



## Modelo Hipátia: análise dos requisitos arquivísticos e computacionais

*Hipátia Model: computational and archival requirements analysis*

**Tatiana Canelhas (1), Tiago Emmanuel Nunes Braga (2)**

(1) Instituto Brasileiro de Informação em Ciência e Tecnologia, SAUS 5, bloco H, Brasília, DF, [tatianapignataro@ibict.br](mailto:tatianapignataro@ibict.br). (2) [tiagobraga@ibict.br](mailto:tiagobraga@ibict.br)

### Resumo

A atuação do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) na temática de preservação digital remonta ao início do século, quando o Instituto acompanhou um movimento internacional em prol do tema (ARELLANO, 2004). Desde então, diversas iniciativas foram empreendidas e em 2018, em parceria com o Tribunal de Justiça do Distrito Federal e Territórios (TJDFT), surgiram os primeiros estudos na direção do que viria a se tornar o Modelo Hipátia; a proposta da Instituição para a preservação digital arquivística (SHINTAKU; BRAGA; DE FARIA OLIVEIRA, 2021). Seguindo a estrutura proposta pelo Conarq (2015), o Modelo Hipátia foi sistematizado em cinco grandes fases: preparação arquivística, preparação computacional, extração de objetos digitais, preservação e disseminação. Este resumo se debruça sobre as duas primeiras fases: preparação arquivística e computacional, no intuito de especificar os requisitos a serem seguidos durante a execução das duas fases citadas. Na perspectiva teórico-metodológica, este trabalho possui uma abordagem exploratória, por analisar os documentos referência na implantação do Repositório Arquivístico Digital Confiável (RDC-Arq), bem como a documentação do Modelo Hipátia. Foram identificadas e consideradas as normativas do Conselho Nacional de Arquivos (Conarq) e do Conselho Nacional de Justiça (CNJ), a documentação e relatórios de pesquisa produzidos pelo IBICT e o código fonte do BarraPres: o barramento utilizado para automatizar a implementação do RDC-Arq na perspectiva do Modelo Hipátia. Enquanto resultado, pode-se considerar duas vertentes analisadas. A primeira, relacionada à preparação arquivística, identificou três principais ações propostas pelo Modelo: Análise do sistema produtor, Identificação dos dados e metadados e Proposição de infraestrutura informacional. Na Análise do sistema produtor, levanta-se todas as possibilidades do trâmite documental, desde sua gênese até a sua movimentação final. Isso significa dizer que o estudo se dá desde a produção documental até sua avaliação final, seguindo com as etapas de arquivamento ou de eliminação do documento. Os dois casos podem ser tratados no projeto, uma vez que o Hipátia recolhe os objetos digitais e seus metadados e os disponibiliza para o ambiente de preservação. A preservação poderá ser de cunho permanente, culminando com o arquivamento, ou de longa temporalidade, quando há a eliminação. Em seguida, se analisa quais metadados fazem parte de quais ações dentro do sistema e por qual meio eles podem ser extraídos: via banco de dados, webservices ou APIs. Todo esse processo é realizado a fim de compor o pacote de transferência inicial, ou seja, antes

do envio para o RDC-Arq. Na identificação dos dados e metadados é preciso estudar e apresentar o ciclo de vida documental, levantando os metadados de gestão que são considerados importantes pela normativa arquivística para serem preservados. Tais informações são imprescindíveis quando se quer atestar a autenticidade documental. Esses metadados de gestão estão referenciados no Modelo de Requisitos para Sistemas de Gestão Arquivística de Documentos (e-Arq Brasil) do CONARQ, que é aplicável aos setores públicos e privados de qualquer esfera governamental. No judiciário, segue-se o Modelo de Requisitos para Sistemas Informatizados de Gestão de Processos e Documentos do Judiciário brasileiro (MoReq-Jus). Por meio desses dois modelos compara-se a sessão de metadados encontrados nos sistemas em estudo com os metadados citados nas normativas. Isso porque “não considerar a preservação dos metadados, impõe ao documento o risco de perda de seu vínculo arquivístico e, portanto, de sua capacidade de ser presumido autêntico” (PRAXEDES, RANGEL, 2018, p. 64). Após concluir as duas ações anteriores, em que se determina quais metadados devem ser extraídos do sistema de gestão e sua agregação junto aos objetos digitais, é criado o dicionário de dados. Tal dicionário é único por sistema e por instituição que for aplicar o modelo de preservação Hipátia, pois os metadados a serem preservados são personalizados para as demandas institucionais, assim como o processo de recuperação no sistema produtor. A terceira ação na preparação arquivística é a Proposição da infraestrutura informacional. Esta ação pode ser entendida como a criação do pacote de transferência inicial estruturada no formato BagIt, desenvolvido em conjunto pela Biblioteca do Congresso dos Estados Unidos e pela Biblioteca Digital da Califórnia. O modelo do *Open Archival Information System* (OAIS) aponta que a forma mais segura de se tramitar objetos digitais entre sistemas é pela codificação de pacotes de informações de arquivo. A estrutura hierárquica de diretórios e arquivos deve ser padronizada de acordo com as especificações do formato e do sistema que irá receber os pacotes. No modelo Hipátia, é prevista atualmente a utilização do Archivematica, mas existe a possibilidade de se adotar outros sistemas, a depender das necessidades da instituição. Por exemplo, se for necessário trabalhar com o empacotamento de processos que são compostos por vários arquivos, o Hipátia pode prever a construção de um *Structmap* personalizado para o agrupamento desses arquivos no sistema de acesso. Já a segunda vertente desta pesquisa focou na etapa de preparação computacional do modelo Hipátia. Foram identificadas as seguintes ações principais relacionadas à etapa: análise do sistema produtor, identificação dos aspectos computacionais, proposição do modelo de extração, análise da arquitetura, proposição da arquitetura computacional, instalação e configuração dos sistemas. Esta etapa ocorre em duas ações paralelas uma vez que a análise da arquitetura independe da análise do sistema produtor. Na análise da arquitetura será identificada a estrutura lógica dos servidores instalados no parque computacional da instituição. Também é feita uma análise da arquitetura de rede adotada, quando são identificados aspectos relacionados à performance e segurança. Por fim, na sequência de ações dessa linha são instalados os três sistemas computacionais adotados pelo modelo: BarraPres, Archivematica e AtoM e configuradas as permissões de acesso para os sistemas e usuários que terão interações no escopo do Modelo. Nas ações derivadas da análise do sistema produtor relacionadas à etapa de preparação computacional é estabelecida uma proposta de extração dos dados e metadados que alimentarão o modelo de preservação digital. Essa extração se dará pelos mecanismos disponíveis atualmente, tal qual APIRest<sup>2</sup>, acesso direto ao banco de dados e acesso ao sistema operacional do computador no qual o sistema produtor de objetos digitais está instalado. Ao final dessa linha de é proposto uma configuração de extração que baliza todo o processo de implementação do RDC-Arq. Os resultados destas duas fases em conjunto é que irão possibilitar que a implementação do RDC-Arq, baseado no

---

<sup>2</sup> Recurso que permite representar parte ou o todo do estado atual de um ambiente em outro por meio de transações computacionais.

modelo Hipátia, ocorra com sucesso. Para que a implementação do modelo RDC-Arq seja bem-sucedida, a realização das duas etapas analisadas depende do suporte direto das áreas de gestão documental e tecnologias da informação. Isso se faz necessário por nestas etapas estar contemplado o levantamento de dados e informações primárias relacionadas à instituição que almeja implementar o modelo Hipátia. Como considerações finais, pode-se afirmar que as etapas de preparação arquivística e computacional previstas no Modelo são fundamentais para se fazer um diagnóstico do ambiente a ser preservado. Elas fornecem os insumos necessários para se realizar o planejamento da implementação de uma estrutura de RDC-Arq, e os produtos derivados dessas etapas são utilizados durante todo o ciclo informacional previsto no modelo Hipátia.

**Palavras-chave:** Modelo Hipátia; Preservação Digital; RDC-Arq; Arquitetura da Informação.

## REFERÊNCIAS

ARELLANO, M. Á. M. Preservação de Documentos Digitais. **Ciência da Informação**, v. 33, n. 2, 10 dez. 2004. DOI 10.18225/ci.inf.v33i2.1043. Disponível em: <http://revista.ibict.br/ciinf/article/view/1043>. Acesso em: 11 jan. 2022.

CONARQ. **Resolução nº 43, de 04 de setembro de 2015**. Disponível em: [https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq\\_diretrizes\\_rdc\\_arq\\_resolucao\\_43.pdf](https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_diretrizes_rdc_arq_resolucao_43.pdf). Acesso em: 11 jan. 2022.

PRAXEDES, K. V.; RANGEL, K. da S. Relações entre o vínculo arquivístico e a autenticidade de documentos nato digitais: alguns apontamentos a respeito dos metadados. **Revista do Arquivo**, São Paulo, Ano II, Nº 6, p. 63-76, 2018.

SHINTAKU, M.; BRAGA, T. E. N.; DE FARIA OLIVEIRA, A. Hipátia: uma ferramenta livre no apoio à preservação digital. **Revista Brasileira de Preservação Digital**, v. 2, p. e021006–e021006, 2021.