

Overview of Systems Description of Web Retrieval Task at the Third NTCIR Workshop

RunID	Subtask	TopicPart	LinkInfo	IndexUnit	IndexTech	IndexStruc	QueryMethod	QueryUnit	IRModel	Ranking
GRACE-LA1-1	I-A1	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA1-2	I-A1	T	link/cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA1-3	I-A1	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA1-4	I-A1	D	link/cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA2-1	I-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA2-2	I-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA2-3	I-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LA2-4	I-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LB-1	I-B	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LB-2	I-B	T	link/cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LB-3	I-B	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-LB-4	I-B	D	link/cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA1-1	II-A1	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA1-2	II-A1	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA1-3	II-A1	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA1-4	II-A1	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA2-1	II-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA2-2	II-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA2-3	II-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SA2-4	II-A2	T[R1]	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SB-1	II-B	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SB-2	II-B	T	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SB-3	II-B	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
GRACE-SB-4	II-B	D	cont	n-gram	normalization	inverted file	automatic	word and phrase	probabilistic	tf-idf, document length
K3100-01	II-A1	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original

QueryExpan	Filtering	SimModel	ClassModel	SearchTime	IndexTime	Note	RunID
pseudo-relevance feedback	no						GRACE-LA1-1
pseudo-relevance feedback	no						GRACE-LA1-2
pseudo-relevance feedback	no						GRACE-LA1-3
pseudo-relevance feedback	no						GRACE-LA1-4
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-LA2-1
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-LA2-2
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-LA2-3
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-LA2-4
pseudo-relevance feedback	no						GRACE-LB-1
pseudo-relevance feedback	no						GRACE-LB-2
pseudo-relevance feedback	no						GRACE-LB-3
pseudo-relevance feedback	no						GRACE-LB-4
no	no						GRACE-SA1-1
pseudo-relevance feedback	no						GRACE-SA1-2
no	no						GRACE-SA1-3
pseudo-relevance feedback	no						GRACE-SA1-4
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-SA2-1
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-SA2-2
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-SA2-3
pseudo-relevance feedback	no	weighted relevance feedback					GRACE-SA2-4
no	no						GRACE-SB-1
pseudo-relevance feedback	no						GRACE-SB-2
no	no						GRACE-SB-3
pseudo-relevance feedback	no						GRACE-SB-4
no query expansion	no filtering method						K3100-01

K3100-02	II-A1	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-03	II-A1	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-04	II-A1	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-05	I-A1	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-06	I-A1	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-07	I-A1	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-08	I-A1	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-09	II-B	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-10	II-B	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-11	II-B	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-12	II-B	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-13	I-B	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-14	I-B	T	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-15	I-B	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
K3100-16	I-B	D	link/cont	n-gram	n-gram	inverted file	automatic	word	Original	Original
NAICR-I-A1-1	I-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback twice)	OKAPI, Rocchio's formula
NAICR-I-A1-2	I-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model	OKAPI
NAICR-I-A1-3	I-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula, GroupRanking
NAICR-I-A1-4	I-A1	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model	OKAPI
NAICR-I-A2-1	I-A2	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model	OKAPI
NAICR-I-A2-2	I-A2	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula
NAICR-I-A2-3	I-A2	T[R1]	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula

no query expansion	no filtering method						K3100-02
no query expansion	no filtering method						K3100-03
no query expansion	no filtering method						K3100-04
no query expansion	no filtering method						K3100-05
no query expansion	no filtering method						K3100-06
no query expansion	no filtering method						K3100-07
no query expansion	no filtering method						K3100-08
no query expansion	no filtering method						K3100-09
no query expansion	no filtering method						K3100-10
no query expansion	no filtering method						K3100-11
no query expansion	no filtering method						K3100-12
no query expansion	no filtering method						K3100-13
no query expansion	no filtering method						K3100-14
no query expansion	no filtering method						K3100-15
no query expansion	no filtering method						K3100-16
no query expansion	pattern filtering (long words etc)						NAICR-I-A1-1
no query expansion	pattern filtering (long words etc)						NAICR-I-A1-2
no query expansion	pattern filtering (long words etc)						NAICR-I-A1-3
no query expansion	pattern filtering (long words etc)						NAICR-I-A1-4
no query expansion	pattern filtering (long words etc)						NAICR-I-A2-1
no query expansion	pattern filtering (long words etc)						NAICR-I-A2-2
no query expansion	pattern filtering (long words etc)	probablistic model (relevance feedback)					NAICR-I-A2-3

NAICR-I-A2-4	I-A2	T[R1]	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback twice)	OKAPI, Rocchio's formula
NAICR-I-B-1	I-B	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula
NAICR-I-B-2	I-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model	OKAPI
NAICR-I-B-3	I-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (make index twice)	OKAPI
NAICR-I-B-4	I-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback once)	OKAPI, Rocchio's formula
NAICR-II-A1-1	II-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback once)	OKAPI, Rocchio's formula
NAICR-II-A1-2	II-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback twice)	OKAPI, Rocchio's formula
NAICR-II-A1-3	II-A1	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula, GroupRanking
NAICR-II-A1-4	II-A1	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula
NAICR-II-A2-1	II-A2	T[R1]	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback once)	OKAPI, Rocchio's formula
NAICR-II-A2-2	II-A2	T[R1]	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula
NAICR-II-A2-3	II-A2	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula
NAICR-II-A2-4	II-A2	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model	OKAPI
NAICR-II-B-1	II-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback once)	OKAPI, Rocchio's formula
NAICR-II-B-2	II-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback twice)	OKAPI, Rocchio's formula
NAICR-II-B-3	II-B	D	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (make index twice)	OKAPI
NAICR-II-B-4	II-B	T	cont	word	morphology(chasen)	inverted file	automatic	word	probablistic model (relevance feedback)	OKAPI, Rocchio's formula

no query expansion	pattern filtering (long words etc)	probablistic model (relevance feedback)					NAICR-I-A2-4
no query expansion	pattern filtering (long words etc)						NAICR-I-B-1
no query expansion	pattern filtering (long words etc)						NAICR-I-B-2
no query expansion	pattern filtering (long words etc)						NAICR-I-B-3
no query expansion	pattern filtering (long words etc)						NAICR-I-B-4
no query expansion	pattern filtering (long words etc)						NAICR-II-A1-1
no query expansion	pattern filtering (long words etc)						NAICR-II-A1-2
no query expansion	pattern filtering (long words etc)						NAICR-II-A1-3
no query expansion	pattern filtering (long words etc)						NAICR-II-A1-4
no query expansion	pattern filtering (long words etc)	probablistic model (relevance feedback)					NAICR-II-A2-1
no query expansion	pattern filtering (long words etc)	probablistic model (relevance feedback)					NAICR-II-A2-2
no query expansion	pattern filtering (long words etc)						NAICR-II-A2-3
no query expansion	pattern filtering (long words etc)						NAICR-II-A2-4
no query expansion	pattern filtering (long words etc)						NAICR-II-B-1
no query expansion	pattern filtering (long words etc)						NAICR-II-B-2
no query expansion	pattern filtering (long words etc)						NAICR-II-B-3
no query expansion	pattern filtering (long words etc)						NAICR-II-B-4

OASIS11	II-A1	D	cont	bi-word and phrases	We did not use any NLP technique. The dynamic window (from 2 to 7 Japanese characters) was shifted through texts.	inverted index	automatic	bi-word and phrases	VSM	tf-idf
OASIS12	II-A1	D	cont	bi-word and phrases	We did not use any NLP technique. The dynamic window (from 2 to 7 Japanese characters) was shifted through texts.	inverted index	automatic	bi-word and phrases	VSM	tf-idf
OKSAT-WEB-F-01	II-A1	TD	cont	n-gram	none	inverted n-gram file	interactive	word and phrase	probabilistic model	tf-idf
OKSAT-WEB-F-02	I-A1	TD	cont	n-gram	none	inverted n-gram file	interactive	word and phrase	probabilistic model	tf-idf
OKSAT-WEB-F-03	II-A1	T	cont	n-gram	none	inverted n-gram file	automatic	word and phrase	probabilistic model	tf-idf
OKSAT-WEB-F-04	I-A1	T	cont	n-gram	none	inverted n-gram file	automatic	word and phrase	probabilistic model	tf-idf
OKSAT-WEB-F-05	II-A1	D	cont	n-gram	none	inverted n-gram file	automatic	word and phrase	probabilistic model	tf-idf
OKSAT-WEB-F-06	I-A1	D	cont	n-gram	none	inverted n-gram file	automatic	word and phrase	probabilistic model	tf-idf
ORGREF-LA1-1(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number
ORGREF-LA1-2(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number (synonyms are counted as a same search term)
ORGREF-LA1-3(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum
ORGREF-LA1-4(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum (synonyms are counted as a same search term)
ORGREF-LA1-5(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, then by the presence of uniq search terms with weights regarding the order in the query (synonyms are counted as a same search term)

no query expansion	No filtering			about 26 hours for all queries	about 1 week	10 index files according directories 00,..09 were produced. Each query were processed by each index. 200 results were obtained from each index. They were merged and sorted. The 1000 best results were presented in the OASIS01.res file.	OASIS11
no query expansion	No filtering			about 42 hours 30 minutes for all queries	about 1 week	10 index files according directories 00,..09 were produced. Each query were processed by each index. 1000 results were obtained from each index. They were merged and sorted. The 1000 best results were presented in the OASIS12.res file.	OASIS12
no query expansion	none			60 seconds	2 hours		OKSAT-WEB-F-01
no query expansion	none			300 seconds	24 hours		OKSAT-WEB-F-02
no query expansion	none			60 seconds	2 hours		OKSAT-WEB-F-03
no query expansion	none			300 seconds	24 hours		OKSAT-WEB-F-04
no query expansion	none			60 seconds	2 hours		OKSAT-WEB-F-05
no query expansion	none			300 seconds	24 hours		OKSAT-WEB-F-06
no	no						ORGREF-LA1-1(*)
no	no						ORGREF-LA1-2(*)
no	no						ORGREF-LA1-3(*)
no	no						ORGREF-LA1-4(*)
no	no						ORGREF-LA1-5(*)

ORGREF-LA1-6(*)	I-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, by the presence of uniq search terms with weights regarding the order in the query, then by the total occurrence number considering the document length (synonyms are counted as a same search term)
ORGREF-LB-1(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number
ORGREF-LB-2(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number (synonyms are counted as a same search term)
ORGREF-LB-3(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum
ORGREF-LB-4(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum (synonyms are counted as a same search term)
ORGREF-LB-5(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, then by the presence of uniq search terms with weights regarding the order in the query (synonyms are counted as a same search term)
ORGREF-LB-6(*)	I-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, by the presence of uniq search terms with weights regarding the order in the query, then by the total occurrence number considering the document length (synonyms are counted as a same search term)
ORGREF-SA1-1(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number
ORGREF-SA1-2(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number (synonyms are counted as a same search term)

no	no						ORGREF-LA1-6(*)
no	no						ORGREF-LB-1(*)
no	no						ORGREF-LB-2(*)
no	no						ORGREF-LB-3(*)
no	no						ORGREF-LB-4(*)
no	no						ORGREF-LB-5(*)
no	no						ORGREF-LB-6(*)
no	no						ORGREF-SA1-1(*)
no	no						ORGREF-SA1-2(*)

ORGREF-SA1-3(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum
ORGREF-SA1-4(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum (synonyms are counted as a same search term)
ORGREF-SA1-5(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, then by the presence of uniq search terms with weights regarding the order in the query (synonyms are counted as a same search term)
ORGREF-SA1-6(*)	II-A1	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, by the presence of uniq search terms with weights regarding the order in the query, then by the total occurrence number considering the document length (synonyms are counted as a same search term)
ORGREF-SB-1(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number
ORGREF-SB-2(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the number of presence of uniq search terms, then by the total occurrence number (synonyms are counted as a same search term)
ORGREF-SB-3(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum
ORGREF-SB-4(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	scoring each query term with tf-idf based measure, summing the scores, then ranking by the sum (synonyms are counted as a same search term)
ORGREF-SB-5(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, then by the presence of uniq search terms with weights regarding the order in the query (synonyms are counted as a same search term)

no	no						ORGREF-SA1-3(*)
no	no						ORGREF-SA1-4(*)
no	no						ORGREF-SA1-5(*)
no	no						ORGREF-SA1-6(*)
no	no						ORGREF-SB-1(*)
no	no						ORGREF-SB-2(*)
no	no						ORGREF-SB-3(*)
no	no						ORGREF-SB-4(*)
no	no						ORGREF-SB-5(*)

ORGREF-SB-6(*)	II-B	T	cont	character for Japanese, word for European lang.	character normalization; modified 2-gram; offset posting	inverted file	automatic	phrase	boolean	ranking by the presence of proximity search results, by the presence of uniq search terms with weights regarding the order in the query, then by the total occurrence number considering the document length (synonyms are counted as a same search term)
UAIFI1	I-A1	D	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI10	II-A1	D	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI11	II-A1	T	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI12	II-A1	T	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI13	II-B	D	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI14	II-B	D	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI15	II-B	T	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI16	II-B	T	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI2	I-A1	D	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI3	I-A1	T	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI4	I-A1	T	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI5	I-B	D	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI6	I-B	D	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI7	I-B	T	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi
UAIFI8	I-B	T	cont	word	morphology	inverted file	automatic	word	probabilistic model	Okapi
UAIFI9	II-A1	D	cont	word and bi-word	morphology	inverted file	automatic	word and bi-word	probabilistic model	Okapi

(*) 'ORGREF-*' are the runs submitted by the members of Executives Committee of Web Retrieval Task at the Third NTCIR Workshop.

no	no						ORGREF-SB-6(*)
No	No						UAIFI1
No	No						UAIFI10
No	No						UAIFI11
No	No						UAIFI12
No	No						UAIFI13
No	No						UAIFI14
No	No						UAIFI15
No	No						UAIFI16
No	No						UAIFI2
No	No						UAIFI3
No	No						UAIFI4
No	No						UAIFI5
No	No						UAIFI6
No	No						UAIFI7
No	No						UAIFI8
No	No						UAIFI9

Subtask: 'I' indicates the result using 100GB data set and 'II' the one using 10GB data set. The detailed task designs are explained in the overview paper.

TopicPart: the part of the topic used. The characters of 'T', 'D' and 'R[*]' respectively indicate TITLE, DESC and the *th documents specified in RDOC.

LinkInfo: how to use link information, e.g., link information only, link and contents information, contents only, etc.

IndexUnit: character, bi-character, word, bi-word, phrase, the name of the HTML tags used, link structure, etc.

IndexTech: the techniques used to process index terms, e.g., morphology, stemming, POS, etc.

IndexStruc: PAT, inverted file, signature file, etc.

QueryMethod: 'automatic' or 'interactive'. The 'automatic' indicates the run without any human intervention during query processing and search. The 'interactive' indicates one of the runs other than 'automatic'

QueryUnit: character, word, phrase, etc.

IRModel: Boolean model, vector space model, probabilistic model, etc.

Ranking: ranking factor for measuring each term, e.g., tf, tf-idf, mutual information, document length, PageRank, etc.

QueryExpan: techniques used to expand query or no query expansion

Filtering: filtering method for distilling junk pages, e.g. Kleinburg and Chakrabarti's topic distillation, Web pages selection using internet directory, SPAM filtering using SPAM-like words or patterns, etc., and the name of resources if used.

SimModel: [This item is only necessary for 'A2: Similarity Retrieval'] techniques used to retrieve similar documents using <RDOC> information, e.g., Rocchio's relevance feedback, probabilistic-model-based relevance feedback, the method using co-cited information and/or co-citing information, the method by retrieving pages on the same sites or directories, etc.

ClassModel: [This item is only necessary for 'C1: Search Results Classification'] techniques used to classify the documents, e.g., text categorization, text clustering, the method using term co-occurrence information within Web page titles, Web communities based on link analysis, the method by grouping pages on the same sites or directories, etc., and whether or not they are performed 'off-line' or 'on-line'.
The 'off-line' indicates that the classification has done before searching. The 'on-line' indicates that classification has done on the searching process.

SearchTime: [This item is optional.] the averaged seconds consumed for searching

IndexTime: [This item is optional.] the averaged hours consumed for indexing

Note: [This item is optional] any additional information

