

## NAME

LevenstheinAutomata

Klassen zur Erstellung und Verwendung von Levenshtein Automaten

## SYNOPSIS

lev\_demo corpus.txt-file

## DESCRIPTION

*corpus.txt* muss eine Datei sein, die pro Zeile genau ein Wort enthält. Für optimale Ergebnisse sollte diese Datei so viele Worte wie möglich enthalten. *lev\_demo* erstellt aus der übergebenen Datei ein Korpus, auf dessen Grundlage nach Worten innerhalb der Levenshtein-Distanz gesucht wird. Anschließend werden die Levenshtein Automaten für die angegebenen Testworte erzeugt - diese Testworte sind das Wort 'badger' mit einer maximalen Levenshtein-Distanz von 1, 'duckling' mit einer maximalen Distanz von 2, sowie 'crocodile' mit einer Distanz von 3. Die entsprechenden Automaten enthalten alle Worte, die diese maximale Levenshtein-Distanz zu dem jeweiligen Wort haben. Die gefundenen Worte werden schließlich gedruckt und die ersten beiden Automaten zusätzlich im dot-Format ausgegeben.

## USAGE

```
g++ -std=c++0x -I include src/lev_demo.cpp -o lev_demo
lev_demo.exe data\web2.txt
dot -Tpdf -O Lev_badger_1.dot
dot -Tpdf -O Lev_duckling_2.dot
```

## FILES

C++11 wird benötigt

(Sowie doxygen und graphviz für die Erstellung der Dokumentation und der Automaten als pdf.)

## SEE ALSO

Demonstration einer möglichen Anwendung von Levenshtein Automaten -  
Eine "Meinten Sie?"-Funktion:

```
g++ -std=c++0x -I include src/didyoumean_demo.cpp -o didyoumean_demo
didyoumean_demo.exe data\web2.txt
```

## AUTHOR

Lisanne Wiengarten

Matrikelnummer 764870