

# rARC – O arquivo da web portuguesa

Criado por [Vítor M.](#) em 22 de Dezembro de 2009 | [27 comentários](#)

## Sabia que a web portuguesa é guardada centralmente para a proteger em caso de catástrofe?

Podemos muitas vezes pensar que existe sempre alguém ou algo a gerir tudo o que se cria, modifica e apaga na web. Embora saibamos que nada é eterno, confiamos que esta aldeia global não está assente sobre frágeis estacas binárias. É aqui que os serviços prestados pelo **Arquivo da Web Portuguesa** ultrapassam o âmbito histórico-cultural da preservação de informação digital.

Este arquivo vai mais além da “perda material”, este arquivo contribui para a expansão do uso do português enquanto língua para comunicação na Web; contribui para a disseminação de conteúdos de interesse às diversas comunidades científicas... são provas devidamente acauteladas.



## O que é o rARC?

A Fundação para a Computação Científica Nacional (FCCN) lançou um projecto de arquivo da web portuguesa. Periodicamente a web portuguesa é recolhida e guardada num arquivo central. Este processo requer uma grande quantidade de espaço em disco para guardar cada recolha.

O rARC é um sistema que permitirá a um utilizador da Internet disponibilizar espaço em disco no seu computador para armazenar cópias de segurança de pequenas partes dos dados armazenados no arquivo central. Para este efeito apenas terá de instalar uma aplicação simples no seu computador.

Pretende-se assim que em caso de destruição do arquivo central, causada por exemplo por uma catástrofe natural, os conteúdos históricos arquivados possam ser salvos a partir das cópias de segurança distribuídas pelos utilizadores.

Configuração do rARC

**Informações do contribuidor**  
É um contribuidor registado do rARC?

Sim  
 Não

Nome:   
(será publicado no sítio Web do rARC)

Endereço de correio electrónico:

Escolher uma palavra-chave:

Confirmar palavra-chave:

< Anterior   Seguinte >   Cancelar

Quais as principais características do rARC?

O rARC deverá ser:

- Escalável  
Numa primeira fase deverá suportar milhares de utilizadores distribuídos pela Internet.
- Seguro  
Os dados contidos numa cópia de segurança não deverão estar acessíveis aos utilizadores do computador.
- Robusto  
Deverá ser robusto contra utilizadores mal intencionados e garantir a integridade das cópias de segurança.
- Fácil de usar  
Os utilizadores deverão conseguir instalar facilmente a aplicação que permitirá disponibilizar espaço para guardar as cópias de segurança.
- Configurável  
Deverá ser fácil de usar e integrar em iniciativas de arquivo independentes.

Serão feitas cópias de segurança de todo o Arquivo?

O principal objectivo do rARC é permitir que mesmo que o arquivo central se perca totalmente, a informação armazenada nos computadores dos utilizadores se salve.

Apenas será possível criar cópias de segurança para todo o arquivo se existir espaço suficiente disponibilizado pelos utilizadores. Quanto maior for a adesão por parte destes utilizadores mais informação poderemos salvar.

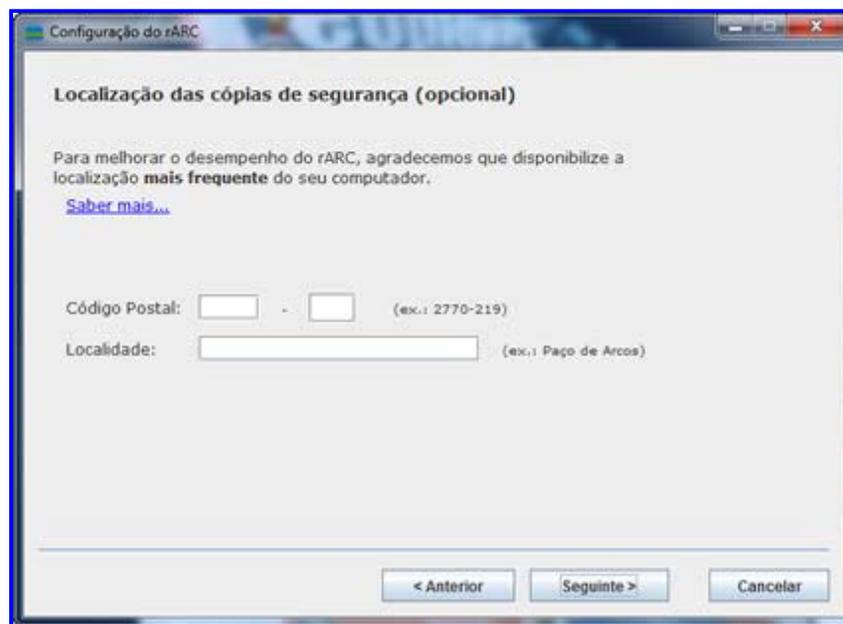
No entanto, mesmo que não seja possível copiar todo o arquivo, o rARC provará ser útil se conseguir salvar pelo menos uma parte da informação em caso de destruição do arquivo central, evitando assim a sua total perda.

### Quem poderá guardar cópias de segurança?

Qualquer indivíduo ou instituição poderá colaborar para a preservação da web portuguesa. Contamos com a colaboração dos cidadãos cientes da importância da necessidade de um esforço comum, assim como das organizações com preocupações e responsabilidades ao nível da preservação da cultura e História de Portugal.

Esperamos também que o rARC venha a ser usado também para troca de cópias de segurança entre arquivos da web. Por exemplo, o Arquivo da Web Portuguesa poderia disponibilizar espaço para armazenar conteúdos de um arquivo da web do Brasil, e este poderia fazer o mesmo pelo AWP.

Assim sendo, em caso de catástrofe, um dos arquivos da web poderia ser reconstruído a partir das cópias de segurança guardadas em países longínquos onde os efeitos da catástrofe não se tivessem feito sentir.



### Quanto espaço tenho de dar?

No mínimo 100 MB, ou seja, o espaço para guardar um ficheiro ARC.

Em Fevereiro de 2008, um computador vulgar vinha equipado com um disco de 320 GB. Uma cópia de segurança de 100 MB de conteúdos arquivados corresponde apenas a 0,03% do espaço em disco.

Quanto mais espaço oferecer, maior será a probabilidade de preservarmos a informação publicada na web para o futuro. Contamos consigo.

### O meu computador vai ficar mais lento?

O programa cliente do rARC que é instalado no computador do utilizador terá um impacto mínimo no seu desempenho. Este programa, após descarregar as cópias de segurança a partir do arquivo central, está a maior parte do tempo inactivo, realizando apenas ligações esporádicas ao arquivo central para verificar a integridade das cópias de segurança.

### Vou ter de manter a minha cópia do arquivo para sempre?

Não.

É natural que passado algum tempo as pessoas mudem de computador, decidam desinstalar aplicações ou simplesmente percam o interesse em colaborar com o projecto rARC.

No entanto, enquanto tiveram as cópias guardadas no seu disco contribuíram para a preservação da História, pois caso tivesse havido um problema com o arquivo central, estas cópias teriam sido muito valiosas para que a informação arquivada não se tivesse perdido.

Poderá reduzir o espaço oferecido quando quiser.

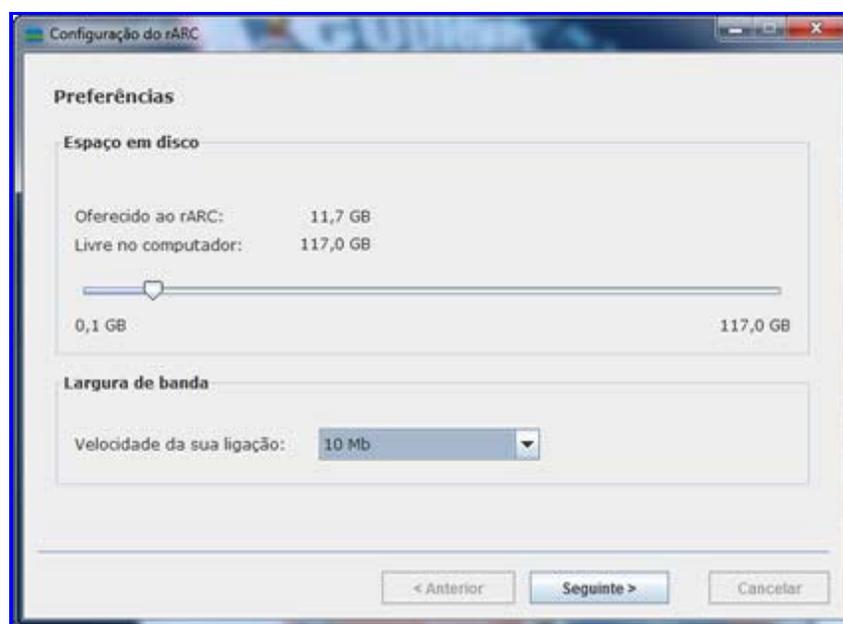
Por favor, não apague os ficheiros relacionados com o rARC directamente.

**Porque é que eu havia de dar espaço do meu computador?**

Pela mesma razão que daria qualquer outra coisa. Essencialmente, esperamos que os cidadãos com consciência da importância da História venham a aderir ao projecto.

A “recompensa” é apenas participar num esforço de preservação da cultura e desenvolvimento do país.

Quando o rARC entrar em produção iremos manter uma lista de todos os colaboradores do projecto, no presente e no passado. Assim como um top dos colaboradores mais antigos e mais generosos. Semanalmente, estará em destaque na página do projecto o nome de um colaborador do projecto escolhido aleatoriamente.



**Como funciona?**

O rARC terá uma arquitectura cliente-servidor. Um utilizador da Internet instala uma aplicação cliente no seu computador. A aplicação cliente comunica com o servidor de modo a efectuar a cópia de segurança, recuperação ou verificação do estado de uma cópia feita previamente.

Sempre que o cliente comunica com o servidor é informado acerca do processo que está a ser executado: cópia de segurança ou recuperação.

Caso o servidor esteja a realizar as cópias de segurança, o cliente irá descarregar informação arquivada.

Caso o servidor esteja a recuperar uma perda de informação, a aplicação cliente irá enviar as suas cópias de

segurança para o servidor.

Periodicamente o cliente comunica com o servidor para verificar o estado das cópias de segurança que detém, permitindo ao servidor verificar a sua integridade.

**Posso contribuir no desenvolvimento do código?**

Sim.

O rARC é um projecto código aberto. Participações técnicas para a detecção e correcção de erros, assim como desenvolvimento de novas funcionalidades são bem vindas.

**O rARC só serve para o Arquivo da Web Portuguesa?**

Não.

O rARC é um projecto de código aberto disponibilizado gratuitamente. Poderá ser usado por qualquer iniciativa de arquivo da web que armazene os conteúdos no formato ARC.

**À escala mundial existe o Internet Archive**

O [Internet Archive](#) recolhe e arquiva conteúdos Web à escala mundial. No entanto, é difícil para uma única organização fazer um arquivo exaustivo de todos os conteúdos publicados na Web, porque esta está em permanente mutação e muitos conteúdos desaparecem antes de poderem ser recolhidos para arquivo.

Acontecimentos históricos de grande importância como o Furacão Katrina originaram acções de arquivo extra-ordinárias por parte do Internet Archive, para que este acontecimento que marcou a história dos Estados Unidos da América ficasse documentado o mais exaustivamente possível.

No entanto, a documentação de acontecimentos históricos de relevância nacional para Portugal não é prioritária para o Internet Archive.

As comunidades de vários países, sensibilizadas para a urgência da preservação da informação publicada na Web de interesse nacional desencadearam iniciativas formais de preservação e catalogação da informação digital.

Decorrem em paralelo iniciativas de arquivo focadas em diferentes conteúdos e todas partilham o objectivo de preservar o conhecimento disponível na Web. Mas mesmo existindo estas várias iniciativas em curso, é difícil conseguir preservar a maioria da informação publicada na Web, devido à sua grande dimensão e ao curto intervalo de tempo em que a informação está disponível para poder ser arquivada.

O arquivo da web requer um esforço à escala mundial.



Licença: Freeware



Sistemas Operativos: Windows/Linux



Download: [rARC](#) [0.97MB]



Homepage: [rARC](#)



Homepage: [Arquivo.pt](http://Arquivo.pt)