## Invention Key Terms

Invention keyterms provide information about the roles played by terms within specific patents. Initially, roles were defined as follows:

* **invention**: a phrase which describes the invention or a hyponym. A hyponym might be a word like “process” or “system” which captures the type of the invention.
* **context**: a phrase which modifies, or narrows down the invention (e.g., indicates scope, function, use, beneficiary of invention).
* **attribute**: property or constituent of the invention
* **related**: conceptually related but not part of the core description

To illustrate this, below is the keyterm analysis for patent US4538295A.xml (entitled “Speech recognition system for an automotive vehicle”).

* invention: speech recognition system
* context: automotive vehicle
* attributes: bandpass filters, speech sound formant frequencies, engine noise, compensator, voltage level, electrical signal, db, notch filters
* related: noise, vehicle engines, hz, speech recognizer, decade increase, frequency

However, in the current instantiation of the keyterm code the number of categories was reduced to three:

* invention
* component or attribute
* contextual term

The module uses a maximum entropy classifier trained on 25 annotated patents from a random sample of English patents. Annotators labeled terms in 35 patents from the computer science domain. Because of the structure of patents, key terms tend to be introduced early and the earlier occurrences tend to have the strongest linguistic diagnostic contexts. For these reasons, we chose to limit our scope to the first thirty noun phrases occurring within the title or abstract sections of each patent (and similarly, when running the classifier we typically run it on the first 30 noun phrases).

The input to the classifier is the same as for the technology score, namely the feature vectors for the terms that are being classified.