

Веб-переводчики для серфинга

РУБРИКА: ТЕМТЕГИ: ПЕРЕВОДЧИК



Наш обзор посвящен лучшим сервисам, предназначенным для автоматического перевода сайтов в ходе серфинга. И пускай у общественности к результату их работы отношение зачастую снисходительно-юмористическое, иногда других вариантов просто нет.

В конце концов, не все же мы полиглоты. Да и технологии на месте не стоят. Посмотрим на их актуальное состояние... При отборе ресурсов для тестирования учитывались следующие обстоятельства. Во-первых, сервисы должны были уметь напрямую переводить на русский язык хотя бы с нескольких иностранных языков. Во-вторых, быть бесплатными. Требования простые, однако они отсеяли немало претендентов.

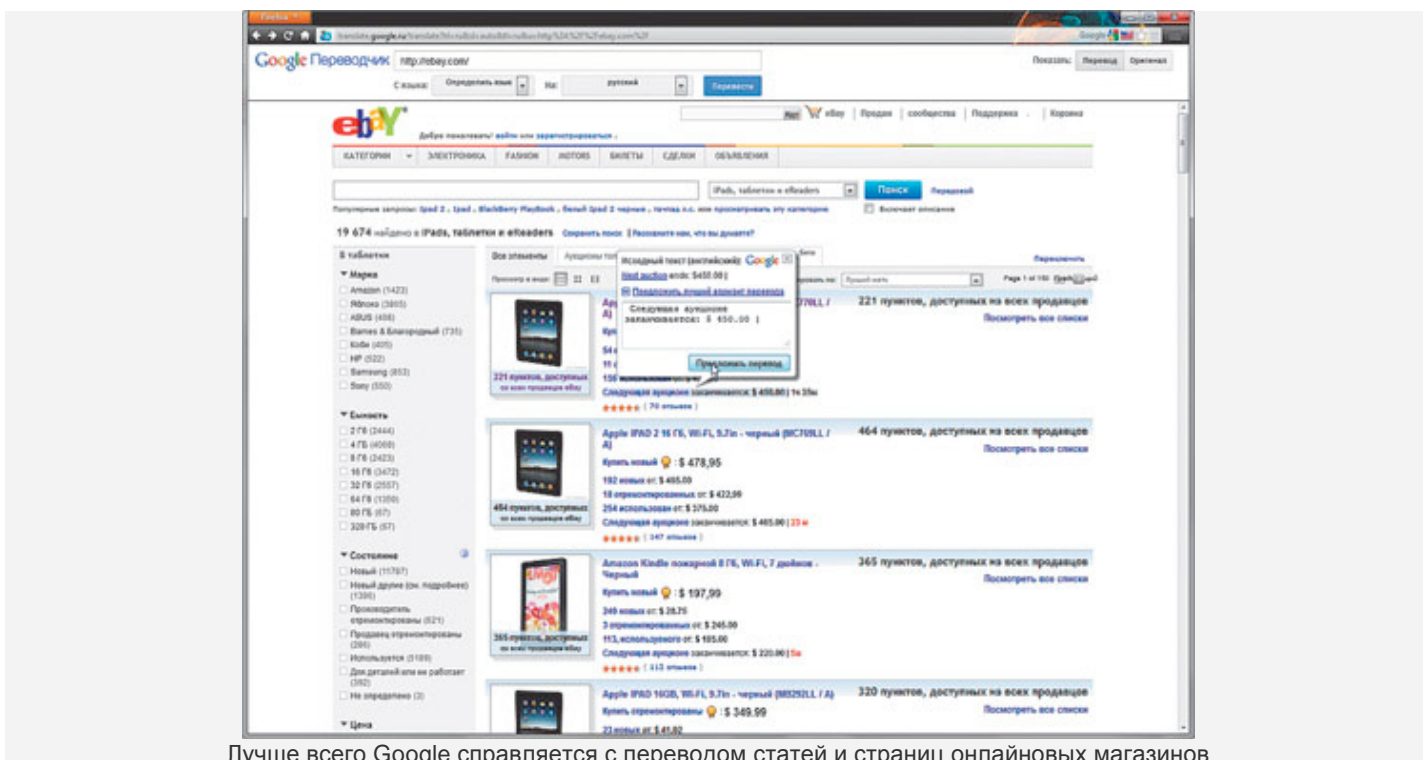
Поскольку мы рассматривали сервисы, заточенные под перевод в ходе серфинга, значение имело корректное отображение сайтов и скорость перевода.

Кроме собственно качества перевода мы использовали и другие критерии оценки. Поскольку мы рассматривали сервисы, заточенные под перевод в ходе серфинга, значение имело корректное отображение сайтов и скорость перевода – никому ведь не охота подолгу ждать загрузки нужного ресурса или глазеть на искореженный дизайн страницы. Обратили мы внимание и на удобство интерфейсов, а также наличие различных вспомогательных инструментов для браузеров.

«Переводчик Google»

Сервис перевода от Google (translate.google.ru), бесспорно, является одним из самых популярных. Его поддержка по умолчанию встроена в браузер Chrome / Chromium. «Переводчик Google» любят и разработчики различных дополнений. Найти плагин или панель инструментов для другого браузера проблемы не составляет. Сильная сторона сервиса Google – поддержка большого количества языков: всего доступно несколько десятков, как европейских, так и азиатских. «Переводчик Google» умеет автоматически распознавать введенные в поле перевода URL.

Действует автоматическое определение языка открытой веб-страницы, причем работает данный инструмент вполне успешно. В то же время попытка перевода веб-страницы, на которой находится текст на нескольких языках (я использовал страничку с параллельными переводами рождественских гимнов), приводит к конфузу.



Предпочтение было отдано английскому тексту, который был обработан корректно. А вот вместо немецкого можно было наблюдать диковатую солянку из оставленных нетронутыми немецких слов, попыток транскрипции и перевода английских аналогов, похожих по написанию.

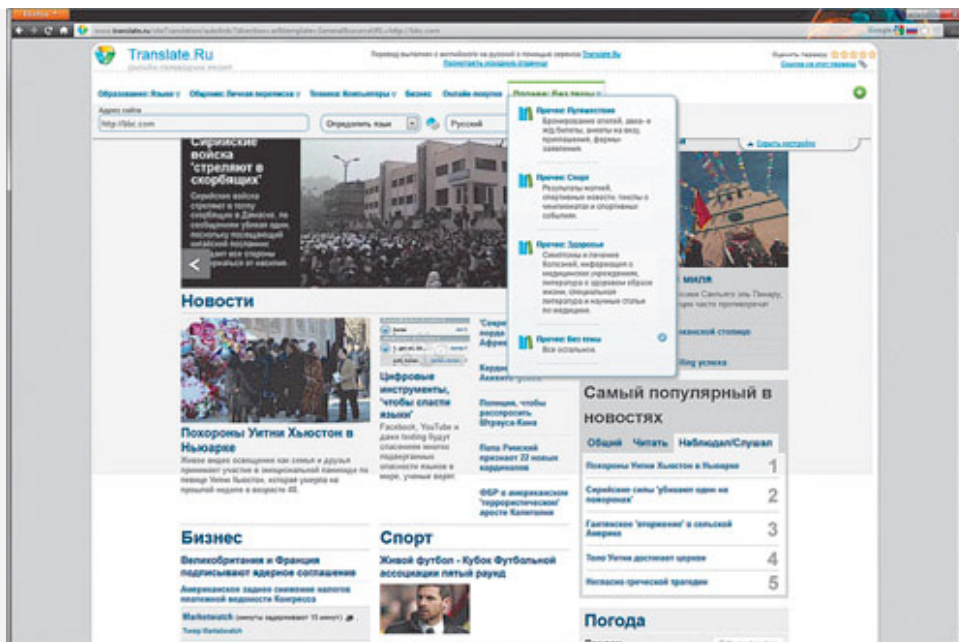
Скорость ресурса вполне комфортна и не отвлекает от серфинга. Верстка страниц сохраняется достаточно аккуратно – сбои происходят нечасто и в действительно сложных случаях. Кроме того, сервис Google умеет переводить текст, выводимый скриптами, например пролистывающими заголовки новостей на порталах.

«Переводчик Google» отображает на страницах небольшую панель инструментов, с помощью которой можно управлять языковыми парами перевода. Фрагменты оригинального текста демонстрируются во всплывающих «пузырьках». Это удобно, если вы знакомы с исходным языком и нуждаетесь в периодических подсказках. В других случаях это просто отвлекает от работы с переведенным сайтом.

Translate.ru

Следующий участник нашего обзора – сервис, разработанный компанией «Промт» (translate.ru), хорошо известной своими программами машинного перевода. Данный сервис не стремится охватить десятки языков, работая только с ведущими европейскими языками: немецким, французским, английским, испанским, итальянским, португальским.

В отличие от «Переводчика Google», Translate.ru умеет использовать специализированные тематические словари для различных сфер деятельности. Всего имеется полтора десятка вариантов, собранных в категории-вкладки. Выбрать нужный словарь можно на вспомогательной панели сервиса.



Сервис Translate.ru построен на платформе «Промт» и предлагает тематические словари и хорош для чтения новостей

Кроме этого, панель Translate.ru предлагает инструменты оценки качества перевода, а также возможность быстрого сохранения постоянной ссылки на перевод, которой можно поделиться с друзьями. Оригинальный вариант страницы можно увидеть только целиком – параллельный вывод текста не предусмотрен. Внешний вид сайтов после перевода вполне приличный, текст в динамических виджетах отображается корректно.

Разработчиками заявлена функция автоматического определения исходного языка, однако она частенько сбивает. Так, Translate.ru откровенно симпатизирует английскому языку, время от времени принимая за него тексты на французском и немецком. Поэтому сервис периодически приходится поправлять вручную. Что касается обработки двуязычных документов, то Translate.ru с ними также не дружит: предпочтение, как и в случае «Переводчика Google», отдается английскому языку, а слова на других языках или вообще не переводятся или обрабатываются как подобные по написанию английские. Детальный разбор таких ошибок смысла не имеет: в итоге все равно получается нечитаемая каша.

«Яндекс.Перевод» Beta

После сервиса Translate.ru логично упомянуть еще один продукт, построенный на технологиях «Промт». Это – новинка «Яндекса» (translate.yandex.ru). В настоящее время продукт еще находится в стадии бета-тестирования, однако уже вполне пригоден для использования. Сейчас он оперирует восемью языками, среди которых – основные европейские, польский, украинский и турецкий. Как видим, состав языков немного отличается от сервиса Translate.ru. Еще одним серьезным отличием версии «Яндекса» от исходного проекта является отсутствие возможности выбора тематических словарей – все проблемы повышения качества перевода целиком возложены на компьютеры.

Онлайновый переводчик «Яндекса» использует платформу «Промт», однако предлагает собственный интерфейс

Функция автоматического определения языка открываемой веб-страницы у сервиса «Яндекса» есть, однако ее при желании можно отключить – соответствующая опция доступна прямо на стартовой странице ресурса. Как и на других сайтах, пары языков можно выставить в меню самостоятельно. При вводе символов в поле перевода задействуется автодополнение как слов, так и веб-адресов.

Скорость перевода в основном хорошая, однако пару раз пришлось столкнуться с сообщениями об ошибке: «Перевод не удалось выполнить в отведенное время», в результате чего приходилось отправлять запрос повторно. Разметка страницы

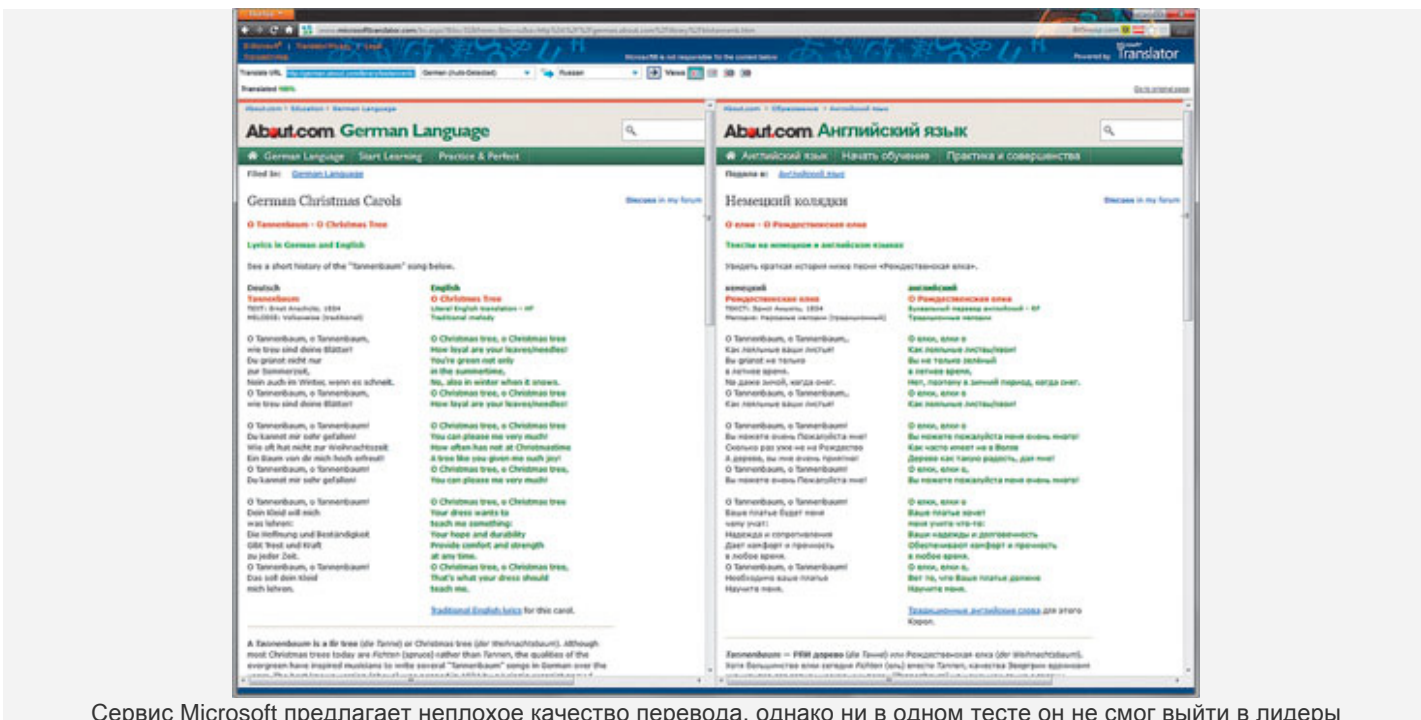
сохраняется, однако текст в виджетах часто остается непереуведенным. Двухязычный текст не распознается, хотя переводчик иногда пытается перевести отдельные слова на втором языке. Как бы то ни было, система корректно работает только с однородным контентом.

Панель инструментов весьма скромна. Нам предлагают выбор языковых пар, а также несколько вариантов демонстрации исходного текста и перевода: по отдельности или совмещенными по горизонтали или вертикали. В отличие от Microsoft Translator, который также имеет подобный режим, при наведении курсора на фрагмент оригинального текста соответствующий переведенный фрагмент не подсвечивается, что несколько затрудняет работу.

Отдельных вспомогательных приложений у «Яндекс.Перевода» пока нет. В то же время все его функции доступны в штатной многофункциональной панели инструментов «Яндекс.Бар».

Microsoft Translator

Относительно недавно своим собственным онлайн-переводчиком обзавелась и корпорация Microsoft (www.microsofttranslator.com). Данный сервис также работает в поисковике Bing, принадлежащем все той же корпорации. Поддерживается почти сорок языков – как европейские, так и азиатские.



Сервис Microsoft предлагает неплохое качество перевода, однако ни в одном тесте он не смог выйти в лидеры

Скорость перевода – на уровне. Неприятных задержек практически нет. Пары языков можно устанавливать вручную или же попробовать положиться на автоматическое определение. Надо заметить, что при открытии двухязычного документа нас ожидает приятный сюрприз: данный сервис различает языки и переводит фрагменты по отдельности. Пусть с пропусками и ошибками, но переводит. А вот с текстом в скриптах Microsoft Translator справляется с трудом, обычно просто его пропускает.

Панель инструментов позволяет управлять режимами параллельной демонстрации оригинального текста и перевода. По умолчанию включен вывод фрагментов исходника при наведении курсора на текст перевода. При желании можно включить обратный режим с демонстрацией оригинальной страницы и фрагментами перевода в «пузырьках». Кроме того, можно разделить экран на две зоны, в которых будут показаны оба варианта.

Вспомогательные программы представлены только букмарклетом, который можно разместить на панели закладок своего «Рабочего браузера». После его установки одним кликом будет достаточно для того, чтобы открытая страница перезагрузилась уже переведенной. Такой букмарклет может трудиться в любом браузере, что достаточно удобно, однако не дает доступа к каким-либо дополнительным функциям.

Методика тестирования

Тестирование сервисов перевода не такая тривиальная задача, поскольку определить критерии качества может быть непросто. Выявление же лучших подразумевает применение сопоставимых критериев. Приемлемым уровнем будем считать вариант, как минимум не вводящий пользователя в заблуждения неправильным переводом отдельных слов, особенно имеющих различные значения. В определенной мере можно оценить и словарную базу проектов – по количеству оставшихся без перевода слов. Особенно это важно для «айтишных» и вообще любых тематических сайтов, на которых часто встречается специфическая лексика и различные неологизмы.

Что же мы будем рассматривать в качестве типовых задач? Первым заданием для участников станет чтение сайтов общего профиля. В качестве тестовых заданий выступают крупные новостные ресурсы. Вторая типовая задача – работа с тематическими статьями. Источником тестовых заданий в этом случае послужат национальные разделы «Википедии». Третье задание – совершение онлайн-покупок. Здесь в качестве материала послужат локализованные версии всемирной интернет-барахолки eBay.

Поскольку наборы доступных пар перевода отличаются, для корректного сравнения мы ограничимся только основными европейскими языками. Оценку будем ставить качественную, просто распределяя сервисы по местам в соответствии с каждым критерием. Если результаты по качеству окажутся близки, несколько сервисов получат одинаковые места. Кто наберет большее количество первых мест – того и признаем победителем.

На старт, внимание, марш!

Итак, стартуем с чтения новостных ресурсов. В принципе, все участники тестирования выдают более-менее приемлемые результаты. Перевод с английского языка (bbc.com), пожалуй, лучше всего получается у Translate.ru. Артефакты, конечно же, присутствуют, однако перевод отдельных слов и особенно структура фраз здесь лучше, чем у конкурентов. Интересно, что версия перевода от «Яндекса», основанного на той же базе, отличается. Сложно однозначно определить, у кого это получается лучше. С одной стороны, в версии Translate.ru оставалось больше пропущенных слов. С другой – построение фраз зачастую было лучше, чем у «Яндекса». Сервисы Google и Microsoft близки по качеству, основные проблемы у них возникают с формами переведенных слов и построением фраз.

Французские новости (www.lemonde.fr) больше удались Translate.ru. Общее качество перевода получилось заметно ниже, чем в случае с английским текстом, однако смысл был достаточно понятен. Сервис «Яндекса» оставлял много слов непереуведенными. «Переводчик Google» справлялся с переводом слов хорошо, однако построение фраз было, мягко говоря, не самым понятным. То же можно сказать и о Microsoft Bing Translator – читать итоговый текст очень трудно. В целом было ясно, что речь идет, допустим, про голосование в Совбезе ООН, а не про нашествие марсианских пришельцев-осьминогов, однако многие детали безвозвратно терялись в мешанине переведенных по отдельности и кое-как приставленных друг к другу слов.

Новости на немецком языке были взяты на сайте www.focus.de. Результаты Translate.ru и «Яндекс.Перевод» совпали дословно. Общее качество приемлемое, хотя и не блестящее – некоторые фрагменты практически непонятны. «Переводчик

Google» лучше всех справился с переводом отдельных слов – пропусков не было, количество ошибок невелико, в то же время с грамматикой он оплошал. Тот же комментарий можно дать и результату работы сервиса от Microsoft – более-менее неплохая работа с отдельными словами, но трудно воспринимаемые предложения.

Перейдем к тематическим статьям, а именно статьям про интернет в различных национальных версиях «Википедии». Качество перевода с английского языка у «Переводчика Google» было весьма хорошим. Текст, несмотря на определенные артефакты и отсутствие изящества в построении фраз, получился понятный и вполне читабельный. Немного уступил здесь по качеству Microsoft Translator – пользоваться полученным результатом тоже можно не перенапрягая мозг, хотя неудачных конструкций здесь встречалось больше, чем у Google. Неправильными в основном были формы глаголов, хватало и других нарушений в словообразовании, что иногда заметно мешало восприятию смысла текста. По качеству перевода Translate.ru и «Яндекс.Перевод» в данном случае заметно отстали.

Откроем немецкую энциклопедическую статью. История с новостными сайтами повторяется: качество перевода немецкого текста у участников обзора в целом заметно ниже, чем английского. Относительно неплохо показал себя Translate.ru – в основном переведенный текст понятен. В то же время результат работы его «брата по разуму» – сервиса «Яндекс.Перевод» – получился достаточно трудным для анализа. Хватало в нем и простых ошибок, например с обработкой артиклей. Недурно выглядел в этом тесте сервис Microsoft Translator. Несмотря на ошибки с переводом отдельных слов и терминов, он выдавал вполне адекватный, пригодный к чтению текст. Не самым лучшим образом справился «Переводчик Google». Качество переведенного текста таково, что без применения психоактивных веществ понятно немногим больше половины текста.

Статья из французской «Википедии» тоже оказалась крепким орешком. Молодцом держался онлайн-переводчик Google»: текст получился достаточно понятный, приходилось пропускать только отдельные фрагменты. Другие участники обзора в этом тесте определенно отставали как по качеству перевода отдельных терминов, так и по стилю итоговых фраз. Translate.ru, что любопытно, не помогло даже включение специализированного словаря «Компьютеры». Радикальных изменений в тексте, к сожалению, не произошло.

Сервисы от «Яндекса» и Microsoft также звезд с неба не хватало, предлагая различные забавные варианты. Так, например, Microsoft Translator перепутал аббревиатуру протокола EGP с обозначением денежной единицы Египта. Проблемы с обработкой аббревиатур и переводом терминологии, равно как и с построением предложений испытывал и «Яндекс».

Остался последний тест – проверка на совместимость с сетевыми аукционами eBay. Как обычно, начнем с английской версии. В общем и целом онлайн-переводчики со своей задачей справляются, вряд ли вы купите топор вместо фиалки, однако и глюков также хватает. Общий недостаток – частенько некорректный перевод названий брендов. И если «Яблоко» в разделе электроники еще вполне понятный вариант, то швейные машинки «Певица» вместо почтенного «Зингера» уже вызывают легкий ступор, да и предложенные названия разделов в духе «Ручной И Закончил Штук» – это сильно!

Не менее важен перевод различных типовых фрагментов текста, например рассказывающих о вариантах доставки или оплаты. Наиболее корректными они получились у «Переводчика Google». У машин на основе «Промт», равно как и у сервиса Microsoft здесь проблем возникало больше. Переход к другим языковым версиям общую расстановку сил не изменил: по качеству лидирует Google, за которым с небольшим отставанием идут остальные претенденты.

Итоги

Общий итог нашего обзора прост: качество машинного перевода постепенно улучшается, и откровенного тихого ужаса,

породившего в свое время немало анекдотов, в обработанных текстах становится все меньше и меньше. Комбинируя использование различных сервисов, можно получить приемлемые результаты во всех типовых заданиях. Пусть онлайн-переводчики и не будут перемахивать через языковые барьеры с грацией призового скакуна, но на роль скромной тяговой лошади для серфера, не обремененного знанием десятка-другого языков, они вполне подойдут.

Качественная оценка точности перевода текста различной тематики и других параметров				
	«Переводчик Google»	Translate.ru	«Яндекс.Перевод» Beta	Microsoft Translator
Английский язык (новости), место	2	1	1	2
Английский язык (статья), место	1	3	3	2
Английский язык (покупка), место	1	2	2	2
Французский язык (новости), место	3	1	2	3
Французский язык (статья), место	1	2	2	2
Французский язык (покупка), место	1	2	2	2
Немецкий язык (новости), место	2	1	1	2
Немецкий язык (статья), место	3	1	3	2
Немецкий язык (покупка), место	1	2	2	2
Дополнения, место	1	2	2	2
Интерфейс страницы перевода, место	1	3	2	2
Стабильность и скорость, место	1	3	4	2
Общая оценка, место	1	2	3	4

Цифры, полученные по итогам тестирования, приведены в таблице. По сумме баллов победителем стоит признать «Переводчик Google»: по количеству первых мест он почти вдвое опережает ближайшего конкурента – Translate.ru. В то же время значительное преимущество ему дает большое количество дополнений, общая скорость работы (сервисы на основе «Промт» довольно задумчивы) и функциональность панели инструментов. Если же учитывать только качество перевода, эти два сервиса приходят к финишу практически ноздря в ноздю.

«Яндекс.Перевод» пока предлагается в бета-версии, будем надеяться, что ко времени релиза он прибавит в скорости и избавится от периодических сбоев по тайм-ауту и других шероховатостей. По качеству он в основном сопоставим с Translate.ru, хотя предлагаемые варианты перевода иногда отличались.

Двойственное впечатление оставил онлайн-переводчик от Microsoft. С одной стороны, он предлагал комфортную скорость работы, довольно приятный интерфейс и конкурентоспособное качество перевода. Однако ни в одном тесте он не смог выйти в лидеры, что и определило его итоговое место. UP