The current literature surounding counterfactuals exposes a number of challenges when dealing with counterfactuals.

The most important disadvantage of counterfactuals is the Rashommon Effect[**molnar2019**]. If all of the counterfactuals are viable, but contradict each other, we have to decide which of the *truths* are worth considering.

This decision reveals the next challenge of evaluation. Although, the criteria can support us with the decision, it remains an open research question *how* to evaluate counterfactuals according to **carvalho˙MachineLearningInterpretability˙2019**[**ca**
So far, no one was able to establish a standardised evaluation protocol[**hsieh˙DiCE4ELInterpret**
Every automated measure comes with implicit assumptions and they cannot guarantee a realistic explanations. Furthermore, we attempt to explain something with – in simple terms – *experiences* that never actually occured. We still need domain experts to assess their *plausibility*.

The generation of counterfactual sequences contribute to both former challenges, due to the combinatorial expansion of the solution space. This problem is common for counterfactual sentence generation and has been adressed within the Natural Language Processing (NLP). However, as process mining data not only consist of discrete objects like *words*, but also event and case features, the problem remains a daunting task. So far, little work has gone into the generation of multivariate counterfactual sequences like process instances.