

4.4. Matching

Uma vez os dados padronizados, criticados e muitas vezes enriquecidos chega-se ao momento em que deve-se comparar e agrupar os registros que se correspondem. Para fazer essa **comparação** ou **matching**, o processo calcula a probabilidade de um registro estar relacionado a outro. Ele envolve uma configuração de quão semelhantes são os registros, usando limiares (limites). Os limites definem qual o valor necessário para que os registros sejam considerados duplicados ou não.

Através dessas regras, é possível realizar uma comparação probabilística e estabelecer notas para validar se duas ou mais ocorrências se tratam de um mesmo registro, mesmo que possuam divergências entre si, como por exemplo **Igor, Ygor ou Higor**. A solução leva em consideração muito mais do que apenas o NOME, pois é possível unificar **homônimos**, nem apenas o CPF pois pode-se unificar **irmãos ou marido e mulher** que compartilharam daquele documento, situações essas bastante comuns encontradas nos sistemas.

Para serem mais assertivos, as regras de matching/comparações, aplicadas na solução MD2 MDM são mais inteligentes e robustas, possuindo inúmeros passos, com diferentes regras que foram sendo desenvolvidas e melhoradas durante anos de prática nos inúmeros projetos entregues. A cada nova release essas regras podem ser revisadas pela equipe de desenvolvimento que estão constantemente revisando e melhorando os motores de qualidade.

Uma vez feita essa comparação, cada registro é agrupado com seus similares e ficam à disposição para seguir a diante na esteira de qualidade.

Revision #13

Created 22 July 2022 15:29:12

Updated 17 August 2022 18:13:30