

Voluntariado Literário

o projecto Página-a-Página

Por:
Rita Farinha



Exponor – 16 Abril 2010

Voluntariado

“Success is a science; if you have the conditions, you get the result.”

Oscar Wilde

O projecto conta com voluntários espalhados um pouco por todo o mundo, principalmente Portugal, Brasil, França e Luxemburgo.

Contexto do Projecto

Projecto Gutenberg (PG)
[www.gutenberg.org/pt]

Projecto Distributed Proofreaders (DP)
[www.pgdp.net]

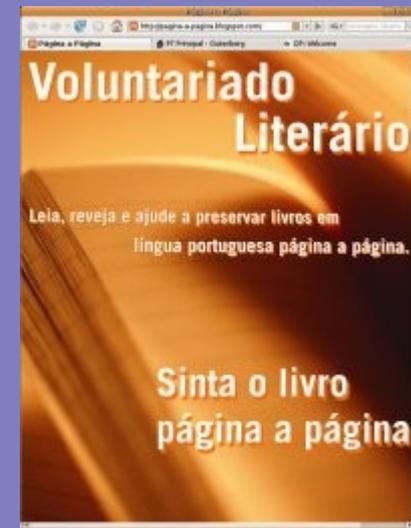
Projecto Página a Página (P-a-P)
[pagina-a-pagina.blogspot.com]

O P-a-P surgiu em 2004 como forma de promover a revisão de obras no domínio público em língua portuguesa

O que é o P-a-P?

Sítio de apoio aos Voluntários de Língua Portuguesa do Projecto DP

- ✓ Divulgação
- ✓ Estatísticas
- ✓ Esclarecimento de dúvidas
- ✓ Newsletters
- ✓ Revisão de obras através do DP



<http://pagina-a-pagina.blogspot.com>

A Revisão

Efectuada totalmente *em linha*, sem *download* de programas

- ✓ Comparação OCR/Imagem
- ✓ Revisão efectuada por pelo menos 5 voluntários

Criação de e-book – txt e html

The screenshot displays a web browser window with the URL `http://www.pgdp.net/c/tools/proofers/proof.php?projectid=projectID4608f3e479942&proj_state=F2.proj_avail`. The main content area shows a scanned image of a page with the text: (50)
suave, e ligeira. A' medida que as Artes hião subindo se foi procurando a variedade. Mas os vestidos tem sido sempre considerados como hum tecido ligeiro, cujas pregas não devião ser nem amontoadas, nem carrichosamente dispersas.

Below the image, the HTML markup for the same text is shown: `suave, e ligeira. Á medida que as Artes hião subindo se foi procurando a variedade. Mas os vestidos tem sido sempre considerados como hum tecido ligeiro, cujas pregas não devião ser nem amontoadas, nem carrichosamente dispersas.`

The interface includes several control buttons: "Save as 'In Progress'", "Save as 'Done' & Proofread Next Page", "Save as 'Done'", "Stop Proofreading", "Switch to Vertical", "Show All Text", "Return Page to Round", and "WordCheck".

At the bottom, there is a status bar with "Page: 126 — F1: [disgrace](#) View: [Project Comments](#) | [Image](#)" and "Image Resize: -25% +25% Original". A toolbar with various icons and a "HELP" button is also visible.

At the very bottom of the browser window, the text "Concluído" is on the left and "DPStat" is on the right.

O porquê do OCR / da revisão

- ✓ Permitir o acesso a pessoas com deficiência visual (através de software que lê txt/HTML)
- ✓ TXT/HTML é mais leve do que obras digitalizadas
- ✓ Aplicação a novos recursos de leitura de e-books (ex: Kindle)
- ✓ Permitir a pesquisa de texto nas obras
- ✓ Possibilidade de extracção de excertos da obra para trabalhos académicos e semelhantes
- ✓ Preservação da língua portuguesa – Facilitar o estudo da língua

Era **urna**
vez...

O OCR automático reconhece apenas 90% dos caracteres (no melhor dos casos), sendo necessária a revisão humana.

O Página-a-Página

Wikipedia

Parceiros

Estatísticas

Objectivos

Curiosidades

Wikipedia

Parcerias

Estatísticas

Objectivos

Curiosidades

Semelhanças:

- ✓ Projecto *em linha*
- ✓ Projecto multi-idioma
- ✓ Utilização simples e intuitiva
- ✓ Sem *downloads* de programas
- ✓ Sistema voluntário/distribuído
- ✓ Promoção de cultura
- ✓ Disponibilização livre de conteúdos



A Wikipédia contém ligações, nas páginas dos autores, para as obras concluídas disponíveis no Projecto Gutenberg.

Wikipedia

Parcerias

Estatísticas

Objectivos

Curiosidades

Entidades:

- ✓ Biblioteca Nacional Digital
- ✓ Intraneia - Sistemas de Informação, Lda
- ✓ Centro de Estudos Ataíde Oliveira (Univ. do Algarve)
- ✓ Projecto Natura (Univ. do Minho)
- ✓ University of Toronto Libraries

Fornecimento de obras

Projectos em parceria

Alojamento de ficheiros

Divulgação



O Projecto Dicionário resultou num Dicionário Online que em breve estará disponível também em formato Wiki [<http://dicionario-aberto.net>]

Wikipedia

Parcerias

Estatísticas

Objectivos

Curiosidades

✓ Números Gerais (a 14 Abr. 2010)

289 Voluntários

419 Obras no PG-pt

597 Participantes

153 Através do P-a-P/DP

12163 Páginas em 2010

Ano	Págs.
2005	17870
2006	33114
2007	70200
2008	50748
2009	44688



Neste momento, estão 13 obras em processamento e mais de 100 obras em espera.

Wikipedia

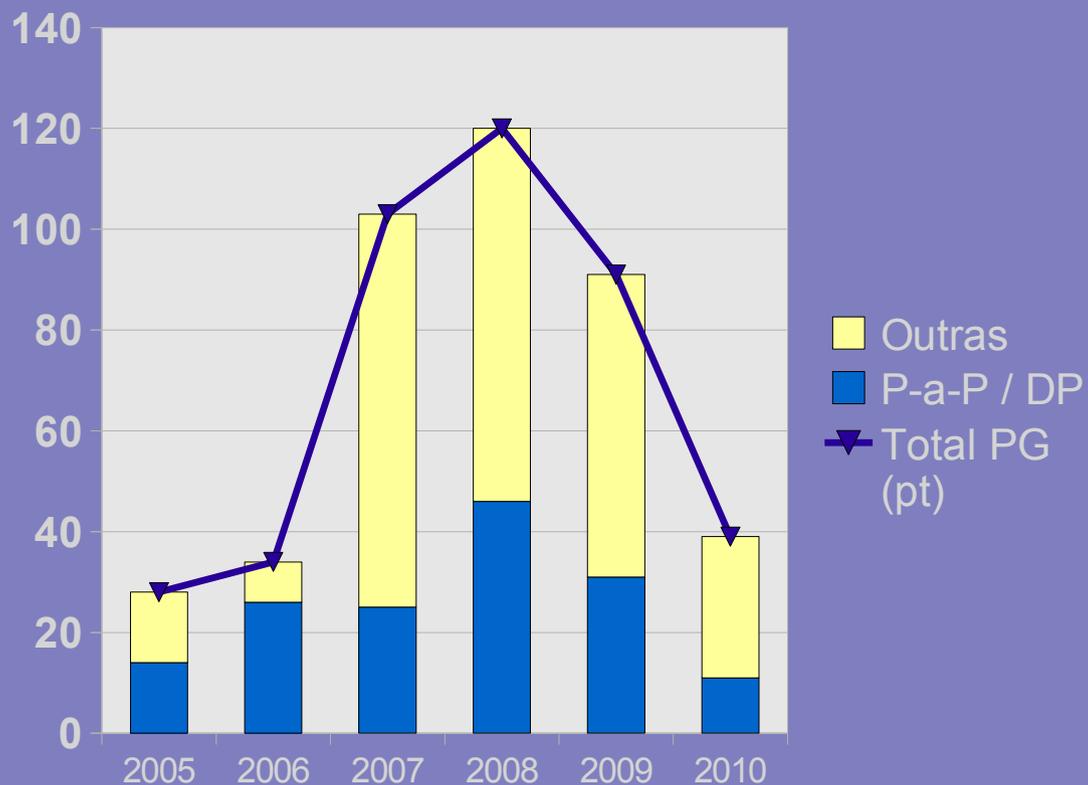
Parcerias

Estatísticas

Objectivos

Curiosidades

✓ Resultados em e-books



Wikipedia

Parcerias

Estatísticas

Objectivos

Curiosidades

Objectivos para 2010

- ✓ 50000 páginas revistas em livros de língua portuguesa (+ 5312 que em 2009)
- ✓ 300 voluntários (+ 30 que em 2009)
- ✓ 50 livros publicados no sítio do Projecto Gutenberg com origem no DP (+ 17 do que os que conseguimos em 2009)

**Início de um grande projecto:
Projecto Periódico**

Wikipedia

Parcerias

Estatísticas

Objectivos

- ✓ A obra em língua portuguesa mais descarregada

**“Noções elementares de archeologia”
(Joaquim Possidónio Narciso da Silva)**

- ✓ O autor lusófono com mais obras disponíveis

Camilo Castelo Branco

Curiosidades

- ✓ A maior obra trabalhada tinha mais de 2000 páginas

**Novo Dicionário da Língua Portuguesa
(1913)**

Em média, por mês contamos com a ajuda de 42 voluntários activos, o que resulta na revisão de cerca de 3600 páginas mensais.

Como Ajudar?

✓ Ler e Rever

1. Registo > 2. Ler Regras > 3. Começar a Rever

Lema : "Rever uma página por dia"

No início, não deverá demorar mais de 12 minutos a rever uma página, tempo que será reduzido com a prática.

✓ Divulgar

Boca a Boca | Blogues | Facebook | Twitter | Imprensa

Disponível para esclarecimentos adicionais.

Obrigada !