NTCIR Workshop　Cross-Lingual IR Task

	Group's ID	
	Run ID	
	1.Overall Approach	
1.1	1)Basic approach	1.query translation 2.document translation 3.other 0.NA (answerd) 99.NA
	2.Query construction	
2.1	1)Auto or manually	1.auto 2.manual 0.NA (answerd) 99.NA
	2)If manually query builder	
2.21	-Domain expert	1.yes 2.no 0.NA (answerd) 99.NA
2.22	-Computer system expert	1.yes 2.no 0.NA (answerd) 99.NA
2.23	-Other	1.yes 2.no 0.NA (answerd) 99.NA
2.3	3)If manually ability Japanese	1.native speaker 2.write academic paper 3.read academic paper
		4.learn more than 3 months 5.can't understanding 6.other 0.NA (answerd) 99.NA
2.4	4)If manually ability English	1.native speaker 2.write academic paper 3.read academic paper
		4.learn more than 3 months 5.can't understanding 6.other 0.NA (answerd) 99.NA
2.5	5)Time to complete query (min)	0.NA (answerd) 99.NA
	6)Method Constructing query	0.NA (answerd) 99.NA
2.61	-Tokenizing	1.uni-gram 2.bi-gram 3.other n-gram 4.word 5.phrase 6.other 0.NA (answerd) 99.NA
2.62	-Phrase identification	1.yes 2.no 0.NA (answerd) 99.NA
2.63	-Syntactic parsing	1.yes 2.no 0.NA (answerd) 99.NA
2.64	-Word sense disambiguation	1.yes 2.no 0.NA (answerd) 99.NA
2.65	-Proper noun identification	1.yes 2.no 0.NA (answerd) 99.NA
2.66	-Auto expansion queries	1.lexical resources 2.automatic relevance feedback 3.other 0.no
2.67	-Auto addition Boolean	1.yes 2.no 0.NA (answerd) 99.NA
2.68	-Other	0.NA (answerd) 99.NA
2.7	7)Spelling checking	1.yes 2.no 0.NA (answerd) 99.NA
2.8	8)Correcting them	1.yes 2.no 0.NA (answerd) 99.NA
	3.Method Query translation	
	1)Multilingual dictionary	
3.11	-externally-constructed	1.yes 2.no 0.NA (answerd) 99.NA
3.12	*Name	0.NA (answerd) 99.NA
3.13	*Size (MB)	0.NA (answerd) 99.NA
3.14	-internally-constructed	1.yes 2.no 0.NA (answerd) 99.NA
3.15	*Source,material,method	0.NA (answerd) 99.NA
3.16	*Size (MB)	0.NA (answerd) 99.NA
3.2	2)Corpus	1.parallel corpus 2.comparable corpus 3.other 0.NA (answerd) 99.NA

	3)Machine translation system	
3.31	-externally-constructed	1.yes 2.no 0.NA (answerd) 99.NA
3.32	*Name	0.NA (answerd) 99.NA
3.33	*Size (MB)	0.NA (answerd) 99.NA
3.34	-internally-constructed	1.yes 2.no 0.NA (answerd) 99.NA
3.35	*Source,material,method	0.NA (answerd) 99.NA
3.36	*Size (MB)	0.NA (answerd) 99.NA
3.4	4)Other	0.NA (answerd) 99.NA
3.5	5)Manual effort involve trans	1.yes 2.no 0.NA (answerd) 99.NA
3.6	6)Query expansion	1.before translation 2.after translation 0.NA (answerd) 99.NA
3.7	7)Method query expansion	1.auto relevance feedback 2.local context analysis 3.global relevance feedback
		4.thesaurus,lexicon 5.other 0.NA (answerd) 99.NA
3.8	8)Disambiguation	1.yes 2.no 0.NA (answerd) 99.NA
	4.Searching	
	4.1.Search times	
4.1.1	1)Run ID	0.NA (answerd) 99.NA
4.1.2	2)Time to search (s)	0.NA (answerd) 99.NA
	4.2.Searching methods	
4.2.1	1)Vector space model	1.yes 2.no 0.NA (answerd) 99.NA
4.2.2	2)Probabilistic model	1.yes 2.no 0.NA (answerd) 99.NA
4.2.3	3)Other	0.NA (answerd) 99.NA
	4.3.Factors in ranking	
4.3.1	1)TF (Term Frequency)	1.yes 2.no 0.NA (answerd) 99.NA
4.3.2	2)IDF (Inverse Doc Frequency)	1.yes 2.no 0.NA (answerd) 99.NA
4.3.3	3)Other term weights	0.NA (answerd) 99.NA
4.3.4	4)Semantic closeness	1.yes 2.no 0.NA (answerd) 99.NA
4.3.5	5)Positional info in doc	1.yes 2.no 0.NA (answerd) 99.NA
4.3.6	6)Syntactic clues	1.yes 2.no 0.NA (answerd) 99.NA
4.3.7	7)Proximity of terms	1.yes 2.no 0.NA (answerd) 99.NA
4.3.8	8)Document length	1.yes 2.no 0.NA (answerd) 99.NA
4.3.9	9)Other	0.NA (answerd) 99.NA
	4.4.Machine information	
4.4.1	1)Machine type for experiment	1.Sun Ultra SPARC 2 2.Sun Ultra 10 3.Sun U2/U1/SS20 4.EQUIUM
		5.Sun Spark Station 5 6.Sun SS UA2 7.Pro2333DL 8.SPARC 20 9.AT compatible
4.4.2	2)Machine d/s	1.dedicated 2.shared 0.NA (answerd) 99.NA
4.4.3	3)Amount hard disk (MB)	0.NA (answerd) 99.NA

4.4.4	4)Amount of RAM (MB)	0.NA (answerd) 99.NA
4.4.5	5)Clock rate CPU (MHz)	0.NA (answerd) 99.NA
	4.5.Others	
4.5.1	1)Brief description other	0.NA (answerd) 99.NA
4.5.2	2)Other	0.NA (answerd) 99.NA
4.5.31	3)Group J native	1.yes 2.no 0.NA (answerd) 99.NA
4.5.32	unerstand J	1.yes 2.no 0.NA (answerd) 99.NA

NTCIR Workshop Cross-Lingual IR Task Summary of System Description Form

Group's ID	Run ID	t	d	n	c	jc	ec	ac	f	a/I	1.1	2.1	2.21	2.22	2.23	2.3	2.4	2.5	2.61	2.62	2.63	2.64	2.65	2.66	2.67	2.68	2.7	2.8	3.11	3.12	3.13	3.14	3.15	3.16	3.2	3.31	3.32	3.33
BKYTR	BKJEMTFU	1	1	1	1	1	0	1	1	a	1	1	0	0	0	0	0	5/60	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	0	1	!	?
BKYTR	BKJEBKFU	1	1	1	1	1	1	1	1	a	1	1	0	0	0	0	0	5/60	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	0	1	!	?
BKYTR	BKJEBDFU	1	1	1	1	1	0	1	1	a	1	1	0	0	0	0	0	5/60	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	0	1	!	?
BKYTR	BKJEECFU	0	0	0	1	1	1	1	0	a	1	99	99	99	99	99	99	5/60	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	0	1	!	?
BKYTR	BKJEBDDS	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	5/60	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	0	1	!	?
CRL	CRL1	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.5	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	?	99	2	0	0
CRL	CRL2	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.5	4	2	2	2	2	0	2	0	2	2	1	EDR	?	2	0	0	99	2	0	0
CRL	CRL11	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.5	4	2	2	2	2	0	2	0	2	2	99	99	99	99	99	99	99	2	0	0
FLAB1	FLAB11	1	1	1	1	1	0	0	0	i	1	2	99	99	99	1	2	99	99	99	99	99	99	99	99	99	99	99	2	0	0	1	99	99	99	99	99	99
FLAB1	FLAB12	1	1	1	1	1	0	0	0	i	1	2	99	99	99	1	2	99	99	99	99	99	99	99	99	99	99	99	2	0	0	1	99	99	99	99	99	99
NTE15	NTE153	0	1	0	0	0	0	0	0	a	3!	1	0	0	0	0	0	0.3	4	2	2	2	2	0	2	0	2	2	2	0	0	2	0	0	1!	2	0	0
NTE15	NTE154	1	1	0	1	1	0	0	0	a	3!	1	0	0	0	0	0	0.3	4	2	2	2	2	0	2	0	2	2	2	0	0	2	0	0	1!	2	0	0
SONIA	SONIA1	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	4	6!	2	2	2	2	3!	2	!	2	2	2	0	0	1	99	460K	1	2	0	0
SONIA	SONIA2	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	4	6!	2	2	2	2	3!	2	!	2	2	2	0	0	1	99	460K	1	2	0	0
SONIA	SONIA3	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	4	6!	2	2	2	2	3!	2	!	2	2	2	0	0	1	99	460K	1	2	0	0
sstut	sstut1	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0	3,5	2!	2	2	2	3!	2	0	2!	2	1	!	!	1	!	!	0	2	0	0
sstut	sstut2	0	1	1	0	0	0	0	0	a	1!	1	0	0	0	0	0	0	3,5	2!	2	2	2	3!	2	0	2!	2	1	!	!	1	!	!	0	2	0	0
TSB	TSB1	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	2~3s	0	0	0	0	0	0	0	0	0	0	99	99	99	99	99	99	99	1	!	!
TSB	TSB2	0	1	0	0	0	0	0	0	a	1	1!	0	0	0	0	0	0	0	0	0	0	0	2!	0	0	0	0	99	99	99	99	99	99	99	1	!	!
TSB	TSB3	0	1	0	0	0	0	0	0	i	1	2	2	2	1!	1	6!	1	0	0	0	0	0	0	0	0	1	1	99	99	99	99	99	99	99	99	99	99
TSB	TSB4	0	1	0	0	0	0	0	0	i	1	2!	2	2	1!	1	6!	0	0	0	0	0	0	2!	0	0	1	1	99	99	99	99	99	99	99	99	99	99
TSB	TSB5	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	99	99	99	99	99	99	1	!	!
TSB	TSB6	0	1	0	0	0	0	0	0	a	1	1!	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	99	99	99	99	99	99	1	!	!
TSTAR	tstar1	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	99	99	99	99
TSTAR	tstar4	0	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	99	99	99	99
TSTAR	tstar7	1	0	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	99	99	99	99
TSTAR	tstar10	1	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	99	99	99	99
TSTAR	tstar21	1	1	1	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	99	99	99	99
TSTAR	tstar2	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar5	0	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar8	1	0	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar11	1	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar22	1	1	1	1	1	1	1	0	a	1	1	0	0	0	0	0	0	4	2	2	2	2	0	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar3	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.8	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar6	0	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.8	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar9	1	0	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.8	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar12	1	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.8	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar17	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar18	0	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar19	1	0	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar20	1	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	0.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar23	1	1	1	1	1	1	1	0	a	1	1	0	0	0	0	0	0.9	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar13	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	2.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar14	0	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	2.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar15	1	0	0	1	1	1	1	0	a	1	1	0	0	0	0	0	2.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
TSTAR	tstar16	1	1	0	1	1	1	1	0	a	1	1	0	0	0	0	0	2.3	4	2	2	1	2	3!	2	0	2	2	2	0	0	1	99	3MB	3!	99	99	99
ULIS	ULIS1	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS2	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS3	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS4	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS5	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS6	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS7	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS8	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS9	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS11	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS12	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS13	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS14	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS15	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS16	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS17	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS18	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
ULIS	ULIS19	0	1	0	0	0	0	0	0	a	1!	1	0	0	0	0	0	0.2	4	2	2	2	2	0	2	0	2	2	2	0	0	1	!	!	99	99	99	99
UMD	umd1	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0.5	0	0	0	0	0	0	0	!	2!	2!	!	0	0	0	0	0	99	99	99	99
UMD	umd2	0	0	1	0	0	0	0	0	a	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
1K	1KE	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	4,5	1	2	2	2	3!	0	0	2	2	2	0	0	1	!	!	99	99	99	99
1K	1KE3	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	4,5	1	2	1	2	3!	0	0	2	2	2	0	0	1	!	!	99	99	99	99
1K	1KE_ij	0	1	0	0	0	0	0	0	a	1	1	0	0	0	0	0	0	4,5	1	2	2	2	3!	0	0	2	2	2	0	0	1	!	!	99	99	99	99

Group's ID	Run ID	3.34	3.35	3.36	3.4	3.5	3.6	3.7	3.8	4.1.2	4.2.1	4.2.2	4.2.3	4.3.1	4.3.2	4.3.3	4.3.4	4.3.5	4.3.6	4.3.7	4.3.8	4.3.9	4.4.1	4.4.2	4.4.3	4.4.4	4.4.5	4.5.1	4.5.2	4.5.31	4.5.32
BKYTR	BKJEMTFU	2	0	0	99	2	0	0	2!	3.50	2	1	0	1	1	0	2	2	2	2	1	!	1	2	40	512	148	99	99	1	1
BKYTR	BKJEBKFU	2	0	0	99	2	0	0	2!	3.05	2	1	0	1	1	0	2	2	2	2	1	!	1	2	40	512	148	99	99	1	1
BKYTR	BKJEBDFU	2	0	0	99	2	0	0	2!	3.05	2	1	0	1	1	0	2	2	2	2	1	!	1	2	40	512	148	99	99	1	1
BKYTR	BKJEECFU	2	0	0	99	2	0	0	2!	1.90	2	1	0	1	1	0	2	2	2	2	1	!	1	2	40	512	148	99	99	1	1
BKYTR	BKJEBDDS	2	0	0	99	2	0	0	2!	2.00	2	1	0	1	1	0	2	2	2	2	1	!	1	2	40	512	148	99	99	1	1
CRL	CRL1	2	0	0	0	2	99	99	2	500.00	2	?	!	2	1	0	2	2	2	2	2	0	2	1	60	1G	?	0	!	1	1
CRL	CRL2	2	0	0	0	2	99	99	2	500.00	2	?	!	2	1	0	2	2	2	2	2	0	2	1	60	1G	?	0	!	1	1
CRL	CRL11	2	0	0	0	2	99	99	2	99.00	2	?	!	2	1	0	2	2	2	2	2	0	2	1	60	1G	?	0	!	1	1
FLAB1	FLAB11	99	99	99	99	99	99	99	99	99.00	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
FLAB1	FLAB12	99	99	99	99	99	99	99	99	99.00	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
NTE15	NTE153	2	0	0	!	2	0	99	99	0.98	1	2	0	1	1	0	2	2	2	1	1	0	6	2	30	1024	296	99	99	1	1
NTE15	NTE154	2	0	0	!	2	0	99	99	0.98	1	2	0	1	1	0	2	2	2	1	1	0	6	2	30	1024	296	99	99	1	1
SONIA	SONIA1	1	!	99	99	2	1	5!	1!	2.80	1	2	0	1	1	0	2	2	2	2	2	0	1	2	50	524	296	0	0	1	0
SONIA	SONIA2	1	!	99	99	2	1	5!	1!	2.80	1	2	0	1	1	0	2	2	2	2	2	0	1	2	50	524	296	0	0	1	0
SONIA	SONIA3	1	!	99	99	2	1	5!	1!	2.80	1	2	0	1	1	0	2	2	2	2	2	0	1	2	50	524	296	0	0	1	0
sstut	sstut1	0	0	0	99	2	0!	5!	1!	5395.30	2	2	!	1	1	0	2	2	2	2	2	!	9	2	8	512	400	!	!	1	0
sstut	sstut2	0	0	0	99	2	0!	5!	1!	13716.44	2	2	!	1	1	0	2	2	2	2	2	!	9	2	8	512	400	!	!	1	0
TSB	TSB1	2	0	0	99	2	0	99	1!	99.00	2	1	0	1	1	BM25	99	2	2	2	1	0	3	2	2500	512	296	!	99	1	0
TSB	TSB2	2	0	0	99	2	2	1	1!	99.00	2	1	0	1	1	BM25	99	2	2	2	1	0	3	2	2500	512	296	!	99	1	0
TSB	TSB3	99	99	99	99	1	0	99	0	99.00	2	1	0	1	1	BM25	99	2	2	2	1	0	3	2	2500	512	296	!	99	1	0
TSB	TSB4	99	99	99	99	1	2	1	0	99.00	2	1	0	1	1	BM25	99	2	2	2	1	0	3	2	2500	512	296	!	99	1	0
TSB	TSB5	2	0	0	99	2	0	99	0	0.10	1	2	0	1	1	99	99	1	2	1	2	0	4	1	2G+2G	320	!	!	99	1	0
TSB	TSB6	2	0	0	99	2	0	99	0	0.10	1	2	0	1	1	99	99	1	1	1	2	0	4	1	2G+2G	320	!	!	99	1	0
TSTAR	tstar1	99	99	99	99	2	0	2	2	4.71	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar4	99	99	99	99	2	0	2	2	6.22	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar7	99	99	99	99	2	0	2	2	5.66	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar10	99	99	99	99	2	0	2	2	6.22	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar21	99	99	99	99	2	0	2	2	8.86	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar2	99	99	99	99	2	0	2	2	5.00	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar5	99	99	99	99	2	0	2	2	5.92	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar8	99	99	99	99	2	0	2	2	5.94	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar11	99	99	99	99	2	0	2	2	6.79	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar22	99	99	99	99	2	0	2	2	13.02	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar3	99	99	99	99	2	2	5!	1	3.00	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar6	99	99	99	99	2	2	5!	1	5.92	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar9	99	99	99	99	2	2	5!	1	5.94	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar12	99	99	99	99	2	2	5!	1	6.79	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar17	99	99	99	99	2	2	5!	1	4.96	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar18	99	99	99	99	2	2	5!	1	6.28	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar19	99	99	99	99	2	2	5!	1	5.66	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar20	99	99	99	99	2	2	5!	1	6.28	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar23	99	99	99	99	2	2	5!	1	10.18	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar13	99	99	99	99	2	2	5!	1	5.09	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar14	99	99	99	99	2	2	5!	1	6.60	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar15	99	99	99	99	2	2	5!	1	5.86	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
TSTAR	tstar16	99	99	99	99	2	2	5!	1	6.45	1	2	0	1	1	0	2	2	2	2	1	0	5	2	8G	32	99	99	99	0	1
ULIS	ULIS1	99	99	99	!	2	0	99	1!	0.22	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS2	99	99	99	!	2	0	99	1!	0.24	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS3	99	99	99	!	2	0	99	1!	0.27	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS4	99	99	99	!	2	0	99	1!	0.31	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS5	99	99	99	!	2	0	99	1!	0.32	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS6	99	99	99	!	2	0	99	1!	0.37	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS7	99	99	99	!	2	0	99	1!	0.40	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS8	99	99	99	!	2	0	99	1!	0.42	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS9	99	99	99	!	2	0	99	1!	0.48	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS11	99	99	99	!	2	0	99	1!	0.22	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS12	99	99	99	!	2	0	99	1!	0.23	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS13	99	99	99	!	2	0	99	1!	0.27	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS14	99	99	99	!	2	0	99	1!	0.31	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS15	99	99	99	!	2	0	99	1!	0.32	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS16	99	99	99	!	2	0	99	1!	0.37	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS17	99	99	99	!	2	0	99	1!	0.45	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS18	99	99	99	!	2	0	99	1!	0.47	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
ULIS	ULIS19	99	99	99	!	2	0	99	1!	0.55	1	2	0	1	1	0	2	2	2	2	2	0	7	2	6.4	192	333	!	!	1	0
UMD	umd1	99	99	99	99	2	0	99	2	1.50	2	1!	0	1	1	0	2	2	2	2	1	0	8	2	22	64	33	99	99	0!	0
UMD	umd2	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
1K	1KE	99	99	99	99	2	!	4	99	99	2	2	?	1	1	2	2	2	2	2	1	0	99	99	99	99	99	99	99	1	0
1K	1KE3	99	99	99	99	2	!	4	99	99	2	2	?	1	1	2	2	2	2	2	1	0	99	99	99	99	99	99	99	1	0
1K	1KE_ij	99	99	99	99	2	!	4	99	99	2	2	?	1	1	2	2	2	2	2	1	0	99	99	99	99	99	99	99	1	0

NTCIR Workshop　Cross-Lingual IR Task

Group's ID	Run ID	Q.No.	Contents
BKYTR	BKJEMTFU BKJEBKFU BKJEBDFU BKJEECFU BKJEBDDS	3.15	We extracted the Japanese and English keyword fields from the documents in the ntc1-je0 collection. The Japanese keywords are paired with the English keywords from the same document in the order in which they appear in the keyword fields. When there are more than one English translations for the same Japanese keyword, the most frequent English translation found in the ntc1-je0 collection is selected as the translaton of a Japanese keyword.
		3.16	373,447 entries in 23 MB
		3.32	GISELLE (a research system from University of Southern California Information Sciences Institute) uses phonetic translation, but does not have a technical term dictionary
		3.8	NO for dictionary based systems,unknown for machine translation
		4.3.9	
CRL	CRL1	3.15	je0のキーワード部分から取得
	CRL1	4.2.3	ロバートソンの式の亜流
	CRL2 CRL11	4.5.2	今回の検索に，jeのデータから知識獲得してその知識を用いて e データを検索するのは，若干クローズドの意味合いがでてくるように思います． 例えば，jeのデータをまるごと知識として扱ってよい場合，je に対して検索を行ないその結果の中で eテキストに該当するものを取り出すという手法も考えられます．この場合クロスリンガル検索といえなくなってきます． (モノリンガル検索と同程度の精度が出てしまいます．) je のデータを検索の知識ベース作成に用いることに関して なんらかの対応が必要に思います．
NTE15	NTE153 NTE154	1.1	対訳コーパスを利用し，類似度計算(タームベクトルの内積)により日本語質問文から英語検索式を自動生成
		3.2	文書単位の対応づけのある対訳コーパス
		3.4	日本語質問文から単語を切り出し，各日本語単語に対応する英単語を取得する．英単語取得には，対訳コーパスから作成したタームベクトルを利用し，ベクトルの内積値が高い単語を訳語として採用する．得られた訳語をOR結合して検索式とする．
SONIA	SONIA1 SONIA2 SONIA3	2.61	<検索要求>の部分のみを辞書を用いて、日本語キーワードにまず分割
		2.66	NTC1-J0 から単語共起頻度をとって、相互情報的基準で閾値以上のを、検索キーワードに追加
		2.68	最後に日本語キーワードを英語に翻訳
		3.35	文章を対象としたシステムではなく、辞書にある訳語候補の中から、入力された複数のキーワードの適した訳語を選択するもの
		3.7	予めコーパスから単語の共起頻度をとっておき、各検索キーワードに相互情報量的な基準で、閾値以上のものを、検索キーに追加。
		3.8	予め、ソース言語とターゲット言語でそれぞれ単語共起頻度データを計数する。可能な訳語候補の組合せのうち、それらの共起頻度を要素とするベクトルの方向が、日本語検索キーワード間の共起頻度を要素とするベクトルの方向に最も近付くような、訳語候補の組合せを選択する。
sstut	sstut1 sstut2	1.1	検索式を全部分文字列に分割し，多言語辞書(同義語を含む)を用いて翻訳した．
		2.62	ただし間接的に，抽出と同等の効果のある方法を使用した．
		2.66	同義語も含むように拡張した．
		2.7	ただし同等の効果がある方法を使用した．
		3.12 3.13	電気・電子・情報用語対訳辞典(日外アソシエーツ) 　　－サイズ 約79,000[語数] 2[MB]: 建築・土木用語対訳辞典(日外アソシエーツ) 約49,000[語数] 1[MB] コンピュータ用語辞典(日外アソシエーツ) 約27,000[語数] 0.7[MB] 25万医学用語大辞典(日外アソシエーツ) 約480,000[語数] 6[MB]: 最新科学技術用語辞典(三修社) 約160,000[語数] 7[MB]
		3.15	上の5つの辞書をつなげて一つの辞書として使用した．

		3.16	ーサイズ 579,116[語数] 17[MB]
		3.6	ただし，翻訳の訳数を複数使用した
		3.7	同義語辞書と等価のものを使用した.
		3.8	対象データが学会論文データベースであることを考慮し，技術用語の辞典を用いて翻訳した
		4.2.3	dpマッチングアルゴリズムを拡張し，文字列単位で比較を行なう.
		4.3.9	それぞれの部分文字列についてtf,idfを計算し，tfが1の場合と，idfが0.05よりも大きい場合は有効な文字列ではないとしてスコアは0，それ以外は文字列の長さにidfを掛けたものをその部分文字列のスコアとする．各ドキュメントにおいてスコアの総和をとったものをそのドキュメントのスコアとし，それによってランク付けする.
		4.5.1	本システムの特色は検索式を単語に分割するということをせずに，全部分文字列に対して同義語も含めて英訳し検索するというものである. dpマッチングを使用した方法で，一つのベースラインを示す目的で参加した.
		4.5.2	sstut1は<検索要求>のみを用いた検索結果. sstut2は<検索要求>をのみ用いて得た検索結果の上位一万件のドキュメントをピックアップし，その後<検索要求>と<検索要求説明>を用いて検索したもの.
TSB	TSB1 TSB2 TSB3 TSB4 TSB5 TSB6	2.1	TSB1: automatic TSB2: automatic+local feedback TSB3: manual TSB4: manual+local feedback TSB5: automatic TSB6: automatic, with syntactic analysis
		2.2	a researcher in IR(TSB3&4)
		2.4	Japanese/English bilingual(TSB3&4)
		2.66	local feedback(TSB2&4)
		3.32	Toshiba ASTRANSAC(TSB1&2) Toshiba The Honyaku Professional V4.0(TSB5&6)
		3.33	140M entries(TSB1&2) 140M entries 70MB(TSB5&6)
		3.8	MT disambiguation(TSB1&2)
		4.4.5	296(TSB1-4) Pentium2 45m(TSB5&6)
		4.5.1	combination of morpheme matching and string matching stemming
TSTAR	tstar1~23	2.66	word co-occurrence(tstar3,6,9,12,13,14,15,16,17,18,19,20,23)
		3.2	Monolingual Corpus: LOB/NACSIS/TREC6 (Disc 4 and 5) (tstar2,5,8,11,22) Monolingual Corpus: LOB(tstar3,6,9,12) Monolingual Corpus: NACSIS-E collection(tstar17,18,19,20) Monolingual Corpus;TREC6 (Disc 4 and 5)(tstar13,14,15,16,23)
		3.7	word co-occurrence(tstar3,6,9,12,13,14,15,16,17,18,19,20,23)
ULIS	ULIS1 ULIS2 ULIS3 ULIS4 ULIS5 ULIS6 ULIS7 ULIS8 ULIS9 ULIS10 ULIS11 ULIS12 ULIS13 ULIS14 ULIS15 ULIS16 ULIS17 ULIS18 ULIS19	1.1	まず、形態素解析によって「検索要求」中の内容語を抽出し、検索タームとする 次に、以下の3つのステップで検索タームを翻訳する ステップ1: 専門用語辞書を用いて辞書引きする ステップ2: ステップ1で失敗した「カタカナ」文字列を「翻字」する ステップ3: ステップ1,2で失敗した単語を一般辞書を用いて辞書引きする 検索タームが複合語の場合は、辞書を用いて語基に分割しながら上記の 3ステップを行う 実行IDと翻訳方式の対応を以下に示す ULIS1, ULIS4, ULIS7, ULIS11, ULIS14, ULIS17: ステップ1のみ ULIS2, ULIS5, ULIS8, ULIS12, ULIS15, ULIS18: ステップ1,2のみ ULIS3, ULIS6, ULIS9, ULIS13, ULIS16, ULIS19: ステップ1～3 全て適用
		3.15	EDR専門用語辞書(情報処理) 2語基からなる複合語対訳に対して、日本語エントリを分割し、語基対訳辞書を作成した
		3.16	2.4万語(0.50MB)
		3.15	EDR日英対訳辞書 1語基エントリ(単純語)のみを抽出した
		3.16	26万語(7.4MB)

		3.4	NACSISコレクションの英語抄録から抽出した英単語bi-gram
		3.8	EDR専門用語辞書から抽出した語基の対応頻度とNACSISコレクションから抽出した英語bi-gramを用いて訳語候補ごとに「確率スコア」を計算し、スコアが上位の候補を検索に用いた 実行IDと訳語候補数との対応を以下に示す ULIS1, ULIS2, ULIS3, ULIS11, ULIS12, ULIS13: 上位1訳語 ULIS4, ULIS5, ULIS6, ULIS14, ULIS15, ULIS16: 上位3訳語 ULIS7, ULIS8, ULIS9, ULIS17, ULIS18, ULIS19: 上位10訳語
		4.5.1	(a)「検索式翻訳」と「検索エンジン」が完全に独立しているので、 メンテナンスや使用モジュールの切替えが容易である (b) 日英／英日双方向の言語横断検索が可能
		4.5.2	実行ID「ULIS n」と「ULIS n + 10」の違い: 専門用語辞書を作成する際の日本語エントリの分割法が異なる
UMD	umd1 umd2	2.6	Our query construction consists of three step: - Fields extraction: In this first step fileds of <topic> (query number) and <description> are extracted from the topics file. The resulted file is of the query numbers together with their descriptions; -Segmentation The above file is passed to JUMAN2.2 for processing. The output file of JUMAN is processed again so that the result is the query number together with their segmented descriptions; - Query translation Query file created in step 2 is passed to our DQT system for translation. The result is queries together with their number in a format accepted by our search system INQUERY;
		2.7	Our DQT automatically translates the queries, so no spelling checking is needed.
		3.1	We use a Japanese/English dictionary called "edict" freely available from Monash University. (64433 entries, 2.6Mb)
		4.2.2	INQUERY
		4.5.31	None of the members in our group can understand Japanese. But we do have native Japanese speakers available on an occasional basis.
1K	1KE 1KE3 1KE_ij	2.66 3.15 3.16	1KE, 1KE3 - ntc1-jeのキーワードから自動構築した多言語クラスタ (37270 entry) 1KE_ij - 市販の情報処理関係辞書4種から作成したクラスタ (20636 entry) 1KE, 1KE3 - ntc1-jeのキーワードから自動構築した多言語クラスタ (37270 entry)
		3.6	翻訳するときに行なった