



Clues

Suggest

Search

Browse

Related

Info

Settings

Help

EPA

Emergencies and Cleanup (1 of 2915)

Pollution Prevention (6722 of 37043)

Conservation (10175 of 10175)

Energy Efficiency (0 of 0)

Fuel Economy (992 of 992)

Pollution Prevention (6735 of 6735)

Renewable Energy (2536 of 3601)

Sustainable Development (806 of 806)

Waste Reduction (1153 of 8012)

Regulatory and Industrial Topics (0 of 4219)

Risk Analysis (754 of 47561)

## Fuel

## Economy

### Clues

Clues are used to describe the language that used in a document that make it about a particular topic.

They may be any number of words up to a maximum of 200 characters per clue. However, most clues will consist of one, two or three words. Use double quotes around single words to disable stemming. Use double quotes around phrases to invoke exact phrase matching.

**Example:** A class called "Global Warming" may have the following clues:

- Global Warming
- Greenhouse Gases
- CO2 Emissions
- Pollution

Clues can also be clicked on to see examples of selected documents containing the clue concept.

**Please note:** Changes made to a class will have no effect unless documents are re-classified using conceptClassifier.

### Clue types

- **Standard** - A single word or multi-word concept or phrase. Use double quotes around the clue to force an exact phrase match.
- **Case Sensitive** - A case sensitive exact phrase match.
- **Metadata** - A clue based on document metadata.
- **Phonetic** - A clue that is matched on a fuzzy basis using a phonetic algorithm.
- **RegEx** - A clue based on a Regular Expression.
- **ClassID** - A list of ClassIDs that must classified before this class can be classified.
- **Class Boost** - A list of ClassIDs that contribute to the Score for this Class.
- **Language** - A list of Language codes, only documents that match one of these languages can be classified against this class.

## Clue scoring

Higher scores indicate a stronger association with the topic.

**Example:** Global Warming with a score of 50 will cause a document with this concept to be matched.

**Example:** Pollution with a score of 20 (on its own) will not be sufficient to cause the document as being about global warming.

Clues can also be assigned a negative value, which will prevent incorrect associations.

**Example:** Noise pollution should not be associated with Global Warming. So Noise pollution would be added with a negative value.

Scores are expressed as percentages of the Threshold, for example if the Threshold is 100 then:

- **100** = guarantees that this term alone will be sufficient to classify the document
- **50** = this term will get half way to the target
- **10** = this term is of low importance but its presence should boost a document score
- **0** = zero weight - use to disable a clue
- **-10** = this term is a small negative indicator
- **-100** = this term is a strong negative indicator
- **-1000** = the presence of this term should force the document to not be classified

## Suggested Clues

Clues can be used to statistically produce a list of suggested clues that can be assigned to the class, via following methods:

**Suggest Clues for whole class:** Click on the 'Suggest Clues' link under the class heading to produce a list of suggestions, based on all existing clues in the class.

**Single Clue:** Click on the 'Suggest' link against each clue to produce a list of suggestions, based on only this clue.

**Search result:** Click on the 'Suggest' link against each search result to produce a list of suggestions, based on the document.