Genic Interaction Extraction Challenge
Named Entities Data Format

1 File Structure

The LLL dictionary format provides a list of named entities (gene/protein names) that contains the Genic Interaction Extraction Challenge sentences. It consists of one entry per line.

2 Field Structure

A tab separates each element of an entry.
The fisrt element of each entry represents the canonical form of the reference entity.
The following elements can be either a synonymous of this reference entity or a typographic variation. It can also be both.

reference (tabulation) synonymous1 (tabulation) synonymous2 (tabulation) ...

3 Example

sigK SigK sigmaK sigma(K)