Appendix I

TIPSTER/MUC-5 INFORMATION EXTRACTION TEST SCORES

This appendix contains the summary score reports for each Tipster system on the Tipster/MUC-5 evaluation. The reports are ordered by language/domain (EJV, JJV, EME, JME) and secondarily by site (in alphabetical order).

Deeper investigation into system performance may be done on the basis of template-by-template score reports, which detail the performance of the system on each article in the test set. The template-by-template score reports are not included in this appendix.

A brief introduction to reading the score reports is presented here. Some information on the performance metrics may also be found in the paper, "Tipster/MUC-5 Information Extraction System Evaluation" by B. Sundheim in this volume. Further description of the score report and further definition and explanation of the scoring categories and performance metrics may be found in the MUC-5 proceedings volume in the paper titled "MUC-5 Evaluation Metrics" by N. Chinchor and B. Sundheim.

The first column in the report, **SLOT**, contains one of the following three types of label:

      A slot name, e.g., **content**
      An object name, e.g., **<template>**
      The word **"subtotals"**

Of the three types of rows, the one of primary interest is the one labeled with a slot name. The rows labeled with an object name may be disregarded. The rows containing subtotals may be of interest to those interested in performance on the various types of objects in the template.

The following tallies for the test set are presented in columns in the score reports:

POS	Number possible (= COR+PAR+INC+MIS), i.e., the number of fills a system should generate. This value is somewhat variable among systems because of alternative cardinalities of slot fills allowed for in the answer key.
ACT	Number "actual" (= COR+PAR+INC+SPU), i.e., the number of fills generated by the system under evaluation
COR	Number correct
PAR	Number partially correct
INC	Number incorrect
SPU	Number spurious
MIS	Number missing
NON	Number "noncommittal," i.e., number of null slot fills generated by system, where the fill is also null in the answer key

The rightmost columns in the report contain the scores for the following extraction metrics:

ERR	Error per response fill	$= (.5PAR+INC+SPU+MIS)/(COR+PAR+INC+SPU+MIS)$
UND	Undergeneration	$= MIS/POS$
OVG	Overgeneration	$= SPU/ACT$
SUB	Substitution	$= (.5PAR+INC)/(COR+PAR+INC)$
REC	Recall	$= (COR+.5PAR)/POS$
PRE	Precision	$= (COR+.5PAR)/ACT$

None of the above metrics includes the **NON** scoring category in its computation. Readers who wish to consider noncommittal fills to be correct may revise the formulas by including **NON** in all denominators and in those numerators where **COR** appears. Note, however, that the high undergeneration scores obtained by most systems on many slots in the evaluation places in question the value of considering noncommittal fills to be correct.

The first two rows at the bottom of the report are the **ALL-OBJECTS** and **MATCHED-ONLY** rows. These offer two options in the way the **ERR, UND, OVG, SUB, REC,** and **PRE** scores are computed.

ALL-OBJECTS gives the official overall tallies and overall **ERR, UND, OVG, SUB, REC,** and **PRE** scores. The tallies in **POS, ACT, COR, PAR, INC, SPU, MIS,** and **NON** columns add up to the **ALL-OBJECTS** values, but not all rows are considered. The rows that are not considered are the object rows (i.e., those where the entry in the **SLOT** column is an object name), the subtotal rows (since they simply sum up the tallies for the scored slots in each object), and the following slot rows in the **<template>** object: **doc-nr, doc-date,** and **doc-source** (these appear only in the EJV and JJV full-template reports).

MATCHED-ONLY contains tallies and scores given a method of scoring that eliminates slot-fill penalties in the case where the system under evaluation failed to generate an expected object and in the case where it generated a spurious object. The tallies in the slot rows do *not* sum up to the totals contained in the **MATCHED-ONLY** row.

The scores for the metrics in the **ALL-OBJECTS** and **MATCHED-ONLY** rows are not computed by averaging the scores in the individual slot rows, but instead are computed on the basis of the tally totals contained in the **ALL-OBJECTS** or **MATCHED-ONLY** row.

The remaining rows at the bottom of the report contain scores for the following metrics:

TEXT-FILTERING	Text filtering (document detection)
RICHNESS-NORMALIZED-ERROR	Richness-normalized error (normalized version of error per response fill)
ERROR-RATE-PER-WORD	Error rate per word
F-MEASURES	F-measures (weighted combination of recall and precision)

The formulation of each of the above metrics is described below.

Text-filtering scores are computed on the basis of relevant/nonrelevant decisions made by the system for the *text as a whole*, not on the basis of tallies in the slot rows. Some of the score category columns are used to contain the tallies for these decisions across the test set: **COR** (decides relevant and relevant is correct), **SPU** (decides relevant and nonrelevant is correct), **MIS** (decides nonrelevant and relevant is correct), and **NON** (decides nonrelevant and nonrelevant is correct). Text-filtering formulas are therefore as follows: **REC = COR/(COR+MIS); PRE = COR/(COR+SPU); UND = MIS/(COR+MIS); OVG = SPU/(COR+SPU).**

Two values of richness-normalized error are computed. The formulas are as follows: **Min-err = Wrong/All-fills, Max-err = Wrong/Req-fills.** There will be a range of fill counts from the minimum number of fills required **(Req-fills)** to the maximum number of fills allowed **(All-fills),** according to the answer key. These values are almost, but not quite, system-independent.

Error rate per word is computed as **Wrong/Word-count.** The scores for Japanese are not accurate.

F-measure scores corresponding to three different weightings of precision relative to recall are computed. The **P&R** value weights precision the same as recall; the **2P&R** value gives twice the weight to precision; the **P&2R** value weights precision half as much as recall. The general formula is $F = (\beta^2+1.0) \times P \times R/(\beta^2 \times P)+R$, where β is the weighting factor.

EJV Summary Scores
BBN PLUM System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-nr	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-date	275	282	273	0	2	7	0	0	3	0	2	1	99	97
doc-source	276	213	207	0	0	6	69	0	26	25	3	0	75	97
content	348	331	253	0	10	68	85	6	39	24	20	4	73	76
subtotals	348	331	253	0	10	68	85	6	39	24	20	4	73	76
<tie-up-relati	348	330	253	0	10	67	85	0	39	24	20	4	73	77
status	348	330	210	0	53	67	85	0	49	24	20	20	60	64
entity	791	861	469	0	55	337	267	0	58	34	39	10	59	54
joint-venture	179	71	35	0	10	26	134	96	83	75	37	22	20	49
ownership	103	99	46	0	2	51	55	150	70	53	52	4	45	46
activity	387	226	115	0	17	94	255	17	76	66	42	13	30	51
subtotals	1808	1587	875	0	137	575	796	263	63	44	36	14	48	55
<entity>	977	831	646	0	51	134	280	0	42	29	16	7	66	78
name	873	737	459	30	111	137	273	17	53	31	18	21	54	64
aliases	359	347	231	5	8	103	115	321	49	32	30	4	65	67
location	323	95	47	10	10	28	256	444	85	79	29	22	16	55
nationality	265	158	82	0	3	73	180	437	76	68	46	4	31	52
type	977	831	605	0	92	134	280	0	46	29	16	13	62	73
entity-rel	1156	940	579	0	133	228	444	0	58	38	24	19	50	62
person	118	16	10	0	1	5	107	567	92	91	31	9	8	62
facility	83	20	3	0	2	15	78	598	97	94	75	40	4	15
subtotals	4154	3144	2016	45	360	723	1733	2384	58	42	23	16	49	65
<entity-relati	436	352	256	0	30	66	150	0	49	34	19	10	59	73
entity1	883	891	464	0	81	346	338	0	62	38	39	15	52	52
entity2	266	85	43	0	13	29	210	97	85	79	34	23	16	50
rel-ent2-to-e	436	352	178	0	108	66	150	0	64	34	19	38	41	50
status	431	352	195	0	90	67	146	0	61	34	19	32	45	55
subtotals	2016	1680	880	0	292	508	844	97	65	42	30	25	44	52
<activity>	386	213	135	0	17	61	234	0	70	61	29	11	35	63
industry	418	176	120	0	8	48	290	0	74	69	27	6	29	68
site	552	67	24	1	15	27	512	29	96	93	40	39	4	36
start	103	0	0	0	0	0	103	98	100	100	*	*	0	*
end	16	0	0	0	0	0	16	128	100	100	*	*	0	*
revenue	23	0	0	0	0	0	23	131	100	100	*	*	0	*
subtotals	1112	243	144	1	23	75	944	386	88	85	31	14	13	59
<industry>	412	176	127	0	11	38	274	0	72	66	22	8	31	72
type	381	174	106	0	26	42	249	0	75	65	24	20	28	61
prod-serv	809	176	67	14	57	38	671	27	91	83	22	46	9	42
subtotals	1190	350	173	14	83	80	920	27	86	77	23	33	15	51
<facility>	164	36	25	0	6	5	133	0	85	81	14	19	15	69
name	36	1	0	0	1	0	35	24	100	97	0	100	0	0
location	142	11	5	1	4	1	132	3	96	93	9	45	4	50
type	163	36	25	0	6	5	132	0	85	81	14	19	15	69
subtotals	341	48	30	1	11	6	299	27	91	88	12	27	9	64
<person>	118	15	11	0	1	3	106	0	91	90	20	8	9	73
name	118	15	8	1	3	3	106	0	93	90	20	29	7	57
entity	118	16	10	0	2	4	106	0	92	90	25	17	8	62
position	136	15	5	0	7	3	124	0	96	91	20	58	4	33
subtotals	372	46	23	1	12	10	336	0	94	90	22	35	6	51
<ownership>	103	78	51	0	2	25	50	0	60	48	32	4	50	65
time	22	0	0	0	0	0	22	45	100	100	*	*	0	*
owned	103	42	11	0	14	17	78	0	91	76	40	56	11	26
capital	28	24	20	1	0	3	7	28	34	25	12	2	73	85
owner%	376	171	78	4	30	59	264	3	82	70	34	28	21	47
subtotals	529	237	109	5	44	79	371	76	82	70	33	29	21	47
<revenue>	23	0	0	0	0	0	23	0	100	100	*	*	0	*
type	23	0	0	0	0	0	23	0	100	100	*	*	0	*
total	15	0	0	0	0	0	15	0	100	100	*	*	0	*
total-eta	11	0	0	0	0	0	11	0	100	100	*	*	0	*
rate	8	0	0	0	0	0	8	0	100	100	*	*	0	*
rate-eta	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	57	0	0	0	0	0	57	0	100	100	*	*	0	*
<time>	149	0	0	0	0	0	149	0	100	100	*	*	0	*
after	16	0	0	0	0	0	16	0	100	100	*	*	0	*
before	27	0	0	0	0	0	27	0	100	100	*	*	0	*
during	91	0	0	0	0	0	91	0	100	100	*	*	0	*
on	19	0	0	0	0	0	19	0	100	100	*	*	0	*
subtotals	153	0	0	0	0	0	153	0	100	100	*	*	0	*

ALL-OBJECTS	12080	7666	4503	67	99	2124	6538	3266	68	54	28	18	38	59
MATCHED-ONLY	6940	6037	4503	67	99	803	1706	3266	41	24	13	13	65	75
TEXT-FILTERING	251	262	237	-	-	25	14	6	-	6	10	-	94	90

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	9667.5	11803	12128	0.7971	0.8191
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	9667.5	92862	0.1041		
	P&R	2P&R	P&2R		
F-MEASURES	45.95	53.07	40.51		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
content	348	331	253	0	10	68	85	6	39	24	20	4	73	76
subtotals	348	331	253	0	10	68	85	6	39	24	20	4	73	76
<tie-up-relati	348	330	253	0	10	67	85	0	39	24	20	4	73	77
status	348	330	210	0	53	67	85	0	49	24	20	20	60	64
entity	791	861	468	0	56	337	267	0	58	34	39	11	59	54
joint-venture	179	71	35	0	10	26	134	96	83	75	37	22	20	49
subtotals	1318	1262	713	0	119	430	486	96	59	37	34	14	54	56
<entity>	977	831	645	0	52	134	280	0	42	29	16	7	66	78
name	873	737	459	30	111	137	273	17	53	31	18	21	54	64
aliases	359	347	230	4	9	104	116	319	50	32	30	4	65	67
location	323	95	47	10	10	28	256	445	85	79	29	22	16	55
nationality	265	158	82	0	4	72	179	436	76	68	46	5	31	52
type	977	831	604	0	93	134	280	0	46	29	16	13	62	73
entity-rel	1156	940	578	0	134	228	444	0	58	38	24	19	50	61
subtotals	3953	3108	2000	44	361	703	1548	1217	56	39	23	16	51	65
<entity-relati	436	352	256	0	30	66	150	0	49	34	19	10	59	73
entity1	883	891	463	0	82	346	338	0	62	38	39	15	52	52
entity2	266	85	43	0	13	29	210	97	85	79	34	23	16	50
rel-ent2-to-e	436	352	178	0	108	66	150	0	64	34	19	38	41	50
status	431	352	195	0	90	67	146	0	61	34	19	32	45	55
subtotals	2016	1680	879	0	293	508	844	97	65	42	30	25	44	52
ALL-OBJECTS	7635	6381	3845	44	94	1709	2963	1416	59	39	27	17	51	61
MATCHED-ONLY	5269	5078	3845	44	94	668	859	1416	35	16	13	12	73	76
TEXT-FILTERING	251	262	237	-	-	25	14	6	-	6	10	-	94	90

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	5477	92862	0.059
	P&R	2P&R	P&2R
F-MEASURES	55.18	58.31	52.37

EJV Summary Scores
GE/CMU Shogun System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-nr	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-date	275	275	275	0	0	0	0	7	0	0	0	0	100	100
doc-source	276	282	276	0	0	6	0	0	2	0	2	0	100	98
content	348	389	289	0	6	94	53	11	35	15	24	2	83	74
subtotals	348	389	289	0	6	94	53	11	35	15	24	2	83	74
<tie-up-relati	348	389	289	0	6	94	53	0	35	15	24	2	83	74
status	348	389	231	0	64	94	53	0	48	15	24	22	66	59
entity	791	934	535	0	94	305	162	0	51	20	33	15	68	57
joint-venture	180	212	101	0	23	88	56	96	62	31	42	18	56	48
ownership	103	122	64	0	4	54	35	180	59	34	44	6	62	52
activity	387	367	214	0	44	109	129	10	57	33	30	17	55	58
subtotals	1809	2024	1145	0	229	650	435	286	53	24	32	17	63	56
<entity>	976	1057	749	0	61	247	166	0	39	17	23	8	77	71
name	872	937	554	44	115	224	159	27	47	18	24	19	66	61
aliases	359	389	232	6	5	146	116	399	53	32	38	3	65	60
location	322	338	140	19	23	156	140	443	69	43	46	18	46	44
nationality	265	212	100	0	11	101	154	509	73	58	48	10	38	47
type	976	1057	716	0	94	247	166	0	41	17	23	12	73	68
entity-rel	1153	1242	747	0	117	378	289	0	51	25	30	14	65	60
person	118	170	58	0	0	112	60	637	75	51	66	0	49	34
facility	84	192	37	0	10	145	37	650	84	44	76	21	44	19
subtotals	4149	4537	2584	69	375	1509	1121	2665	54	27	33	14	63	58
<entity-relati	437	436	311	0	24	101	102	0	42	23	23	7	71	71
entity1	885	982	568	0	97	317	220	0	53	25	32	14	64	58
entity2	268	260	128	0	32	100	108	93	65	40	38	20	48	49
rel-ent2-to-e	437	436	239	0	96	101	102	0	56	23	23	29	55	55
status	432	436	239	0	94	103	99	0	55	23	24	28	55	55
subtotals	2022	2114	1174	0	319	621	529	93	56	26	29	21	58	56
<activity>	387	366	262	0	20	84	105	0	44	27	23	7	68	72
industry	420	424	228	0	35	161	157	5	61	37	38	13	54	54
site	552	826	217	3	71	535	261	83	80	47	65	25	40	26
start	104	143	27	0	22	94	55	126	86	53	66	45	26	19
end	16	3	1	0	0	2	15	251	94	94	67	0	6	33
revenue	23	44	9	0	3	32	11	235	84	48	73	25	39	20
subtotals	1115	1440	482	3	131	824	499	700	75	45	57	22	43	34
<industry>	415	425	271	0	22	132	122	0	50	29	31	8	65	64
type	384	425	227	0	42	156	115	0	58	30	37	16	59	53
prod-serv	847	1193	287	54	220	632	286	1	79	34	53	44	37	26
subtotals	1231	1618	514	54	262	788	401	1	73	32	49	35	44	33
<facility>	164	280	104	0	19	157	41	0	68	25	56	15	63	37
name	36	0	0	0	0	0	36	91	100	100	*	*	0	*
location	142	180	41	10	28	101	63	8	81	44	56	42	32	26
type	163	280	91	0	30	159	42	0	72	26	57	25	56	32
subtotals	341	460	132	10	58	260	141	99	77	41	56	32	40	30
<person>	118	162	71	0	5	86	42	0	65	36	53	6	60	44
name	118	162	64	1	11	86	42	0	68	36	53	15	55	40
entity	118	162	54	0	22	86	42	0	74	36	53	29	46	33
position	136	162	42	0	33	87	61	0	81	45	54	44	31	26
subtotals	372	486	160	1	66	259	145	0	74	39	53	29	43	33
<ownership>	103	122	76	0	2	44	25	0	48	24	36	2	74	62
time	21	2	0	0	1	1	20	65	100	95	50	100	0	0
owned	103	123	38	0	39	46	26	0	74	25	37	51	37	31
capital	28	40	18	2	2	18	6	45	59	21	45	14	68	48
owner%	376	450	199	9	54	184	114	5	64	30	41	22	54	45
subtotals	528	615	255	11	96	249	166	115	66	31	40	28	49	42
<revenue>	23	37	10	0	4	23	9	0	78	39	62	28	43	27
type	23	37	9	0	5	23	9	0	80	39	62	36	39	24
total	15	30	5	1	2	22	7	4	85	47	73	31	37	18
total-eta	11	14	4	0	1	9	6	6	80	54	64	20	36	28
rate	8	7	4	0	1	2	3	6	60	38	28	20	50	57
rate-eta	0	0	0	0	0	0	0	10	*	*	*	*	*	*
subtotals	57	88	22	1	9	56	25	26	80	44	64	30	39	26
<time>	149	137	36	0	31	70	82	0	84	55	51	46	24	26
after	16	13	0	0	3	10	13	0	100	81	77	100	0	0
before	27	23	6	0	1	16	20	0	86	74	70	14	22	26
during	91	74	25	0	7	42	59	0	81	65	57	22	27	34
on	19	32	5	0	0	27	14	0	89	74	84	0	26	16
subtotals	153	142	36	0	11	95	106	0	85	69	67	23	24	25

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	12125	13913	6793	149	124	5405	3621	3996	61	30	39	19	57	49
MATCHED-ONLY	9140	9729	6793	149	124	1624	1039	3996	36	11	17	15	75	70
TEXT-FILTERING	251	262	242	-	-	20	9	11	-	4	8	-	96	92

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	10662.5	11813	12138	0.8784	0.9026
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	10662.5	92862	0.1148		
	P&R	2P&R	P&2R		
F-MEASURES	52.75	50.66	55.02		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
content	348	389	289	0	6	94	53	11	35	15	24	2	83	74
subtotals	348	389	289	0	6	94	53	11	35	15	24	2	83	74
<tie-up-relati	348	389	289	0	6	94	53	0	35	15	24	2	83	74
status	348	389	231	0	64	94	53	0	48	15	24	22	66	59
entity	791	934	535	0	94	305	162	0	51	20	33	15	68	57
joint-venture	180	212	101	0	23	88	56	96	62	31	42	18	56	48
subtotals	1319	1535	867	0	181	487	271	96	52	20	32	17	66	56
<entity>	976	1057	749	0	61	247	166	0	39	17	23	8	77	71
name	872	937	554	43	116	224	159	27	47	18	24	19	66	61
aliases	359	389	232	5	6	146	116	399	54	32	38	3	65	60
location	322	338	140	19	23	156	140	443	69	43	46	18	46	44
nationality	265	212	100	0	11	101	154	509	73	58	48	10	38	47
type	976	1057	716	0	94	247	166	0	41	17	23	12	73	68
entity-rel	1153	1242	747	0	117	378	289	0	51	25	30	14	65	60
subtotals	3947	4175	2489	67	367	1252	1024	1378	51	26	30	14	64	60
<entity-relati	437	436	311	0	24	101	102	0	42	23	23	7	71	71
entity1	885	982	568	0	97	317	220	0	53	25	32	14	64	58
entity2	268	260	128	0	32	100	108	93	65	40	38	20	48	49
rel-ent2-to-e	437	436	239	0	96	101	102	0	56	23	23	29	55	55
status	432	436	239	0	94	103	99	0	55	23	24	28	55	55
subtotals	2022	2114	1174	0	319	621	529	93	56	26	29	21	58	56
ALL-OBJECTS	7636	8213	4819	67	96	2454	1877	1578	52	24	30	16	64	59
MATCHED-ONLY	6041	6122	4819	67	96	627	546	1578	27	9	10	12	80	79
TEXT-FILTERING	251	262	242	-	-	20	9	11	-	4	8	-	96	92

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	5237.5	92862	0.0564
	P&R	2P&R	P&2R
F-MEASURES	61.23	59.93	62.6

267

EJV Summary Scores
NMSU/Brandeis Diderot System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-nr	282	282	282	0	0	0	0	0	0	0	0	0	100	100
doc-date	275	205	204	0	1	0	70	7	26	25	0	0	74	100
doc-source	276	282	276	0	0	6	0	0	2	0	2	0	100	98
content	348	275	189	0	30	56	129	15	53	37	20	14	54	69
subtotals	348	275	189	0	30	56	129	15	53	37	20	14	54	69
<tie-up-relati	348	275	189	0	30	56	129	0	53	37	20	14	54	69
status	348	275	156	0	63	56	129	0	61	37	20	29	45	57
entity	791	419	252	0	80	87	459	0	71	58	21	24	32	60
joint-venture	179	59	20	0	13	26	146	74	90	82	44	39	11	34
ownership	103	27	15	0	5	7	83	123	86	80	26	25	14	56
activity	387	136	75	0	33	28	279	6	82	72	20	30	19	55
subtotals	1808	916	518	0	194	204	1096	203	74	61	22	27	29	56
<entity>	975	490	381	0	50	59	544	0	63	56	12	12	39	78
name	873	452	258	34	95	65	486	7	71	56	14	29	32	61
aliases	359	216	136	2	6	72	215	165	68	60	33	5	38	63
location	323	179	27	11	23	118	262	181	93	81	66	47	10	18
nationality	265	143	53	0	12	78	200	225	84	75	54	18	20	37
type	975	490	375	0	56	59	544	0	64	56	12	13	38	76
entity-rel	1153	619	354	0	97	168	702	0	73	61	27	22	31	57
person	118	63	23	0	1	39	94	314	85	80	62	4	19	36
facility	81	16	2	0	1	13	78	354	98	96	81	33	2	12
subtotals	4147	2178	1228	47	291	612	2581	1246	74	62	28	20	30	57
<entity-relati	437	350	201	0	51	98	185	0	62	42	28	20	46	57
entity1	885	567	291	0	90	186	504	0	73	57	33	24	33	51
entity2	267	58	21	0	17	20	229	74	93	86	34	45	8	36
rel-ent2-to-e	437	350	114	0	138	98	185	0	79	42	28	55	26	32
status	432	350	153	0	96	101	183	0	71	42	29	38	35	44
subtotals	2021	1325	579	0	341	405	1101	74	76	54	30	37	29	44
<activity>	387	135	87	0	28	20	272	0	79	70	15	24	22	64
industry	420	200	88	0	34	78	298	0	82	71	39	28	21	44
site	552	70	17	1	27	25	507	18	97	92	36	61	3	25
start	104	10	1	0	1	8	102	57	99	98	80	50	1	10
end	16	0	0	0	0	0	16	84	100	100	*	*	0	*
revenue	23	0	0	0	0	0	23	86	100	100	*	*	0	*
subtotals	1115	280	106	1	62	111	946	245	91	85	40	37	10	38
<industry>	413	201	113	0	35	53	265	0	76	64	26	24	27	56
type	382	201	104	0	34	63	244	0	77	64	31	25	27	52
prod-serv	818	216	52	8	88	68	670	0	94	82	31	62	7	26
subtotals	1200	417	156	8	122	131	914	0	88	76	31	44	13	38
<facility>	164	29	15	0	5	9	144	0	91	88	31	25	9	52
name	35	0	0	0	0	0	35	12	100	100	*	*	0	*
location	142	8	1	0	2	5	139	3	99	98	62	67	1	12
type	163	29	15	0	5	9	143	0	91	88	31	25	9	52
subtotals	340	37	16	0	7	14	317	15	95	93	38	30	5	43
<person>	118	59	25	0	2	32	91	0	83	77	54	7	21	42
name	118	24	9	1	3	11	105	0	93	89	46	27	8	40
entity	118	59	23	0	4	32	91	0	85	77	54	15	19	39
position	136	53	11	0	17	25	108	0	93	79	47	61	8	21
subtotals	372	136	43	1	24	68	304	0	90	82	50	36	12	32
<ownership>	103	27	20	0	5	2	78	0	81	76	7	20	19	74
time	22	0	0	0	0	0	22	18	100	100	*	*	0	*
owned	103	22	7	0	14	1	82	0	93	80	4	67	7	32
capital	28	17	8	0	8	1	12	1	72	43	6	50	28	47
owner%	376	25	13	1	8	3	354	4	96	94	12	39	4	54
subtotals	529	64	28	1	30	5	470	23	95	89	8	52	5	44
<revenue>	23	0	0	0	0	0	23	0	100	100	*	*	0	*
type	23	0	0	0	0	0	23	0	100	100	*	*	0	*
total	15	0	0	0	0	0	15	0	100	100	*	*	0	*
total-eta	11	0	0	0	0	0	11	0	100	100	*	*	0	*
rate	8	0	0	0	0	0	8	0	100	100	*	*	0	*
rate-eta	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	57	0	0	0	0	0	57	0	100	100	*	*	0	*
<time>	149	10	1	0	3	6	145	0	99	97	60	75	1	10
after	16	0	0	0	0	0	16	0	100	100	*	*	0	*
before	27	0	0	0	0	0	27	0	100	100	*	*	0	*
during	91	7	1	0	3	3	87	0	99	96	43	75	1	14
on	19	0	0	0	0	0	19	0	100	100	*	*	0	*
subtotals	153	7	1	0	3	3	149	0	99	97	43	75	1	14

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	12090	5635	2864	58	81	1609	8064	1821	79	67	28	28	24	51
MATCHED-ONLY	4882	4059	2864	58	81	495	1318	1821	46	27	12	19	59	71
TEXT-FILTERING	251	208	192	-	-	16	59	15	-	24	8	-	76	92

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	10806	11804	12129	0.8909	0.9155

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	10806	92862	0.1164

	P&R	2P&R	P&2R
F-MEASURES	32.64	41.77	26.79

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	282	282	0	0	0	0	0	0	0	0	0	100	100
content	348	275	189	0	30	56	129	15	53	37	20	14	54	69
subtotals	348	275	189	0	30	56	129	15	53	37	20	14	54	69
<tie-up-relati	348	275	189	0	30	56	129	0	53	37	20	14	54	69
status	348	275	156	0	63	56	129	0	61	37	20	29	45	57
entity	791	419	252	0	80	87	459	0	71	58	21	24	32	60
joint-venture	179	59	20	0	13	26	146	74	90	82	44	39	11	34
subtotals	1318	753	428	0	156	169	734	74	71	56	22	27	32	57
<entity>	975	490	381	0	50	59	544	0	63	56	12	12	39	78
name	873	452	258	34	95	65	486	7	71	56	14	29	32	61
aliases	359	216	136	2	6	72	215	165	68	60	33	5	38	63
location	323	179	27	11	23	118	262	181	93	81	66	47	10	18
nationality	265	143	53	0	12	78	200	225	84	75	54	18	20	37
type	975	490	375	0	56	59	544	0	64	56	12	13	38	76
entity-rel	1153	619	354	0	97	168	702	0	73	61	27	22	31	57
subtotals	3948	2099	1203	47	289	560	2409	578	73	61	27	20	31	58
<entity-relati	437	350	201	0	51	98	185	0	62	42	28	20	46	57
entity1	885	567	291	0	90	186	504	0	73	57	33	24	33	51
entity2	267	58	21	0	17	20	229	74	93	86	34	45	8	36
rel-ent2-to-e	437	350	114	0	138	98	185	0	79	42	28	55	26	32
status	432	350	153	0	96	101	183	0	71	42	29	38	35	44
subtotals	2021	1325	579	0	341	405	1101	74	76	54	30	37	29	44
ALL-OBJECTS	7635	4452	2399	47	75	1190	4373	741	72	57	27	26	32	54
MATCHED-ONLY	3647	3277	2399	47	75	384	754	741	40	21	12	16	66	74
TEXT-FILTERING	251	208	192	-	-	16	59	15	-	24	8	-	76	92

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	6402.5	92862	0.0689

	P&R	2P&R	P&2R
F-MEASURES	40.08	47.61	34.62

269

EJV Summary Scores
Umass/Hughes CIRCUS System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	281	281	0	0	0	1	0	0	0	0	0	100	100
doc-nr	282	281	281	0	0	0	1	0	0	0	0	0	100	100
doc-date	275	281	274	0	0	7	1	0	3	0	2	0	100	98
doc-source	276	281	275	0	0	6	1	0	2	0	2	0	100	98
content	349	223	169	0	10	44	170	21	57	49	20	6	48	76
subtotals	349	223	169	0	10	44	170	21	57	49	20	6	48	76
<tie-up-relati	349	237	182	0	10	45	157	0	54	45	19	5	52	77
status	349	167	115	0	33	19	201	0	69	58	11	22	33	69
entity	793	509	304	0	66	139	423	0	67	53	27	18	38	60
joint-venture	180	78	33	0	16	29	131	53	84	73	37	33	18	42
ownership	102	41	27	0	2	12	73	114	76	72	29	7	26	66
activity	388	103	45	0	24	34	319	15	89	82	33	35	12	44
subtotals	1812	898	524	0	141	233	1147	182	74	63	26	21	29	58
<entity>	976	1017	611	0	128	278	237	0	51	24	27	17	63	60
name	873	873	320	84	215	254	254	16	68	29	29	42	41	41
aliases	359	92	66	3	4	19	286	358	82	80	21	8	19	73
location	325	223	91	8	23	101	203	393	78	62	45	22	29	43
nationality	265	190	87	1	9	93	168	418	76	63	49	10	33	46
type	976	878	545	0	120	213	311	0	54	32	24	18	56	62
entity-rel	1155	599	392	0	73	134	690	0	70	60	22	16	34	65
person	118	42	16	0	4	22	98	535	88	83	52	20	14	38
facility	81	0	0	0	0	0	81	577	100	100	*	*	0	*
subtotals	4152	2897	1517	96	448	836	2091	2297	69	50	29	24	38	54
<entity-relati	436	237	175	0	29	33	232	0	63	53	14	14	40	74
entity1	883	518	304	0	76	138	503	0	70	57	27	20	34	59
entity2	266	78	38	0	19	21	209	53	87	78	27	33	14	49
rel-ent2-to-e	436	237	101	0	103	33	232	0	78	53	14	50	23	43
status	431	237	136	0	66	35	229	0	71	53	15	33	32	57
subtotals	2016	1070	579	0	264	227	1173	53	74	58	21	31	29	54
<activity>	387	101	66	0	19	16	302	0	84	78	16	22	17	65
industry	420	88	64	0	8	16	348	0	85	83	18	11	15	73
site	552	0	0	0	0	0	552	9	100	100	*	*	0	*
start	103	0	0	0	0	0	103	48	100	100	*	*	0	*
end	16	0	0	0	0	0	16	61	100	100	*	*	0	*
revenue	23	17	2	0	1	14	20	60	94	87	82	33	9	12
subtotals	1114	105	66	0	9	30	1039	178	94	93	28	12	6	63
<industry>	413	89	66	0	11	12	336	0	84	81	13	14	16	74
type	380	58	41	0	8	9	331	1	89	87	16	16	11	71
prod-serv	807	113	21	26	50	16	710	11	96	88	14	65	4	30
subtotals	1187	171	62	26	58	25	1041	12	94	88	15	49	6	44
<facility>	164	4	1	0	1	2	162	0	99	99	50	50	1	25
name	35	0	0	0	0	0	35	1	100	100	*	*	0	*
location	142	4	1	0	1	2	140	0	99	98	50	50	1	25
type	163	0	0	0	0	0	163	0	100	100	*	*	0	*
subtotals	340	4	1	0	1	2	338	1	100	99	50	50	0	25
<person>	118	195	64	0	11	120	43	0	73	36	62	15	54	33
name	118	164	45	5	17	97	51	0	78	43	59	29	40	29
entity	118	42	16	0	8	18	94	0	88	80	43	33	14	38
position	136	112	46	0	7	59	83	0	76	61	53	13	34	41
subtotals	372	318	107	5	32	174	228	0	80	61	55	24	29	34
<ownership>	102	41	29	0	3	9	70	0	74	69	22	9	28	71
time	21	0	0	0	0	0	21	27	100	100	*	*	0	*
owned	102	17	7	0	7	3	88	0	93	86	18	50	7	41
capital	27	8	4	0	1	3	22	15	87	81	38	20	15	50
owner%	376	124	48	10	30	36	288	1	87	76	29	40	14	43
subtotals	526	149	59	10	38	42	419	43	89	80	28	40	12	43
<revenue>	23	17	2	0	3	12	18	0	94	78	70	60	9	12
type	23	17	2	0	3	12	18	0	94	78	70	60	9	12
total	15	0	0	0	0	0	15	2	100	100	*	*	0	*
total-eta	11	0	0	0	0	0	11	2	100	100	*	*	0	*
rate	8	17	2	0	1	14	5	0	91	62	82	33	25	12
rate-eta	0	0	0	0	0	0	0	2	*	*	*	*	*	*
subtotals	57	34	4	0	4	26	49	6	95	86	76	50	7	12
<time>	149	0	0	0	0	0	149	0	100	100	*	*	0	*
after	16	0	0	0	0	0	16	0	100	100	*	*	0	*
before	27	0	0	0	0	0	27	0	100	100	*	*	0	*
during	91	0	0	0	0	0	91	0	100	100	*	*	0	*
on	19	0	0	0	0	0	19	0	100	100	*	*	0	*
subtotals	153	0	0	0	0	0	153	0	100	100	*	*	0	*

ALL-OBJECTS	12078	5869	3088	137	161	1639	7848	2793	77	65	28	25	26	54
MATCHED-ONLY	5772	4137	3088	137	161	313	1948	2793	48	34	8	17	55	76
TEXT-FILTERING	251	174	164	-	-	10	87	21	-	35	6	-	65	94

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	10560.5	11807	12133	0.8704	0.8944
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	10560.5	92862	0.1137		
	P&R	2P&R	P&2R		
F-MEASURES	35.18	44.39	29.13		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	282	281	281	0	0	0	1	0	0	0	0	0	100	100
content	349	223	169	0	10	44	170	21	57	49	20	6	48	76
subtotals	349	223	169	0	10	44	170	21	57	49	20	6	48	76
<tie-up-relati	349	237	182	0	10	45	157	0	54	45	19	5	52	77
status	349	167	115	0	33	19	201	0	69	58	11	22	33	69
entity	793	509	304	0	66	139	423	0	67	53	27	18	38	60
joint-venture	180	78	33	0	16	29	131	53	84	73	37	33	18	42
subtotals	1322	754	452	0	115	187	755	53	70	57	25	20	34	60
<entity>	976	1017	611	0	128	278	237	0	51	24	27	17	63	60
name	873	873	320	84	215	254	254	16	68	29	29	42	41	41
aliases	359	92	66	3	4	19	286	358	82	80	21	8	19	73
location	325	223	91	8	23	101	203	393	78	62	45	22	29	43
nationality	265	190	87	1	9	93	168	418	76	63	49	10	33	46
type	976	878	545	0	120	213	311	0	54	32	24	18	56	62
entity-rel	1155	599	392	0	73	134	690	0	70	60	22	16	34	65
subtotals	3953	2855	1501	96	444	814	1912	1185	68	48	28	24	39	54
<entity-relati	436	237	175	0	29	33	232	0	63	53	14	14	40	74
entity1	883	518	304	0	76	138	503	0	70	57	27	20	34	59
entity2	266	78	38	0	19	21	209	53	87	78	27	33	14	49
rel-ent2-to-e	436	237	101	0	103	33	232	0	78	53	14	50	23	43
status	431	237	136	0	66	35	229	0	71	53	15	33	32	57
subtotals	2016	1070	579	0	264	227	1173	53	74	58	21	31	29	54
ALL-OBJECTS	7640	4902	2701	96	145	1272	4010	1312	69	52	26	24	36	56
MATCHED-ONLY	4513	3562	2701	96	145	284	1235	1312	43	27	8	16	61	77
TEXT-FILTERING	251	174	164	-	-	10	87	21	-	35	6	-	65	94

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	6163	92862	0.0664
	P&R	2P&R	P&2R
F-MEASURES	43.84	50.44	38.76

271

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-nr	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-date	147	147	147	0	0	0	0	3	0	0	0	0	100	100
doc-source	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	205	246	158	0	8	80	39	1	44	19	32	5	77	64
subtotals	205	246	158	0	8	80	39	1	44	19	32	5	77	64
<tie-up-relati	205	246	158	0	8	80	39	0	44	19	32	5	77	64
status	205	246	151	0	15	80	39	0	47	19	32	9	74	61
entity	440	370	237	0	21	112	182	0	57	41	30	8	54	64
joint-venture	23	21	5	0	2	14	16	138	86	70	67	28	22	24
ownership	31	0	0	0	0	0	31	134	100	100	*	*	0	*
activity	235	92	34	0	24	34	177	14	87	75	37	41	14	37
subtotals	934	729	427	0	62	240	445	286	64	48	33	13	46	58
<entity>	423	316	265	0	20	31	138	0	42	33	10	7	63	84
name	405	295	199	23	45	28	138	3	51	34	9	21	52	71
aliases	174	118	68	3	9	38	94	127	67	54	32	13	40	59
location	118	34	0	0	26	8	92	207	100	78	24	100	0	0
nationality	125	22	14	0	1	7	110	195	89	88	32	7	11	64
type	423	316	262	0	23	31	138	0	42	33	10	8	62	83
entity-rel	495	388	254	0	31	103	210	0	58	42	26	11	51	65
person	75	9	8	0	0	1	67	227	89	89	11	0	11	89
facility	108	52	13	0	11	28	84	206	90	78	54	46	12	25
subtotals	1923	1234	818	26	146	244	933	965	62	48	20	16	43	67
<entity-relati	219	246	161	0	8	77	50	0	46	23	31	5	74	65
entity1	454	370	242	0	18	110	194	0	57	43	30	7	53	65
entity2	37	21	6	0	3	12	28	138	88	76	57	33	16	28
rel-ent2-to-e	219	246	144	0	25	77	50	0	51	23	31	15	66	58
status	218	246	152	0	16	78	50	0	49	23	32	10	70	62
subtotals	928	883	544	0	62	277	322	138	55	35	31	10	59	62
<activity>	235	65	38	0	19	8	178	0	84	76	12	33	16	58
industry	275	64	38	0	19	7	218	0	86	79	11	33	14	59
site	616	1	0	0	0	1	616	0	100	100	100	*	0	0
start	84	0	0	0	0	0	84	25	100	100	*	*	0	*
end	6	0	0	0	0	0	6	37	100	100	*	*	0	*
revenue	18	0	0	0	0	0	18	33	100	100	*	*	0	*
subtotals	999	65	38	0	19	8	942	95	96	94	12	33	4	58
<industry>	255	64	42	0	15	7	198	0	84	78	11	26	16	66
type	253	64	42	0	15	7	196	0	84	77	11	26	17	66
prod-serv	566	0	0	0	0	0	566	0	100	100	*	*	0	*
subtotals	819	64	42	0	15	7	762	0	95	93	11	26	5	66
<facility>	99	46	19	0	6	21	74	0	84	75	46	24	19	41
name	11	17	0	0	3	14	8	12	100	73	82	100	0	0
location	51	10	2	1	4	3	44	5	95	86	30	64	5	25
type	98	46	18	0	7	21	73	0	85	74	46	28	18	39
subtotals	160	73	20	1	14	38	125	17	90	78	52	41	13	28
<person>	75	9	8	0	1	0	66	0	89	88	0	11	11	89
name	75	9	8	0	1	0	66	0	89	88	0	11	11	89
entity	73	9	8	0	1	0	64	0	89	88	0	11	11	89
position	75	9	8	0	1	0	66	0	89	88	0	11	11	89
subtotals	223	27	24	0	3	0	196	0	89	88	0	11	11	89
<ownership>	31	0	0	0	0	0	31	0	100	100	*	*	0	*
time	12	0	0	0	0	0	12	0	100	100	*	*	0	*
owned	30	0	0	0	0	0	30	0	100	100	*	*	0	*
capital	9	0	0	0	0	0	9	0	100	100	*	*	0	*
owner%	76	0	0	0	0	0	76	0	100	100	*	*	0	*
subtotals	127	0	0	0	0	0	127	0	100	100	*	*	0	*
<revenue>	16	0	0	0	0	0	16	0	100	100	*	*	0	*
type	16	0	0	0	0	0	16	0	100	100	*	*	0	*
total	10	0	0	0	0	0	10	0	100	100	*	*	0	*
total-eta	10	0	0	0	0	0	10	0	100	100	*	*	0	*
rate	6	0	0	0	0	0	6	0	100	100	*	*	0	*
rate-eta	15	0	0	0	0	0	15	0	100	100	*	*	0	*
subtotals	57	0	0	0	0	0	57	0	100	100	*	*	0	*
<time>	107	0	0	0	0	0	107	0	100	100	*	*	0	*
after	101	0	0	0	0	0	101	0	100	100	*	*	0	*
before	13	0	0	0	0	0	13	0	100	100	*	*	0	*
during	1	0	0	0	0	0	1	0	100	100	*	*	0	*
on	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	115	0	0	0	0	0	115	0	100	100	*	*	0	*

272

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	6490	3321	2071	27	46	894	4063	1502	72	63	27	14	32	63
MATCHED-ONLY	3146	2565	2071	27	46	272	853	1502	39	27	11	9	66	81
TEXT-FILTERING	145	133	129	-	-	4	16	1	-	11	3	-	89	97

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	5299.5	6466	6633	0.799	0.8196
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	5299.5	1771	2.9924		
	P&R	2P&R	P&2R		
F-MEASURES	42.49	52.71	35.59		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	205	246	158	0	8	80	39	1	44	19	32	5	77	64
subtotals	205	246	158	0	8	80	39	1	44	19	32	5	77	64
<tie-up-relati	205	246	158	0	8	80	39	0	44	19	32	5	77	64
status	205	246	151	0	15	80	39	0	47	19	32	9	74	61
entity	440	370	237	0	21	112	182	0	57	41	30	8	54	64
joint-venture	23	21	5	0	1	15	17	137	87	74	71	17	22	24
subtotals	668	637	393	0	37	207	238	137	55	36	32	9	59	62
<entity>	423	316	265	0	20	31	138	0	42	33	10	7	63	84
name	405	295	199	23	45	28	138	3	51	34	9	21	52	71
aliases	174	118	68	3	9	38	94	127	67	54	32	13	40	59
location	118	34	0	0	26	8	92	207	100	78	24	100	0	0
nationality	125	22	14	0	1	7	110	195	89	88	32	7	11	64
type	423	316	262	0	23	31	138	0	42	33	10	8	62	83
entity-rel	495	388	254	0	31	103	210	0	58	42	26	11	51	65
subtotals	1740	1173	797	26	135	215	782	532	58	45	18	15	46	69
<entity-relati	219	246	161	0	8	77	50	0	46	23	31	5	74	65
entity1	454	370	242	0	18	110	194	0	57	43	30	7	53	65
entity2	37	21	6	0	3	12	28	138	88	76	57	33	16	28
rel-ent2-to-e	219	246	144	0	25	77	50	0	51	23	31	15	66	58
status	218	246	152	0	16	78	50	0	49	23	32	10	70	62
subtotals	928	883	544	0	62	277	322	138	55	35	31	10	59	62
ALL-OBJECTS	3541	2939	1892	26	44	779	1381	808	56	39	26	12	54	65
MATCHED-ONLY	2457	2299	1892	26	44	228	386	808	29	16	10	8	78	83
TEXT-FILTERING	145	133	129	-	-	4	16	1	-	11	3	-	89	97

	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	2415	1771	1.3636		
	P&R	2P&R	P&2R		
F-MEASURES	58.8	62.27	55.69		

273

JJV Summary Scores
GE/CMU Shogun System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-nr	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-date	147	147	147	0	0	0	0	3	0	0	0	0	100	100
doc-source	150	150	119	31	0	0	0	0	10	0	0	10	90	90
content	204	197	170	0	3	24	31	2	25	15	12	2	83	86
subtotals	204	197	170	0	3	24	31	2	25	15	12	2	83	86
<tie-up-relati	204	197	170	0	3	24	31	0	25	15	12	2	83	86
status	204	197	164	0	9	24	31	0	28	15	12	5	80	83
entity	438	458	325	0	24	109	89	0	40	20	24	7	74	71
joint-venture	22	22	9	0	2	11	11	146	73	50	50	18	41	41
ownership	32	16	10	0	1	5	21	141	73	66	31	9	31	62
activity	236	206	124	0	11	71	101	13	60	43	34	8	52	60
subtotals	932	899	632	0	47	220	253	300	45	27	24	7	68	70
<entity>	423	433	357	0	10	66	56	0	27	13	15	3	84	82
name	405	417	296	22	37	62	50	2	34	12	15	14	76	74
aliases	177	226	126	9	6	85	36	171	50	20	38	7	74	58
location	118	101	65	7	3	26	43	240	52	36	26	9	58	68
nationality	126	61	50	0	1	10	75	243	63	60	16	2	40	82
type	423	433	351	0	16	66	56	0	28	13	15	4	83	81
entity-rel	494	502	351	0	32	119	111	0	43	22	24	8	71	70
person	75	71	50	0	0	21	25	285	48	33	30	0	67	70
facility	108	93	23	0	10	60	75	258	86	69	64	30	21	25
subtotals	1926	1904	1312	38	105	449	471	1199	44	24	24	8	69	70
<entity-relati	219	208	174	0	5	29	40	0	30	18	14	3	79	84
entity1	454	469	331	0	24	114	99	0	42	22	24	7	73	70
entity2	37	33	14	0	1	18	22	147	74	59	54	7	38	42
rel-ent2-to-e	219	208	162	0	17	29	40	0	35	18	14	9	74	78
status	218	208	167	0	11	30	40	0	33	18	14	6	77	80
subtotals	928	918	674	0	53	191	201	147	40	22	21	7	73	73
<activity>	234	206	146	0	3	57	85	0	50	36	28	2	62	71
industry	274	208	131	0	18	59	125	0	61	46	28	12	48	63
site	615	536	217	0	29	290	369	154	76	60	54	12	35	40
start	84	49	12	0	8	29	64	80	89	76	59	40	14	24
end	6	0	0	0	0	0	6	144	100	100	*	*	0	*
revenue	18	0	0	0	0	0	18	140	100	100	*	*	0	*
subtotals	997	793	360	0	55	378	582	518	74	58	48	13	36	45
<industry>	255	208	160	0	2	46	93	0	47	36	22	1	63	77
type	253	208	135	0	26	47	92	0	55	36	22	16	53	65
prod-serv	568	417	140	36	120	121	272	22	77	48	29	47	28	38
subtotals	821	625	275	36	146	168	364	22	70	44	27	36	36	47
<facility>	99	76	28	0	8	40	63	0	80	64	53	22	28	37
name	11	16	3	3	1	9	4	19	78	36	56	36	41	28
location	51	35	6	2	10	17	33	6	90	65	48	61	14	20
type	98	76	28	0	8	40	62	0	80	63	53	22	28	37
subtotals	160	127	37	5	19	66	99	25	82	62	52	35	25	31
<person>	75	70	55	0	0	15	20	0	39	27	21	0	73	78
name	75	70	47	7	1	15	20	0	44	27	21	8	67	72
entity	73	70	50	0	5	15	18	0	43	25	21	9	68	71
position	75	70	46	0	9	15	20	0	49	27	21	16	61	66
subtotals	223	210	143	7	15	45	58	0	45	26	21	11	66	70
<ownership>	32	16	10	0	2	4	20	0	72	62	25	17	31	62
time	12	0	0	0	0	0	12	5	100	100	*	*	0	*
owned	31	4	3	0	0	1	28	0	91	90	25	0	10	75
capital	10	6	0	0	4	2	6	4	100	60	33	100	0	0
owner%	80	32	18	1	3	10	58	1	79	72	31	16	23	58
subtotals	133	42	21	1	7	13	104	10	85	78	31	26	16	51
<revenue>	16	0	0	0	0	0	16	0	100	100	*	*	0	*
type	16	0	0	0	0	0	16	0	100	100	*	*	0	*
total	10	0	0	0	0	0	10	0	100	100	*	*	0	*
total-eta	10	0	0	0	0	0	10	0	100	100	*	*	0	*
rate	6	0	0	0	0	0	6	0	100	100	*	*	0	*
rate-eta	15	0	0	0	0	0	15	0	100	100	*	*	0	*
subtotals	57	0	0	0	0	0	57	0	100	100	*	*	0	*
<time>	107	49	15	0	15	19	77	0	88	72	39	50	14	31
after	101	47	15	0	11	21	75	0	88	74	45	42	15	32
before	13	1	0	0	0	1	13	0	100	100	100	*	0	0
during	1	1	0	0	0	1	1	0	100	100	100	*	0	0
on	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	115	49	15	0	11	23	89	0	89	77	47	42	13	31

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	6496	5764	3639	87	48	1577	2309	2223	54	36	27	12	57	64
MATCHED-ONLY	4783	4676	3639	87	48	584	691	2223	31	14	12	10	77	79
TEXT-FILTERING	145	146	143	-	-	3	2	2	-	1	2	-	99	98

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	4390.5	6462	6628	0.6624	0.6794

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	4390.5	1771	2.4791

	P&R	2P&R	P&2R
F-MEASURES	60.07	62.31	58

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	204	197	170	0	3	24	31	2	25	15	12	2	83	86
subtotals	204	197	170	0	3	24	31	2	25	15	12	2	83	86
<tie-up-relati	204	197	170	0	3	24	31	0	25	15	12	2	83	86
status	204	197	164	0	9	24	31	0	28	15	12	5	80	83
entity	438	458	325	0	24	109	89	0	40	20	24	7	74	71
joint-venture	22	22	9	0	2	11	11	146	73	50	50	18	41	41
subtotals	664	677	498	0	35	144	131	146	38	20	21	6	75	74
<entity>	423	433	357	0	10	66	56	0	27	13	15	3	84	82
name	405	417	296	22	37	62	50	2	34	12	15	14	76	74
aliases	177	226	126	9	6	85	36	171	50	20	38	7	74	58
location	118	101	65	7	3	26	43	240	52	36	26	9	58	68
nationality	126	61	50	0	1	10	75	243	63	60	16	2	40	82
type	423	433	351	0	16	66	56	0	28	13	15	4	83	81
entity-rel	494	502	351	0	32	119	111	0	43	22	24	8	71	70
subtotals	1743	1740	1239	38	95	368	371	656	40	21	21	8	72	72
<entity-relati	219	208	174	0	5	29	40	0	30	18	14	3	79	84
entity1	454	469	331	0	24	114	99	0	42	22	24	7	73	70
entity2	37	33	14	0	1	18	22	147	74	59	54	7	38	42
rel-ent2-to-e	219	208	162	0	17	29	40	0	35	18	14	9	74	78
status	218	208	167	0	11	30	40	0	33	18	14	6	77	80
subtotals	928	918	674	0	53	191	201	147	40	22	21	7	73	73
ALL-OBJECTS	3539	3532	2581	38	42	727	734	951	39	21	20	7	73	74
MATCHED-ONLY	2990	3018	2581	38	42	268	240	951	20	8	9	5	87	86
TEXT-FILTERING	145	146	143	-	-	3	2	2	-	1	2	-	99	98

	Wrong	Word-count	Error-rate
ERROR-RATE-PER-WORD	1666	1771	0.9407

	P&R	2P&R	P&2R
F-MEASURES	73.54	73.58	73.5

JJV Summary Scores
GE/CMU TEXTRACT System (Optional Test Run)

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-nr	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-date	147	148	147	0	0	1	0	2	1	0	1	0	100	99
doc-source	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	205	188	173	0	1	14	31	3	21	15	7	0	84	92
subtotals	205	188	173	0	1	14	31	3	21	15	7	0	84	92
<tie-up-relati	205	188	173	0	1	14	31	0	21	15	7	0	84	92
status	205	188	165	0	9	14	31	0	25	15	7	5	80	88
entity	440	439	345	0	10	84	85	0	34	19	19	3	78	78
joint-venture	23	17	6	0	1	10	16	148	82	70	59	14	26	35
ownership	32	28	21	0	0	7	11	139	46	34	25	0	66	75
activity	235	191	150	0	6	35	79	0	44	34	18	4	64	78
subtotals	935	863	687	0	26	150	222	287	37	24	17	4	73	80
<entity>	423	410	352	0	8	50	63	0	26	15	12	2	83	86
name	405	398	317	18	18	45	52	2	28	13	11	8	80	82
aliases	176	155	123	9	2	21	42	191	35	24	14	5	72	82
location	118	89	69	6	3	11	40	254	44	34	12	8	61	81
nationality	126	84	66	1	3	14	56	232	52	44	17	5	53	79
type	423	410	349	0	11	50	63	0	26	15	12	3	82	85
entity-rel	495	456	354	0	23	79	118	0	38	24	17	6	72	78
person	75	73	54	0	0	19	21	283	42	28	26	0	72	74
facility	106	99	22	0	15	62	69	259	87	65	63	40	21	22
subtotals	1924	1764	1354	34	75	301	461	1221	38	24	17	6	71	78
<entity-relati	219	188	171	0	2	15	46	0	27	21	8	1	78	91
entity1	454	439	345	0	8	86	101	0	36	22	20	2	76	78
entity2	37	17	6	0	0	11	31	148	88	84	65	0	16	35
rel-ent2-to-e	219	188	153	0	20	15	46	0	35	21	8	12	70	81
status	218	188	162	0	10	16	46	0	31	21	8	6	74	86
subtotals	928	832	666	0	38	128	224	148	37	24	15	5	72	80
<activity>	235	191	162	0	5	24	68	0	37	29	12	3	69	85
industry	275	206	124	0	21	61	130	2	63	47	30	14	45	60
site	616	532	264	4	49	215	299	160	68	48	40	16	43	50
start	85	43	25	0	7	11	53	91	74	62	26	22	29	58
end	6	0	0	0	0	0	6	157	100	100	*	*	0	*
revenue	18	10	5	0	4	1	9	151	74	50	10	44	28	50
subtotals	1000	791	418	4	81	288	497	561	67	50	36	16	42	53
<industry>	255	206	155	0	7	44	93	0	48	36	21	4	61	75
type	253	206	127	0	34	45	92	0	57	36	22	21	50	62
prod-serv	565	528	126	37	151	214	251	8	81	44	40	54	26	27
subtotals	818	734	253	37	185	259	343	8	75	42	35	43	33	37
<facility>	99	99	39	0	8	52	52	0	74	52	52	17	39	39
name	13	11	6	0	0	5	7	29	67	54	45	0	46	54
location	51	56	9	4	10	33	28	9	87	55	59	52	22	20
type	98	99	38	0	9	52	51	0	75	52	52	19	39	38
subtotals	162	166	53	4	19	90	86	38	78	53	54	28	34	33
<person>	75	73	60	0	0	13	15	0	32	20	18	0	80	82
name	75	73	53	5	2	13	15	0	37	20	18	8	74	76
entity	73	73	53	0	5	15	15	0	40	20	20	9	73	73
position	75	73	55	0	5	13	15	0	38	20	18	8	73	75
subtotals	223	219	161	5	12	41	45	0	38	20	19	8	73	75
<ownership>	32	28	22	0	0	6	10	0	42	31	21	0	69	78
time	12	4	0	0	2	2	10	14	100	83	50	100	0	0
owned	31	28	12	0	9	7	10	0	68	32	25	43	39	43
capital	10	6	4	0	0	2	6	16	67	60	33	0	40	67
owner%	80	56	41	0	7	8	32	0	53	40	14	14	51	73
subtotals	133	94	57	0	18	19	58	30	62	44	20	24	43	61
<revenue>	16	10	5	0	5	0	6	0	69	38	0	50	31	50
type	16	10	5	0	5	0	6	0	69	38	0	50	31	50
total	10	5	5	0	0	0	5	0	50	50	0	0	50	100
total-eta	10	3	0	0	3	0	7	0	100	70	0	100	0	0
rate	6	5	0	0	2	3	4	5	100	67	60	100	0	0
rate-eta	15	3	0	0	3	0	12	0	100	80	0	100	0	0
subtotals	57	26	10	0	13	3	34	5	83	60	12	56	18	38
<time>	108	49	29	0	13	7	66	0	75	61	14	31	27	59
after	102	46	29	0	10	7	63	0	73	62	15	26	28	63
before	13	3	0	0	0	3	13	0	100	100	100	*	0	0
during	1	0	0	0	0	0	1	0	100	100	*	*	0	*
on	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	116	49	29	0	10	10	77	0	77	66	20	26	25	59

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	6501	5726	3861	84	28	1303	2078	2301	50	32	23	12	60	68
MATCHED-ONLY	4939	4939	3861	84	28	606	606	2301	30	12	12	10	79	79
TEXT-FILTERING	145	145	143	-	-	2	2	3	-	1	1	-	99	99

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	3901	6471	6638	0.5877	0.6028
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	3901	1771	2.2027		
	P&R	2P&R	P&2R		
F-MEASURES	63.84	66.37	61.5		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	205	188	173	0	1	14	31	3	21	15	7	0	84	92
subtotals	205	188	173	0	1	14	31	3	21	15	7	0	84	92
<tie-up-relati	205	188	173	0	1	14	31	0	21	15	7	0	84	92
status	205	188	165	0	9	14	31	0	25	15	7	5	80	88
entity	440	439	345	0	10	84	85	0	34	19	19	3	78	78
joint-venture	23	17	6	0	1	10	16	148	82	70	59	14	26	35
subtotals	668	644	516	0	20	108	132	148	34	20	17	4	77	80
<entity>	423	410	352	0	8	50	63	0	26	15	12	2	83	86
name	405	398	317	18	18	45	52	2	28	13	11	8	80	82
aliases	176	155	123	9	2	21	42	191	35	24	14	5	72	82
location	118	89	69	6	3	11	40	254	44	34	12	8	61	81
nationality	126	84	66	1	3	14	56	232	52	44	17	5	53	79
type	423	410	349	0	11	50	63	0	26	15	12	3	82	85
entity-rel	495	456	354	0	23	79	118	0	38	24	17	6	72	78
subtotals	1743	1592	1278	34	60	220	371	679	34	21	14	6	74	81
<entity-relati	219	188	171	0	2	15	46	0	27	21	8	1	78	91
entity1	454	439	345	0	8	86	101	0	36	22	20	2	76	78
entity2	37	17	6	0	0	11	31	148	88	84	65	0	16	35
rel-ent2-to-e	219	188	153	0	20	15	46	0	35	21	8	12	70	81
status	218	188	162	0	10	16	46	0	31	21	8	6	74	86
subtotals	928	832	666	0	38	128	224	148	37	24	15	5	72	80
ALL-OBJECTS	3544	3256	2633	34	19	470	758	978	34	21	14	5	75	81
MATCHED-ONLY	2954	2943	2633	34	19	190	201	978	16	7	6	4	90	90
TEXT-FILTERING	145	145	143	-	-	2	2	3	-	1	1	-	99	99

	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	1364	1771	0.7702		
	P&R	2P&R	P&2R		
F-MEASURES	77.94	79.97	76.01		

277

JJV Summary Scores
NMSU/Brandeis Diderot System

1. Full-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-nr	150	150	150	0	0	0	0	0	0	0	0	0	100	100
doc-date	147	147	147	0	0	0	0	3	0	0	0	0	100	100
doc-source	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	203	219	159	0	5	55	39	3	38	19	25	3	78	73
subtotals	203	219	159	0	5	55	39	3	38	19	25	3	78	73
<tie-up-relati	203	218	159	0	5	54	39	0	38	19	25	3	78	73
status	203	218	153	0	11	54	39	0	40	19	25	7	75	70
entity	436	382	281	0	26	75	129	0	45	30	20	8	64	74
joint-venture	22	15	2	0	3	10	17	143	94	77	67	60	9	13
ownership	31	0	0	0	0	0	31	136	100	100	*	*	0	*
activity	235	134	88	0	13	33	134	10	67	57	25	13	37	66
subtotals	927	749	524	0	53	172	350	289	52	38	23	9	56	70
<entity>	423	345	297	0	5	43	121	0	36	29	12	2	70	86
name	405	341	239	23	35	44	108	0	44	27	13	16	62	73
aliases	174	73	57	6	2	8	109	161	67	63	11	8	34	82
location	118	98	50	3	7	38	58	185	67	49	39	14	44	52
nationality	125	44	35	0	2	7	88	206	73	70	16	5	28	80
type	423	345	292	0	10	43	121	0	37	29	12	3	69	85
entity-rel	494	405	290	0	23	92	181	0	50	37	23	7	59	72
person	75	43	37	0	0	6	38	243	54	51	14	0	49	86
facility	106	4	0	0	2	2	104	248	100	98	50	100	0	0
subtotals	1920	1353	1000	32	81	240	807	1043	53	42	18	9	53	75
<entity-relati	219	206	154	0	6	46	59	0	42	27	22	4	70	75
entity1	454	389	283	0	23	83	148	0	47	32	21	8	62	73
entity2	37	6	2	0	1	3	34	142	95	92	50	33	5	33
rel-ent2-to-e	219	206	144	0	16	46	59	0	46	27	22	10	66	70
status	218	206	145	0	14	47	59	0	45	27	23	9	66	70
subtotals	928	807	574	0	54	179	300	142	48	32	22	8	62	71
<activity>	235	134	116	0	2	16	117	0	54	50	12	2	49	86
industry	274	198	87	0	38	73	149	0	75	54	37	30	32	44
site	617	356	164	4	46	142	403	81	78	65	40	22	27	47
start	84	7	4	0	0	3	80	69	95	95	43	0	5	57
end	6	0	0	0	0	0	6	113	100	100	*	*	0	*
revenue	18	0	0	0	0	0	18	107	100	100	*	*	0	*
subtotals	999	561	255	4	84	218	656	370	79	66	39	25	26	46
<industry>	255	198	110	0	40	48	105	0	64	41	24	27	43	56
type	253	198	102	0	47	49	104	0	66	41	25	32	40	52
prod-serv	569	58	11	1	35	11	522	0	98	92	19	76	2	20
subtotals	822	256	113	1	82	60	626	0	87	76	23	42	14	44
<facility>	99	8	4	0	2	2	93	0	96	94	25	33	4	50
name	9	8	0	0	0	8	9	0	100	100	100	*	0	0
location	51	2	0	0	0	2	51	2	100	100	100	*	0	0
type	98	8	4	0	2	2	92	0	96	94	25	33	4	50
subtotals	158	18	4	0	2	12	152	2	98	96	67	33	2	22
<person>	75	43	40	0	0	3	35	0	49	47	7	0	53	93
name	75	38	33	0	2	3	40	0	58	53	8	6	44	87
entity	73	43	37	0	3	3	33	0	51	45	7	8	51	86
position	75	42	34	0	5	3	36	0	56	48	7	13	45	81
subtotals	223	123	104	0	10	9	109	0	55	49	7	9	47	84
<ownership>	31	0	0	0	0	0	31	0	100	100	*	*	0	*
time	12	0	0	0	0	0	12	0	100	100	*	*	0	*
owned	30	0	0	0	0	0	30	0	100	100	*	*	0	*
capital	9	0	0	0	0	0	9	0	100	100	*	*	0	*
owner%	76	0	0	0	0	0	76	0	100	100	*	*	0	*
subtotals	127	0	0	0	0	0	127	0	100	100	*	*	0	*
<revenue>	16	0	0	0	0	0	16	0	100	100	*	*	0	*
type	16	0	0	0	0	0	16	0	100	100	*	*	0	*
total	10	0	0	0	0	0	10	0	100	100	*	*	0	*
total-eta	10	0	0	0	0	0	10	0	100	100	*	*	0	*
rate	6	0	0	0	0	0	6	0	100	100	*	*	0	*
rate-eta	15	0	0	0	0	0	15	0	100	100	*	*	0	*
subtotals	57	0	0	0	0	0	57	0	100	100	*	*	0	*
<time>	107	7	4	0	2	1	101	0	96	94	14	33	4	57
after	101	5	4	0	1	0	96	0	96	95	0	20	4	80
before	13	2	0	0	0	2	13	0	100	100	100	*	0	0
during	1	0	0	0	0	0	1	0	100	100	*	*	0	*
on	0	0	0	0	0	0	0	0	*	*	*	*	*	*
subtotals	115	7	4	0	1	2	110	0	96	96	28	20	3	57

	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
ALL-OBJECTS	6479	4093	2737	37	43	947	3333	1849	63	51	23	12	42	67
MATCHED-ONLY	3968	3364	2737	37	43	315	919	1849	36	23	9	10	69	82
TEXT-FILTERING	145	130	128	-	-	2	17	3	-	12	2	-	88	98

RICHNESS-NORMALIZED-ERROR	Wrong 4670.5	Req-fills 6454	All-fills 6619	Min-err 0.7056	Max-err 0.7237
ERROR-RATE-PER-WORD	Wrong 4670.5	Word-count 1771	Error-rate 2.6372		
F-MEASURES	P&R 52.13	2P&R 60.29	P&2R 45.91		

2. Core-Template Test

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	150	150	150	0	0	0	0	0	0	0	0	0	100	100
content	203	219	159	0	5	55	39	3	38	19	25	3	78	73
subtotals	203	219	159	0	5	55	39	3	38	19	25	3	78	73
<tie-up-relati	203	218	159	0	5	54	39	0	38	19	25	3	78	73
status	203	218	153	0	11	54	39	0	40	19	25	7	75	70
entity	436	382	281	0	26	75	129	0	45	30	20	8	64	74
joint-venture	22	15	2	0	3	10	17	143	94	77	67	60	9	13
subtotals	661	615	436	0	40	139	185	143	46	28	23	8	66	71
<entity>	423	345	297	0	5	43	121	0	36	29	12	2	70	86
name	405	341	239	23	35	44	108	0	44	27	13	16	62	73
aliases	175	73	58	6	2	7	109	161	66	62	10	8	35	84
location	118	98	50	3	7	38	58	185	67	49	39	14	44	52
nationality	125	44	35	0	2	7	88	206	73	70	16	5	28	80
type	423	345	292	0	10	43	121	0	37	29	12	3	69	85
entity-rel	494	405	290	0	23	92	181	0	50	37	23	7	59	72
subtotals	1740	1306	964	32	79	231	665	552	50	38	18	9	56	75
<entity-relati	219	206	154	0	6	46	59	0	42	27	22	4	70	75
entity1	454	389	283	0	23	83	148	0	47	32	21	8	62	73
entity2	37	6	2	0	1	3	34	142	95	92	50	33	5	33
rel-ent2-to-e	219	206	144	0	16	46	59	0	46	27	22	10	66	70
status	218	206	145	0	14	47	59	0	45	27	23	9	66	70
subtotals	928	807	574	0	54	179	300	142	48	32	22	8	62	71
ALL-OBJECTS	3532	2947	2133	32	38	604	1189	840	48	34	20	8	61	73
MATCHED-ONLY	2599	2447	2133	32	38	150	302	840	22	12	6	6	83	88
TEXT-FILTERING	145	130	128	-	-	2	17	3	-	12	2	-	88	98

ERROR-RATE-PER-WORD	Wrong 1987	Word-count 1771	Error-rate 1.122	
F-MEASURES	P&R 66.34	2P&R 70.14	P&2R 62.93	

279

EME Summary Scores
BBN PLUM System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	300	300	300	0	0	0	0	0	0	0	0	0	100	100
content	609	590	436	0	7	147	166	39	42	27	25	2	72	74
subtotals	609	590	436	0	7	147	166	39	42	27	25	2	72	74
<entity>	554	494	326	0	65	103	163	0	50	29	21	17	59	66
name	499	494	285	26	63	120	125	0	52	25	24	20	60	60
location	231	59	26	11	6	16	188	216	87	81	27	27	14	53
nationality	76	55	20	2	7	26	47	270	79	62	47	28	28	38
type	554	494	325	0	66	103	163	0	50	29	21	17	59	66
subtotals	1360	1102	656	39	142	265	523	486	58	38	24	19	50	61
<microelectron	610	590	437	0	7	146	166	0	42	27	25	2	72	74
process	610	590	373	0	71	146	166	0	51	27	25	16	61	63
developer	288	107	30	0	8	69	250	271	92	87	64	21	10	28
manufacturer	239	430	119	0	57	254	63	191	76	26	59	32	50	28
distributor	233	106	48	0	18	40	167	242	82	72	38	27	21	45
purchaser	425	236	87	0	43	106	295	158	84	69	45	33	20	37
subtotals	1795	1469	657	0	197	615	941	862	73	52	42	23	37	45
<layering>	145	127	94	0	3	30	48	0	46	33	24	3	65	74
type	145	127	89	1	7	30	48	0	49	33	24	8	62	70
film	85	53	30	0	0	23	55	55	72	65	43	0	35	57
temperature	4	2	0	0	1	1	3	91	100	75	50	100	0	0
device	62	26	8	0	3	15	51	61	90	82	58	27	13	31
equipment	98	74	33	0	22	19	43	16	72	44	26	40	34	44
subtotals	394	282	160	1	33	88	200	223	67	51	31	17	41	57
<lithography>	213	220	153	0	14	53	46	0	42	22	24	8	72	70
type	213	220	122	5	40	53	46	0	53	22	24	25	58	56
granularity	45	7	0	0	1	6	44	118	100	98	86	100	0	0
device	92	99	46	0	6	47	40	80	67	43	47	12	50	46
equipment	177	225	75	0	55	95	47	12	72	26	42	42	42	33
subtotals	527	551	243	5	102	201	177	210	66	34	36	30	46	44
<etching>	60	73	42	0	0	31	18	0	54	30	42	0	70	58
type	60	73	37	0	5	31	18	0	59	30	42	12	62	51
etchant	6	3	0	0	0	3	6	37	100	100	100	*	0	0
film	27	22	6	2	0	14	19	25	83	70	64	12	26	32
granularity	11	14	1	0	2	11	8	32	95	73	78	67	9	7
temperature	1	1	0	0	0	1	1	41	100	100	100	*	0	0
device	40	26	12	0	3	11	25	20	76	62	42	20	30	46
equipment	46	43	19	0	6	18	21	9	70	46	42	24	41	44
subtotals	191	182	75	2	16	89	98	164	73	51	49	18	40	42
<packaging>	193	170	120	0	8	42	65	0	49	34	25	6	62	70
type	193	170	100	5	23	42	65	0	56	34	25	20	53	60
pitch	14	15	0	0	3	12	11	103	100	78	80	100	0	0
material	92	100	44	0	3	53	45	51	70	49	53	6	48	44
p-l-count	53	52	32	0	2	18	19	82	55	36	35	6	60	62
units	58	33	20	2	2	9	34	84	69	59	27	12	36	64
bonding	33	45	16	0	4	25	13	94	72	39	56	20	48	36
device	173	124	46	0	24	54	103	11	80	60	44	34	26	37
equipment	0	4	0	0	0	4	0	116	100	*	100	*	*	0
subtotals	616	543	258	7	61	217	290	541	69	47	40	20	42	48
<device>	194	187	105	0	19	63	70	0	59	36	34	15	54	56
function	194	187	101	3	20	63	70	0	60	36	34	17	53	55
size	112	98	48	0	5	45	59	53	69	53	46	9	43	49
device-speed	25	19	6	0	2	11	17	91	83	68	58	25	24	32
subtotals	331	304	155	3	27	119	146	144	65	44	39	15	47	51
<equipment>	350	199	160	0	16	23	174	0	57	50	12	9	46	80
name	132	56	18	5	14	19	95	79	86	72	34	44	16	37
manufacturer	268	141	58	0	50	33	160	17	81	60	23	46	22	41
modules	45	11	3	0	0	8	42	140	94	93	73	0	7	27
type	350	199	145	3	28	23	174	0	61	50	12	17	42	74
throughput	17	1	0	0	0	1	17	152	100	100	100	*	0	0
status	373	199	151	0	25	23	197	0	62	53	12	14	40	76
wafer-size	49	7	5	0	0	2	44	146	90	90	28	0	10	71
subtotals	1234	614	380	8	117	109	729	534	71	59	18	24	31	62
ALL-OBJECTS	7057	5637	3020	65	33	1850	3270	3203	66	46	33	19	43	54
MATCHED-ONLY	4686	4292	3020	65	33	739	1133	3203	44	24	17	14	65	71
TEXT-FILTERING	220	244	203	-	-	41	17	39	-	8	17	-	92	83

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	5854.5	7485	7822	0.7485	0.7822
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	5854.5	125680	0.0466		
	P&R	2P&R	P&2R		
F-MEASURES	48.09	51.55	45.07		

280

EME Summary Scores
GE/CMU Shogun System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	300	300	300	0	0	0	0	0	0	0	0	0	100	100
content	608	633	440	0	4	189	164	32	45	27	30	1	72	70
subtotals	608	633	440	0	4	189	164	32	45	27	30	1	72	70
<entity>	553	560	320	0	90	150	143	0	54	26	27	22	58	57
name	497	553	284	22	84	163	107	0	55	22	29	24	59	53
location	231	179	57	11	19	92	144	177	81	62	51	28	27	35
nationality	78	11	4	2	0	5	72	285	94	92	45	17	6	45
type	553	560	320	0	90	150	143	0	54	26	27	22	58	57
subtotals	1359	1303	665	35	193	410	466	462	61	34	31	24	50	52
<microelectron	609	633	441	0	3	189	165	0	45	27	30	1	72	70
process	609	633	372	0	72	189	165	0	53	27	30	16	61	59
developer	294	201	56	0	25	120	213	268	86	72	60	31	19	28
manufacturer	240	476	120	0	55	301	65	148	78	27	63	31	50	25
distributor	236	359	96	0	28	235	112	187	80	47	65	22	41	27
purchaser	426	461	124	0	78	259	224	131	82	52	56	39	29	27
subtotals	1805	2130	768	0	258	1104	779	734	74	43	52	25	42	36
<layering>	145	166	109	0	4	53	32	0	45	22	32	4	75	66
type	145	166	104	1	8	53	32	0	47	22	32	8	72	63
film	84	93	25	5	8	55	46	49	80	55	59	28	33	30
temperature	4	6	2	0	0	4	2	105	75	50	67	0	50	33
device	63	56	27	0	3	26	33	69	70	52	46	10	43	48
equipment	98	77	49	0	15	13	34	25	56	35	17	23	50	64
subtotals	394	398	207	6	34	151	147	248	61	37	38	15	53	53
<lithography>	212	203	147	0	7	49	58	0	44	27	24	4	69	72
type	212	203	109	4	41	49	58	0	57	27	24	28	52	55
granularity	45	80	25	2	2	51	16	97	73	36	64	10	58	32
device	92	183	50	0	11	122	31	50	77	34	67	18	54	27
equipment	176	172	101	0	24	47	51	11	55	29	27	19	57	59
subtotals	525	638	285	6	78	269	156	158	64	30	42	22	55	45
<etching>	60	58	39	0	0	19	21	0	51	35	33	0	65	67
type	60	58	36	0	3	19	21	0	54	35	33	8	60	62
etchant	6	2	1	0	0	1	5	36	86	83	50	0	17	50
film	27	28	8	2	0	18	17	24	80	63	64	10	33	32
granularity	11	16	3	0	1	12	7	33	87	64	75	25	27	19
temperature	1	2	0	0	1	1	0	37	100	0	50	100	0	0
device	40	31	14	0	1	16	25	19	75	62	52	7	35	45
equipment	45	30	22	0	3	5	20	7	56	44	17	12	49	73
subtotals	190	167	84	2	9	72	95	156	68	50	43	10	45	51
<packaging>	193	206	120	0	10	76	63	0	55	33	37	8	62	58
type	193	206	92	9	29	76	63	0	64	33	37	26	50	47
pitch	14	7	2	0	2	3	10	113	88	71	43	50	14	28
material	92	58	35	0	2	21	55	63	69	60	36	5	38	60
p-l-count	53	62	37	0	0	25	16	82	52	30	40	0	70	60
units	57	42	19	2	1	20	35	82	74	61	48	9	35	48
bonding	33	29	17	0	1	11	15	99	61	45	38	6	52	59
device	173	173	61	0	26	86	86	18	76	50	50	30	35	35
equipment	0	0	0	0	0	0	0	120	*	*	*	*	*	*
subtotals	615	577	263	11	61	242	280	577	69	46	42	20	44	46
<device>	193	240	127	0	23	90	43	0	55	22	38	15	66	53
function	193	240	122	5	23	90	43	0	56	22	38	17	64	52
size	112	96	62	0	3	31	47	66	57	42	32	5	55	64
device-speed	25	162	19	0	1	142	5	85	89	20	88	5	76	12
subtotals	330	498	203	5	27	263	95	151	65	29	53	12	62	41
<equipment>	352	297	228	0	14	55	110	0	44	31	18	6	65	77
name	133	153	48	15	14	76	56	94	73	42	50	28	42	36
manufacturer	273	286	119	0	71	96	83	6	68	30	34	37	44	42
modules	45	32	8	0	2	22	35	199	88	78	69	20	18	25
type	352	297	198	19	25	55	110	0	49	31	18	14	59	70
throughput	17	28	7	0	5	16	5	202	79	29	57	42	41	25
status	375	297	211	0	31	55	133	0	51	35	18	13	56	71
wafer-size	51	41	22	0	0	19	29	200	68	57	46	0	43	54
subtotals	1246	1134	613	34	148	339	451	701	60	36	30	21	50	56
ALL-OBJECTS	7072	7478	3528	99	26	3039	2633	3219	65	37	41	19	50	48
MATCHED-ONLY	4993	5303	3528	99	26	1136	826	3219	42	16	21	14	72	67
TEXT-FILTERING	219	257	208	-	-	49	11	32	-	5	19	-	95	81

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	6533.5	7489	7821	0.8354	0.8724
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	6533.5	125680	0.052		
	P&R	2P&R	P&2R		
F-MEASURES	49.18	48.37	50.01		

EME Summary Scores
NMSU/Brandeis Diderot System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	300	299	299	0	0	0	1	0	0	0	0	0	100	100
content	606	425	308	0	11	106	287	46	57	47	25	3	51	72
subtotals	606	425	308	0	11	106	287	46	57	47	25	3	51	72
<entity>	553	403	243	0	85	75	225	0	61	41	19	26	44	60
name	498	403	206	26	76	95	190	0	63	38	24	29	44	54
location	231	118	32	3	16	67	180	119	89	78	57	34	14	28
nationality	77	0	0	0	0	0	77	216	100	100	*	*	0	*
type	553	400	243	0	82	75	228	0	61	41	19	25	44	61
subtotals	1359	921	481	29	174	237	675	335	69	50	26	28	36	54
<microelectron	607	425	309	0	11	105	287	0	57	47	25	3	51	73
process	607	425	245	0	75	105	287	0	66	47	25	23	40	58
developer	294	156	53	0	16	87	225	152	86	76	56	23	18	34
manufacturer	237	253	49	0	45	159	143	100	88	60	63	48	21	19
distributor	235	96	34	0	30	32	171	164	87	73	33	47	14	35
purchaser	421	293	64	0	76	153	281	76	89	67	52	54	15	22
subtotals	1794	1223	445	0	242	536	1107	492	81	62	44	35	25	36
<layering>	145	103	69	0	4	30	72	0	60	50	29	5	48	67
type	145	103	61	5	7	30	72	0	64	50	29	13	44	62
film	85	31	18	0	0	13	67	36	82	79	42	0	21	58
temperature	4	0	0	0	0	0	4	68	100	100	*	*	0	*
device	63	25	12	0	4	9	47	46	83	75	36	25	19	48
equipment	98	42	31	0	4	7	63	13	70	64	17	11	32	74
subtotals	395	201	122	5	15	59	253	163	72	64	29	12	32	62
<lithography>	211	172	117	0	15	40	79	0	53	37	23	11	55	68
type	211	138	74	0	34	30	103	0	69	49	22	31	35	54
granularity	45	3	0	0	2	1	43	91	100	96	33	100	0	0
device	91	58	25	0	8	25	58	66	78	64	43	24	27	43
equipment	175	125	72	0	15	38	88	15	66	50	30	17	41	58
subtotals	522	324	171	0	59	94	292	172	72	56	29	26	33	53
<etching>	60	43	27	0	2	14	31	0	64	52	32	7	45	63
type	60	26	15	0	1	10	44	0	78	73	38	6	25	58
etchant	6	0	0	0	0	0	6	24	100	100	*	*	0	*
film	27	10	4	1	1	4	21	18	85	78	40	25	17	45
granularity	11	1	0	0	0	1	11	22	100	100	100	*	0	0
temperature	1	0	0	0	0	0	1	27	100	100	*	*	0	*
device	40	11	6	0	2	3	32	17	86	80	27	25	15	54
equipment	45	23	13	0	2	8	30	7	75	67	35	13	29	56
subtotals	190	71	38	1	6	26	145	115	82	76	37	14	20	54
<packaging>	191	88	57	0	8	23	126	0	73	66	26	12	30	65
type	191	51	26	5	11	9	149	0	86	78	18	32	15	56
pitch	14	0	0	0	0	0	14	53	100	100	*	*	0	*
material	92	44	25	0	1	18	66	26	77	72	41	4	27	57
p-1-count	52	24	18	0	1	5	33	36	68	63	21	5	35	75
units	57	0	0	0	0	0	57	43	100	100	*	*	0	*
bonding	34	0	0	0	0	0	34	48	100	100	*	*	0	*
device	172	93	35	0	12	46	125	9	84	73	49	26	20	38
equipment	1	8	1	0	0	7	0	52	88	0	88	0	100	12
subtotals	613	220	105	5	25	85	478	267	84	78	39	20	18	49
<device>	194	184	97	0	16	71	81	0	63	42	38	14	50	53
function	194	184	91	6	16	71	81	0	64	42	38	17	48	51
size	114	58	29	0	3	26	82	49	79	72	45	9	25	50
device-speed	25	7	2	0	1	4	22	91	93	88	57	33	8	28
subtotals	333	249	122	6	20	101	185	140	71	56	40	16	38	50
<equipment>	349	198	153	0	9	36	187	0	60	54	18	6	44	77
name	132	49	23	3	4	19	102	89	84	77	39	18	18	50
manufacturer	267	153	52	0	40	61	175	15	84	66	40	43	19	34
modules	44	0	0	0	0	0	44	143	100	100	*	*	0	*
type	349	198	128	19	15	36	187	0	64	54	18	15	39	69
throughput	17	0	0	0	0	0	17	144	100	100	*	*	0	*
status	372	198	140	0	22	36	210	0	66	56	18	14	38	71
wafer-size	49	1	1	0	0	0	48	138	98	98	0	0	2	100
subtotals	1230	599	344	22	81	152	783	529	74	64	25	20	29	59
ALL-OBJECTS	7042	4233	2136	68	17	1396	4205	2259	74	60	33	24	31	51
MATCHED-ONLY	3630	3142	2136	68	17	537	1025	2259	48	28	17	17	60	69
TEXT-FILTERING	220	206	172	-	-	34	48	46	-	22	16	-	78	83

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	6268	7469	7802	0.8034	0.8392
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	6268	125680	0.0499		
	P&R	2P&R	P&2R		
F-MEASURES	38.49	45.26	33.49		

EME Summary Scores
UMass/Hughes CIRCUS System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
\<template\>	300	300	300	0	0	0	0	0	0	0	0	0	100	100
content	605	318	225	0	4	89	376	24	68	62	28	2	37	71
subtotals	605	318	225	0	4	89	376	24	68	62	28	2	37	71
\<entity\>	549	647	256	0	129	262	164	0	68	30	40	34	47	40
name	495	630	203	38	119	270	135	2	71	27	43	38	45	35
location	231	52	18	3	8	23	202	157	92	87	44	33	8	38
nationality	78	34	12	0	1	21	65	225	88	83	62	8	15	35
type	549	631	251	0	128	252	170	0	69	31	40	34	46	40
subtotals	1353	1347	484	41	256	566	572	384	74	42	42	35	37	37
\<microelectron	605	318	225	0	4	89	376	0	68	62	28	2	37	71
process	605	347	190	0	39	118	376	0	74	62	34	17	31	55
developer	289	144	29	0	15	100	245	124	92	85	69	34	10	20
manufacturer	235	363	56	0	31	276	148	89	89	63	76	36	24	15
distributor	232	347	52	0	29	266	151	86	90	65	77	36	22	15
purchaser	418	226	69	0	20	137	329	77	88	79	61	22	16	30
subtotals	1779	1427	396	0	134	897	1249	376	85	70	63	25	22	28
\<layering\>	145	149	93	0	5	51	47	0	52	32	34	5	64	62
type	145	149	85	5	8	51	47	0	55	32	34	11	60	59
film	84	51	24	0	7	20	53	49	77	63	39	22	28	47
temperature	4	0	0	0	0	0	4	90	100	100	*	*	0	*
device	63	67	19	0	6	42	38	54	82	60	63	24	30	28
equipment	98	30	20	0	5	5	73	29	80	74	17	20	20	67
subtotals	394	297	148	5	26	118	215	222	71	54	40	16	38	51
\<lithography\>	212	134	93	0	7	34	112	0	62	53	25	7	44	69
type	212	132	64	8	26	34	114	0	72	54	26	31	32	52
granularity	45	13	2	1	6	4	36	66	95	80	31	72	6	19
device	90	113	44	0	3	66	43	30	72	48	58	6	49	39
equipment	175	129	50	0	27	52	98	6	78	56	40	35	28	39
subtotals	522	387	160	9	62	156	291	102	76	56	40	29	32	42
\<etching\>	60	63	24	0	9	30	27	0	73	45	48	27	40	38
type	60	54	14	0	15	25	31	0	84	52	46	52	23	26
etchant	6	0	0	0	0	0	6	19	100	100	*	*	0	*
film	27	32	8	0	5	19	14	12	83	52	59	38	30	25
granularity	11	0	0	0	0	0	11	17	100	100	*	*	0	*
temperature	1	0	0	0	0	0	1	24	100	100	*	*	0	*
device	41	29	14	0	1	14	26	9	74	63	48	7	34	48
equipment	45	12	7	0	2	3	36	4	85	80	25	22	16	58
subtotals	191	127	43	0	23	61	125	85	83	65	48	35	22	34
\<packaging\>	192	197	120	0	11	66	61	0	53	32	34	8	62	61
type	192	197	92	1	38	66	61	0	64	32	34	29	48	47
pitch	14	0	0	0	0	0	14	114	100	100	*	*	0	*
material	91	37	26	0	1	10	64	62	74	70	27	4	28	70
p-l-count	51	0	0	0	0	0	51	88	100	100	*	*	0	*
units	57	0	0	0	0	0	57	93	100	100	*	*	0	*
bonding	34	0	0	0	0	0	34	105	100	100	*	*	0	*
device	172	249	89	0	17	143	66	16	72	38	57	16	52	36
equipment	0	1	0	0	0	1	0	119	100	*	100	*	*	0
subtotals	611	484	207	1	56	220	347	597	75	57	45	21	34	43
\<device\>	192	215	119	0	19	77	54	0	56	28	36	14	62	55
function	192	198	118	1	12	67	61	0	54	32	34	10	62	60
size	114	70	36	0	7	27	71	44	74	62	38	16	32	51
device-speed	25	11	2	0	0	9	23	105	94	92	82	0	8	18
subtotals	331	279	156	1	19	103	155	149	64	47	37	11	47	56
\<equipment\>	348	175	136	0	10	29	202	0	64	58	16	7	39	78
name	132	80	17	8	16	39	91	52	88	69	49	49	16	26
manufacturer	269	484	80	0	44	360	145	1	87	54	74	35	30	16
modules	44	0	0	0	0	0	44	126	100	100	*	*	0	*
type	348	163	111	7	17	28	213	0	70	61	17	15	33	70
throughput	17	0	0	0	0	0	17	129	100	100	*	*	0	*
status	371	175	129	0	17	29	225	0	68	61	16	12	35	74
wafer-size	49	0	0	0	0	0	49	125	100	100	*	*	0	*
subtotals	1230	902	337	15	94	456	784	433	80	64	50	23	28	38
ALL-OBJECTS	7016	5568	2156	72	22	2666	4114	2372	77	59	48	24	31	39
MATCHED-ONLY	3663	3722	2156	72	22	1140	1081	2372	54	30	31	15	60	59
TEXT-FILTERING	219	243	186	-	-	57	33	24	-	15	23	-	85	76

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	7490	7449	7779	0.9628	1.0055
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	7490	125680	0.0596		
	P&R	2P&R	P&2R		
F-MEASURES	34.84	37.42	32.59		

JME Summary Scores
BBN PLUM System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	100	100	100	0	0	0	0	0	0	0	0	0	100	100
content	123	109	77	0	0	32	46	23	50	37	29	0	63	71
subtotals	123	109	77	0	0	32	46	23	50	37	29	0	63	71
<entity>	121	111	71	0	8	32	42	0	54	35	29	10	59	64
name	120	111	64	2	13	32	41	0	57	34	29	18	54	58
location	58	0	0	0	0	0	58	36	100	100	*	*	0	*
nationality	36	0	0	0	0	0	36	50	100	100	*	*	0	*
type	121	111	71	0	8	32	42	0	54	35	29	10	59	64
subtotals	335	222	135	2	21	64	177	86	66	53	29	14	40	61
<microelectron	124	109	78	0	0	31	46	0	50	37	28	0	63	72
process	124	109	68	0	10	31	46	0	56	37	28	13	55	62
developer	63	56	16	0	7	33	40	44	83	63	59	30	25	28
manufacturer	85	169	42	0	9	118	34	10	79	40	70	18	49	25
distributor	80	83	31	0	6	46	43	26	75	54	55	16	39	37
purchaser	25	11	6	0	0	5	19	61	80	76	45	0	24	54
subtotals	377	428	163	0	32	233	182	141	73	48	54	16	43	38
<layering>	44	30	23	0	1	6	20	0	54	45	20	4	52	77
type	44	30	21	2	1	6	20	0	56	45	20	8	50	73
film	13	0	0	0	0	0	13	20	100	100	*	*	0	*
temperature	5	0	0	0	0	0	5	21	100	100	*	*	0	*
device	13	2	1	0	0	1	12	15	93	92	50	0	8	50
equipment	39	34	10	0	6	18	23	1	82	59	53	38	26	29
subtotals	114	66	32	2	7	25	73	57	76	64	38	20	29	50
<lithography>	51	53	36	0	3	14	12	0	45	24	26	8	70	68
type	51	53	33	0	6	14	12	0	49	24	26	15	65	62
granularity	23	0	0	0	0	0	23	20	100	100	*	*	0	*
device	35	7	4	0	1	2	30	12	89	86	28	20	11	57
equipment	51	74	30	0	9	35	12	0	65	24	47	23	59	40
subtotals	160	134	67	0	16	51	77	32	68	48	38	19	42	50
<etching>	17	11	5	0	2	4	10	0	76	59	36	28	29	45
type	17	11	3	0	4	4	10	0	86	59	36	57	18	27
etchant	0	0	0	0	0	0	0	5	*	*	*	*	*	*
film	4	0	0	0	0	0	4	4	100	100	*	*	0	*
granularity	2	0	0	0	0	0	2	4	100	100	*	*	0	*
temperature	0	0	0	0	0	0	0	5	*	*	*	*	*	*
device	6	3	2	0	0	1	4	2	71	67	33	0	33	67
equipment	10	14	4	0	1	9	5	0	79	50	64	20	40	28
subtotals	39	28	9	0	5	14	25	20	83	64	50	36	23	32
<packaging>	12	15	8	0	0	7	4	0	58	33	47	0	67	53
type	12	15	8	0	0	7	4	0	58	33	47	0	67	53
pitch	0	0	0	0	0	0	0	8	*	*	*	*	*	*
material	4	13	3	0	0	10	1	4	78	25	77	0	75	23
p-l-count	6	5	5	0	0	0	1	3	17	17	0	0	83	100
units	0	0	0	0	0	0	0	8	*	*	*	*	*	*
bonding	0	0	0	0	0	0	0	8	*	*	*	*	*	*
device	14	15	5	0	3	7	6	0	76	43	47	38	36	33
equipment	0	0	0	0	0	0	0	8	*	*	*	*	*	*
subtotals	36	48	21	0	3	24	12	39	65	33	50	12	58	44
<device>	49	19	12	0	3	4	34	0	77	69	21	20	24	63
function	49	19	12	0	3	4	34	0	77	69	21	20	24	63
size	34	0	0	0	0	0	34	5	100	100	*	*	0	*
device-speed	4	0	0	0	0	0	4	10	100	100	*	*	0	*
subtotals	87	19	12	0	3	4	72	15	87	83	21	20	14	63
<equipment>	110	78	55	0	4	19	51	0	57	46	24	7	50	70
name	27	10	6	1	1	2	19	38	78	70	20	19	24	65
manufacturer	91	89	31	0	10	48	50	1	78	55	54	24	34	35
modules	15	0	0	0	0	0	15	51	100	100	*	*	0	*
type	110	78	53	0	6	19	51	0	59	46	24	10	48	68
throughput	3	0	0	0	0	0	3	55	100	100	*	*	0	*
status	110	78	49	0	10	19	51	0	62	46	24	17	44	63
wafer-size	12	0	0	0	0	0	12	49	100	100	*	*	0	*
subtotals	368	255	139	1	27	88	201	194	69	55	34	16	38	55
ALL-OBJECTS	1639	1309	655	5	6	535	865	607	70	53	41	15	40	50
MATCHED-ONLY	987	955	655	5	6	220	252	607	46	26	23	10	67	69
TEXT-FILTERING	68	66	57	-	-	9	11	23	-	16	14	-	84	86

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	1516.5	1664	1749	0.8671	0.9114
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	1516.5	1784	0.8501		
	P&R	2P&R	P&2R		
F-MEASURES	44.61	47.82	41.8		

284

JME Summary Scores
GE/CMU Shogun System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	100	100	100	0	0	0	0	0	0	0	0	0	100	100
content	123	134	94	0	2	38	27	21	42	22	28	2	76	70
subtotals	123	134	94	0	2	38	27	21	42	22	28	2	76	70
<entity>	121	131	91	0	9	31	21	0	40	17	24	9	75	69
name	121	131	77	3	20	31	21	0	48	17	24	22	65	60
location	58	47	25	4	4	14	25	42	62	43	30	18	46	57
nationality	36	19	14	0	4	1	18	61	62	50	5	22	39	74
type	121	131	91	0	9	31	21	0	40	17	24	9	75	69
subtotals	336	328	207	7	37	77	85	103	49	25	23	16	63	64
<microelectron	124	134	94	0	2	38	28	0	42	22	28	2	76	70
process	124	134	84	0	12	38	28	0	48	22	28	12	68	63
developer	63	91	23	0	9	59	31	43	81	49	65	28	36	25
manufacturer	83	141	43	0	15	83	25	12	74	30	59	26	52	30
distributor	80	138	45	0	9	84	26	20	72	32	61	17	56	33
purchaser	25	36	13	0	1	22	11	62	72	44	61	7	52	36
subtotals	375	540	208	0	46	286	121	137	68	32	53	18	55	38
<layering>	44	57	36	0	1	20	7	0	44	16	35	3	82	63
type	44	57	32	2	3	20	7	0	48	16	35	11	75	58
film	13	2	0	0	1	1	12	28	100	92	50	100	0	0
temperature	5	5	2	0	0	3	3	31	75	60	60	0	40	40
device	13	9	6	0	0	3	7	24	62	54	33	0	46	67
equipment	39	57	20	0	13	24	6	0	68	15	42	39	51	35
subtotals	114	130	60	2	17	51	35	83	63	31	39	23	54	47
<lithography>	51	47	35	0	1	11	15	0	44	29	23	3	69	74
type	51	47	30	0	6	11	15	0	52	29	23	17	59	64
granularity	24	36	8	5	2	21	9	15	77	38	58	30	44	29
device	35	39	20	0	2	17	13	7	62	37	44	9	57	51
equipment	51	43	32	0	2	9	17	0	47	33	21	6	63	74
subtotals	161	165	90	5	12	58	54	22	58	34	35	14	57	56
<etching>	17	15	9	0	1	5	7	0	59	41	33	10	53	60
type	17	15	6	0	4	5	7	0	73	41	33	40	35	40
etchant	0	0	0	0	0	0	0	9	*	*	*	*	*	*
film	4	1	1	0	0	0	3	8	75	75	0	0	25	100
granularity	2	7	0	2	0	5	0	5	86	0	71	50	50	14
temperature	0	0	0	0	0	0	0	9	*	*	*	*	*	*
device	6	3	2	0	0	1	4	4	71	67	33	0	33	67
equipment	10	11	7	0	1	3	2	0	46	20	27	12	70	64
subtotals	39	37	16	2	5	14	16	35	68	41	38	26	44	46
<packaging>	12	15	10	0	0	5	2	0	41	17	33	0	83	67
type	12	15	10	0	0	5	2	0	41	17	33	0	83	67
pitch	0	3	0	0	0	3	0	10	100	*	100	*	*	0
material	4	6	3	0	0	3	1	7	57	25	50	0	75	50
p-l-count	5	5	5	0	0	0	0	5	0	0	0	0	100	100
units	0	0	0	0	0	0	0	10	*	*	*	*	*	*
bonding	0	0	0	0	0	0	0	10	*	*	*	*	*	*
device	14	11	7	0	0	4	7	0	61	50	36	0	50	64
equipment	0	0	0	0	0	0	0	10	*	*	*	*	*	*
subtotals	35	40	25	0	0	15	10	52	50	28	38	0	71	62
<device>	49	56	37	0	1	18	11	0	45	22	32	3	76	66
function	49	56	34	3	1	18	11	0	47	22	32	6	72	63
size	33	31	22	0	2	7	9	11	45	27	22	8	67	71
device-speed	4	11	3	0	1	7	0	31	73	0	64	25	75	27
subtotals	86	98	59	3	4	32	20	42	49	23	33	8	70	62
<equipment>	110	115	79	0	2	34	29	0	45	26	30	2	72	69
name	27	21	16	0	0	5	11	54	50	41	24	0	59	76
manufacturer	89	112	44	0	20	48	25	0	68	28	43	31	49	39
modules	15	4	3	0	0	1	12	74	81	80	25	0	20	75
type	110	115	74	0	7	34	29	0	49	26	30	9	67	64
throughput	3	8	1	0	0	7	2	73	90	67	88	0	33	12
status	110	115	74	0	7	34	29	0	49	26	30	9	67	64
wafer-size	12	4	1	0	0	3	11	68	93	92	75	0	8	25
subtotals	366	379	213	0	34	132	119	269	57	32	35	14	58	56
ALL-OBJECTS	1635	1851	972	19	17	703	487	764	58	30	38	14	60	53
MATCHED-ONLY	1261	1345	972	19	17	231	147	764	34	12	17	12	78	73
TEXT-FILTERING	68	77	66	-	-	11	2	21	-	3	14	-	97	86

RICHNESS-NORMALIZED-ERROR	Wrong 1356.5	Req-fills 1664	All-fills 1749	Min-err 0.7756	Max-err 0.8152
ERROR-RATE-PER-WORD	Wrong 1356.5	Word-count 1784	Error-rate 0.7604		
F-MEASURES	P&R 56.31	2P&R 54.29	P&2R 58.49		

285

JME Summary Scores
GE/CMU TEXTRACT System (Optional Test Run)

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	100	100	100	0	0	0	0	0	0	0	0	0	100	100
content	122	85	70	0	0	15	52	22	49	43	18	0	57	82
subtotals	122	85	70	0	0	15	52	22	49	43	18	0	57	82
<entity>	117	112	81	0	6	25	30	0	43	26	22	7	69	72
name	116	112	70	4	13	25	29	0	49	25	22	17	62	64
location	56	39	21	5	4	9	26	40	64	46	23	22	42	60
nationality	36	31	20	0	4	7	12	57	53	33	22	17	56	64
type	117	112	81	0	6	25	30	0	43	26	22	7	69	72
subtotals	325	294	192	9	27	66	97	97	50	30	22	14	60	67
<microelectron	123	85	70	0	0	15	53	0	49	43	18	0	57	82
process	123	85	68	0	2	15	53	0	51	43	18	3	55	80
developer	62	52	18	0	3	31	41	37	81	66	60	14	29	35
manufacturer	83	101	44	0	7	50	32	11	67	38	50	14	53	44
distributor	80	77	42	0	1	34	37	20	63	46	44	2	52	54
purchaser	23	42	10	0	3	29	10	46	81	43	69	23	43	24
subtotals	371	357	182	0	16	159	173	114	66	47	44	8	49	51
<layering>	44	30	24	0	1	5	19	0	51	43	17	4	54	80
type	44	30	23	0	2	5	19	0	53	43	17	8	52	77
film	13	10	0	0	4	6	9	15	100	69	60	100	0	0
temperature	5	6	0	0	2	4	3	19	100	60	67	100	0	0
device	13	11	4	0	3	4	6	13	76	46	36	43	31	36
equipment	39	24	17	0	3	4	19	3	60	49	17	15	44	71
subtotals	114	81	44	0	14	23	56	50	68	49	28	24	38	54
<lithography>	51	34	27	0	0	7	24	0	53	47	20	0	53	79
type	51	34	22	0	5	7	24	0	62	47	20	18	43	65
granularity	24	14	10	0	2	2	12	13	62	50	14	17	42	71
device	35	23	15	0	4	4	16	5	62	46	17	21	43	65
equipment	51	33	25	0	2	6	24	0	56	47	18	7	49	76
subtotals	161	104	72	0	13	19	76	18	60	47	18	15	45	69
<etching>	16	10	9	0	0	1	7	0	47	44	10	0	56	90
type	16	10	7	1	1	1	7	0	56	44	10	17	47	75
etchant	0	2	0	0	0	2	0	7	100	*	100	*	*	0
film	4	0	0	0	0	0	4	7	100	100	*	*	0	*
granularity	2	1	1	0	0	0	1	8	50	50	0	0	50	100
temperature	0	0	0	0	0	0	0	9	*	*	*	*	*	*
device	5	3	2	0	0	1	3	5	67	60	33	0	40	67
equipment	9	6	6	0	0	0	3	2	33	33	0	0	67	100
subtotals	36	22	16	1	1	4	18	38	59	50	18	8	46	75
<packaging>	12	11	8	0	0	3	4	0	47	33	27	0	67	73
type	12	11	7	0	1	3	4	0	53	33	27	12	58	64
pitch	0	0	0	0	0	0	0	8	*	*	*	*	*	*
material	4	6	3	0	0	3	1	5	57	25	50	0	75	50
p-l-count	7	4	4	0	0	0	3	3	43	43	0	0	57	100
units	0	0	0	0	0	0	0	8	*	*	*	*	*	*
bonding	0	0	0	0	0	0	0	8	*	*	*	*	*	*
device	14	5	5	0	0	0	9	0	64	64	0	0	36	100
equipment	0	0	0	0	0	0	0	8	*	*	*	*	*	*
subtotals	37	26	19	0	1	6	17	40	56	46	23	5	51	73
<device>	48	39	27	0	7	5	14	0	49	29	13	20	56	69
function	48	39	24	3	7	5	14	0	52	29	13	25	53	65
size	34	29	22	0	3	4	9	2	42	26	14	12	65	76
device-speed	4	4	3	0	1	0	0	23	25	0	0	25	75	75
subtotals	86	72	49	3	11	9	23	25	47	27	12	20	59	70
<equipment>	109	63	53	0	2	8	54	0	55	50	13	4	49	84
name	27	17	15	0	0	2	12	33	48	44	12	0	56	88
manufacturer	89	91	43	0	7	41	39	1	67	44	45	14	48	47
modules	15	0	0	0	0	0	15	49	100	100	*	*	0	*
type	109	63	52	0	3	8	54	0	56	50	13	5	48	82
throughput	3	3	2	0	0	1	1	50	50	33	33	0	67	67
status	109	63	49	0	6	8	54	0	58	50	13	11	45	78
wafer-size	12	11	7	0	2	2	3	43	50	25	18	22	58	64
subtotals	364	248	168	0	18	62	178	176	60	49	25	10	46	68
ALL-OBJECTS	1616	1289	812	13	10	363	690	580	59	43	28	12	51	63
MATCHED-ONLY	1015	1058	812	13	10	160	117	580	30	12	15	9	81	77
TEXT-FILTERING	67	75	64	-	-	11	3	22	-	4	15	-	96	85

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	1160.5	1641	1725	0.6728	0.7072
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	1160.5	1784	0.6505		
	P&R	2P&R	P&2R		
F-MEASURES	56.35	60.43	52.79		

JME Summary Scores
NMSU/Brandeis Diderot System

SLOT	POS	ACT	COR	PAR	INC	SPU	MIS	NON	ERR	UND	OVG	SUB	REC	PRE
<template>	100	100	100	0	0	0	0	0	0	0	0	0	100	100
content	122	92	69	0	3	20	50	27	51	41	22	4	56	75
subtotals	122	92	69	0	3	20	50	27	51	41	22	4	56	75
<entity>	117	91	71	0	6	14	40	0	46	34	15	8	61	78
name	116	91	61	4	12	14	39	0	52	34	15	18	54	69
location	56	34	20	1	4	9	31	30	68	55	26	18	37	60
nationality	36	8	4	0	0	4	32	48	90	89	50	0	11	50
type	117	91	71	0	6	14	40	0	46	34	15	8	61	78
subtotals	325	224	156	5	22	41	142	78	57	44	18	13	49	71
<microelectron	123	92	70	0	2	20	51	0	51	41	22	3	57	76
process	123	92	65	0	7	20	51	0	54	41	22	10	53	71
developer	62	29	11	0	1	17	50	38	86	81	59	8	18	38
manufacturer	81	83	34	0	5	44	42	11	73	52	53	13	42	41
distributor	79	28	19	0	1	8	59	25	78	75	28	5	24	68
purchaser	23	19	11	0	1	7	11	50	63	48	37	8	48	58
subtotals	368	251	140	0	15	96	213	124	70	58	38	10	38	56
<layering>	44	28	23	0	1	4	20	0	52	45	14	4	52	82
type	44	28	22	1	1	4	20	0	53	45	14	6	51	80
film	13	1	0	0	0	1	13	17	100	100	100	*	0	0
temperature	5	0	0	0	0	0	5	21	100	100	*	*	0	*
device	13	1	1	0	0	0	12	16	92	92	0	0	8	100
equipment	39	21	14	0	3	4	22	1	67	56	19	18	36	67
subtotals	114	51	37	1	4	9	72	55	70	63	18	11	33	74
<lithography>	51	37	29	0	0	8	22	0	51	43	22	0	57	78
type	51	37	23	0	6	8	22	0	61	43	22	21	45	62
granularity	23	0	0	0	0	0	23	18	100	100	*	*	0	*
device	35	9	4	0	2	3	29	9	89	83	33	33	11	44
equipment	51	40	27	0	1	12	23	0	57	45	30	4	53	68
subtotals	160	86	54	0	9	23	97	27	70	61	27	14	34	63
<etching>	16	17	8	0	0	9	8	0	68	50	53	0	50	47
type	16	17	4	1	3	9	8	0	82	50	53	44	28	26
etchant	0	0	0	0	0	0	0	8	*	*	*	*	*	*
film	4	2	1	0	0	1	3	7	80	75	50	0	25	50
granularity	2	0	0	0	0	0	2	6	100	100	*	*	0	*
temperature	0	0	0	0	0	0	0	8	*	*	*	*	*	*
device	5	1	1	0	0	0	4	4	80	80	0	0	20	100
equipment	9	8	5	0	0	3	4	2	58	44	38	0	56	62
subtotals	36	28	11	1	3	13	21	35	76	58	46	23	32	41
<packaging>	12	10	8	0	0	2	4	0	43	33	20	0	67	80
type	12	10	8	0	0	2	4	0	43	33	20	0	67	80
pitch	0	0	0	0	0	0	0	8	*	*	*	*	*	*
material	4	4	2	0	0	2	2	5	67	50	50	0	50	50
p-l-count	5	3	3	0	0	0	2	4	40	40	0	0	60	100
units	0	0	0	0	0	0	0	8	*	*	*	*	*	*
bonding	0	0	0	0	0	0	0	8	*	*	*	*	*	*
device	14	5	5	0	0	0	9	0	64	64	0	0	36	100
equipment	0	0	0	0	0	0	0	8	*	*	*	*	*	*
subtotals	35	22	18	0	0	4	17	41	54	48	18	0	51	82
<device>	48	16	14	0	0	2	34	0	72	71	12	0	29	88
function	48	16	11	3	0	2	34	0	75	71	12	11	26	78
size	33	5	5	0	0	0	28	2	85	85	0	0	15	100
device-speed	3	0	0	0	0	0	3	12	100	100	*	*	0	*
subtotals	84	21	16	3	0	2	65	14	80	77	10	8	21	83
<equipment>	109	69	56	0	4	9	49	0	52	45	13	7	51	81
name	27	9	8	0	1	0	18	40	70	67	0	11	30	89
manufacturer	87	57	34	0	9	14	44	2	66	50	24	21	39	60
modules	15	0	0	0	0	0	15	53	100	100	*	*	0	*
type	109	69	51	0	9	9	49	0	57	45	13	15	47	74
throughput	3	0	0	0	0	0	3	55	100	100	*	*	0	*
status	109	69	52	0	8	9	49	0	56	45	13	13	48	75
wafer-size	12	0	0	0	0	0	12	48	100	100	*	*	0	*
subtotals	362	204	145	0	27	32	190	198	63	52	16	16	40	71
ALL-OBJECTS	1606	979	646	10	8	240	867	599	65	54	24	12	40	66
MATCHED-ONLY	949	793	646	10	8	82	238	599	37	25	10	8	68	82
TEXT-FILTERING	67	62	56	-	-	6	11	27	-	16	10	-	84	90

	Wrong	Req-fills	All-fills	Min-err	Max-err
RICHNESS-NORMALIZED-ERROR	1195	1641	1725	0.6928	0.7282
	Wrong	Word-count	Error-rate		
ERROR-RATE-PER-WORD	1195	1784	0.6698		
	P&R	2P&R	P&2R		
F-MEASURES	50.37	58.95	43.97		

287