

Custom Dictionary Lingvo

Custom dictionary (пользовательский словарь) - словарь, формируемый самим пользователем как дополнение к стандартному словарю, поставляемому в комплекте программного обеспечения.

Пользовательские DSL-словари

Существует два способа создания пользовательских словарей для Lingvo:

1. Создание словаря «на лету» непосредственно из оболочки Lingvo, с карточками простой структуры, имеющими только заголовок и перевод. Возможности визуальной и содержательной разметки для таких карточек отсутствуют.
2. Создание словаря с использованием языка **DSL (Dictionary Specification Language)**.

Карточки **DSL-словаря**, могут быть максимально похожи по структуре на карточки системных словарей Lingvo. Карточки могут содержать все необходимые элементы оформления (bold, italic, underline, цвет текста, отступы), ссылки на мультимедийные (звуковые и графические) файлы-иллюстрации, перекрестные ссылки между карточками одного или нескольких словарей. Возможна разметка содержимого карточек по зонам перевода, комментариев и примеров.

Любой пользователь словаря может создать и распространять свой собственный **DSL-словарь**, т.к. компилятор языка DSL входит в состав ABBYY Lingvo (начиная с 6 версии):

DSL целесообразно использовать для создания словарей большого объема.



Dictionary Specification Language

Dictionary Specification Language (DSL) – язык для разметки словарных статей Lingvo, разработанный компанией ABBYY.

Основное назначение языка **DSL** - описание того, как показывать словарную статью. В языке **DSL** есть понятия "курсив", "ссылка", "подстатья" и т.д.

Язык **DSL** не определяет в точности структуру карточки или возможные типы информации. В языке нет понятия "синоним", "лексическое значение" или "синтаксическая модель".

Команды DSL

Все команды имеют вид «[a]», где a - конкретная команда (один или несколько символов).

Соответствующая ей команда выключения - «[/a]».

1	[b], [/b]	полужирный текст
2	[i], [/i]	курсивный текст
3	[u], [/u]	подчеркнутый текст
4	[c], [/c]	цветной текст.
5	[*], [/*]	выделение текста в качестве вторичного отображения
6	[m], [/m]	устанавливается сдвиг абзаца относительно левого края карточки, N - произвольная цифра от 0 до 9. Абзац после этой команды до эксплицитной команды отмены сдвига («[/m]») в текущей карточке будет сдвинут от левого края на число пробелов, кратное числу N в команде.
7	[trn], [/trn]	выделение зоны перевода
8	[ex], [/ex]	выделение зоны примера
9	[com], [/com]	выделение зоны комментариев
10	[!trs], [/!trs]	исключить из индексирования внутри зоны
11	[s],[/s]	выделение зоны мультимедиа (используется для вставки картинок или звуковых файлов)

12	<code>[url],[/url]</code>	ссылка на web-страницу
13	<code>[p], [/p]</code>	выделение помет (при нажатии на помету появляется разъясняющий текст)
14	<code>[!],[/!]</code>	выделение ударной гласной в слове текста карточки. можно изменить цвет ударных гласных на закладке Вид карточки (меню Сервис>Настройки...)
15	<code>[lang][lang]</code>	язык слова или словосочетания; указывается для слов внутри карточки, язык которых отличается от языка перевода (в этом случае для слов будет доступен полнотекстовый поиск, перевод из карточки, показ грамматических форм и т.д.) Например: <code>[lang name="Russian"]</code> .
16	<code>[ref][ref]</code>	ссылка на карточку в том же словаре, к которому относится карточка, содержащая эту ссылку (эквивалентна заключению заголовка карточки, на которую ставится ссылка, в двойные угловые скобки - <<>>
17	<code>[sub][sub]</code>	нижний индекс
18	<code>[sup][sup]</code>	верхний индекс
<p>Замечание: при использовании команд <code>[ref][ref]</code> необходимо убедиться, что карточка, на которую указывает ссылка, действительно существует, поскольку при компиляции словаря проверка наличия соответствующей карточки не осуществляется.</p> <p>Sic! Вложение зон индексации и команд форматирования текста одного типа недопустимо по правилам языка DSL 😊.</p>		

Цвет шрифта в DSL-словаре

Цветной текст в DSL-словаре можно установить в контейнере тегов `[c] ... [/c]`;

Если имя цвета не указано, то используется цвет, установленный в настройках ABBYY Lingvo.

Имя цвета пишется без кавычек.

Пробелы перед командами и после них присутствуют только тогда, когда они реально нужны для отображения.

В результате можно форматировать даже часть слова.

Если пробел стоит и перед командой, и после нее, то в окончательном отображении будет стоять только один пробел.

Пример:

excellent dictionary

[u][i]превосходный[/i] [b]словарь[c red]a[/c][b]/b[/u]

Текст «превосходный» будет выделен курсивом, текст «словарь» — полужирным, ударная буква «а» в слове словарь будет выделена красным, а вся карточка («превосходный словарь») будет подчеркнута.

Sic! Границы способов форматирования не могут пересекаться. Если какая-либо команда была включена раньше другой, она должна быть выключена позже.

Возможные цвета:

Aliceblue, antiquewhite, aqua, aquamarine, azure, beige, bisque, blanchalmond, blue, blueviolet, brown, burlywood, cadetblue, chartreuse, chocolate, coral, cornsilk, crimson, cyan, darkblue, darkcyan, darkgoldenrod, darkgray, darkgreen, darkkhaki, darkmagenta, darkolivegreen, darkorange, darkorchid, darkred, darksalmon, darkseagreen, darkslateblue, darkslategray, darkturquoise, darkviolet, deeppink, deepskyblue, dimgray, dodgerblue, firebrick, floralwhite, forestgreen, fuchsia, galnsboro, ghostwhite, gold, goldenrod, gray, green, greenyellow, honeydew, hotpink, indianred, indigo, ivory, khaki, lavender, lavenderblush, lawngreen, lemonchiffon, lightblue, lightcoral, lightcyan, lightgoldenrodyellow, lightgreen, lightpink, lightsalmon, lightseagreen, lightskyblue, lightslategray, lightsteelblue, lightyellow, lime, limegreen, linen, magenta, maroon, mediumaquamarine, mediumblue, mediumorchid, mediumpurple, mediumseagreen, mediumslateblue, mediumspringgreen, mediumturquoise, mediumvioletred, midnightblue, mintcream, mistyrose, moccasin, navajowhite, navy, oldlace, olive, olivedrab, orange, orangered, orchid, palegoldenrod, palegreen, paleturquoise, palevioletred, papayawhip, peachpuff, peru, pink, plum, powderblue, purple, red, rosybrown, royalblue, saddlebrown, salmon, sandybrown, seagreen, seashell, sienna, silver, skyblue, slateblue, slategray, snow, springgreen, steelblue, tan, teal, thistle, tomato, turquoise, violet, wheat, white, whitesmoke, yellow, yellowgreen.

Алгоритм создания DSL-словаря

1	в текстовом редакторе создайте файл словаря, используя команды языка DSL.	это может быть MS Word, Блокнот или другой редактор
---	---	---

2	Сохраните текст словаря в кодировке ANSI или Unicode.	осуществляется заданием соответствующей кодировки в окне сохранения документа
3	Измените расширение текстового файла словаря с .txt на .dsl	например, в Проводнике Windows, в Total Commander или убрав флажок с опции "Скрывать расширения для зарегистрированных типов файлов" в настройках вида каталога: Свойства папки...\Вид\....)
4	Скомпилируйте полученный DSL-словарь .	компилятор входит в состав пакета программ Lingvo и запускается из Главного меню (Пуск\ Все программы\ABBYY Lingvo\DSL Compiler)

Некоторые поддерживаемые языки

Язык	Имя языка
Африкаанс	Afrikaans
Баскский	Basque
Белорусский	Belarusian
Английский	English
Французский	French
Немецкий (старая орфография)*	German
Немецкий	GermanNewSpelling
Китайский	Chinese
Китайский (традиционное начертание)	ChinesePRC
Латинский	Latin
Польский	Polish
Русский	Russian
Суахили	Swahili
Турецкий	Turkish
Украинский	Ukrainian

При создании объемных словарей удобнее использовать базу данных или текст со специальным форматированием, заменяющим теги языка DSL; а затем из этого формата преобразовать в DSL-формат.

Для преобразования doc-файлов в dsl-файлы удобен специальный макрос [DSLConverter](#).

Этот макрос стал победителем конкурса, организованного ABBYY по инициативе пользователей.

Автор [DSLConverter](#) – **Назар Кравцов** (пользователь Lingvo из Львова, создатель десятков словарей для Lingvo, сделанных в формате DSL).

Можно набрать текст словаря в редакторе MS Word отформатировав его в точном соответствии с правилами [DSL Converter Description](#)), а затем, используя макрос, преобразовать полученный **doc-файл** в **dsl-файл**.

[Здесь](#) документ с DSL Converter Description

[Здесь](#) шаблон .dot с макросом [DSLConverter](#)

ABBYY Lingvo Content

30.04.2010 Компания ABBYY представила **ABBYY Lingvo Intranet Server**, корпоративную версию электронного словаря ABBYY Lingvo, которая дополняется **ABBYY Lingvo Content** - визуальным редактором (WYSIWYG-редактором) для составления и развития терминологических и корпоративных **DSL-словарей**.

ABBYY Lingvo Content – клиент-серверное приложение для создания, редактирования и публикации словарей в бумажном и электронном виде (на сайтах, на корпоративных порталах, в виде приложений для PC, Mac, смартфонов и мобильных устройств).

На компьютеры пользователей устанавливается приложения для редактирования словарных статей, а все данные хранятся на едином сервере (пользователи могут работать над одним словарем, находясь в разных уголках мира, просматривать словарные статьи всего словаря или нескольких словарей и обмениваться комментариями. Если одна из статей взята на редактирование пользователем, то она автоматически блокируется для редактирования другими пользователями, что позволяет избежать одновременного внесения изменений в статью).

История словаря хранит данные о различных типах изменений в словарной статье. Все виды действий со статьей фиксируются в истории, а версии статьи хранятся на сервере. Можно вернуться к предыдущей версии статьи или восстановить удаленные статьи.

В **ABBYY Lingvo Content** все статьи чётко структурированы, все данные в словаре систематизированы и единообразны. Разметка словарной статьи на различные зоны также позволяет:

- Осуществлять продвинутый поиск по словарю;
- Задавать различные варианты дизайна для словарных зон;
- Управлять экспортом словаря;
- Создавать малые словари из больших.

ABBYY Lingvo Content

включает фильтр статей позволяющий создавать выборки статей по различным параметрам (например: статьи, у которых встречается слово «добро» в примерах; все статьи, у которых более 5-ти

значений; все статьи, имеющие синонимы; все статьи, относящиеся к определенной тематике или начинающиеся с одной буквы). Статьи, найденные при помощи параметров фильтрации, можно назначить для обработки одному редактору, сохранить как выборку или как отдельный словарь.

Словари, созданные в **ABBYY Lingvo Content** можно быстро публиковать в виде бумажных, электронных, online или мобильных словарей. Форматы экспорта:



- XML и TBX – популярные форматы для обмена словарями и глоссариями, публикации словарей;
- HTML – для публикации словарей в интернете;
- RTF и DOCX – для подготовки макета бумажного словаря;
- DSL и LSD – форматы, используемые в электронном словаре ABBYY Lingvo.

Структура карточки DSL-словаря

1. **Заглавное слово**

Пишется с первой (**Sic!**) позиции новой строки.

В заголовке карточки могут встречаться алфавитные символы русского, английского, немецкого, французского и других языков, цифры, пробел, дефис, "'" и {}-используется для выделения в заголовке несортируемой части.



2. **Тело карточки**

Тело карточки записывается со следующей строки после заголовка.

Каждая строка в ней должна отступать от первой позиции хотя бы на один пробел!

Тело карточки простирается до заголовка следующей карточки (признаком его является символ в первой позиции, отличный от пробела или табуляции) или до конца файла.

В карточке могут выделяться следующие зоны:

- пометы (по желанию);
- зона перевода;
- зона комментариев (по желанию);
- зона примеров (по желанию);
- картинки, звук (по желанию).
- ссылки на web-страницы (по желанию).

Замечание: Порядок, в котором расположены зоны, необязателен: пометы, картинки, URL-ссылки не имеют строго фиксированного места в теле карточки.

Если тело карточки слишком длинное, можно уменьшить количество информации, одновременно отображаемой на экране:

часть карточки (например, словосочетания с заглавным словом) помечается как вторичное отображение и показывается, только если пользователь этого потребует.

В качестве вторичного отображения может быть задан любой фрагмент карточки (кроме заголовка);

Количество таких фрагментов также может быть любым. В карточке ABBYY Lingvo они могут показываться или не показываться вместе.

Оболочка ABBYY Lingvo при показе вторичного отображения использует шрифт другого цвета.

Мультимедийные вставки

Звуковой файл или файл, содержащий изображение, должен быть размещен либо в папке ABBYY Lingvo, либо в папке словаря. При изображении карточки в ABBYY Lingvo на месте данной записи будет находиться значок. При нажатии на этот значок будет запущен звуковой файл или в отдельном окне откроется картинка.

Чтобы вставить картинку в текст карточки:

- В dsl-тексте карточки укажите имя файла с изображением внутри зоны мультимедиа [s].

Например: [s]mypicture.bmp[/s].

Допустимые форматы картинок:

BMP, Bitmaps (*.bmp),

PCX (*.pcx)

DCX (*.dcx)

JPEG (*.jpg)

TIFF (*.tif)

Чтобы вставить звук в карточку:

- В dsl-тексте карточки укажите имя звукового файла внутри зоны мультимедиа [s].

Например: [s]mysound.wav[/s].

Допустимый формат звуковых файлов:

Wave Sound (*.wav)

Чтобы вставить в карточку видео:

- В dsl-тексте карточки укажите имя файла с изображением внутри зоны мультимедиа [s].

Например: [s]movie.avi[/s].

Этот файл должен располагаться в папке словаря или в папке ABBYY Lingvo.

Это просто невыносимо....

Lfer: Составляю словарь для лингво. как явствует из справки, текст словаря нужно готовить в текстовом документе, который затем конвертируется в dsl. поэтому я либо пишу сразу в текстовом, либо в ворде, а потом копирую в txt. в обоих случаях табуляция в документе, которую надо расставлять перед текстом словарной статьи, постоянно пропадает! и карточки словаря оказываются перепутаны. сколько бы я ее ни исправлял, она при каждом открытии документа теряется, причем не сразу вся, а кусками. так просто невозможно работать, можете что-нибудь посоветовать?

Re: В Блокноте действительно табуляция может пропадать, т.к. он имеет свойство переформатировать абзацы, расставляя переносы на новую строку по-другому. Некоторые текстовые редакторы при нажатии клавиши TAB вообще вставляют определенное количество пробелов (это характерно для текстовых редакторов для программистов). В Word такой проблемы не может быть. Используйте режим отображения скрытых символов (на панели инструментов),

чтобы увидеть вставляемые Вами знаки табуляции. Думаю, что вся проблема с Вашим текстовым редактором.

Re: Прямо в ворде сохраните .txt как Unicode, а лучше возьмите Notepad2 или TextEd. А чтобы не запутаться в Unicode откройте Блокнот и создайте пустой файл Введите имя файла и сохраните кодировка Unicode. Потом измените разшерение .txt на .dsl

Языки словарей

В Lingvo DSL-компилятор позволяет создавать словари на 30 языках:

- в оболочке Lingvo поддержаны отображение всех символов этих языков и сортировка слов по алфавиту (важно правильно указать исходный язык и язык перевода словаря),
- для ряда языков реализована морфологическая поддержка и возможность просмотра\добавления парадигм (словоформ).

Язык	Сортировка	Морфология	Парадигмы
Английский	•	•	•
Африкаанс	•		
Баскский	•		
Белорусский	•		
Болгарский	•	•	
Венгерский	•	•	
Голландский	•	•	
Датский	•	•	
Индонезийский	•		
Испанский (международная сортировка)	•	•	•
Испанский (традиционная сортировка)	•	•	•
Итальянский	•	•	•
Китайский	•		

Китайский (традиционное начертание)	•		
Латинский	•		
Немецкий (Новая Орфография)	•	•	•
Немецкий	•	•	•
Норвежский (Букмол)	•	•	
Норвежский (Нюнорск)	•	•	
Польский	•	•	
Португальский	•	•	
Русский	•	•	•
Сербский (кириллица)	•		
Суахили	•		
Турецкий	•	•	
Украинский	•	•	•
Финский	•	•	
Французский	•	•	•
Чешский	•	•	
Шведский	•	•	
<p>Можно создавать словари для языков за пределами «официального» списка (армянский, иврит, японский, арабский и многие другие).</p> <p>Для создания словарей на других языках в разметке можно указать либо английский язык, либо любой другой, например суахили. При втором варианте не будет «вмешиваться» языковая морфология и сортировка.</p>			

