

## **Ferramentas informáticas de apoio à tradução: uma seleção das mais importantes e gratuitas à disposição dos tradutores**

Kelson dos Santos Araujo - kelsondossantosaraujo@gmail.com  
Doutorando em Educação a Distância e e-Learning - Universidade Aberta de Portugal

Ana Maria de Jesus Ferreira Nobre - Ana.Nobre@uab.pt  
Docteur D'État, PhD  
DEED - Dept.<sup>o</sup> de Educação e Ensino a Distância - Universidade Aberta de Portugal

### **Resumo:**

Existem muitas e variadas ferramentas computadorizadas de apoio à tradução disponíveis aos tradutores atualmente. O presente artigo estabelece uma diferenciação conceitual entre os principais tipos dessas ferramentas, entre gratuitas e pagas. Em seguida, respondemos a pergunta: "Qual é a ferramenta computadorizada de apoio à tradução mais importante para os tradutores?" Para isso, apresentamos os resultados de um inquérito por questionário com tradutores membros da lista de discussão eletrônica, Trad-Prt, do 'Yahoo Groups', para a língua portuguesa, bem como com um grupo de Tradutores provenientes da rede social Facebook, 'Tradutores com Vida'. Elencámos 11 ferramentas computadorizadas de apoio à tradução mais importantes. Estas deveriam compor um conjunto de ferramentas quase obrigatório para todos os Tradutores que desejam ter um lugar no mercado competitivo dos dias de hoje. Por último, objetos de aprendizagem (OA) sobre a utilização dessas 11 ferramentas foram desenvolvidos no âmbito de um estudo de Doutoramento pela Universidade Aberta de Portugal (UAb). Ao final do ano de 2017, esses onze OA estarão disponíveis para a comunidade acadêmica, a partir do repositório da UAb, na forma de recursos educacionais abertos.

### **Abstract:**

There is a number of different computer-aided translation (CAT) tools available to translators today. This article establishes a conceptual differentiation between the main types of CAT tools - freeware and paid ones. Following, we answer the question: 'Which is the most important CAT tool for translators?' For this, we show the results of a survey from Translators - members of Yahoo Groups' Trad-Prt electronic mailing list for the Portuguese language - and from a Facebook group called 'Tradutores com Vida'. The results show a list of 11 most important, freeware, CAT tools. These should be an almost obligatory set of tools for all Translators whom wish to have their place in today's competitive market. Finally, learning objects (LO) about these 11 CAT tools have been developed in the frame of a PhD study from Open University of Portugal (UAb). By the end of 2017, these 11 LO will be available for the academic community from UAb's repository in the form of open educational resources.

### **Resumen:**

Hay una serie de diferentes herramientas de traducción asistida por computadora (CAT, por sus siglas en inglés) disponibles para los tradutores de hoy. Este artículo establece una diferenciación conceptual entre los principales tipos de herramientas CAT 'freeware' y bajo licencia con pago. A continuación, respondemos a la pregunta: '¿Cuál es la herramienta CAT más importante para los tradutores?' Para ello, mostramos los resultados de una encuesta realizada con tradutores miembros de la lista de correo electrónico Trad-Prt, de Yahoo Groups, para la lengua portuguesa, incluyendo un segundo grupo de tradutores de la red social Facebook, 'Tradutores com Vida'. Los resultados muestran una lista de las 11 más importantes herramientas CAT del tipo 'freeware'. Estas deben formar un conjunto de herramientas casi obligatorio para todos los tradutores que deseen tener su lugar en el mercado competitivo actual. Por último, han sido desarrollados objetos de aprendizaje (LO, por sus siglas en inglés) sobre estas 11 herramientas CAT en el marco de un estudio de doctorado de la Universidad Abierta de Portugal (UAb). A finales del año de 2017, estos 11 LO estarán disponibles para la comunidad académica a partir del repositorio de la UAb en la forma de recursos educativos abiertos.

## INTRODUÇÃO

Já há cerca de dez anos, Elimam afirmou que um Tradutor não é mais um profissional que fica sentado em frente a uma mesa com uma caneta e folhas de papel e com vários dicionários ao alcance da mão, mas sim que ele se tornou alguém que utiliza um computador com muitos dicionários eletrônicos e de acesso via Internet, para além de outros programas informáticos que o auxiliam na sua lida diária com documentos dos clientes em muitos formatos diferentes (Elimam, 2007: 4).

Conforme similarmente expresso pela Tradutora Tamara Barile numa revista especializada no setor:

"[...] seria inconcebível ver, hoje, alguém trabalhando como trabalhávamos há 30 anos: usando o computador sem utilizar plenamente seus recursos (ou seja, utilizando-o como uma máquina de escrever), deixando de aproveitar plenamente o que a Internet tem a oferecer de bom ou, ainda, perdendo contatos com clientes e colegas por se recusar a utilizar o formato eletrônico de comunicação. [...] Todas essas mudanças exigem do tradutor um esforço constante de atualização e exigem que ele esteja aberto às novas formas de trabalho e de contatos profissionais." (Barile, 2006: 133-134)

De facto, existem muitas e variadas ferramentas informáticas de apoio à tradução disponíveis aos tradutores atualmente. Essas ferramentas visam auxiliar os Tradutores modernos a resolverem os problemas impostos pelas variadas e complexas exigências dos clientes. Alguns desses problemas, com que muitos tradutores de textos escritos deparam-se corriqueiramente, podem ser encontrados na leitura das mensagens diárias da lista eletrônica de discussão Trad-Prt, alojada na plataforma 'Yahoo Groups', disponível em <http://br.groups.yahoo.com/group/trad-prt/>, da qual extraímos uma pequena amostra (mantendo adequadamente um sigilo dos nomes dos autores):

"Alguém pode me dar uma dica para contar as palavras/caracteres num documento pdf que um cliente me mandou para orçar? [...] Há alguma outra forma de contar os caracteres ou vou ter que pedir para o cliente mandar de novo em formato Word, mas sem os caracteres de formatação?"

"PessoALL: Eu recebi uma proposta de tradução de um site, mas tenho que baixar todo o site para meu PC. Parece já ter rolado na lista informações sobre ferramentas que copiam o site todo; andei procurando, mas não encontrei. Agradeço se puderem me informar."

"Alguém conhece um software que permita contar palavras de um website? Procurei em discussões antigas e encontrei uma sugestão para sacar o código html e colar página a página no bloco de notas, mas isto parece-me muito moroso só para fazer um orçamento..."

"Olá. Recebi um documento grande em pdf e preciso fazer orçamento[,] mas não consigo convertê-lo para poder fazer em Word, por exemplo. Alguém teria uma solução para esta situação?"

No decorrer deste artigo, vamos apresentar ferramentas informáticas que visam solucionar os supracitados problemas, bem como outros tantos enfrentados pelos colegas rotineiramente. Vamos mostrar também o resultado de um estudo que indica a ferramenta informática de apoio ao tradutor considerada a mais importante de todas, junto com outras dez coadjuvantes no auxílio à nossa nobre profissão de derrubar as barreiras linguísticas entre os povos.

## **TIPOS DE FERRAMENTAS INFORMÁTICAS DE APOIO À TRADUÇÃO**

As ferramentas informáticas de apoio à tradução podem ser geralmente definidas como um programa ou uma aplicação (Brasil: aplicativo) de computador que auxilia na tradução de documentos de texto de maneira mais eficiente e rápida do que no caso de usar meios dos quais ainda se valem alguns tradutores 'à moda antiga': papel, caneta e/ou máquina de escrever (datilografia), para além do próprio cérebro, não eletrônico, naturalmente.

Em função desta definição geral, as ferramentas informáticas de apoio à tradução podem ser classificadas em três tipos principais, quais são:

**1) Programas/aplicações informáticas ou sítios eletrônicos na Internet utilizados pelos tradutores, mas cujas funções também são úteis aos profissionais de outros setores.**

Exemplos: reconhecimento óptico de caracteres, motores de busca, tais como Google, Bing, IntelliWebSearch e aplicações informáticas que convertem diversos formatos de ficheiros de texto e imagem (TXT, DOC, RTF, PDF, JPG, TIFF etc.), editam e criam legendas de vídeos e filmes, entre outros. Claro está que este tipo de ferramenta informática não é exclusivamente voltado para utilização por tradutores, mas os profissionais de todos os outros domínios do conhecimento também se valem dessas ferramentas para as suas mais diversas finalidades de estudo, trabalho e entretenimento.

No que respeita ao supracitado tipo 1), as ferramentas informáticas mais utilizadas pelos tradutores incluem importantes fontes de consulta virtual (*on-line*), tais como: periódicos publicados para domínios do conhecimento especializados, glossários, *corpora* linguísticos, tradutores eletrónicos (Google Translator, Bing Translator, Lingoos, entre outros), sem falar nos inúmeros dicionários monolíngues, bilíngues e multilíngues disponíveis em linha, fontes estas que, basicamente, apresentam uma forma de acesso bastante intuitiva, com base na linguagem de marcação de hipertexto e de outras formas modernas de programação direcionadas para a rede mundial de computadores.

Num amplo escopo, estas fontes são recursos e não aplicações informáticas do tipo que se podem descarregar e instalar no computador do utilizador para a sua posterior utilização autónoma (*off-line*). Fundamentam-se num funcionamento em linha, em motores especializados de busca e em recursos voltados para tradutores, tais como: léxicos (EUR-Lex [<http://eur-lex.europa.eu/homepage.html?locale=pt>] e IATE - InterActive Terminology for Europe [<http://iate.europa.eu/>], *thesaurus* (listas de sinónimos, agrupados por conceitos), conjugadores de verbos como aquele disponibilizado em linha pelo Jornal Le Figaro [<http://leconjugueur.lefigaro.fr>], gramáticas etc.).

**2) Programas/aplicações informáticas ou sítios eletrónicos na Internet exclusivamente dedicados à tradução propriamente dita, valendo-se de memórias de tradução e da sua gestão.**

Exemplos: SDL/Trados, WordFast, Déjà Vu, MemoQ, CafeTrans, Across, Star Transit, Alchemy Catalyst, MetaTaxis, OmegaT, WordFisher, ForeignDesk, Felix, entre outros.

De facto, este tipo de ferramentas representa os 'cavalos de batalha' de todo tradutor moderno. Ao passo que ainda possa haver alguns tradutores que resistem à sua utilização, estes já não têm espaço no atual mercado de trabalho dinâmico em que vivemos, nomeadamente, no setor de traduções técnicas. Tal como no caso dos escritores/autores de livros que ainda criam as suas obras literárias utilizando as antigas máquinas de escrever, os tradutores que não utilizam uma ferramenta informática de memória de tradução estão restritos a nichos do mercado, nos quais a dinâmica de apresentação de resultados não tem a mesma velocidade exigida pela maior fatia do mercado da tradução comercial e/ou técnica.

Conforme apontado por DePalma *et al.* (2016), no Relatório Anual sobre o Mercado de Serviços Linguísticos de Tradução, Localização e Interpretação, elaborado pela empresa Common Sense Advisory, o faturamento do setor em 2016 ficou à volta de 40 mil milhões (Brasil: 40 bilhões) de dólares americanos, num mercado altamente competitivo em que os tradutores devem estar sempre atentos ao trinómio, preço/qualidade/celeridade, nos tratos mantidos com os seus clientes. Em todos estes três fatores, as ferramentas informáticas de apoio à tradução são fulcrais para que os tradutores prestem serviços profissionais de excelência.

De várias maneiras distintas, as ferramentas de memória de tradução são as que tratam diretamente dos textos fontes. Contudo, como os textos fontes podem ser entregues ao tradutor em diferentes formatos que não são passíveis de utilização direta nas supracitadas ferramentas (no anacrónico papel, ou em ficheiros de imagem ou PDF, por exemplo), estas podem exigir todo um trabalho de preparação dos textos fontes para que o trabalho de tradução propriamente dito flua de maneira produtiva e eficiente. Nesse trabalho de preparação à tradução, enquadram-se as ferramentas informáticas do terceiro e último tipo apontado a seguir, as quais igualmente dão um

importante contributo para o bom atendimento das exigências de preço, qualidade e celeridade de entrega das traduções aos clientes.

### **3) Programas/aplicações informáticas ou sítios eletrónicos na Internet exclusivamente dedicados a auxiliarem na preparação de textos para tradução.**

Alguns exemplos desse tipo de ferramentas são: aplicações informáticas que extraem texto de ficheiros em PDF, que alinham textos bilíngues para a criação de unidades de tradução e alimentação dos programas de memória de tradução, que agregam *corpora* de textos bilíngues para consulta terminológica, bem como sítios eletrónicos na Internet que excluem códigos de HTML, retornos de carro e outros símbolos que prejudicam a adequada segmentação do texto para os programas de memória de tradução, entre outros.

Este supracitado tipo 3), nomeadamente, as ferramentas informáticas de auxílio à preparação de textos para tradução, com utilização gratuita, será o alvo da nossa consideração alargada. Este tipo 3) abrange programas/aplicações computadorizadas e/ou sítios eletrónicos da Internet que não utilizam/gerem memórias de tradução, mas a sua principal função (ou funções) consiste em fornecerem apoio ou, no jargão informático, suporte ao trabalho das ferramentas do, anteriormente citado, tipo 2).

### **QUAL É A FERRAMENTA INFORMÁTICA DE APOIO À TRADUÇÃO MAIS IMPORTANTE PARA OS TRADUTORES?**

Para respondermos esta pergunta, o grau de importância de diferentes ferramentas informáticas de apoio à tradução foi avaliado perante duas comunidades em linha de tradutores de/para a língua portuguesa, por meio de um inquérito por questionário, alojado na plataforma LimeSurvey da Universidade Aberta de Portugal (<http://inqueritos.lead.uab.pt>). O inquérito em causa foi iniciado em 16 de fevereiro de 2016 e encerrado em 4 de abril do mesmo ano.

A seguir, apresentamos os resultados gerais do inquérito que teve, como público-alvo, o universo de tradutores membros da lista de discussão eletrónica, Trad-Prt, do 'Yahoo Groups', para a língua portuguesa, bem como de um segundo grupo de Tradutores provenientes da rede social Facebook, 'Tradutores com Vida'.

Identificámos e elencámos, com os valores percentuais respectivos, 11 ferramentas informáticas de apoio à tradução tidas como as mais importantes pelo público-alvo do inquérito por questionário, conforme a seguir:

1º lugar: ferramenta que extrai/converte o texto de ficheiros/arquivos no formato PDF tornando-o, assim, passível de tradução por meios eletrónicos. Exemplo: Xpdf, PDF to Word, entre outras (54,88%);

2º lugar: ferramenta que conta palavras e caracteres em ficheiros/arquivos nos formatos RTF, MSWord, TXT etc., para fornecer orçamentos automaticamente. Exemplo: FreeBudget (34,15%);

3º lugar: ferramenta que busca termos, expressões e segmentos em múltiplas referências/corpora bilíngues, em diversos formatos (texto tabulado, memórias TMX, txt, ttx, rtf, doc, itd e csv). Exemplo: ApSIC Xbench (31,71%);

4º lugar: ferramenta que mescla o conteúdo de vários ficheiros/arquivos de texto pequenos em somente um, para facilitar sua tradução em programas do tipo Trados, MemoQ, WordFast, DéjàVu e outros. Exemplo: A.F.7 Merge Your Files (25,51%);

5º lugar: ferramenta que extrai o texto de glossários bilíngues no formato CSV (valores separados por vírgulas) e converte-o para o formato TMX, compatível com programas de memória de tradução anteriormente citados. Exemplo: CSV to TMX Converter (20,73%, empate técnico com a ferramenta infra);

6º lugar: ferramenta para gestão/administração de projetos/clientes, criação de orçamentos e faturas, entre outras tarefas específicas para tradutores. Exemplo: GlobalSight (20,73%, empate técnico com a ferramenta supra);

7º lugar: ferramenta que elimina as marcas de 'nova linha' (PT-PT) ou 'retornos de carro' (PT-BR) ('hard returns' ou 'line feeds', em inglês), os marcadores de resposta em e-mails (>), códigos de HTML e outros símbolos de texto prejudiciais à segmentação correta dos programas de memória de tradução. Exemplos: eCleaner/Codezapper (19,51%);

8º lugar: ferramenta que converte diversos tipos de unidades de medida (centímetros/polegadas, Kg/libras, Celsius/Fahrenheit etc.) para melhor localização da tradução. Exemplo: Convert (by Joshua F. Madison) (17,07%);

9º lugar: ferramenta que copia todo o conteúdo de um 'site' para navegação/tradução 'off-line', cálculo de orçamento de tradução, entre outras finalidades. Exemplo: WebReaper (14,63%);

10º lugar: ferramenta que gere/administra memórias de tradução, fazendo a troca de um idioma para outro, entre outros recursos de manutenção. Exemplo: ENLASO tools (9,76%);

11º lugar: ferramenta que extrai somente o texto de páginas no formato HTML para sua tradução. Exemplo: CatsCraddle (8,54%).

Em seguida, apresentamos uma Tabela que lista os diferentes sítios na Internet donde poderão ser descarregadas gratuitamente as supracitadas aplicações informáticas de apoio à tradução.



Tabela 1 - Localizadores na Internet donde podem ser descarregadas gratuitamente as aplicações informáticas de apoio à tradução.

Ferramenta	Descrição	Disponível em
Xpdf	Ferramenta que extrai/converte o texto de ficheiros/arquivos no formato PDF tornando-o, assim, passível de tradução por meios eletrónicos	<a href="https://tinyurl.com/mgb9a">https://tinyurl.com/mgb9a</a>
FreeBudget	Ferramenta que conta palavras e caracteres em ficheiros/arquivos nos formatos RTF, MSWord, TXT etc., para fornecer orçamentos automaticamente	<a href="https://tinyurl.com/zl4ylh5">https://tinyurl.com/zl4ylh5</a>
ApSIC Xbench	Ferramenta que busca termos, expressões e segmentos em múltiplas referências/corpora bilíngues, em diversos formatos (texto tabulado, memórias TMX, txt, ttx, rtf, doc, itd e csv)	<a href="http://www.apsic.com">www.apsic.com</a>
A.F.7 Merge Your Files	Ferramenta que mescla o conteúdo de vários ficheiros/arquivos de texto pequenos em somente um, para facilitar sua tradução em programas do tipo Trados, MemoQ, WordFast, DéjàVu e outros	<a href="https://tinyurl.com/z82pr6y">https://tinyurl.com/z82pr6y</a>
CSV to TMX Converter	Ferramenta que extrai o texto de glossários bilíngues no formato CSV (valores separados por vírgulas) e converte-o para o formato TMX, compatível com programas de memória de tradução anteriormente citados	<a href="https://tinyurl.com/z8aoe3a">https://tinyurl.com/z8aoe3a</a>
GlobalSight	Ferramenta para gestão/administração de projetos/clientes, criação de orçamentos e faturas, entre outras tarefas específicas para tradutores	<a href="http://www.globalsight.com/">http://www.globalsight.com/</a>
E-cleaner	Ferramenta que elimina as marcas de 'nova linha' (PT-PT) ou 'retornos de carro' (PT-BR) ('hard returns' ou 'line feeds', em inglês), os marcadores de resposta em e-mails (>), códigos de HTML e outros símbolos de texto prejudiciais à segmentação correta dos programas de memória de tradução	<a href="https://tinyurl.com/houkcnp">https://tinyurl.com/houkcnp</a>
Convert	Ferramenta que converte diversos tipos de unidades de medida (centímetros/polegadas, Kg/libras, Celsius/Fahrenheit etc.) para melhor localização da tradução	<a href="https://tinyurl.com/3avu8sg">https://tinyurl.com/3avu8sg</a>
WebReaper	Ferramenta que copia todo o conteúdo de um 'site' para navegação/tradução 'off-line', cálculo de orçamento de tradução, entre outras finalidades	<a href="http://www.webreaper.net">www.webreaper.net</a>
ENLASO Tools	Ferramenta que gere/administra memórias de tradução, fazendo a troca de um idioma para outro, entre outros recursos de manutenção	<a href="https://tinyurl.com/j297m2q">https://tinyurl.com/j297m2q</a>
CatsCradle	Ferramenta que extrai somente o texto de páginas no formato HTML para sua tradução	<a href="https://tinyurl.com/osfpbes">https://tinyurl.com/osfpbes</a>

## OBJETOS DE APRENDIZAGEM SOBRE FERRAMENTAS INFORMÁTICAS DE APOIO À TRADUÇÃO

As ferramentas listadas na Tabela 1 deveriam compor um conjunto quase obrigatório para todos os Tradutores que desejam ter um lugar no mercado competitivo dos dias de hoje. Infelizmente, no que respeita ao ensino da sua utilização por parte das Universidades a todos aqueles que decidem cursar o Ensino Superior em Línguas, com especialização em Tradução, ainda há um longo caminho a percorrer para que os estudantes tirem a sua graduação dominando adequadamente a utilização dessas ferramentas.

Conforme evidenciado por alguns comentários extraídos da já mencionada lista de discussão eletrónica, Trad-Prt, do Yahoo Groups:

"[...] parece que os professores das escolas de tradução ignoram solenemente o mercado que seus alunos vão enfrentar aqui fora e os preparam somente para a tradução dita literária que representa uma ínfima parte do mercado [...]."

"Este debate demonstrou como muitas escolas de tradução estão totalmente desligadas da realidade. Não entendo como um formando em tradução não receba ao menos instrução básica acerca das ferramentas informatizadas com que irá trabalhar no futuro. Computador é atualmente, não somente para o tradutor, mas também para dezenas de outros profissionais, o que o carro é para o vendedor que faz visitas aos clientes; não se concebe um vendedor que visite seus clientes fazê-lo usando o transporte público."

"Quem pensa que universidade ensina o suficiente para sermos bem-sucedidos no mercado, ou é muito ingênuo ou está cometendo um dos maiores enganos já vistos entre novatos. Precisamos aprender e apreender o que é ensinado, sempre e \*para sempre\*, em todos os ambientes possíveis." O problema é que vocês só descobrem isso quando terminam o curso de tradução onde ouviram durante 4 anos que um glossário -- como a cruz e o catecismo dos jesuítas que se embrenhavam na selva -- resolveriam os seus problemas."

Tendo-se em conta essa necessidade do ensino prático das ferramentas informáticas de apoio à tradução, objetos de aprendizagem (OA) sobre a utilização dessas onze ferramentas, identificadas como importantes no inquérito ao público-alvo pertinente ao campo de conhecimento em causa, foram desenvolvidos no âmbito de um estudo de Doutorado em Educação a Distância e E-Learning pela Universidade Aberta de Portugal (UAb). Ao final do ano de 2017, esses onze OA estarão disponíveis para a comunidade académica, a partir do repositório da UAb, na forma de recursos educacionais abertos (<https://repositorioaberto.uab.pt/>).

Figura 1 - Logotipo do Repositório Institucional da Universidade Aberta de Portugal



Os OA foram criados por meio do programa informático eXe Learning, versão 2, um editor de textos no formato XHTML - Extended Hypertext Markup Language (Linguagem Estendida de Marcação de Hipertexto) e disponível gratuitamente em: <http://exelearning.net>.

Figura 2 - Logotipo da aplicação de autoria/edição de conteúdos educativos eletrônicos eXe Learning



A interface dos OA foi finalizada com o auxílio da aplicação de autoria de conteúdos para a Web, KompoZer, versão 0.7.10, compilação 2007083, disponível gratuitamente em: <http://kompozer.sourceforge.net>. Os protótipos dos Objetos de Aprendizagem passaram por uma avaliação da usabilidade e da reusabilidade por parte de peritos no setor em causa (Hix & Hartson, 1993; Nielsen, 1993; Whitehead, 2006) e foram aperfeiçoados concordemente às observações e sugestões oferecidas.

Esses OA poderão, então, servir como modelos para o desenvolvimento paulatino de outros OA sobre as diversas aplicações informáticas específicas para tradutores, tendo-se em conta as possíveis dificuldades da introdução imediata do seu ensino, como no caso de uma extensa atualização dos conteúdos curriculares. Essas dificuldades podem ser causadas pelas diferentes curvas de aprendizagem de utilização das ferramentas por parte dos Professores, por recursos informáticos disponíveis de maneira limitada ou deficiente nas Instituições de Ensino Superior, entre outros fatores restritivos à efetiva implementação do ensino das ferramentas informáticas de apoio à tradução na prática pedagógica regular dos Cursos de Línguas.

A seguir, apresentamos as capturas dos ecrãs principais de que se compõe o Objeto de Aprendizagem sobre a ferramenta informática de apoio à tradução considerada a mais importante de todas, lembrando que existem outras com função semelhante, igualmente gratuitas ou sob pagamento de uma licença de utilização. A função da ferramenta é que está em destaque: extração de textos de ficheiros em PDF (formato de documento portátil).


Figura 3 (coletivamente) - Objeto de Aprendizagem sobre a Ferramenta Xpdf

Início

**Ferramenta para extração de textos de ficheiros no formato PDF**

Olá, Tradutor(a)!

Seus clientes já lhe enviaram alguma vez textos no formato PDF para serem traduzidos?




O formato PDF (Portable Document File, ou Ficheiro de Documento Portátil) não permite a utilização direta de programas de memória de tradução, tais como: Trados/SDLX, Wordfast, MemoQ, DéjàVu, entre outros.

Neste estudo prático, vamos aprender como extrair o texto de um ficheiro PDF para, em seguida, poder traduzi-lo por meio de um qualquer programa de memória de tradução e, assim, realizar o trabalho com qualidade e celeridade!

Lembre-se: quando necessário, basta fazer clique nas hiperligações disponíveis e poderá avançar e voltar durante o percurso deste estudo dirigido.


**Este material ficará disponível no Repositório Educacional Aberto para a sua posterior consulta, quando desejar.**

**Bom proveito e bom estudo!**





Sob licença [Licença Creative Commons Atribuição Não-Comercial Compartilha Igual 4.0](#)

Objetivo




**Objetivo**

Converter o texto contido num ficheiro no formato PDF para o formato TXT (texto puro), de forma a tornar possível traduzi-lo posteriormente mediante um programa informático de memória de tradução.



Sob licença [Licença Creative Commons Atribuição Não-Comercial Compartilha Igual 4.0](#)



Requisitos



**Requisitos**

- Ferramenta informática de auxílio à tradução, Xpdf, para extrair e converter o texto no formato fechado à edição/tradução, em PDF, para o formato aberto TXT (texto puro).
- Ficheiro PDF de exemplo para a conversão: Amostra.pdf, que contém o texto do documento "Introducing IT in Translator Training: Experiences from the COLC Project", de Jaatinen H. & Jääskeläinen R. (2003).

As atividades descritas a seguir fornecerão todas as condições e as instruções necessárias para a instalação e a utilização da ferramenta informática e do ficheiro PDF em causa.



Sob licença [Licença Creative Commons Atribuição Não-Comercial Compartilha Igual 4.0](#)

Figura 3 (continuação) - Objeto de Aprendizagem sobre a Ferramenta Xpdf

Atividade 1

**Instalar a aplicação Xpdf:**

Antes de começar a atividade, por favor, leia as instruções até o fim.




Faça clique [aqui \(xpdf.zip\)](#) e copie o ficheiro, compactado no formato zip, para uma pasta no seu computador.

Em seguida, tome nota do percurso completo até o ficheiro em causa, para que não se esqueça para onde o copiou. Se o seu sistema operativo abrir automaticamente o ficheiro zip, copie os ficheiros descompactados para a pasta da sua preferência.

Por exemplo: C:\Xpdf\xpdf.zip (para facilitar a compreensão, a pasta 'Xpdf' foi criada diretamente na raiz 'C:\', mas poderá utilizar qualquer outro sítio apropriado no disco duro do computador, conforme instruído pelo sistema operativo, tais como: 'C:\Programas\Xpdf', 'C:\Ficheiros de programas\Xpdf', 'C:\Programs\Xpdf', 'C:\Program Files\Xpdf', entre outras opções).

No caso de o computador ainda não ter executado a supracitada operação, abra o ficheiro zip utilizando o seu descompactador preferido (Winzip, WinRar etc.) e guarde todos os ficheiros resultantes na mesma pasta que escolheu utilizar.

Se tiver dúvidas durante os procedimentos em causa, retorne a esta página sempre que for necessário.



Sob licença [Licença Creative Commons Atribuição Não Comercial Compartilha Igual 4.0](#)

Atividade 2

**Converter o texto de PDF para TXT**

Após descompactar os ficheiros da ferramenta Xpdf para a pasta da sua preferência, vamos agora ao processo, propriamente dito, de conversão do texto contido no documento com o formato PDF para o formato TXT.

Faça o seguinte:

- Faça clique [aqui \(Amostra.pdf\)](#) e copie o ficheiro de exemplo da conversão para a pasta Xpdf que foi criada na Atividade 1. Caso tenha esquecido onde essa pasta está no computador, consulte os seus apontamentos da Atividade 1
- Faça clique no botão **Iniciar** do Windows e, em seguida, em **Executar**. Na caixa de diálogo **Executar**, faça clique no botão **Procurar**. Na caixa de diálogo **Procurar**, utilize o explorador de ficheiros do Windows para ir até a pasta onde copiou os ficheiros da aplicação Xpdf, conforme realizado na Atividade 1.
- Localize o ficheiro executável, **pdftotext.exe**, e faça clique duplo sobre o nome do ficheiro.
- Na linha de comandos que aparece na caixa de diálogo **Executar** (C:\Xpdf\pdftotext.exe), acrescente os seguintes parâmetros de processo ao final da linha: **-layout Amostra.pdf**.

O comprimento total da linha de comando deverá ficar assim, conforme ilustrado infra e nas figuras:

C:\Xpdf\pdftotext.exe -layout Amostra.pdf



**Observação:** a ferramenta Xpdf não é capaz de extrair textos contidos em ficheiros PDF gerados a partir de imagens ou protegidos por palavra-passe, quer seja esta para abrir elou modificar o conteúdo.

- Na caixa de diálogo **Executar**, faça clique em **OK** (**Aceitar**). A ferramenta criará automaticamente o ficheiro Multimedia.bit (já no formato de texto puro, procurando manter o 'layout' (diagramação/composição) original do ficheiro PDF. A ferramenta Xpdf criará o ficheiro de texto na mesma pasta C:\Xpdf\ (obviamente, desde que esteja a utilizá-la para o efeito deste exemplo).

**PRONTO! O texto já pode ser traduzido por meio da sua aplicação favorita de memória de tradução! Problema resolvido!**



Sob licença [Licença Creative Commons Atribuição Não Comercial Compartilha Igual 4.0](#)

Para os restantes dez Objetos de Aprendizagem sobre as demais ferramentas informáticas de apoio à tradução que foram objeto deste nosso estudo, adotámos igualmente um desenho instrucional na forma de estudo dirigido e que é composto pelas fases de Introdução, Objetivo, Requisitos e Atividades (tantas quantas forem necessárias para o alcance do Objetivo proposto).

A aplicação eXe Learning atribui a estas fases a nomenclatura de *iDevices* (dispositivos instrucionais) e as elaborou em função do formato educacional padrão SCORM (*Shareable Content Object Reference Model*, ou Modelo de Referência de Objetos para Conteúdos Partilháveis), para facilitar a partilha desses conteúdos em sistemas de ensino e aprendizagem pela via eletrónica (*e-learning*).

Conforme o modo escolhido para a criação e o encapsulamento dos dispositivos instrucionais (*iDevices*), proporcionado pela ferramenta eXe Learning, um objeto de aprendizagem pode conter somente alguns *iDevices*, ou tantos quantos forem necessários, de acordo com o objetivo pedagógico do(a) professor(a) no que respeita à construção eficaz do conhecimento para os seus estudantes.

Tendo-se em conta as características de suficiente simplicidade, pretendidas para os nossos OA, estes foram criados com apenas quatro tipos de dispositivos instrucionais, já mencionados: Introdução, Objetivo, Pré-Requisitos e Atividades. No desenvolvimento dos OA, não incluímos *iDevices* voltados para qualquer tipo de Avaliação do ensino-aprendizagem, posto que esta faceta pedagógica deverá ser desenhada e realizada em função de fatores muito específicos que cada professor precisará de observar individualmente, no seguimento das exigências da própria Instituição de Ensino Superior para a qual trabalha.

Tendo-se em conta que uma significativa parte do trabalho do tradutor moderno baseia-se na intensa utilização da Internet, os OA propostos incluem na sua estrutura simulações realísticas das mais diversas situações práticas com que se depararão os futuros profissionais tradutores no seu trabalho em ambiente virtual. Os Objetos de Aprendizagem serão disponibilizados gratuitamente à comunidade académica, com as ressalvas impostas pela Licença Creative Commons, no Repositório de Recursos Educacionais Abertos da Universidade Aberta de Portugal.

Para além dessa facilidade de utilização (e reutilização), a maior parte das aplicações e ferramentas informáticas tratadas pelos OA desenvolvidos também não exige a compra de licenças nem o emprego de computadores potentes para o seu funcionamento. Isto igualmente facilitará a adoção do seu ensino na prática pedagógica dos Professores que assim o pretenderem.

## CONCLUSÃO

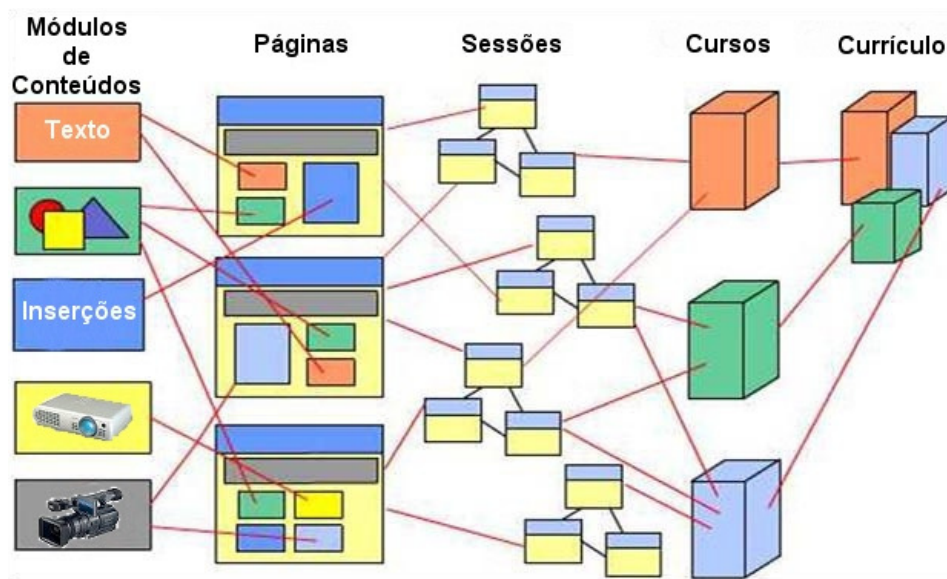
Novos paradigmas educativos emergem da aplicação pedagógica das Novas Tecnologias de Informação e Comunicação no ambiente escolar, com vistas a uma criação flexível e colaborativa do conhecimento voltada para as necessidades dos utilizadores sublinhando, portanto, a importância do sujeito no desenvolvimento da aprendizagem (Dias *et al.*, 2004). Conforme apontam estes mesmos autores:

"As exigências dos novos ambientes de formação baseados na Web implicam uma infoliteracia quer no plano dos utilizadores, quer no plano dos formadores. Para estes últimos, os ambientes virtuais de aprendizagem na Web, implicam também o domínio das tecnologias de educação na Web, que vão desde a concepção dos contextos e situações autênticas de aprendizagem às estratégias de organização da flexibilidade cognitiva das representações; desde o acompanhamento do formando ao seu envolvimento no grupo virtual; e desde a tutoria na aprendizagem individual à gestão do desenvolvimento da aprendizagem colaborativa." (Dias *et al.*, 2004: 2)

O alcance exitoso desta infoliteracia, identificada por estes autores como uma exigência básica, tanto para o corpo discente como para o docente, impõe um desafio a ser ultrapassado. Tendo-se em conta as características intrínsecas das ferramentas informáticas de apoio à tradução e do ambiente virtual em que se efetua o trabalho dos tradutores modernos, o ensino por meio dos Objetos de Aprendizagem em causa torna-se particularmente suscetível de inserção num programa de educação e ensino a distância, nomeadamente mediante ferramentas de ensino eletrónico (*e-Learning*), implantadas e disponibilizadas no contexto das plataformas educativas, com a possibilidade de integrarem-se nas redes internacionais de Recursos Educacionais Abertos e Repositórios relacionados.

Pela consulta desses Recursos Educacionais Abertos e transferência dos OA a partir dos Repositórios afins, o trabalho dos docentes na preparação das aulas será também em muito facilitado, pois os OA poderão ser utilizados e reutilizados em função das necessidades, das competências informáticas dos corpos discente e docente, bem como dos problemas levantados pelas próprias classes dos cursos de Línguas - Tradução.

Figura 4. Mapa conceitual sobre o processo de criação de Objetos de Aprendizagem para sua integração no currículo de ensino



Assim, o presente artigo apresenta uma proposta de implementação facilitada do uso dos Objetos de Aprendizagem em disciplinas específicas sobre as Ferramentas Informáticas de Apoio à Tradução, para uma renovação das grelhas curriculares dos Cursos Superiores de Línguas tradicionais, presenciais, bem como de educação e ensino a distância e *e-Learning*, conforme ilustrado na sua generalidade pela Figura 4 - desenho adaptado de Santally & Senteni (2005).

Esta maior facilidade de utilização inovadora dos OA na sua prática pedagógica está relacionada ao facto de que os docentes não terão a necessidade de serem 'gurus' da Informática para tirarem bom proveito desses Objetos de Aprendizagem e poderão concentrar-se na sua atividade precípua que é a do ensino. Para os estudantes, os benefícios advirão de adquirirem competências que irão ser realmente de utilidade prática no mercado, no âmbito da sua futura profissão de tradutor com formação Superior. Deste modo, não haverá para estes recém-formados a necessidade de pagarem por cursos extracurriculares cuja existência está diretamente relacionada a essa lacuna pedagógica deixada pelas Instituições de Ensino Superior, como evidenciado inicialmente nos 'bate-papos' informais nas listas de discussão e redes sociais de tradutores.



## REFERÊNCIAS BIBLIOGRÁFICAS

Barile, T. (2006). As Novas Tecnologias e a Necessária Adaptação do Tradutor. Mesa redonda do 7<sup>o</sup> Prêmio União Latina/CBL de Tradução Especializada (2006): Seminário organizado na Bienal Internacional do Livro. *In* Ana Julia Perrotti-Garcia, *Confluências - Revista de Tradução Científica e Técnica*, nº 4., maio de 2006, São Paulo, Brasil, págs. 133-134.

DePalma, D. A. *et al.* (2016). *The Language Services Market: 2016 - Annual Review of the Services and Technology Industry That Supports the Delivery of Translation, Localization, and Interpreting*, June 2016 by Common Sense Advisory, Inc., Cambridge, Massachusetts, USA

Dias, P. M. B. S. *et al.* (2004). *e-Learning para e-formadores: Formação de Docentes Universitários*. Actas da Conferência eLES'04: eLearning no Ensino Superior. ISBN 972-789-134-9. Universidade de Aveiro, Portugal.

Elimam, A. S. (2007). The Impact of Translation Memory Tools on the Translation Profession. *In* *Translation Journal - Translators and Computers Section*, Volume 11, Nr. 1, January 2007, pág. 4.

Hix, D. & Hartson, H. R. (1993). *Developing User Interfaces: Ensuring Usability Through Product and Process*. Nova Iorque: John Wiley & Sons, EUA.

Nielsen, J. (1993). *Usability Engineering*. Nova Jérsei: Academic Press Inc., ISBN 0-12-518405-0, San Diego, California, EUA.

Santally, M. I. & Senteni, A. (2005). A Learning Object Approach to Personalized Web-based Instruction. *In* *European Journal of Open, Distance and E-Learning*.

Whitehead, C. C. (2006). Evaluating web page and web site usability. *In* *Proceedings of the ACM Southeast Regional Conference 2006*. Nova Iorque: Association for Computing Machinery, EUA, págs. 788-789.