Extract and clean tools data to acquire tokens

↓

Create documents-tokens matrices

↓

Compute similarity matrices

↓

Learn weights on similarity matrices

↓

Compute weighted average of similarity matrices