1 The evaluation scores for different DBpedia categories of Seen dataset:

==

Comparing the evaluation scores between two model: (flat, modify baseline) according to different DBpedia categories, which are: Astronaut,City,University,Food,SportsTeam.
We mean by "flat" one model from the extended system which contains 3 model:flat,str1,str2
Where we evaluate just this model in term of different DBpedia categories, because the result of other models i.e. str1&str2 for the whole dataset seen and unseen not very good comparing by this model
We mean by "Modify baseline" the modified baseline of webnlg challeng

1.1 Table of Evaluation in term of BLEU

Category Name	Model	BLEU
	Modify Baseline	50.02
Astronaut	Flat	**45.38**
	Modify Baseline	46.92
City	Flat	**47.81**
	Modify Baseline	39.35
Food	Flat	**37.77**
	Modify Baseline	43.20
SportTeam	Flat	**45.02**
	Modify Baseline	54.45
University	Flat	**63.10**

1.2 Table of Evaluation in term of METEOR

Category Name	Model	METEOR
	Modify Baseline	0.37
Astronaut	Flat	**0.38**
	Modify Baseline	0.46
City	Flat	**0.50**
	Modify Baseline	0.34
Food	Flat	**0.37**
	Modify Baseline	0.37
SportTeam	Flat	**0.39**
	Modify Baseline	0.37
University	Flat	**0.44**

1.3 Table of Evaluation in term of TER

Category Name	Model	TER
Astronaut	Modify Baseline	0.42
	Flat	**0.45**
City	Modify Baseline	0.39
	Flat	**0.39**
Food	Modify Baseline	0.52
	Flat	**0.51**
SportTeam	Modify Baseline	0.48
	Flat	**0.45**
University	Modify Baseline	0.41
	Flat	**0.32**