

**ИСПОЛЬЗОВАНИЕ ИНТЕРНЕТ-СЕРВИСОВ ДЛЯ
АВТОМАТИЗАЦИИ УТОЧНЕНИЯ КОРРЕКТНОСТИ СЛОВ,
ОПРЕДЕЛЕННЫХ ОШИБОЧНЫМИ, ПРИ ПРОВЕРКЕ
ОРФОГРАФИИ В ПРОГРАММЕ
«ЛОКАЛИЗАТОР ИСХОДНОГО КОДА»**

*Федорин Денис Николаевич
Заведующий сектором интернет- технологий,
Филиал Воронежского государственного технического университета
в городе Борисоглебске,
РФ, Воронежская обл., г. Борисоглебск
E-mail: fedrin8989@mail.ru*

**USE OF INTERNET SERVICES FOR AUTOMATION OF
CORRECTNESS OF WORDS AT SPELL CHECKING IN THE
PROGRAM
"LOCALIZER SOURCE CODE»**

Fedorin Denis
Head of Internet technologies sector,
Branch of Voronezh state technical University in the town of Borisoglebsk,
Russia, Voronezh region, Borisoglebsk

АННОТАЦИЯ

В статье рассмотрены проблемы проверки орфографии, заключающиеся в сложности проверки редких терминов и имен; рассмотрены основные пути уточнения их корректности, применяемыми пользователями в таких случаях, и предложено решение по автоматизации проверки орфографии на примере использования программы «Локализатор исходного кода» с применением популярных интернет - сервисов.

ABSTRACT

The article deals with the problems of spelling, namely, the complexity of verification of rare terms and names; the main way of clarification of their correctness, users have applied in such cases, and the proposed automation solution checker on a example using program "Source Localizer" using popular Internet services.

Ключевые слова: орфография; проверка; слово; текст; интернет - сервисы.

Keywords: spelling; verification; word; text; Internet services.

Программные продукты, работающие с текстом для проверки правописания текста, обычно используют различные модули проверки орфографии, например, Hunspell [1] или средства проверки, встроенные в популярные пакеты, такие как, MS Office [2] или Apache Open Office [3].

При проверке орфографии ряд слов может быть ошибочно причислен к ошибкам, например: термины, имена или названия, которые отсутствуют в словарях модулей проверки орфографии.

В таком случае можно обратиться к проверке таких слов через сервисы в сети Интернет. Простейшим способом проверки слова является поиск проверяемого слова в поисковой системе, например, Google, в которой можно определить корректно ли слово или содержит ошибку.

Кроме такого простого подхода к проверке можно воспользоваться широко известными сервисами как Wikipedia или специализированными MultiTran [4] и пр.

На рисунке 1 показаны этапы обычной проверки орфографии пользователем:

1. Исходный текст до проверки орфографии.
2. Автоматическая проверка орфографии выявляет слова с ошибками.
3. Пользователь проводит проверку слов, ранее указанных как ошибочные, и убирает из них корректные слова, являющиеся терминами, именами и т. п.

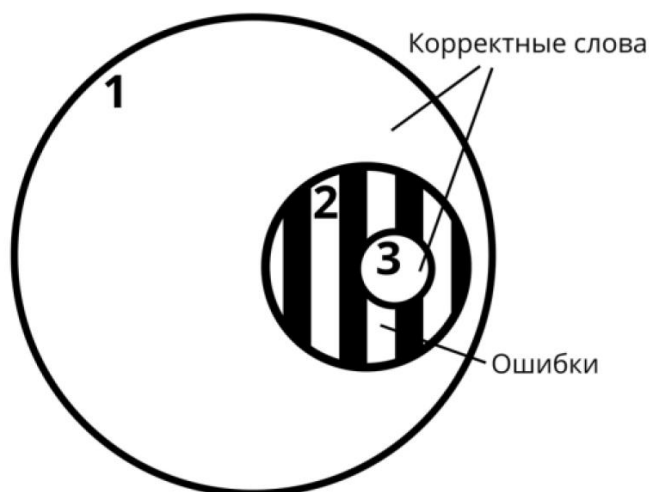


Рисунок 1. Порядок проверки орфографии в тексте

Для повышения скорости проверки текста на этапе 3 можно провести автоматизацию проверки слова через интернет - сервисы с помощью программы «Локализатор исходного кода». Для этого следует подготовить специально сформированный адрес сервиса с

интересуемым словом, в результате чего в браузере будет загружена страница с отчетом по слову для интересующего интернет - сервиса.

Для автоматизации такой проверки в программе «Локализатор исходного кода» [5, 6], подготовлен ряд шаблонов адресных строк, в которых специальная метка на этапе проверки заменяется проверяемым словом.

При подстановке слова в шаблон адресной строки интернет-сервиса следует учитывать, что в ряде случаев требуется преобразование его в шестнадцатеричные коды, и соответственно предложены метки {word} (обычное слово) и {wordhex} (слово в шестнадцатеричном представлении). Шаблоны адресных строк и соответствующие им интернет - сервисы представлены в таблице 1.

Таблица 1.

Список интернет- сервисов и шаблонов адресных строк

Название интернет-сервиса	Шаблон адресной строки
Проверка слова на сайте Gramota.ru («Грамота»)	http://www.gramota.ru/slovari/dic/?word= {wordhex} &all=x
Поиск в Google по сайту русской версии Wikipedia	https://www.google.com/search?q={wordhex} +site:ru.wikipedia.org
Проверка поисковой системой Yandex	http://yandex.ru/yandsearch?text={wordhex}
Проверка в сервисе MultiTran.ru	http://www. multitran. ru/c/m.exe ?CL= 1 &s= {wordhex} &I1 = 1
Проверка поисковой системой Google	https://www.google.ru/search?q={wordhex}
Проверка по орфографическому словарю русского языка на сайте TextoLogia.ru	http://orf.textologia.ru/search/?q=530&sch={wordhex}

В программе «Локализатор исходного кода» предложенный подход реализован через построение html-элементов с подготовленными запросами для каждого проверяемого слова, что хорошо видно на рисунке 2 на примере выбора сервиса для проверки слова «чапельник».

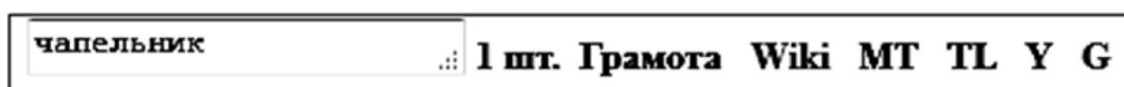


Рисунок 2. Контекстное меню для слова с ошибкой

В результате выбора сервиса «Грамота» формируется строка:
<http://gramota.ru/slovari/dic/?word=%F7%E0%EF%E5%EB%FC%ED%E8%EA&all=x>

Если выбрать «G», то будет сформирована строка:
<https://www.google.ru/search?q=%F7%E0%EF%E5%EB%FC%ED%E8%EA>

Пользователю достаточно навести указатель на проверяемое слово и выбрать из списка один из интернет - сервисов для проверки слова, после чего будет создана новая вкладка в браузере со сформированной строкой запроса, в которой появится результат проверки слова по выбранному сервису. Пример результата проверки слова «чапельник» представлен на рисунке 3.



Рисунок 3. Результаты поиска слова «чапельник» в сервисе Google (в меню - элемент «G»)

В случае с использованием поисковой системы Google можно не только проверить слово и прочитать его определение, но и увидеть поясняющие изображения к проверяемому слову.

Рассмотренное в статье использование предлагаемых шаблонов в программе «Локализатор исходного кода» значительно расширяет возможности программных продуктов и повышает скорость проверки орфографии за счет автоматизации использования популярных и специализированных интернет - сервисов для проверки неизвестных слов или слов со спорным написанием.

Список литературы:

1. Hunspell: About [Электронный ресурс]. — Режим доступа. — URL: <https://hunspell.github.io>
2. Check spelling and grammar in Office 2010 and later [Электронный ресурс]. -Режим доступа. - URL: <https://support.office.com/en-us/article/Check-spelling-and-grammar-in-Office-2010-and-later-5cdeced7-d81d-47de-9096-efd0ee909227>
3. Spell Check - Apache OpenOffice Wiki [Электронный ресурс]. – Режим доступа. - URL: https://wiki.openoffice.org/wiki/Documentation/Writer_for_Students/Auto_Spell_Check.