

# Categorization Methods: Manual

- Determine the category based on rules that are carefully designed to reflect the domain knowledge about the categorization problem
- Works well when
  - The categories are very well defined
  - Categories are easily distinguished based on surface features in text (e.g., special vocabulary is known to only occur in a particular category)
  - Sufficient domain knowledge is available to suggest many effective rules
- Problems
  - Labor intensive → doesn't scale up well
  - Can't handle uncertainty in rules; rules may be inconsistent → not robust
- Both problems can be solved/alleviated by using machine learning