



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

---

**Da pena ao BIT: processamento arquivístico dos documentos digitais de  
Rodrigo de Souza Leão**

**Jorge Phelipe Lira de Abreu**

**Eixo temático:** Preservação das coleções físicas e digitais

**Palavras-chave:** Arquivo pessoal digital. Preservação digital. Arranjo. Descrição.

### **Introdução**

Em meados de 2013 aprovou-se candidatura para o concurso de seleção de bolsistas do Programa de Incentivo à Produção do Conhecimento Técnico e Científico na Área da Cultura da Fundação Casa de Rui Barbosa para o projeto de organização do arquivo de digital do escritor, jornalista, músico e artista plástico Rodrigo de Souza Leão (1965 - 2009). Trata-se do primeiro arquivo de natureza digital recebido pela instituição. O projeto tinha por finalidade identificar, organizar, descrever e submeter a documentação às medidas de preservação e segurança, a fim de garantir seu acesso.

A produção documental de Rodrigo de Souza Leão se insere no bojo dos avanços das tecnologias de informação e comunicação da década de 1990. Os registros até então em meios convencionais se alteram essencialmente e assumem novas características, produzidos em ambientes eletrônicos, são registrados em suportes magnéticos e ópticos, em formato digital. Portanto, o acervo, doado oficialmente ao Arquivo-Museu de Literatura Brasileira (AMLB) da Fundação Casa de Rui Barbosa (FCRB) em novembro de 2012, é composto, em sua maioria (98%), por documentos nascidos digitais. O arquivo evidencia o percurso pessoal, a sistematização da trajetória profissional, a atuação literária, jornalística e artística do titular.

O desenvolvimento do trabalho permitiu constatar que a formação de arquivos



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

---

em ambiente digital apresenta uma série de vantagens na produção, transmissão e acesso, mas, por outro lado, implica em documentos altamente sensíveis e manipuláveis, além de sujeitos à rápida obsolescência tecnológica e à fragilidade do suporte. Nesse contexto, princípios e conceitos arquivísticos como proveniência, ordem original, organicidade, confiabilidade e autenticidade ganham contornos que suscitam questões, especialmente quando levantadas a partir de um arquivo pessoal, terreno marcado pelo hibridismo semiótico e que nem sempre é processado conforme a teoria arquivística.

### **Relato da experiência**

A iniciativa de doação do arquivo de Rodrigo de Souza Leão ao AMLB partiu da própria família, por intermédio do poeta Ramon Mello, curador da obra do escritor. Inicialmente a instituição cogitou não aceitá-lo em virtude da ausência de infraestrutura tecnológica para organizar e preservar arquivos digitais. Entretanto, reconheceu-se que, além de ser um acervo representativo da Literatura brasileira contemporânea, sua aquisição significaria justamente a criação de uma demanda por essa infraestrutura tecnológica a qual instituições de guarda de acervo não podem mais ignorar.

O arquivo chegou ao AMLB em quatro tipos de mídias de transporte: disquetes (121), CDs (18); DVDs (4) e pen drive (1). Nessas mídias os documentos apresentavam-se em diversos formatos, tais como doc, pdf, eml e html (texto); bitmap, gif e jpeg (imagem fixa); mp3 e wma (áudio); wmv, avi e flv (imagem em movimento). Além do material digital, recebeu-se também uma pequena quantidade de documentos em papel (2%) e dois objetos. Em que pese a desproporção entre documentos digitais e em papel, o acervo se caracteriza como híbrido.

### **Processamento do arquivo**

Sob coordenação de Rosely Curi Rondinelli, o processamento do arquivo de Rodrigo



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

de Souza Leão deu-se no período compreendido entre agosto de 2013 e julho de 2015 e foi composto de seis etapas: recolhimento e armazenamento, identificação de proveniência e autoria, conversão para formatos de preservação, elaboração do arranjo, descrição e acesso.

### **Recolhimento e armazenamento**

Procedeu-se primeiramente ao recolhimento das mídias de transporte doadas e ao acondicionamento das mesmas. Em seguida, providenciou-se a criação, no servidor da FCRB, de um espaço específico para o arquivo com acesso restrito aos arquivistas envolvidos no trabalho. Fez-se a transferência do conteúdo das mídias de transporte para o espaço específico no servidor e, por medida de segurança, para um HD externo. Esse conjunto documental foi, então, denominado “original” e a partir dele, criou-se uma cópia de trabalho.

Ressalta-se que o servidor de uma instituição, bem como um HD externo, mesmo que robusto, não são locais adequados para a preservação de arquivos digitais. O empreendimento mais apropriado é a implantação de um Repositório Arquivístico Digital Confiável (RDC-Arq). Entretanto, a implantação de um RDC-Arq demanda o envolvimento de especialistas em preservação de documentos digitais, bem como recursos financeiros e tecnológicos dos quais a instituição não dispunha. Nesse aspecto, trabalhou-se com o viável e não com o cenário ideal.

### **Identificação de proveniência e autoria**

No início da identificação percebeu-se que no arquivo havia documentos que não provinham unicamente do titular. Mediante análise dos metadados automáticos relativos à data e autoria e do conteúdo, foi possível identificar documentos que haviam sido colecionados, pelo curador da obra de Rodrigo, no desejo de reunir em um único conjunto documental tudo o que dissesse respeito ao titular. Esse reconhecimento resultou na depuração da proveniência do arquivo bem como na



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

criação da Coleção Rodrigo de Souza Leão, feita por seu curador.

Paralelamente à questão da proveniência, constatou-se também um problema quanto à autoria dos documentos e que se configura como uma peculiaridade dos arquivos pessoais digitais. Trata-se do compartilhamento do computador por toda a família Souza Leão. Essa prática familiar resultava no comprometimento da autoria dos documentos ali produzidos uma vez que não havia o cuidado da troca de *login* a cada mudança de usuário. Mais uma vez, o critério de análise dos metadados relativos à data e autor, somados ao conteúdo literário, foi utilizado para a inferência sobre a autoria desses documentos digitais.

### **Conversão para formatos de preservação**

A conversão de formatos é uma ação fundamental no ambiente digital tanto para efeitos de preservação como de acesso, uma vez que existem formatos que permitem a fácil alteração do documento e outros muito pesados que dificultam sua visualização. A adesão a formatos de preservação e acesso precisa se dar com base em informações sobre questões de propriedade do formato o qual tem que ser aberto no que se refere à sua especificação e livre de direito autoral. Há ainda casos em que o formato é proprietário mas a especificação é aberta, como o *PDF/A* cujo proprietário é a empresa *Adobe* mas a especificação é pública e livre de direito autoral. No arquivo de Rodrigo de Souza Leão adotou-se os seguintes formatos de preservação: *PDF/A* para documento textual, *TIF* e *JPEG* para documento de imagem fixa, *WAVE* para documento de áudio e *MPEG* para documento de imagem em movimento.

Durante o processo manual de conversão de formatos, percebeu-se que a ação implicava na perda dos metadados originais dos documentos. Assim, por exemplo, um documento originalmente em *DOC*, datado de 18/09/2004, e convertido para *PDF/A*, teve a data original alterada para a data da conversão. Ressalta-se que tal descompasso não ocorreria se o arquivo fosse gerenciado em um RDC-Arq. Entretanto, solucionou-se com a aquisição do programa *MS Word Extract Document*



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

*Properties*, pelo qual os metadados originais dos documentos textuais foram extraídos e originaram uma planilha a qual foi acrescida com a equivalência dos metadados arquivísticos e de preservação.

### **Elaboração do arranjo**

O arranjo de um acervo arquivístico é um procedimento bastante caro à Arquivologia. De acordo com Douglas e MacNeil (2009, p. 27) “os princípios do arranjo são colocados sob a presunção de uma afinidade entre os documentos e seu produtor no qual o arranjo dos documentos atua como um tipo de espelho da entidade que os produziu”. Nesse contexto, entendemos que, ao reunirmos os fatores arquivos pessoais e a criação artística, os quais permeiam o arquivo de Rodrigo, a presunção e não a certeza dessa afinidade é ainda maior.

Inicialmente adotou-se um arranjo que mesclava espécies documentais e temas, tradicionalmente utilizado pelo AMLB. Entretanto, com o reconhecimento de que um arranjo que tem por base as funções e atividades do titular do arquivo é mais coerente com a teoria arquivística e que as instituições arquivísticas que custodiam arquivos pessoais encontram-se num processo de atualização de suas práticas, procedeu-se a elaboração, em novo projeto, de um novo arranjo capaz de sistematizar a trajetória de Rodrigo de Souza Leão numa estrutura multinível.

### **Descrição e acesso**

Mesmo antes de sofrer qualquer intervenção arquivística, a liberação do acesso foi possível exatamente por se tratar de um acervo digital. Assim, num primeiro momento, usando apenas os recursos do sistema operacional *Windows*, todo acervo foi disponibilizado na sala de consultas no “modo somente leitura” o qual permite o acesso aos documentos sem qualquer possibilidade de cópia ou alteração.

A partir do levantamento dos elementos formais e de conteúdo dos documentos encontra-se em elaboração um inventário sumário por meio do qual os



**3º INTEGRAR - Congresso Internacional de Arquivos,  
Bibliotecas, Centros de Documentação e Museus  
“PRESERVAR PARA AS FUTURAS GERAÇÕES”**

---

conjuntos documentais são descritos em consonância com o arranjo multinível mencionado acima, a ser inserido na base de dados descritiva Sophia, adquirida recentemente pela FCRB.

### **Considerações finais**

O trabalho de organização e preservação do arquivo digital de Rodrigo de Souza Leão se mostrou tão complexo quanto instigante na medida em que traduziu-se em uma oportunidade de aprendizado e de revisão do estatuto epistemológico da Arquivologia, a reforçar a importância da contribuição de outras áreas de conhecimento, tais como Diplomática e Informática.

### **Referências**

CAMARGO, Ana Maria de Almeida. Arquivos pessoais são arquivos. *Revista do Arquivo Público Mineiro*. Belo Horizonte, nº 2, p. 26-39, 2009.

CONARQ. *e-ARQ Brasil: modelos e requisitos para sistemas informatizados de gestão arquivística de documentos*. Rio de Janeiro: Arquivo Nacional, 2011.

DOUGLAS, Jennifer; MACNEIL, Heather. Arranging the Self: Literary and Archival Perspectives on Writers' Archives. *Archivara*, Ottawa, n. 67, p. 26–39, 2009.

---

### **Informações do autor**

**Jorge Phelipe Lira de Abreu**

Programa de Pós-Graduação em Gestão de Documentos e Arquivos/UNIRIO

Email: [j.phelipeabreu@gmail.com](mailto:j.phelipeabreu@gmail.com)

