Log Preparation Steps for Extremely Large logs – Splitting Procedure

(In our case: BPI Challenge 2019)

For preliminary splitting in subprocesses the following steps were followed.

The following steps were followed in **Disco:**

1) Filtering by attributes
2) Filtering by (case) Item Category
3) Apply them and create a copy

The following logs were split:

1) Consignment
2) 2 way matching
3) 3 way matching invoice receipt before GR
4) 3 way matching invoice receipt after GR

In the following steps we removed the incomplete traces based on endpoints filter:

1) We lack knowledge of the start and end activities so we choose based on mainstream behavior (removing those with less frequency comparatively)
2) The following steps combined with the semantically meaningful activity name were followed.
3) In addition to the above two steps if the case item category is 3 way matching invoice receipt after GR then " Record Goods Receipt " cannot be an end activity.

As a result, we filter by the following start and end activities:

	Allowed start activities	*Allowed end activities*
3-way matching, invoice after GR	Create Purchase order item, Create Purchase Requisition Item	Clear Invoice, Set Payment Block
3-way matching, invoice before GR	Create Purchase Order Item	Clear Invoice
2-way matching	Vendor creates invoice	Record Invoice Receipt, Clear Invoice

Consignment	Create Purchase Order Item, Create Purchase Order Requisition Item	Record Goods Receipt, Delete Purchase Order Item

The following steps were followed in **Disco:**

1) Filtering by endpoint
2) Select start and end activities
3) Apply them and create a copy