

Modelo de arquivamento de páginas web para Portais de Periódicos: uma iniciativa experimental do Portal de Periódicos da UNICAMP

GILDENIR CAROLINO SANTOS

Universidade Estadual de Campinas

gilldenir@unicamp.br

<https://orcid.org/0000-0002-4375-6815>

JULIANO SIMIÃO CARDOZO

Universidade Estadual de Campinas

ppec@unicamp.br

RESUMO

Eixo temático: Infraestrutura tecnológica.

O avanço tecnológico provocou transformações no campo da informação, e com isso ocorreu uma grande explosão bibliográfica, que demandou tempo para organização e tratamento de documentos, além de modificar as formas de preservação e apropriação de conhecimentos para investir na preservação digital de páginas web. Nesse aspecto, o presente trabalho tem como objetivo apresentar uma iniciativa independente do projeto de preservação das páginas web do Portal de Periódicos Eletrônicos Científicos (PPEC) da Universidade Estadual de Campinas (UNICAMP), vinculado ao Sistema de Bibliotecas (SBU). O PPEC participa de redes de preservação, tais como Rede Carinia (Ibict) e a PKP PN, para buscar e garantir o acesso aberto aos seus periódicos. A partir deste contexto, para descrever a experiência de preservação de documentos, utilizou-se a metodologia experimental, com embasamento bibliográfico, para o arquivamento das páginas web do PPEC, por meio da ferramenta

Conifer, desenvolvida pela Rhizome, uma instituição sem fins lucrativos que apoia a arte e a cultura digital nata. O PPEC efetuou o cadastro no Conifer, e obteve 20 GB de espaço para hospedagem gratuita. O trabalho de captura é realizado por um bolsista sob a coordenação do bibliotecário responsável pelo PPEC. Atualmente estão preservadas as páginas dos 32 títulos de periódicos, além do site de gerenciamento de todo o PPEC, o software Open Journal System – OJS (PKP). A realização desta metodologia pelo Conifer, permite a captura das páginas para o arquivamento com a inclusão básica de metadados pré-definidos no sistema (data de captura, título da página, URL, session ID e tipo de browser). Como resultados, obtivemos até o momento, o arquivamento de 3 títulos de periódicos e das página principais do PPEC. O formato para download das páginas é o WARC, padrão de arquivo de container para armazenar conteúdo web em seu contexto original, mantido pelo International Internet Preservation Consortium (IIPC). Considera-se que este trabalho demandará tempo para sua finalização, mas já é perceptível sua relevância para a preservação e o arquivamento das páginas web do PPEC.

PALAVRAS CHAVE

Preservação web; arquivamento de páginas web; Conifer (ferramenta de arquivamento); portais de periódicos.

Web preservation; web archiving; Connifer (archiving tool); journal portals.



Modelo de arquivamento de páginas web para Portais de Periódicos: uma iniciativa experimental do Portal de Periódicos da UNICAMP

SANTOS, Gildenir Carolino. Universidade Estadual de Campinas, SBU/PPEC, Brasil

CARDOZO, Juliano. Universidade Estadual de Campinas, SBU/PPEC, Brasil

Palavras-chaves

Preservação web. Arquivamento de páginas web. Conifer (Ferramenta de arquivamento). Portais de periódicos.

Eixo temático

Infraestrutura tecnológica

Resumo

Introdução: O avanço tecnológico provocou transformações no campo da informação, e com isso ocorreu uma grande explosão bibliográfica, que demandou tempo para organização e tratamento de documentos, além de modificar as formas de preservação e apropriação de conhecimentos para investir na preservação digital de páginas web. **Objetivo:** Nesse aspecto, o presente trabalho tem como objetivo apresentar uma iniciativa independente do projeto de preservação das páginas web do Portal de Periódicos Eletrônicos Científicos (PPEC) da Universidade Estadual de Campinas (UNICAMP), vinculado ao Sistema de Bibliotecas (SBU). O PPEC participa de redes de preservação, tais como Rede Carinia (Ibict) e a PKP PN, para buscar e garantir o acesso aberto aos seus periódicos. **Metodologia:** A partir deste contexto, para descrever a experiência de preservação de documentos, utilizou-se a metodologia experimental, com embasamento bibliográfico, para o arquivamento das páginas web do PPEC, por meio da ferramenta Conifer,

desenvolvida pela Rhizome, uma instituição sem fins lucrativos que apoia a arte e a cultura digital nata. O PPEC efetuou o cadastro no Conifer, e obteve 20 gb de espaço para hospedagem gratuita. O trabalho de captura é realizado por um bolsista sob a coordenação do bibliotecário responsável pelo PPEC.

Atualmente estão preservadas as páginas dos 32 títulos de periódicos, além do site de gerenciamento de todo o PPEC, o software Open Journal System – OJS (PKP). A realização desta metodologia pelo Conifer, permite a captura das páginas para o arquivamento com a inclusão básica de metadados pré-definidos no sistema (data de captura, título da página, URL, session ID e tipo de browser). **Resultados:** Como resultados, obtivemos até o momento, o arquivamento de 3 títulos de periódicos e das páginas principais do PPEC. O formato para download das páginas é o WARC, padrão de arquivo de container para armazenar conteúdo web em seu contexto original, mantido pelo International Internet Preservation Consortium (IIPC). **Conclusão:** Considera-se que este trabalho demandará tempo para sua finalização, mas já é perceptível sua relevância para a preservação e o arquivamento das páginas web do PPEC.

Licença

