

	BLEU ↑	BERT Score ↑	Rouge-L ↑	Similarity	Parascore, w=0.05	Parascore (ours)
DiffuSeq	0.2971	0.8094	0.5345	0.7959	0.820	0.839
SeqDiffuSeq	0.2984	0.8054	0.5363			
ARDiffusion	0.1923	0.8013	0.5492	0.8032	0.830	0.831
SEDD	0.3274	0.8499	0.5744	1		
LDM	0.3202	0.8486	0.5794	0.8681	0.874	0.887
TEncDM	0.317	0.838	0.570			
GPT2	0.265	0.811	0.509	0.833	0.846	0.868

Method	Mauve ↑	Div ↑	Mem ↓	PPL ↓
Dataset	95.7 ± 0.9	41.8 ± 0.6	36.0 ± 0.4	21.7
TEncDM	79.0 ± 2.9	29.5 ± 0.2	43.8 ± 0.2	29.1 ± 0.9
SEDD	59.8 ± 5.2	37.6 ± 0.3	32.5 ± 0.2	70.8 ± 2.0
SeqDiffuSeq	1.3	3.7	66.3	29.3 ± 0.5
LD4LG	63.3	34.8	43.2	26.5 ± 0.3
ARDiffusion	6.6	10.1	54.0	41.8 ± 0.2
DiffuSeq	8.6	12.4	51.6	50.5 ± 0.3
GPT Neo	72.0 ± 4.7	25.2 ± 0.4	46.9 ± 0.5	20.5 ± 0.1
GPT2	78.9 ± 0.9	25.8 ± 0.1	45.5 ± 0.2	

Метод	QQP	XSUM	ParaDetox	Wiki-Auto	Squad	Dialogues	Mean Rank
Diffuseq	4	6	4	6	6	6	5.2
SeqDiffuseq	5	3	3	4	3	1	3
AR-Diffusion	6	4	6	5	4	2	4.5
SEDD	1	5	5	2	5	5	3.7
LDM	2	2	1	3	2	3	2.2
TEncDM	3	1	2	1	1	4	2.0