

	CNN			Daily Mail		
	train	dev	test	train	valid	test
# documents	90,266	1,220	1,093	196,961	12,148	10,397
# queries	380,298	3,924	3,198	879,450	64,835	53,182
# max entities	527	187	396	371	232	245
# avg entities	26.4	26.5	24.5	26.5	25.5	26.0
# avg tokens	762	763	716	813	774	780
word count		118,497			208,045	