Generator	Dimension	Model Property Factual	CREATED Diversity	Plausibility	Proximity	Sparsity	D4EL Diversity	Plausibility	Proximity	Sparsity
CBG-CBGW-IM	Activity	0	0.972000	1.000000	0.063218	0.046497	0.000000	1.000000	0.000000	0.000000
		1	0.976000	1.000000	0.059401	0.129841	0.000000	1.000000	0.000000	0.214286
		2	0.976800	1.000000	0.086785	0.077784	0.000000	1.000000	0.000000	0.000000
		3	0.972800	1.000000	0.058631	0.053378	0.000000	1.000000	0.000000	0.055556
		4	0.976800	1.000000	0.059257	0.050848	0.000000	1.000000	0.000000	0.000000
		5	0.978400	1.000000	0.047923	0.132316	0.000000	1.000000	0.000000	0.052632
	Resource	0	0.980000	1.000000	0.506296	0.242407	0.000000	0.000000	0.277778	0.111111
		1	0.980000	1.000000	0.511616	0.249226	0.000000	1.000000	0.642857	0.214286
		2	0.979200	1.000000	0.527694	0.180480	0.000000	1.000000	0.642857	0.142857
		3	0.978400	1.000000	0.479959	0.222916	0.000000	1.000000	0.500000	0.222222
		4	0.980000	1.000000	0.446653	0.193787	0.000000	1.000000	0.409091	0.181818
		5	0.980000	1.000000	0.438579	0.201254	0.000000	0.000000	0.473684	0.157895
ES-EGW-CBI-ES-UC3-SBM-RR-IM	Activity	0	0.000000	1.000000	0.055556	0.000000	0.000000	1.000000	0.000000	0.000000
		1	0.320000	0.000000	0.062500	0.362500	0.000000	1.000000	0.000000	0.214286
		2	0.112800	0.000000	0.000000	0.347143	0.000000	1.000000	0.000000	0.000000
		3	0.039200	0.000000	0.165556	0.001111	0.000000	1.000000	0.000000	0.055556
		4	0.000000	1.000000	0.045455	0.045455	0.000000	1.000000	0.000000	0.000000
		5	0.000000	0.000000	0.052632	0.157895	0.000000	1.000000	0.000000	0.052632
	Resource	0	0.000000	0.000000	0.666667	0.277778	0.000000	0.000000	0.277778	0.111111
		1	0.640800	0.000000	0.875000	0.550000	0.000000	1.000000	0.642857	0.214286
		2	0.792000	0.000000	0.714286	0.490000	0.000000	1.000000	0.642857	0.142857
		3	0.936000	0.000000	0.832222	0.294444	0.000000	1.000000	0.500000	0.222222
		4	0.420000	0.000000	0.636364	0.227273	0.000000	1.000000	0.409091	0.181818
		5	0.554400	0.000000	0.631579	0.315789	0.000000	0.000000	0.473684	0.157895
ES-EGW-CBI-RWS-OPC-SBM-FSR-IM	Activity	0	0.000000	0.000000	0.055556	0.055556	0.000000	1.000000	0.000000	0.000000
		1	0.000000	0.000000	0.000000	0.357143	0.000000	1.000000	0.000000	0.214286
		2	0.000000	0.000000	0.000000	0.571429	0.000000	1.000000	0.000000	0.000000
		3	0.000000	1.000000	0.111111	0.166667	0.000000	1.000000	0.000000	0.055556
		4	0.000000	1.000000	0.090909	0.045455	0.000000	1.000000	0.000000	0.000000
		5	0.000000	0.000000	0.052632	0.210526	0.000000	1.000000	0.000000	0.052632
	Resource	0	0.000000	0.000000	0.666667	0.444444	0.000000	0.000000	0.277778	0.111111
		1	0.000000	0.000000	0.785714	0.428571	0.000000	1.000000	0.642857	0.214286
		2	0.000000	0.000000	0.714286	0.714286	0.000000	1.000000	0.642857	0.142857
		3	0.000000	0.000000	0.777778	0.388889	0.000000	1.000000	0.500000	0.222222
		4	0.584800	0.000000	0.772727	0.181818	0.000000	1.000000	0.409091	0.181818
		5	0.000000	0.000000	0.684211	0.210526	0.000000	0.000000	0.473684	0.157895
RG-RGW-IM	Activity	0	0.980000	0.000000	0.021696	0.000000	0.000000	1.000000	0.000000	0.000000
		1	0.980000	0.000000	0.031518	0.000000	0.000000	1.000000	0.000000	0.214286
		2	0.980000	0.000000	0.021023	0.000000	0.000000	1.000000	0.000000	0.000000
		3	0.980000	0.000000	0.034096	0.000000	0.000000	1.000000	0.000000	0.055556
		4	0.980000	0.000000	0.039796	0.000000	0.000000	1.000000	0.000000	0.000000
		5	0.980000	0.000000	0.042676	0.000000	0.000000	1.000000	0.000000	0.052632
	Resource	0	0.000000	0.000000	0.122198	0.000000	0.000000	0.000000	0.277778	0.111111
		1	0.000000	0.000000	0.151669	0.000000	0.000000	1.000000	0.642857	0.214286
		2	0.000000	0.000000	0.151084	0.000000	0.000000	1.000000	0.642857	0.142857
		3	0.000000	0.000000	0.130005	0.000000	0.000000	1.000000	0.500000	0.222222
		4	0.000000	0.000000	0.092543	0.000000	0.000000	1.000000	0.409091	0.181818
		5	0.000000	0.000000	0.121023	0.000000	0.000000	0.000000	0.473684	0.157895

Table 1: A comparison between our model and D4EL

Table 1 shows how each model performs under the evaluation metrics chosen by Hsieh, Moreira, and Ouyang. All of them apply separately to the sequence of resources and the sequence of activities. Each evaluation metric is the mean across all counterfactual results per model.

First, plausibility measures whether the sequence of activities or resources was found in the data—next, proximity is the normalised euclidian similarity between two sequences. The third is sparsity, computed using the normalised Levenshtein similarity.

We see that the evolutionary models are often comparable and sometimes even better than the DiCE4EL solution by Hsieh, Moreira, and Ouyang. We see that, for instance, for proximity. If the proximity of our model is lower than the proximity of the DiCE4EL solution, we can say that our models are, on average, closer to the factual. Similar holds for sparsity. We see this behaviour for both evolutionary generators. However, the Casebased-Search Generator also displays better proximity and sparsity scores than DiCE4EL. Only the Random-Search Generator appears to show worse results.