

Проверка правописания и автокоррекция

"Candidate for a Pullet Surprise"

I have a spelling checker,
It came with my PC.
It plane lee marks four my revue
Miss steaks aye can knot sea.

Eye ran this poem threw it,
Your sure reel glad two no.
Its vary polished in it's weigh.
My checker tolled me sew.

A checker is a bless sing,
It freeze yew lodes of thyme.
It helps me right awl stiles two reed,
And aides me when eye rime.

Типы алгоритмов автокоррекции

Фонетические

Математические

Статистические

Контекстно-чувствительные

Soundex

- Первая буква сохраняется
- В остальной части слова:
 - Гласные (aeiouy) а также (h,w) выбрасываем
 - Каждую согласную заменяем на цифру от 1 до 6
 - b, f, p, v - на 1
 - c, g, j, k, q, s, x, z - на 2
 - d, t - на 3
 - l - на 4
 - m, n - на 5
 - r - на 6
 - Избавляемся от повторов
- Обрезаем до первых четырёх символов

Metaphone

- Взять первый символ или заменить его по правилам
PK, KN, GN => N AE => E WR => R
WH => H X => S
- В остальной части слова:
 - Избавляемся от повторов
 - F J L M N R остаются
 - В остается только если не после символа M
 - замена C на
 - X если в -CIA-, -CH-
 - S если в -CI-, -CE-, -CY-
 - удалить если -SCI-, -SCE-, -SCY-
 - иначе на K

Metaphone

- В остальной части слова:
 - замена D на
J если в -DGE- -DGI- -DGY-
иначе на K
 - замена G на
F если в -GH и не B--GH, D--GH, -H--GH, -H---GH
Удалить если в -GNED, -GN, -DGE-, -DGI-, -DGY-
J если в -GE-, -GI-, -GY- and не GG
иначе K
 - H остается если перед гласной и не после C,G,P,S,T
 - K остается если не перед C
 - P заменяется на P кроме случая PH в котором на F
 - Q заменяется на K

Metaphone

- В остальной части слова:
 - S заменяется на X в -SH-, -SIO- or -SIA- иначе на S
 - T заменяется на
 - X в -TIA- или -TIO-
 - 0 если перед H
 - удаляется если в -TCH-
 - иначе T
 - V => F, Z => S
 - W остается только если оно перед гласной
 - Y остается только если после него гласная
 - X => S в конце слова, иначе KS

Расстояние Левенштейна

Расстояние между словами на основе
количества вставок
количества удалений

Контекстно-зависимые алгоритмы



