



Modelo de Preservação Hipatia: metodologia de estudo de metadados para extração



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES

Dra Ívina Flores Melo
Dr Tiago Braga
Ms. Tatiana Canelhas

O que é o IBICT?

Instituto Brasileiro de Informação em Ciência e Tecnologia vinculado ao
MCTI
(Instituto de pesquisa)

Missão

Promover a competência, o desenvolvimento de recursos e a infraestrutura de informação em ciência e tecnologia para a produção, socialização e integração do conhecimento científico e tecnológico.



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES

Parceiros - Rede Cariniana



Ecosistema de preservação Digital



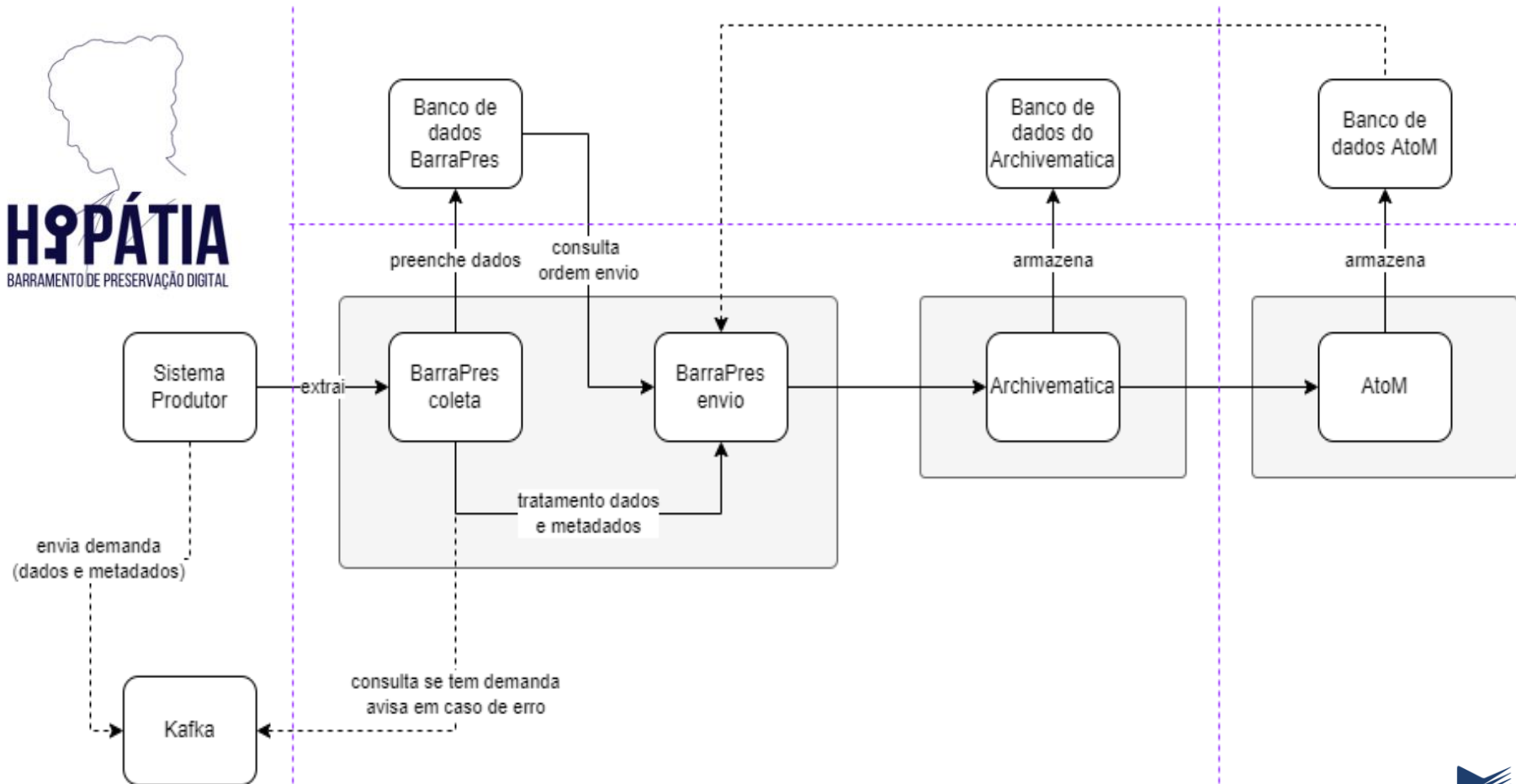
Cariniana



HOPÁTIA

MODELO DE PRESERVAÇÃO

- Concebido em 2018 a partir de uma parceria com o TJDFT com fins a preservar os Dje;
- Visa a preservação de objetos digitais e seus metadados (identificação, descritivos e preservação);
- Aderente ao Modelo OAIS;
- Conjugado de tecnologias (BarraPres/Archivematica/AToM);
- Barramento em Python 3 com uso das bibliotecas (np, numpy, schedule, pyodbc, zeeb, datetime, requests).





Nosso objetivo

Descrever a metodologia proposta pelo modelo Hipatia para se fazer os estudos de metadados.

Caso da pesquisa ocorrida em parceria com o Arquivo Nacional.

As etapas:

1. Acesso ao sistema produtor e sua documentação, se for o caso;
2. Estudo do escopo e da produção informacional;
3. Mapeamento da localização de dados, informações e documentos;
4. Elaboração do dicionário de dados;
5. Configuração do BarraPres;
6. Extração de dados e criação do PTI
7. Envio do PTI para o ambiente de Preservação.

Acesso ao sistema produtor e sua documentação, se for o caso e o estudo do escopo e da produção informacional

- Infraestrutura computacional, como códigos, banco de dados e diretórios de arquivos;
- Mapeamento de todo o ciclo de vida da produção informacional.

Mapeamento da localização de dados, informações e documentos e elaboração de dicionário de dados;

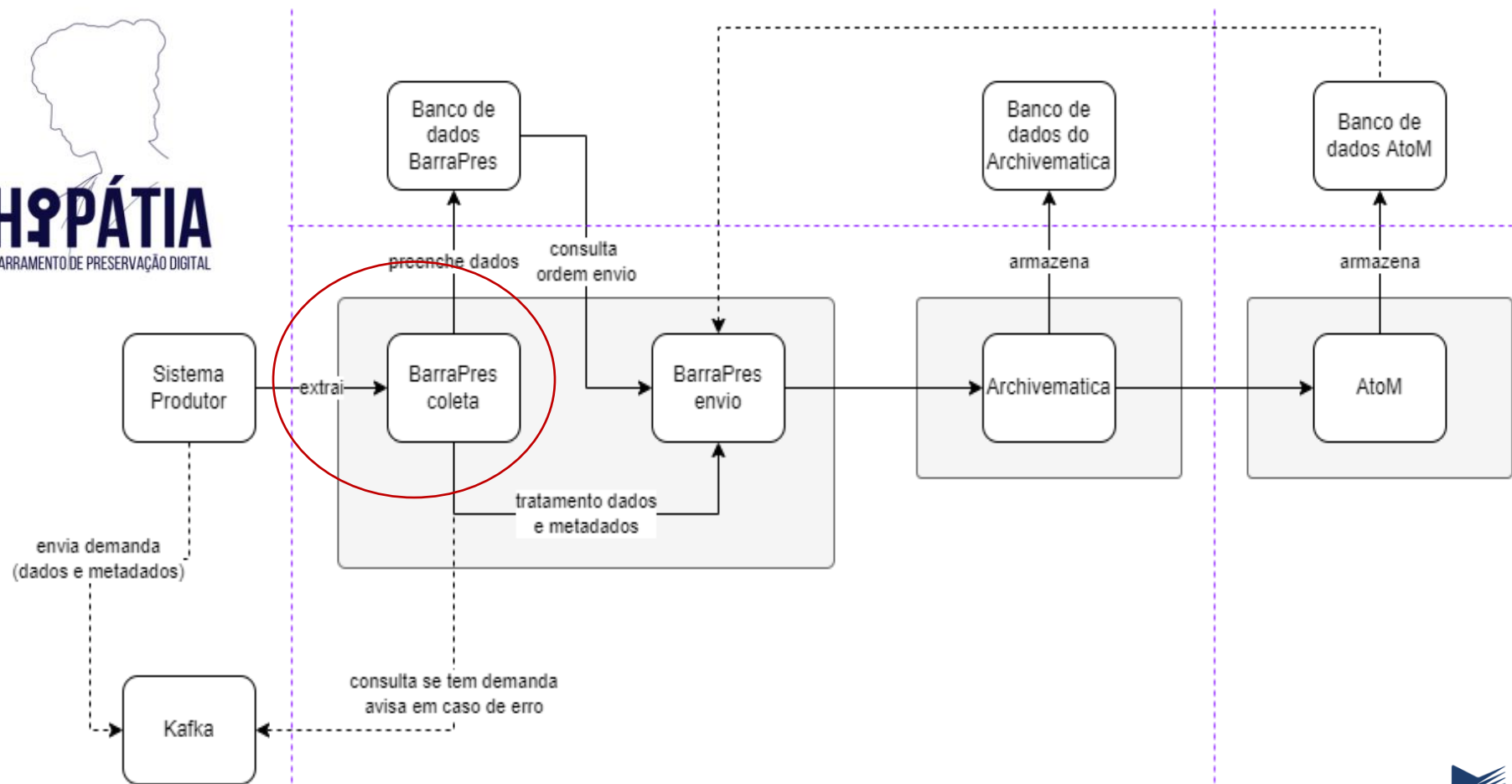
- Mapeamento da localização dos dados, informações e documentos em servidores locais, também conhecidos como *Local Filesystems*, em *Network Filesystems* (NFS) ou até codificados em banco de dados, do tipo *blob*;
- Documento contendo o endereçamento dos dados e informações de maneira que possa ser configurado no barramento para fins de extração.

Consulta ao banco SEI/AN

```
SELECT distinct c.sigla,
               c.nome,
               case
                 p.sta_participacao
                 when 'I' then 'Interessado'
                 when 'D' then 'Destinatário'
                 when 'R' then 'Remetente'
                 when 'A' then 'Acesso Externo'
                 else p.sta_participacao
               end as "destinatario"
FROM participante p
   INNER JOIN contato c on p.id_contato = c.id_contato
WHERE sta_participacao = 'I'
   and ID_PROTOCOLO in (
     SELECT p2.id_protocolo
     FROM protocolo p
        inner join rel_protocolo_protocolo rpp on p.id_protocolo =
            rpp.id_protocolo_1
        inner join protocolo p2 on rpp.id_protocolo_2 = p2.id_protocolo
     WHERE p.id_protocolo = id_protocolo //recuperado em consulta anterior
   );
```

MER do banco SEI/AN





Dados extraídos e sistematizados pelo BarraPres

📁 0700351-25.2021.8.07.0001

📁 metadata

📄 metadata.csv

📁 objects

📄 72685803-Petição Inicial.html

📄 72687601-Certidao.html

📄 72687602-Certidao.html

📄 72687603-Certidao.html

Extração de dados e criação do PTI

Envio do PTI para o ambiente de Preservação.

- o Barrapres migra a pasta que foi salva com os dados que serão preservados para o formato *BagIt**

****BagIt é uma padrão convencionado pela Library of Congress para empacotar diretórios de arquivos gerando e registrando checksums para cada arquivo armazenado em uma bag,***

Pacote de transferência Inicial

📁 0700351-25.2021.8.07.0001

📁 data

📁 metadata

📄 metadata.csv

📁 objects

📄 72685803-Petição Inicial.html

📄 72687601-Certidao.html

📄 72687602-Certidao.html

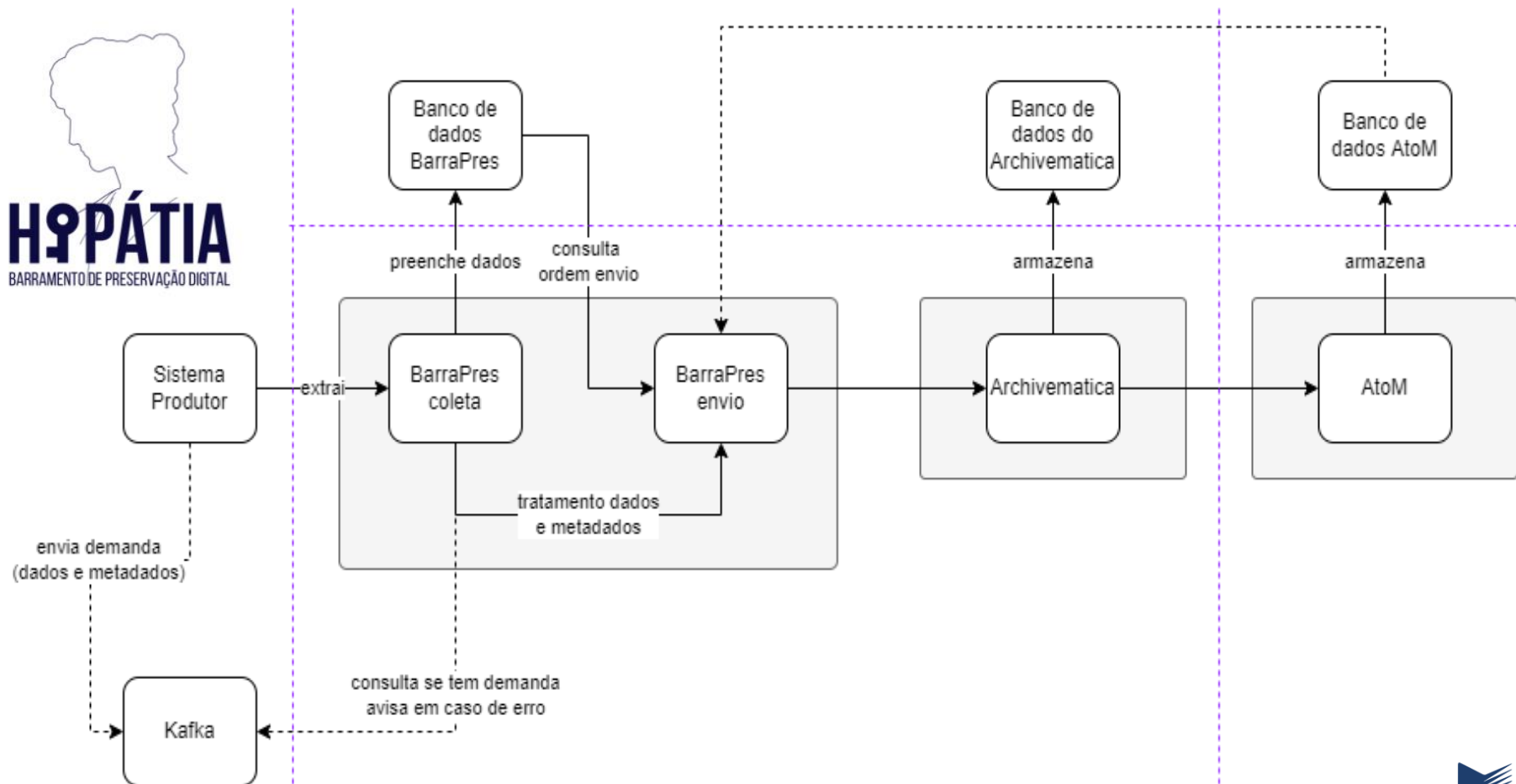
📄 72687603-Certidao.html

📄 bag-info.txt

📄 bagit.txt

📄 manifest-sha256.txt

📄 tagmanifest-sha256.txt



Considerações finais

Por meio do Hipatia, foi possível vislumbrar que o Modelo de referência OAIS pode se tornar realidade corrente nos cenários informacionais do Brasil e se apresenta como uma solução para atender a demanda latente de implementação da preservação digital.

Referências

BRAGA, Tiago E N. O modelo Hipatia: a proposta do IBICT para a preservação digital arquivística. In: **Hipatia: modelo de preservação para repositórios arquivísticos digitais confiáveis**. Brasília: Ibict, 2022. Disponível em: <<http://labcotec.ibict.br/omp/index.php/edcotec/catalog/book/livroHipatia>>. Acesso em: 30 ago 2022.

CCSDS – Consultative Committee for Space Data Systems. **Reference Model for an Open archive Information System (OAIS)**. Washington: CCSDS Secretariat, June 2012. Disponível em: . Acesso em: 09 set 2022.

CONSELHO NACIONAL DE ARQUIVOS. **Modelo de requisitos para sistemas informatizados de gestão arquivística de documentos: e-ARQ Brasil**. Versão 2. Rio de Janeiro: Arquivo Nacional, 2022. Disponível em: <<https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/EARQV203MAI2022.pdf>>. Acesso em 09 set 2022.

LIMA, Rodrigues de Souza, A. H., OLIVEIRA, A. F., D'AVILA, R. T., CHAVES, E. P. da S. S. (2014). O modelo de referência OAIS e a preservação digital distribuída. **Ciência Da Informação**, 41(1). <https://doi.org/10.18225/ci.inf.v41i1.1352> Acesso em: 07 set 2022

SARAMAGO, Maria Lurdes. Metadados para preservação digital e aplicação do modelo OAIS. In: **Actas do congresso nacional de bibliotecários, arquivistas e documentalistas**. 2004.

ivinamelo@ibict.br
tatianacanelhas@ibict.br
tiagobraga@ibict.br

