Data cleaning process visualised