Using text mining to maintain translation memories

Andraž Repar,
International Postgraduate School and Iolar
Poster number: 40
• Translation memory – database of translations
• In use since 1990s – large quantities
• Sentence-aligned parallel corpus
• Difficult to maintain – no functionality to group according to domain
• Large translation memory supposed to contain Legal, Finance and Marketing content
  • Create an ontology using OntoGen
  • Use ontology topics as class labels for classification of new content