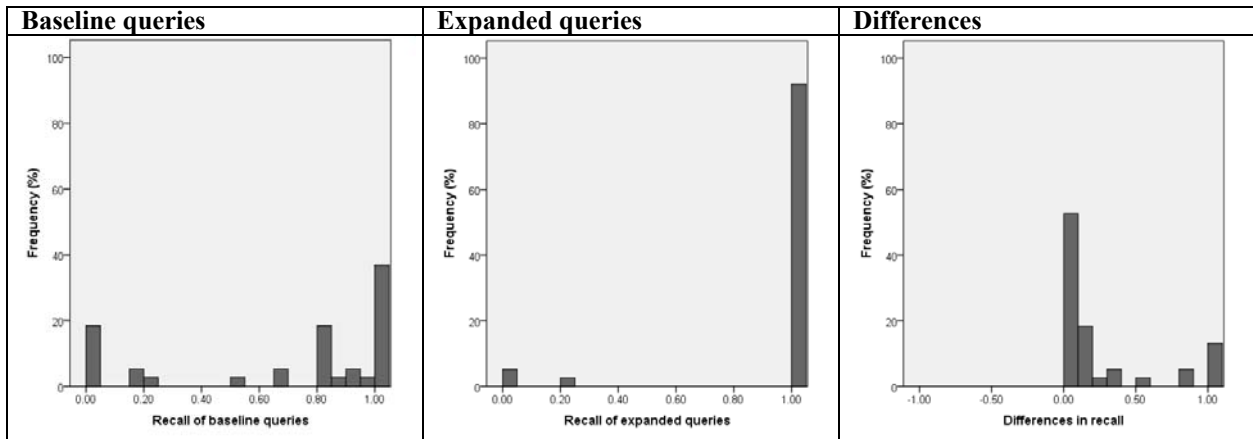
Table 97: Recall values for the expansion of project queries with preferred labels

Baseline queries	Expanded queries	Differences
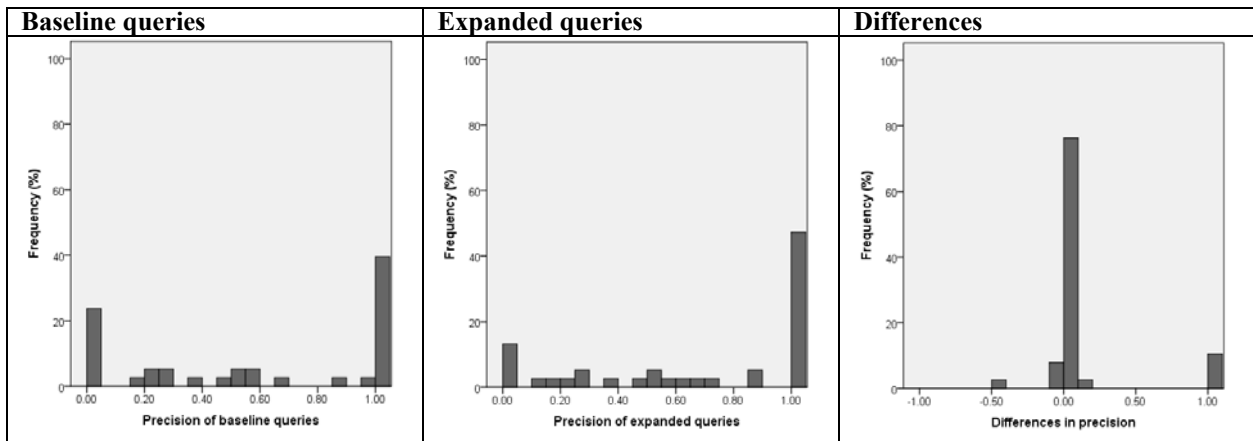
Table 98: Precision values for the expansion of project queries with preferred labels

Baseline queries	Expanded queries	Differences
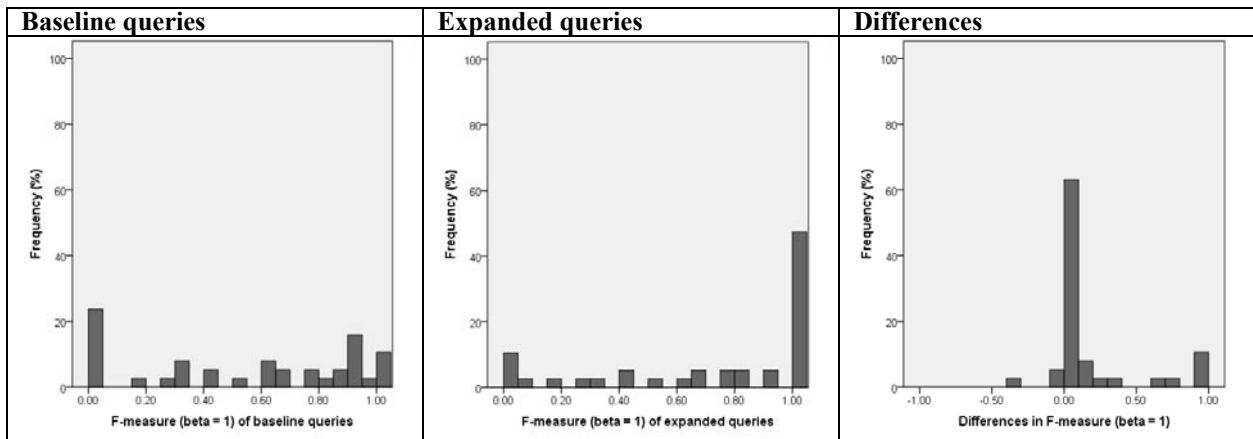
Table 99: $F_{\beta=1}$-measure values for the expansion of organization queries with preferred labels

Baseline queries	Expanded queries	Differences
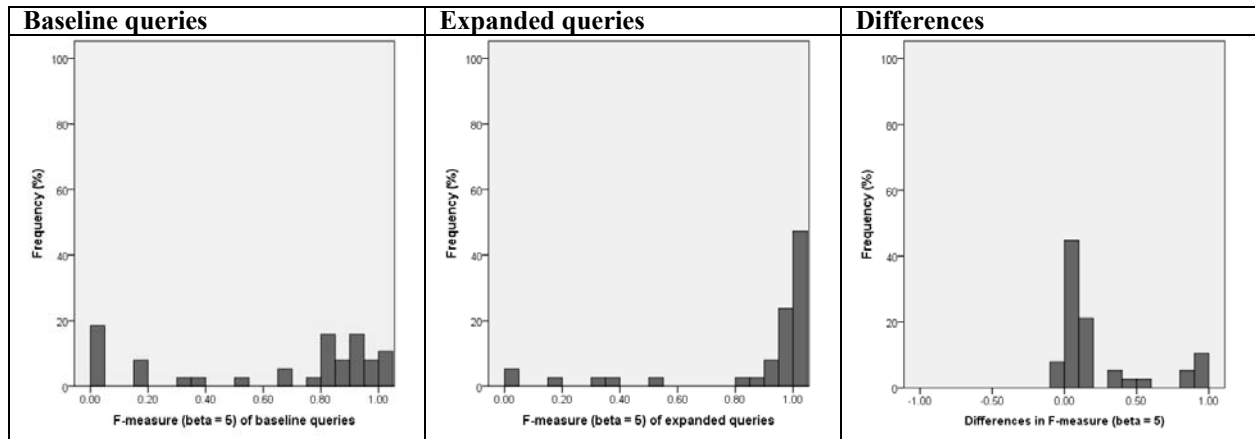
Table 100: $F_{\beta=5}$-measure values for the expansion of project queries with preferred labels

Baseline queries	Expanded queries	Differences
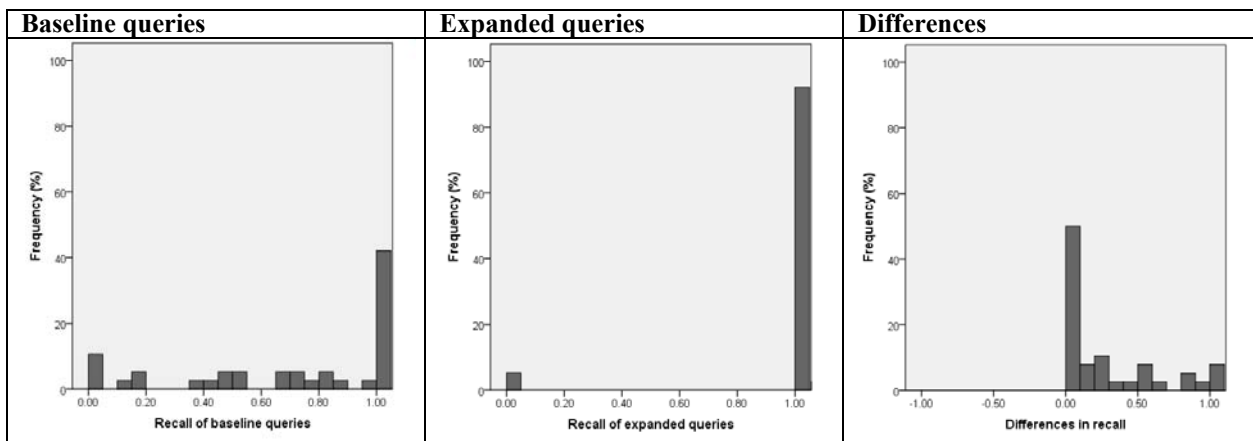
Table 101: Recall values for the expansion of project queries with alternative labels

Baseline queries	Expanded queries	Differences
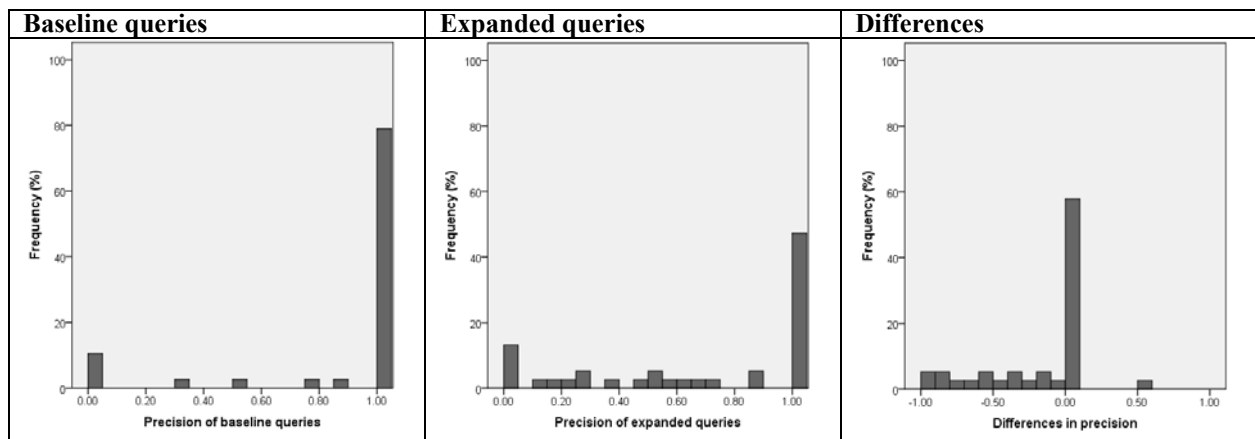
Table 102: Precision values for the expansion of project queries with alternative labels

Baseline queries	Expanded queries	Differences
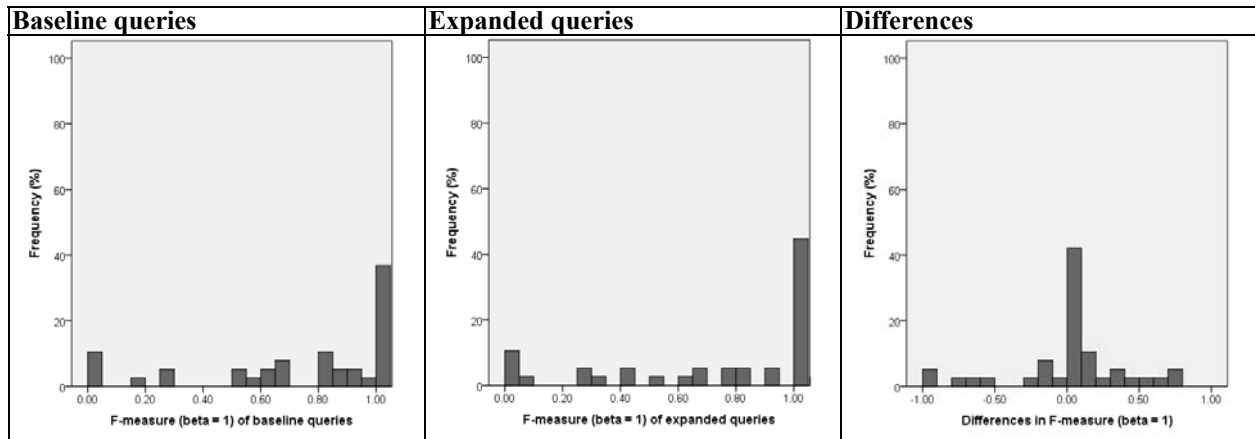
Table 103: $F_{\beta=1}$-measure values for the expansion of project queries with alternative labels

Baseline queries	Expanded queries	Differences
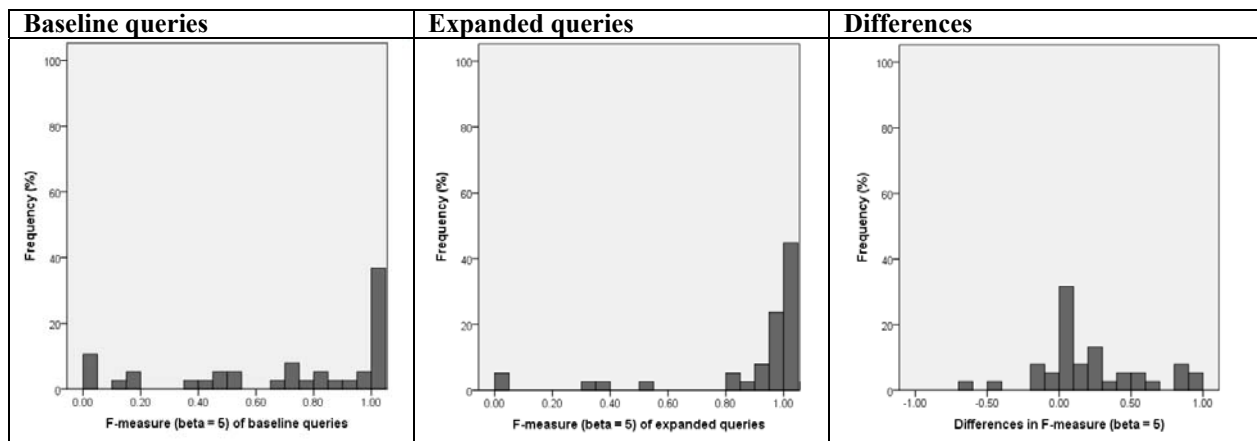
Table 104: $F_{\beta=5}$-measure values for the expansion of project queries with alternative labels

Baseline queries	Expanded queries	Differences
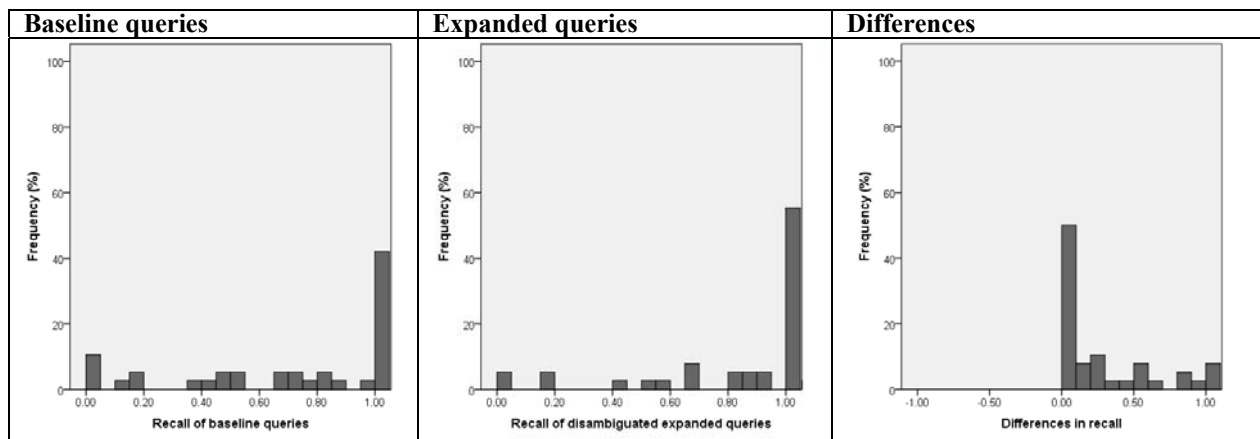
Table 105: Recall values for the disambiguated expansion of project queries with alternative labels and context information

Baseline queries	Expanded queries	Differences
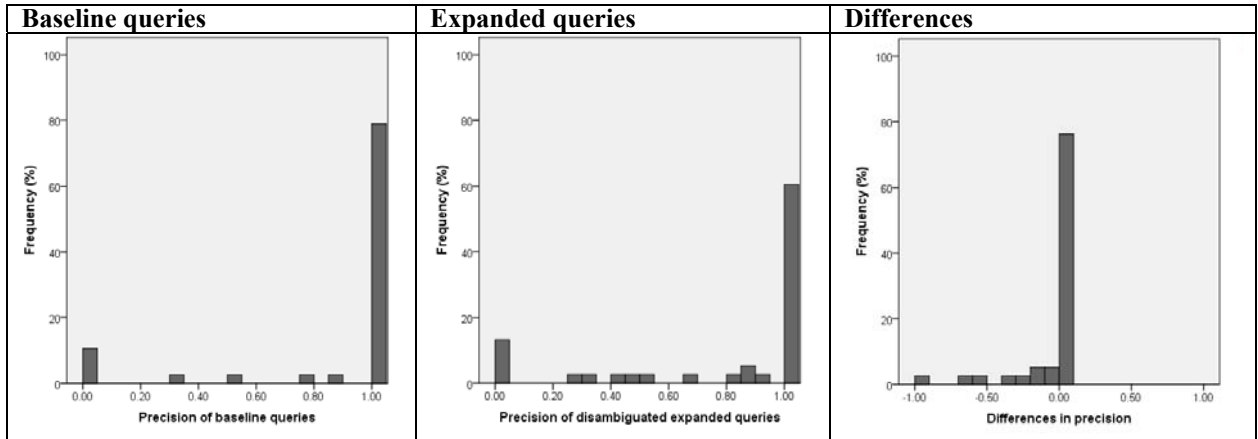
Table 106: Precision values for the disambiguated expansion of project queries with alternative labels and context information

Baseline queries	Expanded queries	Differences
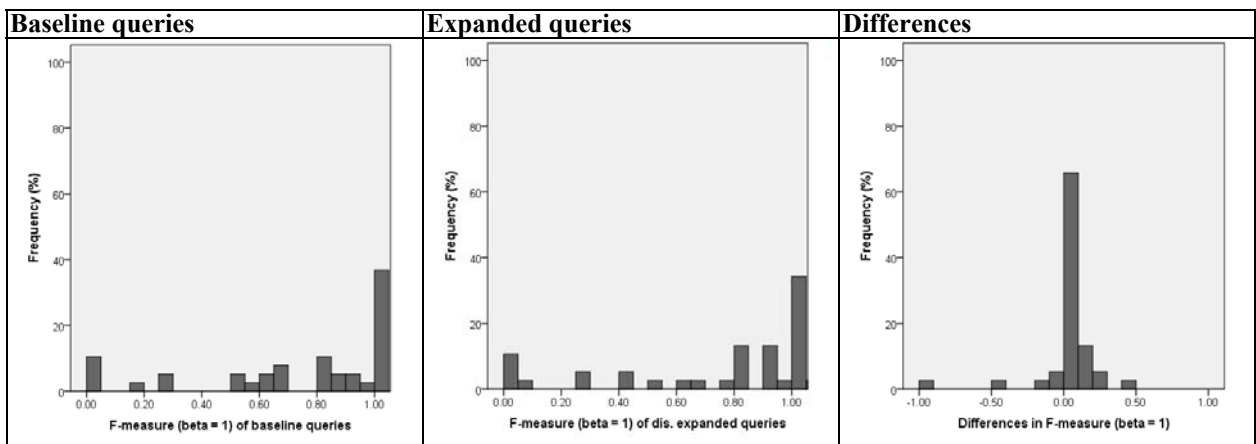
Table 107: $F_{\beta=1}$-measure values for the disambiguated expansion of project queries with alternative labels and context information

Baseline queries	Expanded queries	Differences
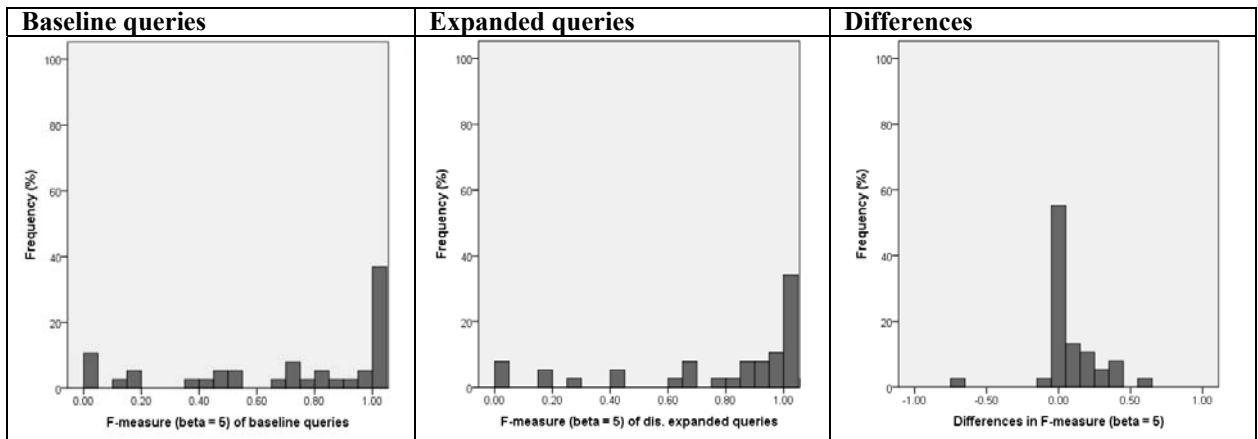
Table 108: $F_{\beta=5}$-measure values for the disambiguated expansion of project queries with alternative labels and context information

Appendix H   Expansion Effects for Organization Queries

Baseline queries	Expanded queries	Differences
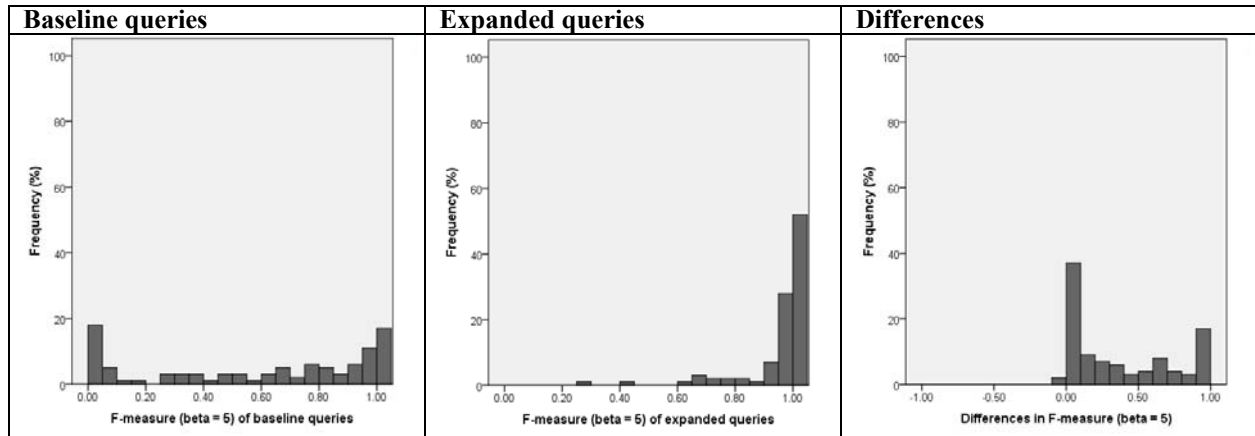
Table 109: $F_{\beta=5}$-measure values for the expansion of organization queries with preferred labels

Baseline queries	Expanded queries	Differences
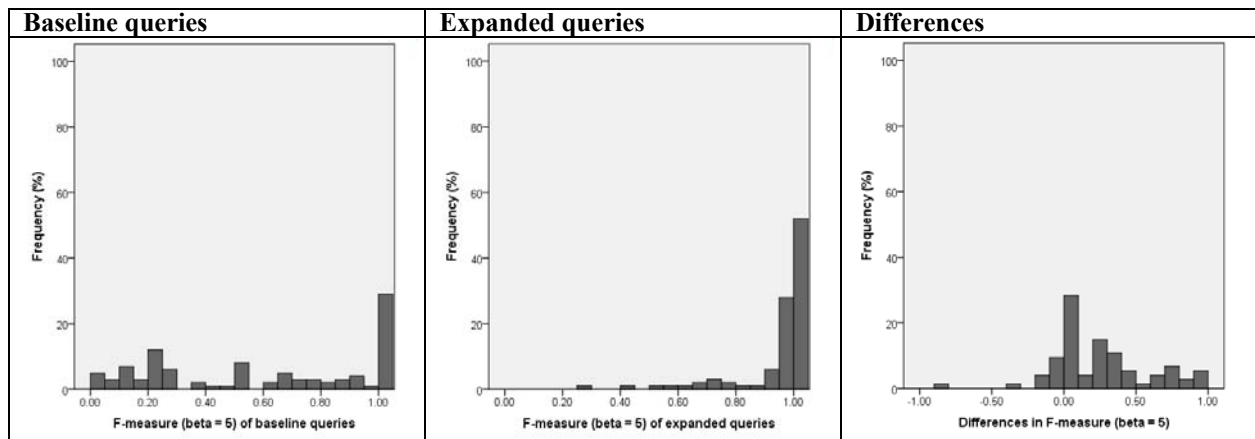
Table 110: $F_{\beta=5}$-measure values for the expansion of organization queries with alternative labels

Baseline queries	Expanded queries	Differences
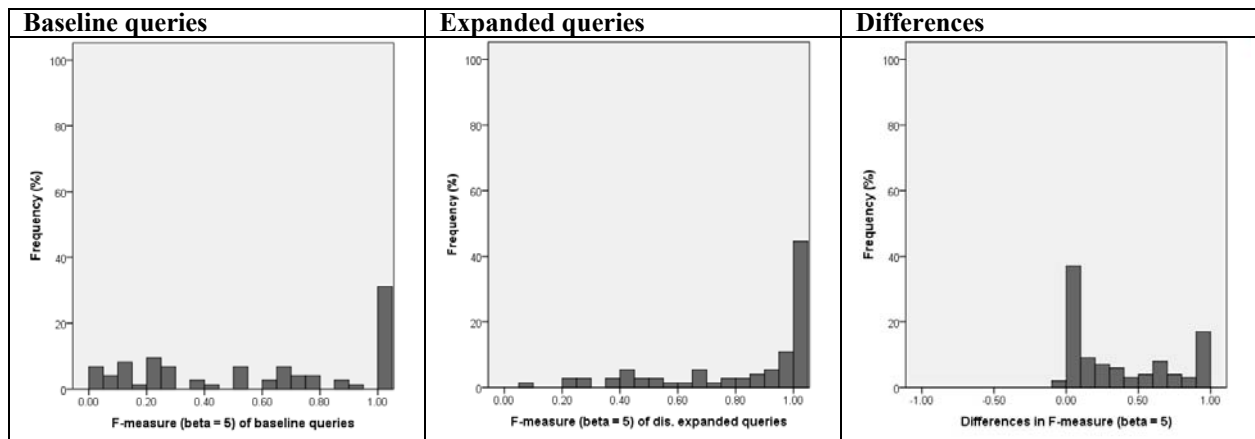
Table 111: $F_{\beta=5}$-measure values for the disambiguated expansion of organization queries with alternative labels and context information

Appendix I   Categorization of Expansion Effects for Organization and Project Queries

<table>
<thead>
<tr><th rowspan="2"></th><th rowspan="2"></th><th colspan="3">Recall</th></tr>
<tr><th>Increase</th><th>Neutral</th><th>Decrease</th></tr>
</thead>
<tbody>
<tr><td rowspan="3">Precision</td><td>Increase</td><td>33%</td><td>0%</td><td>0%</td></tr>
<tr><td>Neutral</td><td>35%</td><td>31%</td><td>0%</td></tr>
<tr><td>Decrease</td><td>0%</td><td>0%</td><td>1%</td></tr>
</tbody>
</table>

Table 112: Effects of expanding organization queries with preferred labels

<table>
<thead>
<tr><th rowspan="2"></th><th rowspan="2"></th><th colspan="3">Recall</th></tr>
<tr><th>Increase</th><th>Neutral</th><th>Decrease</th></tr>
</thead>
<tbody>
<tr><td rowspan="3">Precision</td><td>Increase</td><td>0%</td><td>23.68%</td><td>0%</td></tr>
<tr><td>Neutral</td><td>28.95%</td><td>36.84%</td><td>0%</td></tr>
<tr><td>Decrease</td><td>2.63%</td><td>7.89%</td><td>0%</td></tr>
</tbody>
</table>

Table 113: Effects of expanding project queries with preferred labels
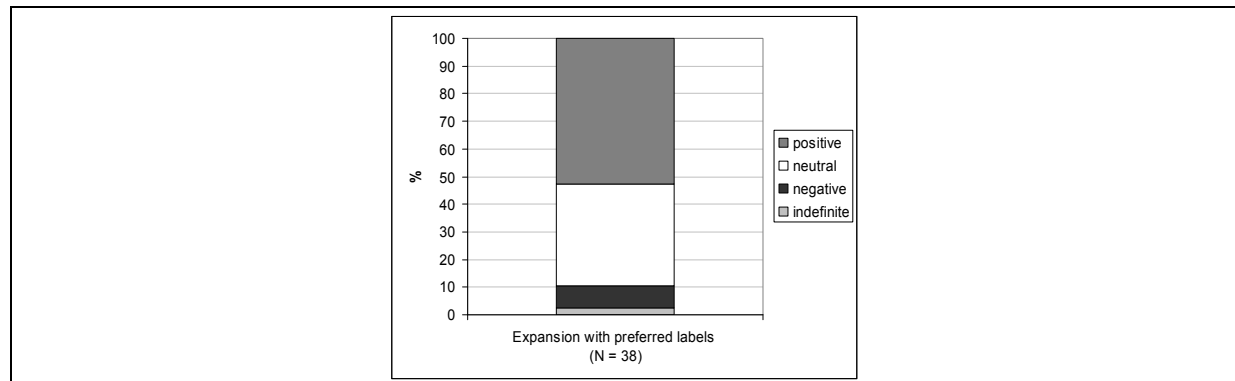
Figure 75: Overall effects of expanding project queries with preferred labels

<table>
<thead>
<tr><th rowspan="2"></th><th rowspan="2"></th><th colspan="3">Recall</th></tr>
<tr><th>Increase</th><th>Neutral</th><th>Decrease</th></tr>
</thead>
<tbody>
<tr><td rowspan="3">Precision</td><td>Increase</td><td>1%</td><td>0%</td><td>0%</td></tr>
<tr><td>Neutral</td><td>47%</td><td>5%</td><td>0%</td></tr>
<tr><td>Decrease</td><td>22%</td><td>23%</td><td>2%</td></tr>
</tbody>
</table>

Table 114: Effects of expanding organization queries with alternative labels

<table>
<thead>
<tr><th rowspan="2"></th><th rowspan="2"></th><th colspan="3">Recall</th></tr>
<tr><th>Increase</th><th>Neutral</th><th>Decrease</th></tr>
</thead>
<tbody>
<tr><td rowspan="3">Precision</td><td>Increase</td><td>5.26%</td><td>0%</td><td>0%</td></tr>
<tr><td>Neutral</td><td>2.63%</td><td>28.95%</td><td>0%</td></tr>
<tr><td>Decrease</td><td>21.05%</td><td>18.42%</td><td>0%</td></tr>
</tbody>
</table>

Table 115: Effects of expanding project queries with alternative labels
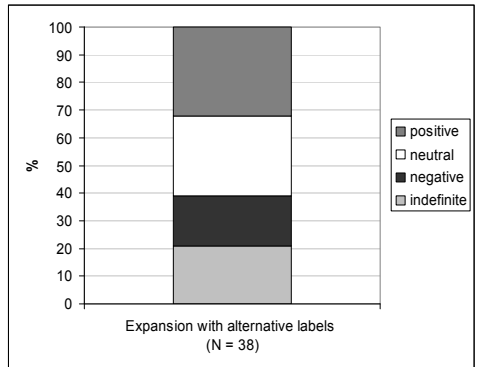
Figure 76: Overall effects of expanding project queries with alternative labels

		Recall		
		Increase	**Neutral**	**Decrease**
Precision	**Increase**	1.35%	0%	0%
	Neutral	50%	25.68%	0%
	Decrease	6.76%	14.86%	0%

Table 116: Effects of expanding organization queries with alternative labels and context information

		Recall		
		Increase	**Neutral**	**Decrease**
Precision	**Increase**	5.26%	0%	0%
	Neutral	28.95%	42.11%	0%
	Decrease	10.53%	10.53%	2.63%

Table 117: Effects of expanding project queries with alternative labels and context information
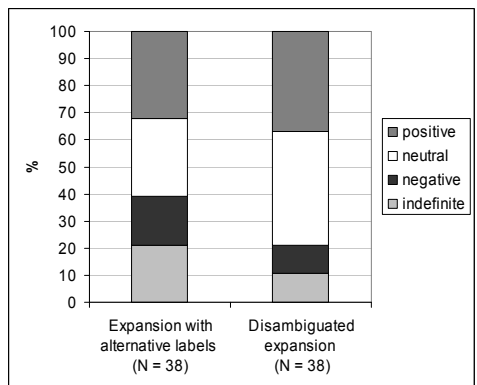
Figure 77: Overall effects of expanding project queries with alternative labels and context information

Appendix J   Sources for the Derivation of Retrieval Topics

Id	Topic definition	Source

1	Decisions of the *KMK*[75] on educational standards for language classes	Logfile:
		 `Freitext:  KMK   [and]`
		 `AND   Schlagwörter:    Beschluss   [and]`
		 `AND  Titel:      fremdsprachenuntericht  [and]`

2	Survey of projects for promoting reading in schools	Logfile:
		 `AND  Titel:     Praxis Deutsch    [and]`
		 `AND  Freitext:  Leseförderprojekte     [and]`

3	The influence of PISA results on elementary pedagogics	Logfile: `AND  Titel:     Alemantarpädagogik nach Pisa`
4	Measurements for language support for migrants in preschool in Germany	User request in the Information Center's Eduserver wiki[76]

5	Introduction of geometric figures in class	Logfile:
		 `Einfache Suche:  Einführung    geometrischer`
		 `Körper     [and]`

6	Survey of projects of the IPN *(Leibniz Institute for Science and Mathematics Education)* about the topic of chemistry classes	Logfile:
		 `AND   Schlagwörter:    IPN  [and]`
		 `AND  Titel:     Projekt    [and]`

7	Open learning in foreign language classes in secondary school	Logfile: `Freitext:   Freiarbeit im Englischunterricht`

Table 118: Sources for the derivation of retrieval topics

---

[75] Abbreviation of *Kultusministerkonferenz*: the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany.
[76] http://wiki.bildungsserver.de/infoboerse/index.php/Bildungssystem_Deutschland.

Appendix K   Topic Definitions

ID	Task
1	Imagine you were working on a research paper about the following topic: **Decisions of the *KMK*[77] on educational standards for language classes**. In the given time, please collect as many documents as possible that are relevant for working on the topic. Those documents are relevant that name, describe or refer to one or more decisions of the KMK on educational standards for language classes.
2	Imagine you were working on a research paper about the following topic: **Survey of projects for promoting reading in schools** For the survey you want to assemble information about as many distinct projects as possible. In the given time, please collect documents that name, describe or refer to as many different projects on reading promotion as possible. Relevant are documents that refer to one or more projects
3	Imagine you were working on a research paper about the following topic: **The influence of PISA results on elementary pedagogics** In the given time, please collect as many documents as possible that are relevant for working on the topic. Those documents are relevant that describe or refer to the influence of the results of one or more PISA studies on elementary pedagogics. These may be empirical studies or theoretic essays.
4	Imagine you were working on a research paper about the following topic: **Measurements for language support for migrants in preschool in Germany** In the given time, please collect as many documents as possible that are relevant for working on the topic. Those documents are relevant that name or describe one or more measurements for language support for migrants in preschool. These may be project reports, empirical studies or theoretic essays.
5	Imagine you were working on a research paper about the following topic: **Introduction of geometric figures in class** In the given time, please collect as many documents as possible that are relevant for working on the topic. Those documents are relevant that discuss the introduction of one or more geometric figures in class. These may be teaching suggestions, empirical studies, field reports or theoretic essays.
6	Imagine you were working on a research paper about the following topic: **Survey of projects of the IPN *(Leibniz Institute for Science and Mathematics Education)* about the topic of chemistry classes** For the survey you want to assemble information about as many distinct projects as possible that primarily or secondarily deal with chemistry classes. In the given time, please collect documents that name, describe or refer to as many such projects as possible. Relevant are documents that refer to one or more projects
7	Imagine you were working on a research paper about the following topic: **Open learning in foreign language classes in secondary school** In the given time, please collect as many documents as possible that are relevant for working on the topic. Those documents are relevant that discuss open learning in foreign language classes in secondary school. These may be teaching suggestions, empirical studies, field reports or theoretic essays.

Table 119: Topic definitions

---

[77] Abbreviation of *Kultusministerkonferenz*: the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany.

Appendix L   Statistics of Collected Documents in Valid Search Sessions

B: Baseline system, **A:** Automatic system, **INT:** Interactive system

Topic	System	Collected		Distinct collected		Relevant collected		Distinct relevant collected		Partly relevant collected		Distinct partly relevant collected		Irrelevant collected		Distinct irrelevant collected		Number of valid search sessions
1	B	85	241	44	80	36	121	16	28	21	53	8	16	28	67	20	36	19
	A	86		46		45		22		14		7		27		17		18
	INT	70		33		40		16		18		8		12		9		19
2	B	147	404	61	100	109	283	35	46	7	18	4	7	31	103	22	47	19
	A	139		64		89		29		7		4		43		31		19
	INT	118		52		85		30		4		3		29		19		18
3	B	79	302	36	97	44	173	16	38	13	62	5	17	22	67	15	42	18
	A	110		56		71		29		22		13		17		14		18
	INT	113		62		58		25		27		14		28		23		19
4	B	111	387	39	82	88	331	24	53	19	43	11	21	4	13	4	8	20
	A	153		55		129		36		16		12		8		7		19
	INT	123		40		114		32		8		7		1		1		19
5	B	112	312	57	99	50	122	26	35	34	122	11	24	28	68	20	40	19
	A	118		56		46		24		50		15		22		17		19
	INT	82		37		26		8		38		12		18		17		18
6	B	86	250	20	41	28	90	5	16	32	93	8	11	26	67	7	14	18
	A	113		22		37		8		51		9		25		5		19
	INT	51		28		25		13		10		5		16		10		18
7	B	65	213	31	69	35	116	8	15	15	58	9	21	15	39	14	33	20
	A	72		31		39		7		19		10		14		14		19
	INT	76		27		42		11		24		9		10		7		18

Table 120: Number and relevance of collected documents by system version

Appendix M   Tool for Relevance Judgments

Please enter your relevance judgment for the selected document here

Imagine you were working on a research paper about the following topic: **The influence of PISA results on elementary pedagogics**

In the given time, please collect as many documents as possible that are relevant for working on the topic.

Those documents are relevant that describe or refer to the influence of the results of one or more PISA studies on elementary pedagogics. These may be empirical studies or theoretic essays.

Autor:	Klieme, Eckhard
Titel:	Fragestellungen, zentrale Befunde und Konsequenzen der Studie "Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten".
Jahr:	2003
Schlagwörter:	Schulsystem; Schüler; Schülerleistung; Steuerung; Soziokulturelle Bedingung; Internationaler Vergleich; Leistung; Struktur; PISA (Programme for International Student Assessment)
Abstract:	Es wird zunächst begründet, warum kurz nach Erscheinen der internationalen PISA-Studie ein vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten vorgenommen wurde. Sodann werden Forschungsmethoden vorgestellt. Bei der Darstellung der Befunde wird der mögliche Einfluss der soziokulturellen Hintergründe, der Struktur des Schulsystems, sowie des Steuerungssystems auf die Schülerleistungen in den jeweiligen Ländern besonders hervorgehoben. (DIPF/Mar.).
Dokumenttyp:	Zeitschriftenaufsatz
Zeitschrift:	Trends in Bildung international
Jahrgang:	3
Heft:	7
Seitenzahl:	5 S.
Sprache:	deutsch
URL:	http://www1.dipf.de/publikationen/tibi/tibi7_klieme.pdf

Please enter your relevance judgment

○   Relevant
○   Partly relevant
○   Irrelevant

Save

Figure 78: Screenshot of the tool for assigning relevance judgments

Appendix N   Examples of Implemented Expansion Relations

Query (DE)	Expansion Relation	Expansion Mode	Type of Expansion Term	Expansion Term(s)
G: "Chemischer Stoff E: "chemical product"	rescon:prefLabelConcept	Automatic	String	G: Chemikalie E: chemical
G: Chemikalie E: chemical	rescon:altLabelConcept	Automatic	String	G: "Chemischer Stoff" E: "chemical product"
G: Pflegeeltern E: "foster parents"	rescon:synonymous	Automatic	Concept	G: Pflegefamilie E: "foster family"
G: Propagandafilm E: "propaganda film"	rescon:component	Automatic	Concept	G: (Propaganda AND Film) E: (propaganda AND film)
G: Psychologie E: psychology	skos:narrower	Interactive	Concept	G: Berufspsycholoige E: "occupational psychology"
G: Mikrochemie E: microchemistry	skos:broader	Interactive	Concept	G: Chemie E: chemistry
G: Schulsport E: "school sport"	skos:related	Interactive	Concept	G: Sportunterricht E: "physical education"
G: Therapie E: therapy	rescon:belongsToCompound	Interactive	Concept	G: Sprachtherapie E: "speech therapy"
G: Ganztagsschule E: "all-day school"	rescon:isSubjectOfProject	Interactive	Project	G: Transfer-21 E: Transfer-21
G: Klima E: climate	rescon:isInterestOfOrganization	Interactive	Organization	G: "Potsdam- Institut fpr Klimafolgenforschung" E: "Potsdam Institute for Climate Impact Research"

Table 121: Examples of implemented expansion relations for concept queries in the automatic and interactive modes

Query (DE)	Expansion Relation	Expansion Mode	Type of Expansion Term	Expansion Term(s)
G: BLM E: BLM	rescon:prefLabelOrganization	Automatic	String	G: "Bayerische Landeszentrale für neue Medien" E: "Bavarian regulatory authority for commercial broadcasting"
G: "Bayerische Landeszentrale für neue Medien" E: "Bavarian regulatory authority for commercial broadcasting"	rescon:altLabelOrganization	Disambiguated expansion with topic interests	String	G: BLM E: BLM
G: Hochschulrektorenkonferenz E: "German Rectors' Conference"	rescon:organizationHasPart	Interactive	Organization	G: "Kompetenzzentrum Bologna" E: "Bologna Centre"
G: "Kompetenzzentrum Bologna" E: "Bologna Centre"	rescon:partOfOrganization	Interactive	Organization	G: Hochschulrektorenkonferenz E: "German Rectors' Conference"
G: "Deutsches Institut für Internationale Pädagogische Forschung" E: "German Institute for International Educational Research"	rescon:organizationHasInterest	Interactive	Concept	G: "Vergleichende Erziehungswissenschaft" E: "comparative educational research"
G: "Bundesministerium für Bildung und Forschung" E: "Federal Ministry of Education and Research"	rescon:organizationCarriesOutProject	Interactive	Project	G: "Chemie im Kontext" E: "Chemistry in Context"

Table 122: Examples of implemented expansion relations for organization queries in the automatic and interactive modes

Query (DE)	Expansion Relation	Expansion Mode	Type of Expansion Term	Expansion Term(s)
G: "Jürgen Baumert" E: "Jürgen Baumer"	rescon:prefLabelPerson	Automatic	String	G: "Baumert, Jürgen" E: "Baumert, Jürgen"
G: "Baumert, Jürgen" E: "Baumert, Jürgen"	rescon:altLabelPerson	Automatic	String	G: "Jürgen Baumert" E: "Jürgen Baumer"
G: "Jürgen Wolff" E: "Jürgen Wolff"	rescon:hasResearchInterest	Interactive	Concept	G: Religionspädagogik EN_ "religious education"
G: "Eckhard Klieme" E: "Eckhard Klieme"	rescon:hasWorkInterest	Interactive	Concept	G: Bildungsqualität E: "quality of education"
G: "Eckhard Klieme" E: "Eckhard Klieme"	rescon:personWorksAtProject	Interactive	Project	G: "Studie zur Entwicklung von Ganztagsschulen" E: "Study on the Development of All-day Schools"
G: "Eckhard Klieme" E: "Eckhard Klieme"	rescon:personHasAffiliation	Interactive	Organization	G: "Deutsches Institut für Internationale Pädagogische Forschung" E: "German Institute for International Educational Research"

Table 123: Examples of implemented expansion relations for person queries in the automatic and interactive modes

Query (DE)	Expansion Relation	Expansion Mode	Type of Expansion Term	Expansion Term(s)
G: ISB E: ISB	rescon;prefLabelProject	Automatic	String	G: "Staatsinstitut für Schulqualität und Bildungsforschung München" E: „State Institute for School Quality and Educational Research Munich"
G: "Staatsinstitut für Schulqualität und Bildungsforschung München" E: „State Institute for School Quality and Educational Research Munich"	rescon:altLabelProject	Disambiguated expansion with topic interests	String	G: ISB E: ISB
G: "Bildung für eine nachhaltige Entwicklung" E: "Education for Sustainable Development"	rescon:hasPredecessorProject	Interactive	Project	G: "Transfer-21" E: "Transfer-21"
G: "Transfer-21" E: "Transfer-21"	rescon:hasSuccessorProject	Interactive	Project	G: "Bildung für eine nachhaltige Entwicklung" E: "Education for Sustainable Development"
G: "SINUS-Transfer-Modell" E: "SINUS-Transfer-Model"	rescon:projectHasPart	Interactive	Project	G: "SINUS-Transfer Thüringen" E: "SINUS-Transfer Thuringia"
G: "SINUS-Transfer Thüringen" E: "SINUS-Transfer Thuringia"	rescon:partOfProject	Interactive	Project	G: "SINUS-Transfer-Modell" E: "SINUS-Transfer-Model"
G: "Physik im Kontext" E: "Physics in Context"	rescon:projectHasSubject	Interactive	Concept	G: Unterrichtskonzeption E: "teaching concepts"
G: "Chemie im Kontext" E: "Chemistry in Context"	rescon:projectIsCarriedOutByOrganization	Interactive	Organization	G: "Bundesministerium für Bildung und Forschung" E: "Federal Ministry of Education and Research"

Table 124: Examples of implemented expansion relations for project queries in the automatic and interactive modes

Appendix O U-Test Results for the Comparison by Subjective Ratings

	Automatic	Interactive
Baseline	z = -0.264 p = 0.792 (ns) MR_B = 21.91 MR_{AUT} = 21.05	z = -1.983 p = 0.047 (*) MR_B = 24.73 MR_{INT} = 17.95
Automatic		z = -1.735 p = 0.083 (ns) MR_{AUT} = 23.40 MR_{INT} = 17.60

Table 125: Handling of the system

	Automatic	Interactive
Baseline	z = -2.213 p = 0.027 (*) MR_B = 25.14 MR_{AUT} = 17.50	z = -1.243 p = 0.214 (ns) MR_B = 19.36 MR_{INT} = 23.85
Automatic		z = -2.700 p = 0.007 (**) MR_{AUT} = 15.88 MR_{INT} = 25.13

Table 126: Time for familiarizing with the system

	Automatic	Interactive
Baseline	z = -0.794 p = 0.427 (ns) MR_B = 18.88 MR_{AUT} = 21.31	z = -0.924 p = 0.355 (ns) MR_B = 20.83 MR_{INT} = 17.85
Automatic		z = -1.365 p = 0.172 (ns) MR_{AUT} = 20.11 MR_{INT} = 15.76

Table 127: Perceived success in task completion

	Automatic	Interactive
Baseline	z = -0.423 p = 0.672 (ns) MR_B = 21.69 MR_{AUT} = 20.27	z = -1.116 p = 0.264 (ns) MR_B = 19.14 MR_{INT} = 22.95
Automatic		z = -1.513 p = 0.130 (ns) MR_{AUT} = 17.95 MR_{INT} = 23.05

Table 128: Formulation of complex queries

	Automatic	Interactive
Baseline	z = -1.190 p = 0.234 (ns) MR_B = 19.55 MR_{AUT} = 23.65	z = -0.086 p = 0.932 (ns) MR_B = 20.86 MR_{INT} = 21.16
Automatic		z = -0.994 p = 0.320 (ns) MR_{AUT} = 21.63 MR_{INT} = 18.29

Table 129: Formulation of effective queries

	Automatic	Interactive
Baseline	z = -1.141 p = 0.254 (ns) MR_B = 23.45 MR_{AUT} = 19.35	z = -2.041 p = 0.041 (*) MR_B = 25.00 MR_{INT} = 17.65
Automatic		z = -0.482 p = 0.630 (ns) MR_{AUT} = 21.35 MR_{INT} = 19.65

Table 130: Usage of document terms for query reformulations

	Automatic	Interactive
Baseline	z = -0.281 p = 0.779 (ns) MR_B = 20.91 MR_{AUT} = 20.00	z = -0.259 p = 0.796 (ns) MR_B = 19.86 MR_{INT} = 19.00
Automatic		z = -0.075 p = 0.940 (ns) MR_{AUT} = 17.61 MR_{INT} = 17.38

Table 131: Satisfaction with the ranking

	Automatic	Interactive
Baseline	z = -0.224 p = 0.822 (ns) MR_B = 20.86 MR_{AUT} = 20.11	z = -0.030 p = 0.976 (ns) MR_B = 20.55 MR_{INT} = 20.45
Automatic		z = -0.252 p = 0.801 (ns) MR_{AUT} = 19.11 MR_{INT} = 19.89

Table 132: Satisfaction with query formulation functionalities

	Automatic	Interactive
Baseline	z = -0.712 p = 0.477 (ns) MR_B = 19.81 MR_{AUT} = 22.25	z = -0.698 p = 0.485 (ns) MR_B = 22.19 MR_{INT} = 19.75
Automatic		z = -1.508 p = 0.132 (ns) MR_{AUT} = 23.03 MR_{INT} = 17.98

Table 133: Traceability of the query interpretation

	Interactive
Automatic	z = -1.522 p = 0.128 (ns) MR_B = 20.97 MR_{AUT} = 16.52

Table 134: Helpfulness of automatic query expansion

Appendix P   Problem Causes

Problem	Baseline system	Automatic system	Interactive system
Not specified	31.82%	15%	45%
Parentheses	22.73%	25%	10%
Search syntax (e.g. operators)	18.18%	20%	10%
Marking of selected or viewed documents	4.55%	15%	---
No problems	---	5%	15%
Unclassified problems	4.55%	10%	5%
Parser mistake	2	5%	---
Positive remarks	4.55%	10%	---
System navigation	---	15%	---
Missing spelling correction	4.55%	5%	---
Empty result sets	2	---	---
Appearance of the term *IPN* as an editor's note in result documents	4.55%	---	---
Bug: Disappearance of search field	4.55%	---	---
Insufficient time for task completion	---	5%	---
Loading times	---	5%	---
Missing opportunity to select several documents at once	---	5%	---
Missing marking of entered query terms in texts	4.55%	---	---
Missing functionality to search by author/editor	---	5%	---
Missing storage of scroll position	4.55%	---	---
Quality of the relevance ranking	4.55%	---	---
Short period of adjustment	---	---	5%
System complexity	---	---	5%
Task remembrance	---	---	5%
Unclear meaning of buttons	---	---	5%

Table 135: Portions of users mentioning problem causes by system version

Appendix Q   Chi-Square Test Results on the Helpfulness of Query Expansion Functionalities

N: sample size; **df:** degrees of freedom; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Type of expansion term	N	df	x^2	p
Broader terms	20	1	5.000	0.013 (*)
Narrower terms	20	2	19.900	< 0.001 (***)
Related terms	20	1	7.200	0.004 (**)
Organization names	20	2	16.300	< 0.001 (***)
Compounds	20	1	9.800	0.001 (***)
Project names	20	2	24.100	< 0.001 (***)
Organization-related information	20	1	7.200	0.004 (**)
Project-related information	20	2	19.900	< 0.001 (***)

Table 136: Chi-Square Test results on the helpfulness of interactive expansion terms

N: sample size; **df:** degrees of freedom; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Type of expansion term	N	df	x^2	p
Expansion by broadening (OR operator)	20	1	12.800	< 0.001 (***)
Expansion by specification (AND operator)	20	2	28.900	< 0.001 (***)
Expansion by replace	20	1	9.800	0.001 (***)

Table 137: Chi-Square Test results on the helpfulness of reformulation buttons

N: sample size; **df:** degrees of freedom; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Type of expansion term	N	df	x^2	p
Disambiguation functionality	20	1	16.200	< 0.001 (***)

Table 138: Chi-Square Test results on the helpfulness of the disambiguation functionality

Appendix R   System Comparisons by Recall

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	19	0.065	0.063	0.035	18	0.086	0.083	0.052	19	0.073	0.065	0.069
2	19	0.128	0.062	0.133	19	0.108	0.058	0.089	18	0.106	0.078	0.089
3	18	0.064	0.044	0.053	18	0.104	0.059	0.105	19	0.080	0.053	0.079
4	20	0.082	0.038	0.074	19	0.126	0.066	0.111	19	0.111	0.052	0.111
5	19	0.075	0.063	0.057	19	0.069	0.062	0.057	18	0.041	0.034	0.029
6	18	0.222	0.149	0.214	19	0.286	0.135	0.286	18	0.119	0.123	0.143
7	20	0.117	0.052	0.133	19	0.137	0.065	0.133	18	0.156	0.082	0.133

Table 139: Location parameters of the systems' recall value distributions

Topic	System	Statistics	df	Significance
1	B	0.215	19	0.021
	AUT	0.233	18	0.011
	INT	0.195	19	0.055
2	B	0.155	19	0.200
	AUT	0.152	19	0.200
	INT	0.144	18	0.200
3	B	0.215	18	0.027
	AUT	0.157	18	0.200
	INT	0.214	19	0.022
4	B	0.204	20	0.029
	AUT	0.219	19	0.017
	INT	0.112	19	0.200
5	B	0.190	19	0.069
	AUT	0.184	19	0.089
	INT	0.352	18	0.000
6	B	0.203	18	0.048
	AUT	0.224	19	0.013
	INT	0.256	18	0.003
7	B	0.280	20	0.000
	AUT	0.320	19	0.000
	INT	0.227	18	0.015

Table 140: Results of Kolmogorov-Smirnov Test for normality for topic-specific recall values

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	$H1_3$ $M_{AUT} > M_B$	$H0_3$ $M_{AUT} \leq M_B$	$H1_4$ $M_{INT} > M_B$	$H0_4$ $M_{INT} \leq MR_B$	$H1_5$ $M_{INT} > M_{AUT}$	$H0_5$ $M_{INT} \leq M_{AUT}$
1	p = 0.253 (ns)		p = 0.372 (ns)			p = 0.367 (ns)
2		p = 0.154 (ns)		p = 0.116 (ns)		p = 0.380 (ns)
3	p = 0.013 (*)		p = 0.145 (ns)			p = 0.098 (ms)
4	p = 0.012 (*)		p = 0.020 (*)			p = 0.319 (ns)
5		p = 0.361 (ns)		p = 0.035 (*)		p = 0.071 (ms)
6	p = 0.114 (ns)			p = 0.016 (*)		p = 0.001 (***)
7	p = 0.094 (ms)		p = 0.037 (ms)		p = 0.220 (ns)	

Table 141: U-Test results for the comparisons by recall

Appendix S System Comparisons by Recall (Loose Relevance Judgments)
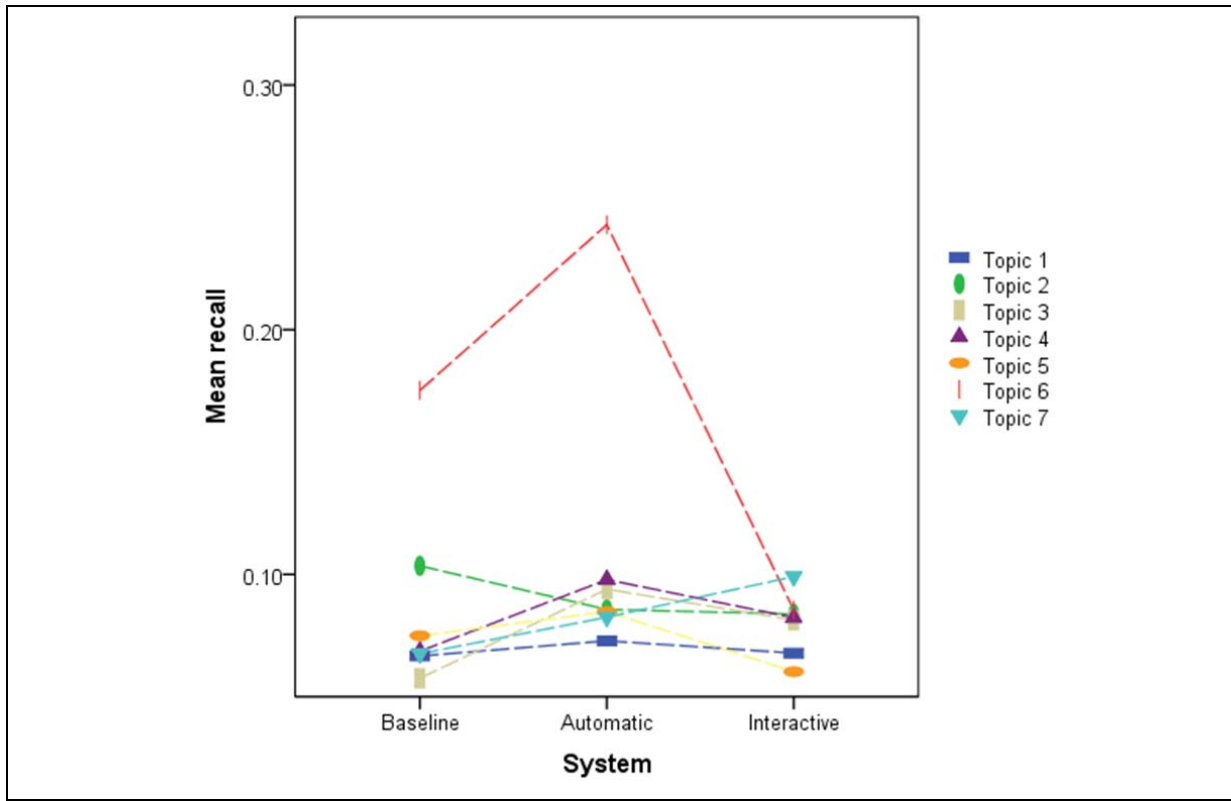
Figure 79: Topic-specific mean recall values by system

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	H1$_3$ $M_{AUT} > M_B$	H0$_3$ $M_{AUT} \leq M_B$	H1$_4$ $M_{INT} > M_B$	H0$_4$ $M_{INT} \leq MR_B$	H1$_5$ $M_{INT} > M_{AUT}$	H0$_5$ $M_{INT} \leq M_{AUT}$
1		p = 0.409 (ns)	p = 0.500 (ns)		p = 0.457 (ns)	
2	p = 0.500 (ns)			p = 0.296 (ns)		p = 0.300 (ns)
3	p = 0.006 (**)		p = 0.070 (ms)			p = 0.160 (ns)
4	p = 0.041 (*)		p = 0.096 (ms)			p = 0.319 (ns)
5	p = 0.200 (ns)			p = 0.166 (ns)		p = 0.031 (*)
6	p = 0.098 (ms)			p = 0.007 (**)		p < 0.001 (***)
7	p = 0.184 (ns)		p = 0.030 (*)		p = 0.136 (ns)	

Table 142: U-Test results for the comparisons by recall (loose relevance judgments)

Appendix T System Comparisons by the Number of Collected Documents

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	19	4.47	2.611	4.00	18	4.78	4.195	4.00	19	3.68	3.198	3.00
2	19	7.74	3.397	7.00	19	7.32	3.513	7.00	18	6.56	3.899	6.50
3	18	4.39	2.253	4.00	18	6.11	2.349	6.50	19	5.95	2.549	6.00
4	20	5.55	2.645	5.00	19	8.05	4.882	6.00	19	6.47	2.913	7.00
5	19	5.89	3.213	5.00	19	6.16	2.522	6.00	18	4.56	2.175	4.00
6	18	4.78	3.422	4.00	19	5.84	3.023	5.00	19	2.83	2.333	2.50
7	20	3.25	1.743	3.00	19	3.79	2.299	3.00	18	4.22	2.211	4.00

Table 143: Location parameters of the systems' distributions of the numbers of collected documents

Topic	System	Statistics	df	Significance
1	B	0.256	19	0.002
	AUT	0.257	18	0.003
	INT	0.174	19	0.131
2	B	0.144	19	0.200
	AUT	0.160	19	0.200
	INT	0.154	18	0.200
3	B	0.180	18	0.129
	AUT	0.147	18	0.200
	INT	0.281	19	0.000
4	B	0.221	20	0.012
	AUT	0.217	19	0.019
	INT	0.113	19	0.200
5	B	0.294	19	0.000
	AUT	0.127	19	0.200
	INT	0.156	18	0.200
6	B	0.252	18	0.004
	AUT	0.189	19	0.074
	INT	0.140	18	0.200
7	B	0.257	20	0.001
	AUT	0.194	19	0.058
	INT	0.151	18	0.200

Table 144: Results of K-S Test for normality for the number of collected documents

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	$H1_9$ $M_{AUT} > M_B$	$H0_9$ $M_{AUT} \leq M_B$	$H1_{10}$ $M_{INT} > M_B$	$H0_{10}$ $M_{INT} \leq MR_B$	$H1_{11}$ $M_{INT} > M_{AUT}$	$H0_{11}$ $M_{INT} \leq M_{AUT}$
1	p = 0.4764 (ns)			p = 0.145 (ns)		p = 0.231 (ns)
2		p = 0.330 (ns)		p = 0.192 (ns)		p = 0.340 (ns)
3	p = 0.012 (*)		p = 0.014 (*)			p = 0.306 (ns)
4	p = 0.045 (*)		p = 0.114 (ns)			p = 0.299 (ns)
5	p = 0.161 (ns)			p = 0.103 (ns)		p = 0.018 (*)
6	p = 0.086 (ms)			p = 0.028 (*)		p = 0.001 (***)
7	p = 0.194 (ns)		p = 0.063 (ms)		p = 0.225 (ns)	

Table 145: U-Test results for the system comparisons by the numbers of collected documents

Appendix U   System Comparisons by the DistinctDocuments Measure

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

$H1_9$ $M_{AUT} > M_B$	$H0_9$ $M_{AUT} \leq M_B$	$H1_{10}$ $M_{INT} > M_B$	$H0_{10}$ $M_{INT} \leq M_B$	$H1_{11}$ $M_{INT} > M_{AUT}$	$H0_{11}$ $M_{INT} \leq M_{AUT}$
$p = 0.090$ (ms)		$p = 0.305$ (ns)			$p = 0.071$ (ms)

Table 146: U-Test results for the system comparisons by the DistinctDocuments measure

Appendix V   System Comparisons by Precision

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	19	0.340	0.283	0.333	18	0.504	0.379	0.500	19	0.474	0.341	0.500
2	19	0.745	0.203	0.765	19	0.647	0.226	0.667	18	0.650	0.342	0.806
3	18	0.587	0.287	0.550	18	0.600	0.242	0.667	19	0.533	0.261	0.600
4	20	0.821	0.186	0.817	19	0.878	0.167	1.000	19	0.886	0.255	1.000
5	19	0.451	0.306	0.400	19	0.356	0.271	0.333	18	0.340	0.256	0.292
6	18	0.336	0.193	0.367	19	0.420	0.250	0.333	18	0.274	0.304	0.225
7	20	0.636	0.296	0.500	19	0.589	0.293	0.600	18	0.573	0.307	0.667

Table 147: Location parameters of the systems' precision value distributions

Task	System	Statistics	df	Significance
1	B	0.201	19	0.041
	AUT	0.131	18	0.200
	INT	0.215	19	0.021
2	B	0.106	19	0.200
	AUT	0.123	19	0.200
	INT	0.204	18	0.046
3	B	0.148	18	0.200
	AUT	0.164	18	0.200
	INT	0.170	19	0.153
4	B	0.282	20	0.000
	AUT	0.294	19	0.000
	INT	0.357	19	0.000
5	B	0.123	19	0.200
	AUT	0.116	19	0.200
	INT	0.138	18	0.200
6	B	0.161	18	0.200
	AUT	0.267	19	0.001
	INT	0.205	18	0.043
7	B	0.241	20	0.004
	AUT	0.185	19	0.085
	INT	0.175	18	0.150

Table 148: Results of K-S Test for normality for precision values

M: Mean; **p:** two-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	$H1_6$ $M_{AUT} \neq M_B$	$H1_7$ $M_{INT} \neq M_B$	$H8_1$ $M_{INT} \neq M_{AUT}$
1	p = 0.166 (ns)	p = 0.186 (ns)	p = 0.805 (ns)
2	p = 0.253 (ns)	p = 0.714 (ns)	p = 0.553 (ns)
3	p = 0.763 (ns)	p = 0.725 (ns)	p = 0.307 (ns)
4	p = 0.362 (ns)	p = 0.100 (ns)	p = 0.380 (ns)
5	p = 0.412 (ns)	p = 0.329 (ns)	p = 0.760 (ns)
6	p = 0.795 (ns)	p = 0.240 (ns)	p = 0.064 (ms)
7	p = 0.989 (ns)	p = 0.756 (ns)	p = 0.890 (ns)

Table 149: U-Test results for the system comparisons by precision

Appendix W System Comparisons by Precision (Loose Relevance Judgments)
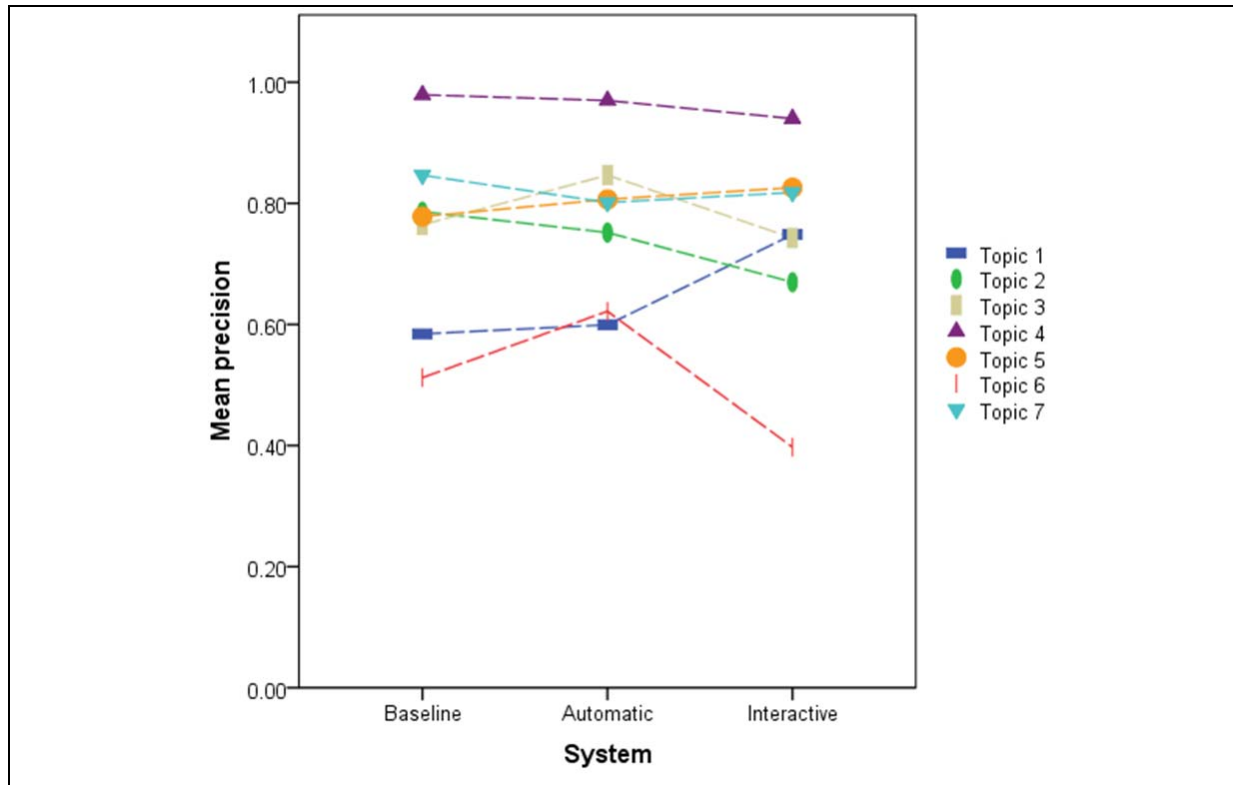
Figure 80: Topic-specific mean precision values by system

M: Mean; **p:** two-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	H1$_6$ $M_{AUT} \neq M_B$	H1$_7$ $M_{INT} \neq M_B$	H8$_1$ $M_{INT} \neq M_{AUT}$
1	p = 0.705 (ns)	p = 0.974 (ns)	p = 0.507 (ns)
2	p = 1.000 (ns)	p = 0.592 (ns)	p = 0.591 (ns)
3	p = 0.357 (ns)	p = 0.835 (ns)	p = 0.259 (ns)
4	p = 0.383 (ns)	p = 0.936 (ns)	p = 0.448 (ns)
5	p = 0.976 (ns)	p = 0.447 (ns)	p = 0.370 (ns)
6	p = 0.334 (ns)	p = 0.306 (ns)	p = 0.057 (ms)
7	p = 0.538 (ns)	p = 0.919 (ns)	p = 0.556 (ns)

Table 150: U-Test results for the system comparisons by precision (loose relevance judgments)

Appendix X   System   Comparisons   by   Pre-Click Confidence

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	18	0.698	0.219	0.636	19	0.657	0.289	0.697	18	0.856	0.204	1.000
2	19	0.866	0.187	0.900	19	0.821	0.195	0.857	16	0.806	0.142	0.826
3	18	0.783	0.175	0.775	18	0.783	0.166	0.778	19	0.783	0.167	0.818
4	20	0.712	0.209	0.697	19	0.831	0.158	0.900	18	0.856	0.117	0.885
5	19	0.768	0.179	0.800	19	0.808	0.166	0.833	18	0.900	0.141	1.000
6	18	0.657	0.246	0.725	19	0.783	0.169	0.750	15	0.778	0.238	0.800
7	20	0.659	0.256	0.691	19	0.745	0.280	0.800	17	0.833	0.157	0.875

Table 151: Location parameters of the systems' pre-click confidence value distributions

Topic	System	Statistics	df	Significance
1	B	0.146	19	0.200
	AUT	0.125	18	0.200
	INT	0.371	18	0.000
2	B	0.185	19	0.087
	AUT	0.190	19	0.071
	INT	0.172	16	0.200
3	B	0.170	18	0.180
	AUT	0.131	18	0.200
	INT	0.158	19	0.200
4	B	0.116	20	0.200
	AUT	0.197	19	0.051
	INT	0.175	18	0.150
5	B	0.113	19	0.200
	AUT	0.139	19	0.200
	INT	0.316	18	0.000
6	B	0.163	18	0.200
	AUT	0.111	19	0.200
	INT	0.224	15	0.041
7	B	0.162	20	0.176
	AUT	0.182	19	0.100
	INT	0.150	17	0.200

Table 152: Results of K-S Test for normality for the pre-click confidence measure

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p > 0.001$

Topic	$H1_{15}$ $M_{AUT} > M_B$	$H0_{15}$ $M_{AUT} < M_B$	$H1_{16}$ $M_{INT} > M_B$	$H0_{16}$ $M_{INT} < MR_B$	$H1_{17}$ $M_{INT} > M_{AUT}$	$H0_{17}$ $M_{INT} < M_{AUT}$
1		p = 0.421 (ns)	p = 0.012 (*)		p = 0.012 (*)	
2		p = 0.378 (ns)		p = 0.198 (ns)		p = 0.258 (ns)
3	p = 0.487 (ns)		p = 0.386 (ns)		p = 0.363 (ns)	
4	p = 0.028(ms)		p = 0.011 (*)		p = 0.439 (ns)	
5	p = 0.218 (ns)		p = 0.009 (**)		p = 0.028 (*)	
6	p = 0.076 (ns)		p = 0.096 (ms)			p = 0.3898 (ns)
7	p = 0.109 (ns)		p = 0.016 (*)		p = 0.230 (ns)	

Table 153: U-Test results for the system comparisons by pre-click confidence

Appendix Y System Comparisons by the Number of Query Reformulations

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	19	5.84	3.53	5.00	18	5.11	2.78	4.00	19	5.89	3.20	6.00
2	19	1.63	1.21	1.00	19	2.21	1.40	2.00	18	2.61	2.00	2.00
3	18	4.06	2.53	3.50	18	4.28	2.30	4.50	19	5.26	3.46	4.00
4	20	4.30	2.64	3.50	19	2.05	1.51	1.00	19	3.16	2.04	3.00
5	19	3.11	2.38	2.00	19	3.00	1.86	3.00	18	4.11	2.68	4.00
6	18	3.83	3.31	2.50	19	3.21	2.86	3.00	18	5.89	3.22	4.50
7	20	6.15	3.47	6.00	19	6.00	3.09	6.00	18	5.67	3.58	5.00

Table 154: Topic-specific distributions of the numbers of reformulations by system

Topic	System	Statistics	df	Significance
1	B	0.173	19	0.138
	AUT	0.211	18	0.034
	INT	0.118	19	0.200
2	B	0.225	19	0.012
	AUT	0.191	19	0.065
	INT	0.175	18	0.148
2	B	0.175	18	0.148
	AUT	0.156	18	0.200
	INT	0.274	19	0.001
2	B	0.189	20	0.060
	AUT	0.284	19	0.000
	INT	0.234	19	0.007
2	B	0.205	19	0.034
	AUT	0.184	19	0.089
	INT	0.203	18	0.048
2	B	0.210	18	0.035
	AUT	0.149	19	0.200
	INT	0.222	18	0.020
2	B	0.153	20	0.200
	AUT	0.184	19	0.089
	INT	0.241	18	0.007

Table 155: Results of K-S Test for normality for the number of reformulations

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p < 0.001$

Topic	$H1_{19}$ $M_B > M_{AUT}$	$H0_{19}$ $M_B \leq M_{AUT}$	$H1_{20}$ $M_{INT} > M_B$	$H0_{20}$ $M_{INT} \leq MR_B$	$H1_{21}$ $M_{INT} > M_{AUT}$	$H0_{21}$ $M_{INT} \leq M_{AUT}$
1	p = 0.329 (ns)		p = 0.442 (ns)		p = 0.209 (ns)	
2		p = 0.100 (ms)	p = 0.060 (ms)		p = 0.321 (ns)	
3		p = 0.339 (ns)	p = 0.109 (ns)		p = 0.274 (ns)	
4	p = 0.001 (***)			p = 0.069 (ms)	p = 0.023 (*)	
5	p = 0.418 (ns)		p = 0.084 (ms)		p = 0..064 (ms)	
6	p = 0.307(ns)		p = 0.013 (*)		p = 0.005 (**)	
7	p = 0.422 (ns)			p = 0.319 (ns)		p = 0.276 (ns)

Table 156: U-Test results for the system comparisons by the number of query reformulations

Appendix Z   Term Tactics by System Version

	Topic	Manual		Interactive	
		Replace	OR	Replace	OR
Baseline	1	22	0		
Automatic		14	0		
Interactive		6	3	5	37
Baseline	2	3	0		
Automatic		5	1		
Interactive		1	0	2	14
Baseline	3	22	1		
Automatic		32	6		
Interactive		3	3	0	40
Baseline	4	26	2		
Automatic		13	1		
Interactive		8	1	3	31
Baseline	5	14	0		
Automatic		15	5		
Interactive		6	0	2	17
Baseline	6	13	1		
Automatic		19	1		
Interactive		11	1	3	36
Baseline	7	61	1		
Automatic		62	2		
Interactive		12	8	3	34

Table 157: Absolute frequency of term tactics by system

Appendix AA System Comparisons by Query Complexity

Topic	B				AUT				INT			
	N	M	SD	Med	N	M	SD	Med	N	M	SD	Med
1	19	0.000	0.000	0.000	18	0.014	0.059	0.000	19	0.466	0.347	0.333
2	19	0.000	0.000	0.000	19	0.018	0.076	0.000	18	0.360	0.427	0.292
3	18	0.028	0.118	0.000	18	0.111	0.280	0.000	19	0.632	0.760	0.333
4	20	0.033	0.149	0.000	19	0.031	0.093	0.000	19	0.340	0.450	0.000
5	19	0.000	0.000	0.000	19	0.018	0.076	0.000	18	0.306	0.466	0.000
6	18	0.022	0.066	0.000	20	0.000	0.000	0.000	18	0.699	0.661	0.667
7	20	0.050	0.163	0.000	19	0.048	0.116	0.000	18	0.718	0.575	0.667

Table 158: Topic-specific distributions of the numbers of expansion terms per facet by system

Topic	System	Statistics	df	Significance
1	B	0.000	0	0.000
	AUT	0.538	18	0.000
	INT	0.175	19	0.127
2	B	0.000	0	0.000
	AUT	0.538	19	0.000
	INT	0.249	18	0.004
3	B	0.538	18	0.000
	AUT	0.466	18	0.000
	INT	0.192	19	0.062
4	B	0.515	20	0.000
	AUT	0.525	19	0.000
	INT	0.235	19	0.007
5	B	0.000	0	0.000
	AUT	0.538	19	0.000
	INT	0.355	18	0.000
6	B	0.518	18	0.000
	AUT	0.000	0	0.000
	INT	0.213	18	0.030
7	B	0.520	20	0.000
	AUT	0.504	19	0.000
	INT	0.176	18	0.146

Table 159: Results of K-S Test for normality for the average number of expansion terms per facet

M: Mean; **p:** one-tailed probability value; ***:** $p < 0.05$; ****:** $p < 0.01$; *****:** $p > 0.001$

Topic	$H1_{22}$ $M_{INT} > M_B$	$H0_{22}$ $M_{INT} < MR_B$	$H1_{23}$ $M_{INT} > M_{AUT}$	$H0_{23}$ $M_{INT} < M_{AUT}$
1	p < 0.001 (***)		p < 0.001 (***)	
2	p < 0.001 (***)		p < 0.001 (***)	
3	p <= 0.001 (***)		p = 0.003 (**)	
4	p = 0.002 (**)		p = 0.005 (**)	
5	p = 0.002 (**)		p = 0.006 (**)	
6	p < 0.001 (***)		p < 0.001 (***)	
7	p < 0.001 (***)		p < 0.001 (***)	

Table 160: U-Test results for the system comparisons by the number of expansion terms per facet