

Estudo para implementação do Modelo Hipátia de Preservação Digital no Tribunal de Contas da União

Relatório Técnico Parcial 1- Meta 1



PRESIDÊNCIA DA REPÚBLICA

Luiz Inácio Lula da Silva
Presidente da República

Geraldo José Rodrigues Alckmin Filho
Vice-Presidente da República

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Luciana Santos
Ministra da Ciência, Tecnologia e Inovação

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Tiago Emmanuel Nunes Braga
Diretoria

Carlos André Amaral de Freitas
Coordenação de Administração - COADM

Ricardo Medeiros Pimenta
Coordenação de Ensino e Pesquisa em Informação para a Ciência e Tecnologia - COEPI

Henrique Denes Hilgenberg Fernandes
Coordenação de Planejamento, Acompanhamento e Avaliação - COPAV

Cecília Leite Oliveira
Coordenação-Geral de Informação Tecnológica e Informação para a Sociedade - CGIT

Washington Luís Ribeiro de Carvalho Segundo
Coordenação-Geral de Informação Científica e Técnica - CGIC

Hugo Valadares Siqueira
Coordenação-Geral de Tecnologias de Informação e Informática - CGTI

Milton Shintaku
Coordenação de Tecnologias para Informação - COTEC

Alexandre Faria de Oliveira
Coordenação de Governança em Tecnologias para Informação e Comunicação - COTIC

©2024 INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Este Relatório é licenciado sob uma licença [Creative Commons Attribution 4.0](#), sendo permitida a reprodução parcial ou total desde que mencionada a fonte.

EQUIPE TÉCNICA

Diretor do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)
Tiago Emmanuel Nunes Braga

Coordenador-Geral de Tecnologias de Informação e Informática (CGTI)
Hugo Valadares Siqueira

Coordenador do Projeto
Alexandre Faria de Oliveira

Autores

Alexandre Faria de Oliveira
Ívina Flores Melo
Cássio Felipe de Oliveira Pires
Marcos Sigismundo da Silva
Marcelo Moro Brondani
Marilete da Silva Pereira
Raul Neres Carlos
Daniel Monteiro
Tauane Fonseca Esteves

Normalização

Tauane Fonseca Esteves
Denise Oliveira de Araujo
Marilete da Silva Pereira

Diagramação

Alisson Eugênio Costa

Este Relatório Técnico é um produto do projeto Implementação de modelo de preservação Hipátia no Tribunal de Contas da União (TCU).

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia e Inovações.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Sumário

1 INTRODUÇÃO.....	5
2 OBJETIVOS.....	7
2.1 OBJETIVO GERAL.....	7
2.2 OBJETIVOS ESPECÍFICOS.....	7
3.1 Diagnóstico e proposição de aspectos necessários para preservação do sistema e-TCU.....	8
3.1.1. <i>Alinhamento inicial e conceitual sobre a Preservação digital e o Modelo Hipátia</i>	8
3.1.1.2 <i>Efetuar estudo diagnóstico quanto ao nível de maturidade em Preservação Digital.....</i>	9
3.1.2 <i>Estruturar a aplicação do modelo de preservação Hipátia.....</i>	12
3.1.2.1 <i>Mapeamento e estudo de metadados e formatos para o sistema e-TCU</i>	12
3.1.2.2 <i>Validação de metadados e formatos delimitados para o sistema e-TCU</i>	18
3.1.2.3 <i>Delimitar parâmetros de configuração de softwares de preservação e do RDC-Arq.....</i>	21
3.1.2.4 <i>Delimitar parâmetros de configuração de softwares de acesso.....</i>	22
CONSIDERAÇÕES FINAIS.....	24
REFERÊNCIAS.....	25
APÊNDICE A - Apresentações relacionadas ao Modelo Hipátia.....	26
APÊNDICE B - Modelo de Maturidade em Preservação Digital.....	107
APÊNDICE C - Relatório Modelo de Maturidade em Preservação Digital.....	133
APÊNDICE D - Modelo para Dicionário de Dados e-TCU.....	155
APÊNDICE E - Capacitação ao TCU.....	189
APÊNDICE F - Formulário de Processing Configuration.....	198

1 INTRODUÇÃO

No Brasil, a adequada gestão documental por parte dos órgãos públicos é uma exigência legal, prevista em diversos diplomas normativos, entre os quais se destacam a Constituição Federal de 1988, que, por meio do § 2º de seu art. 216 e a Lei Federal n. 8.159, de 8 de janeiro de 1991.

Nessa perspectiva, a proteção e a adequada gestão do patrimônio documental público, mais do que uma mera previsão legal, é um dever dos gestores públicos. Isso porque a perda, o extravio ou a destruição indiscriminada de documentos podem acarretar danos irreparáveis e insegurança jurídica não só à própria Administração Pública, mas também à sociedade, que poderá ver comprometidos os direitos dos cidadãos, a produção de conhecimento, a preservação da memória e o registro/divulgação da história. Os documentos arquivísticos devem, portanto, ser preservados pelo tempo necessário ao cumprimento de suas finalidades.

É importante salientar que a legislação federal considera como arquivísticos os documentos produzidos ou recebidos por pessoa natural, ou jurídica, independentemente do suporte da informação. Fica evidente que a necessidade de preservação não é restrita aos documentos em papel, abrangendo todo e qualquer suporte, inclusive os digitais, o que configura um enorme desafio para os órgãos responsáveis pela gestão de acervos documentais.

Diante do desafio supramencionado, uma das atividades imprescindíveis diz respeito à adoção de estratégias de preservação digital em longo prazo, preferencialmente apoiadas em padrões abertos, que garantam a manutenção da integridade, confiabilidade e acesso extemporâneo. Uma das iniciativas de preservação digital adotadas por diversas instituições mundo afora, e recomendada no Brasil pelo Conselho Nacional de Arquivos (Conarq), órgão vinculado ao Ministério da Justiça e Segurança Pública, é a implementação de repositórios arquivísticos digitais confiáveis (RDC-Arq), fundamentada pela Resolução nº 43, de 4

de setembro de 2015, a qual, além de prever diretrizes para a implementação de RDC-Arq, também cobre o recolhimento de documentos arquivísticos digitais para instituições arquivísticas dos órgãos e entidades integrantes do Sistema Nacional de Arquivos (SINAR).

Um RDC-Arq é um ambiente seguro projetado para gerenciar documentos digitais, incorporando políticas de gestão para enfrentar os desafios de preservação e acessibilidade. Ele consiste em uma complexa estrutura que engloba *hardware*, *software*, metadados e procedimentos, indo além do simples armazenamento tecnológico. Esses repositórios interagem com sistemas adicionais para garantir a manutenção, preservação e acesso aos documentos digitais da instituição, mantendo padrões de segurança, confiabilidade e autenticidade.

Ainda que previsto na Resolução nº 43/2015¹, até 2018 não existia no Brasil uma solução prática e baseada em padrões abertos para a efetiva implementação de RDC-Arq. Nesse sentido, por meio do desenvolvimento de pesquisa técnico-científica, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), em parceria com o Tribunal de Justiça do Distrito Federal e dos Territórios (TJDFT), iniciou, em 2018, uma investigação que resultou na proposição do Hipátia, modelo de preservação digital para implementação de RDC-Arq.

O modelo em questão foi proposto a partir dos preceitos de referência do *Open Archival Information System* (OAIS) e considera, para preservação, as seguintes etapas: estudo do sistema produto; extração; empacotamento; preservação; e disseminação. O estudo dos sistemas produtores é crucial para a implementação do modelo, face às particularidades de cada instituição. Em alinhamento às normas e recomendações arquivísticas internacionais, adota-se o Archivematica como sistema de preservação e o AToM como ferramenta de disseminação.

Com base nos parágrafos preambulares, este relatório é apresentado como resultado parcial da parceria estabelecida entre o Ibict e o TCU para a implementação de um RDC-Arq, a fim de viabilizar a preservação digital dos

¹ Atualmente, Resolução nº 51, de 25 de agosto de 2023.

documentos arquivísticos armazenados no sistema e-TCU. Para tanto, ao longo das seções que se seguem, são descritas as atividades previstas na primeira etapa da Meta 1, isto é, os objetivos específicos (OE) 1 e 2, conforme acordado no plano de trabalho construído em conjunto pelas equipes do Ibict e TCU.

2 OBJETIVOS

Os objetivos que norteiam o desenvolvimento do projeto de pesquisa proposto encontram-se descritos a seguir:

2.1 OBJETIVO GERAL

Norteadado pela meta de integração do RDC-Arq ao sistema e-TCU, o objetivo geral deste relatório compreende implementar o modelo de preservação Hipátia definido para o sistema e-TCU controle externo.

2.2 OBJETIVOS ESPECÍFICOS

Especificamente, objetiva:

- a) Diagnosticar aspectos relevantes para a proposição e implementação de um RDC-Arq no TCU; e
- b) Estruturar a aplicação do modelo de preservação Hipátia.

3 META 1: INTEGRAÇÃO DO RDC-ARQ AO SISTEMA E-TCU

A Meta 1 do projeto objetiva diagnosticar os aspectos relevantes para a implementação eficiente do RDC-Arq no TCU, assegurando a integridade e a acessibilidade dos documentos digitais ao longo do tempo.

3.1 Diagnóstico e proposição de aspectos necessários para preservação do sistema e-TCU

Nesta subseção, será apresentado um diagnóstico detalhado e propostas para a preservação do sistema e-TCU. O objetivo é avaliar o nível de maturidade em preservação digital e identificar os requisitos técnicos e organizacionais necessários para implementar práticas eficazes, alinhadas ao Modelo Hipátia.

3.1.1. Alinhamento inicial e conceitual sobre a Preservação digital e o Modelo Hipátia

O Seminário de Alinhamento foi estruturado em sessões que abordaram conhecimentos de preparação informacional e computacional. As sessões ocorreram na primeira quinzena de março de 2024 e foram conduzidas pelos pesquisadores Cássio Pires, Ívina Flores, Daniel Monteiro, Marcos Sigismundo e Marcelo Brondani. Os temas discutidos estão detalhados no Apêndice A deste documento. Informações sobre os treinamentos estão apresentadas no quadro abaixo.

Quadro 1 - Reuniões de Seminário de Alinhamento

Assunto	Data	Horário de início	Número de participantes
1º Seminário de Alinhamento	04/03/2024	17:13	17
2º Seminário de Alinhamento	12/03/2024	14:11	14

Definições de infraestrutura computacional	13/03/2024	18:04	14
Softwares do Modelo Hipátia	15/03/2024	14:05	14

Fonte: Elaborado pelos autores (2024).

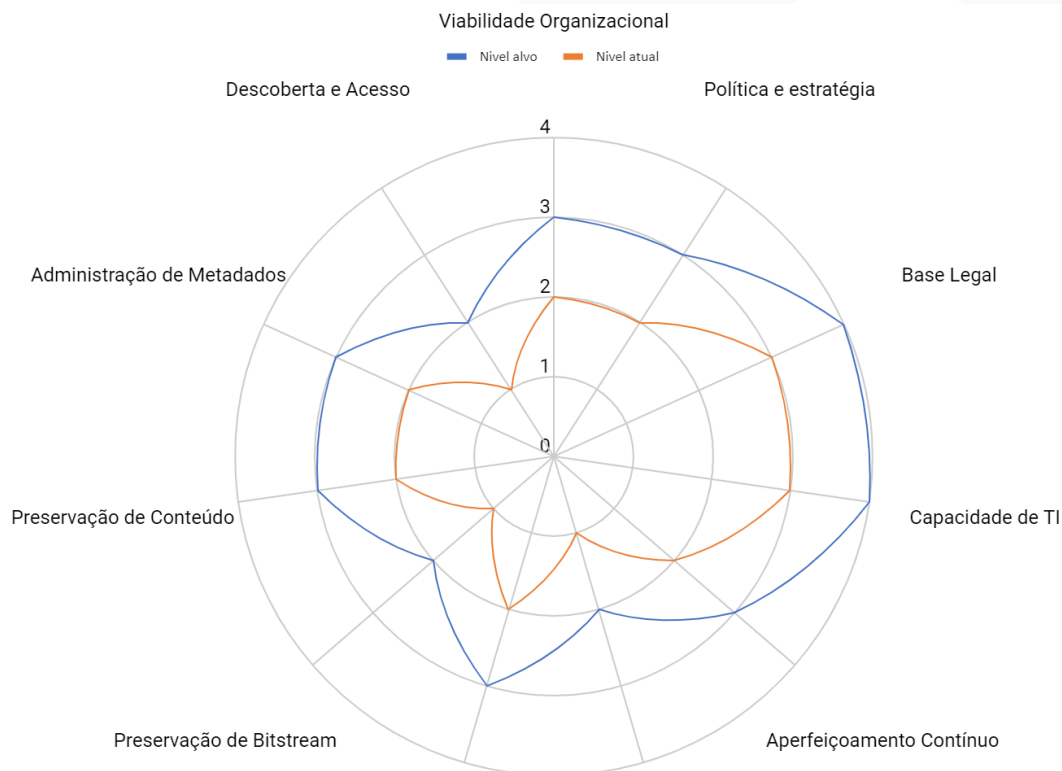
Como resultado, encaminharam-se ao TCU as necessidades em torno dos acessos e ambientes para instalação e desenvolvimento. Na área de gestão de documentos, agendou-se o início das tratativas para estudos de formatos e metadados, com resultados e entregas conforme se segue.

3.1.1.2 Efetuar estudo diagnóstico quanto ao nível de maturidade em Preservação Digital

A metodologia de mensuração de níveis de maturidade utilizada no modelo Hipátia é o RAM-DPC. Os estudos iniciais no TCU, conduzidos pela pesquisadora Ívina Flores, partiram de uma aplicação inicial do RAM-DPC a partir da análise de documentos enviados pelo Tribunal, das questões relativas aos ambientes tecnológicos e de dados observados e coletados nas reuniões do projeto. Após a aplicação inicial, foi realizada uma reunião em que se apresentou a metodologia utilizada no Hipátia e o resultado preliminar na análise de maturidade, disponível no Apêndice B deste documento.

Em seguida, foi enviado o DPC RAM - Modelo de Maturidade para que a equipe do TCU avaliasse o resultado preliminar. Além disso, foi encaminhado o documento de Seções do DPC RAM para auxiliar na avaliação, ambos incluídos no Apêndice C. A equipe do TCU efetuou a avaliação de maturidade, reforçando alguns níveis e alterando outros, resultando no seguinte gráfico:

Figura 1 – Matriz de Maturidade em Preservação Digital



Fonte: Elaborado pelos autores, (2024).

Como apontado no Relatório Modelo de Maturidade, disponível no Apêndice D, para determinar o nível médio de maturidade da instituição, conforme o gráfico fornecido pelo RAM DPC, podemos calcular a média dos níveis atuais para cada área avaliada.

Os níveis atuais são:

- Viabilidade Organizacional: 2
- Política e Estratégia: 2
- Base Legal: 1.5
- Capacidade de TI: 2
- Aperfeiçoamento Contínuo: 1
- Comunidade: 1

- Aquisição, Transferência e Incorporação: 1
- Preservação de Bitstream: 0
- Preservação de Conteúdo: 1.5
- Administração de Metadados: 2
- Descoberta e Acesso: 2.5

Somando esses valores: $2 + 2 + 1.5 + 2 + 1 + 1 + 1 + 0 + 1.5 + 2 + 2.5 = 16.5$

Dividindo pelo número de áreas (11): $16.5 / 11 \approx 1.5$

Portanto, o nível médio de maturidade da instituição é aproximadamente 1.5. Os resultados da análise de maturidade em preservação digital usando o RAM DPC mostram uma discrepância significativa entre os níveis atuais e os níveis alvo em várias áreas críticas. Ainda, com base no relatório e no RAM-DPC, foram atribuídos os níveis tais como descritos abaixo:

Quadro 2 - Modelo RAM-DPC (março 2024-dez 2024)

CAPACIDADE ORGANIZACIONAL		
	Nível Atual (março/2024)	Nível Almejado (dez/ 2024)
A. Viabilidade Organizacional: Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.	2 - Básico	3 - Gerenciado
B. Políticas e Estratégia: Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.	2 - Básico	3 - Gerenciado
C. Base Legal: Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.	3 - Gerenciado	4 - Otimizado
D. Capacidade de TI: Recursos de tecnologia da informação para apoiar atividades de preservação digital.	3 - Gerenciado	4 - Otimizado
E. Aperfeiçoamento Contínuo: Processos para avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.	2 - Básico	3 - Gerenciado
F. Comunidade: Engajamento e contribuição para a comunidade mais ampla de preservação digital.	1 - Conscientização	2 - Básico

CAPACIDADES DE SERVIÇO		
	Nível Atual	Nível Almejado
G. Aquisição, Transferência e Incorporação: Processos para adquirir ou transferir conteúdo e incorporá-lo a um arquivo digital.	2 - Básico	3 - Gerenciado
H. Preservação de Bitstream: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.	1 - Conscientização	2 - Básico
I. Preservação de Conteúdo: Processos para preservar o sentido ou funcionalidade do conteúdo digital e garantir seu acesso e usabilidade contínuos no futuro.	2 - Básico	3 - Gerenciado
J. Gerenciamento de Metadados: Processos para criar e manter metadados suficiente para dar suporte a preservação, gerenciamento e uso de conteúdo digital preservado.	2 - Básico	3 - Gerenciado
K. Descoberta e Acesso: Processos para permitir a descoberta de conteúdo digital e prover acesso aos usuários.	1 - Conscientização	2 - Básico

Fonte: Elaborado pelos autores (2024).

A implementação do Modelo Rápido de Avaliação da *Digital Preservation Coalition* (RAM-DPC) no TCU proporcionou uma visão abrangente do estado atual da preservação digital dentro da instituição. A avaliação também destacou a importância de processos de melhoria contínua e do engajamento comunitário para o avanço das práticas de preservação.

3.1.2 Estruturar a aplicação do modelo de preservação Hipátia

Essa seção detalha os passos necessários para a implementação do Modelo Hipátia, abrangendo o mapeamento e estudo de metadados e formatos dos documentos geridos pelo sistema e-TCU, a validação de metadados e formatos delimitados e, por fim, os parâmetros de configuração de *softwares* de preservação e acesso.

3.1.2.1 Mapeamento e estudo de metadados e formatos para o sistema e-TCU

O objetivo deste estudo foi identificar os diferentes formatos de arquivo e os metadados associados aos documentos e processos, com vistas a compreender melhor a estrutura e as necessidades de preservação digital dos documentos geridos por esse sistema. O escopo do estudo foi delimitado aos documentos e processos geridos pelo sistema e-TCU, considerando tanto os formatos de arquivo quanto os metadados descritores utilizados pelo sistema. A delimitação incluiu: tipos de documentos eletrônicos armazenados, formatos de arquivo (E.G., PDF, DOCX, XML), metadados associados a cada tipo de documento.

A coleta de dados foi realizada nas seguintes etapas:

1. **Revisão de fontes:** realização de um levantamento inicial dos tipos de documentos e formatos de arquivo utilizados no sistema e-TCU com base em entrevistas, reuniões técnicas e documentos enviados pelo TCU.
2. **Entrevistas e reuniões técnicas:** condução de entrevistas e reuniões técnicas com os responsáveis pela gestão documental do TCU para que fosse possível compreender as práticas vigentes e coletar informações detalhadas sobre os formatos de arquivo e metadados utilizados.
3. **Documentação enviada pelo TCU:** análise da documentação fornecida pelo TCU, incluindo manuais e procedimentos operacionais relacionados à gestão de documentos digitais.
4. **Análise técnica:** realização de análise detalhada da documentação técnica em gestão de documentos, com foco nos procedimentos e práticas do sistema e-TCU.
5. **Estudo do rol de metadados do e-Arq Brasil:** estudo do rol de metadados e-Arq Brasil para modelar os esquemas de metadados para o TCU com vistas às melhores práticas. Essa revisão teve o objetivo de identificar, no e-TCU, os metadados elegíveis para a preservação.

A partir da aplicação das etapas descritas, foi criada uma planilha formatada com o rol de metadados do e-ARQ Brasil, indicando a obrigatoriedade ou não de

cada metadado, bem como a justificativa para seu caráter mandatório. Esse documento foi compartilhado com a equipe do TCU, incluindo os caminhos dos metadados na base do sistema e-TCU.

Após diversas reuniões realizadas nos meses de março e abril, às terças-feiras, às 11h, foi delimitado o método de integração e extração dos metadados. Durante o processo de seleção, foram discutidos pontos como a relevância ou redundância de certos metadados, e foi realizada uma interpretação conceitual conjunta dos metadados dos componentes digitais. O mapeamento dos metadados do e-TCU envolveu um processo colaborativo com a equipe técnica do TCU, que incluiu a análise dos caminhos dos metadados na base do sistema e a realização de reuniões para discussão da relevância e aplicabilidade de cada metadado.

O resultado desse estudo foi uma planilha abrangente, que lista os formatos de arquivo e os metadados utilizados no sistema e-TCU (Apêndice E). Essa planilha serve como base de referência para futuras ações de preservação digital e gestão documental no âmbito do sistema e-TCU.

Em reflexão sobre a modelagem dos metadados, entendeu-se que os metadados descritos no e-ARQ Brasil estão intrinsecamente vinculados a funcionalidades e requisitos essenciais à gestão e preservação de documentos arquivísticos digitais. A ausência de alguns metadados no e-TCU indica uma necessidade de evolução e adaptação do sistema para atendimento pleno das diretrizes e dos requisitos estabelecidos pelo e-ARQ Brasil. Essa necessidade de evolução pode envolver a atualização de funcionalidades existentes, a implementação de novos módulos ou a integração de soluções tecnológicas adicionais. A conformidade com os metadados recomendados pelo e-ARQ Brasil é fundamental para assegurar a interoperabilidade com outros sistemas, a resiliência a mudanças tecnológicas e a proteção contínua do patrimônio documental.

Em adição ao estudo dos metadados do e-ARQ, o esquema de metadados do e-TCU foi incrementado para melhor atender às necessidades de preservação digital e gestão documental. A seguir, detalhamos os metadados adicionados:

Quadro 3 - Proposta de Metadados

	Proposta de Metadados	Sugestão IBICT
Processo	Relator	NOVO
Processo	Representante do MPTCU	NOVO
Processo	Parte Interessada	1.18
	Representantes Legais	NOVO
	Órgão Interessado	1.18
Processo	Unidade Jurisdicionada	NOVO
Processo	Responsáveis	NOVO
Processo	Legitimados	NOVO
Processo	Se interesse do Congresso Nacional	NOVO
Processo	Se Urgente	NOVO
Processo	Se Histórico	Selo Histórico (Lista de classe)
Processo	Se Repercussão na Mídia	
Documento/Processo	Data de autuação	1.32. Data de produção
Processo	Tipo	NOVO
Processo	Assunto do processo	NOVO
Processo	Número	NOVO
Processo	Unidade Responsável Técnica (histórico)	NOVO
Processo	Processos apensados	NOVO
Processo	Processos apartados	NOVO
Documento/processo	Confidencialidade	1.30. Níveis de acesso
Processo	Estado	NOVO
Processo	Situação	NOVO
Processo	Deliberações	NOVO

Fonte: Elaborado pelos autores (2024).

Metadados como "Representante do MPTCU", "Órgão Interessado" e "Unidade Jurisdicionada" permitem uma contextualização mais precisa dos processos, enquanto metadados como "Se Histórico" e "Se Repercussão na Mídia" ajudam a

identificar documentos de interesse especial. Além disso, a adição de metadados relacionados à confidencialidade, ao estado e à situação dos processos fortalece a segurança e a gestão do acesso aos documentos. Esses metadados são centrais para garantir que a informação sensível seja adequadamente protegida e que os processos possam ser monitorados e gerenciados com eficácia ao longo do tempo.

O estudo de formatos no sistema e-TCU iniciou-se com a identificação e o levantamento dos formatos de arquivo atualmente em uso. Essa fase envolveu a análise de documentos geridos pelo sistema, incluindo HTML, DOC, PDF, XLSX e JPEG. Foram realizadas reuniões técnicas com os responsáveis, além da revisão de documentos técnicos fornecidos pela instituição.

A segunda fase consistiu na avaliação da adequação desses formatos para preservação digital a longo prazo. Sugere-se que os arquivos em formato PDF sejam convertidos (ação de migração de formato) para PDF/A-1b. Em relação ao formato Excel (XLSX), na perspectiva da preservação de longo prazo, os riscos de dificuldade de acesso à informação por obsolescência de formato são baixos porque o formato é amplamente utilizado. No entanto, por serem formatos desenvolvidos pela Microsoft e produzidos, via de regra, por *softwares* privados, sugere-se a investigação da possibilidade de conversão (isto é, migração de formato) para arquivo no padrão ODF (Open Document Format), baseado na ISO 26300:2015. Essa migração de formato teria como objetivo adicionar ao Pacote de Informação de Arquivamento (AIP) uma nova representação para preservação de longo prazo, se tecnicamente viável, pois não se trata de ação mandatória.

Adicionalmente, foi elaborada uma tabela na qual são detalhados os formatos de arquivo, suas recomendações para preservação e acesso, sendo observadas variações nos tamanhos dos arquivos ao serem transferidos para o sistema de preservação digital Archivematica, reforçando a importância de uma seleção criteriosa de formatos.

O quadro a seguir apresenta uma proposta para a implementação dos formatos de arquivo do sistema e-TCU, incluindo as recomendações para preservação e acesso:

Quadro 4 - Propostas de formatos para implementação no e-TCU

Media Type	Formato	Preservação	Acesso
Portable Document Format	Portable Document Format - PDF	PDF/A-1b	Formato Original
Microsoft Excel	Excel	PDF/A-1b ou formato original	Formato Original
Hyper Text Markup Language	Html	Formato Original	Formato Original
Áudio	MPEG Audio	É tratado como vídeo	É tratado como vídeo
	MPEG 1/2 Audio Layer 3 - MP3	WAV	MP3
	OGG	Formato Original	Formato Original
	MP4 Apple Lossless Audio Codec	É tratado como vídeo	É tratado como vídeo
	Vorbis-OGG	Formato Original	Formato Original
Imagem Rasterizada	Portable Network Graphics - PNG	TIFF Descompactado	JPEG (qualidade: 60)
	JPEG File Interchange Format	TIFF Descompactado	JPEG (qualidade: 60)
Vídeo	MPEG Video -codec libx264 -px: yuv420p	MKV	Formato Original
	OGG	Formato Original	Formato Original
	MPEG-4 Media File	MKV	Formato Original
	Quicktime - MOV	Formato Original	Formato Original
	Windows Media Video - WMV	MKV	MP4

Fonte: Elaborado pelos autores (2024).

Ao transferir arquivos para o sistema de preservação digital Archivematica, observou-se variações nos tamanhos dos arquivos originais e de suas versões derivadas para preservação e acesso. O quadro a seguir ilustra essas variações:

Quadro 5 - Análise das variações de tamanho dos arquivos digitais

Tipo	Tamanho Original (KB)	Derivada Preservação (KB)	Formato Acesso	Derivada Acesso (KB)	AIP.7z (KB)	DIP.rar (KB)
Áudio	1063	4673	MP3	638	4493	800
	1064	1064	OGG	1064	1094	1190
	697	697	OGG	697	720	830
Imagem	87	598	JPEG	42	622	230
	984	3051	JPEG	49	2312	250
Vídeo	446	446	MOV	446	457	550
	3042	36555	MP4	3042	33214	3250
	1282	4684	MPEG	1282	5580	1140
	1695	1695	OGG	1695	1733	1910
	1567	24796	MP4	1032	18746	1180

Fonte: Elaborado pelos autores (2024).

Esse aumento de tamanho ocorre devido aos processamentos e à agregação de padrões de preservação necessários à garantia da integridade e da longevidade dos arquivos digitais.

3.1.2.2 Validação de metadados e formatos delimitados para o sistema e-TCU

A análise inicial da extração de dados dos documentos indicou que os campos sem informações não forneceram dados úteis para análise, como é possível identificar no quadro abaixo. Os campos identificados como "valor+parâmetro" são metadados que requerem qualificação e descrição mais detalhada. Em contrapartida, alguns metadados já são considerados adequados.

Quadro 6 - Avaliação inicial dos metadados

Metadados	Análise
"COD_DOC_GESTAO":692048	Valor+parâmetro
,"NUM_DOC_FORMATADO":"45.402.542-0"	OK

,"COD_PROCESSO":45402545	Valor+parâmetro
,"NUM_TC_FORMATADO":001.077\2011-2"	OK
,"ID_VOLUME":n\u00E3o se aplica"	-
,"NUM_VOLUME":n\u00E3o se aplica"	-
,"STATUS":a definir"	-
,"VERSAO":backlog"	-
,"TITULO":a definir"	-
,"DESCRICAO":a definir"	-
,"ASSUNTO":CRIANDO A NOTA FISCAL ELETR\u00D4NICA NF-E, 55, EM SUBSTITUI\u00C7\u00C3O \u00C1 NOTA FISCAL MODELO 1 OU 1-A, APARTIR DE 1\u00BA DE DEZEMBRO DE 2010"	OK
,"AUTOR":a definir"	-
,"DESTINATARIO":a definir"	-
,"ORIGINADOR":a definir"	-
,"INDEFINIDO":a definir"	-
,"INTERESSADOS":backlog"	-
,"ID_COMP_DIGITAL":a definir"	-
,"GENERO":n\u00E3o se aplica"	Valor+parâmetro
,"ESPECIE":n\u00E3o se aplica"	Valor+parâmetro
,"TIPO":Of\u00E9dio de Requisi\u00E7\u00E3o"	OK
,"IDIOMA":a definir"	-
,"QTT_FOLHAS":a definir"	-
,"NUMERO_PECA_PROCESSO":1	OK

,"SE_TEM_ANEXO": "N\u00E3o"	Qualificar para sim/não e indicar número de anexos
,"ANOTACAO": "a definir"	-
,"RESP_ACAO": "a definir"	-
,"OUTROS_DOCS": "a definir"	-
,"CONFIDENCIALIDADE": "backlog"	-
,"DESCCLASSIFICACAO": "backlog"	-
,"DATA_HORA": "backlog"	-
,"LOCAL": "a definir"	-
,"CLASSE": "backlog"	-
,"DESTINACAO": "a definir"	-
,"PRAZO": "a definir"	-
,"LOCALIZACAO": "a definir"	-
,"COD_DOC_TRAMITAVEL": 45402542	Valor+parâmetro

Fonte: Elaborado pelos autores (2024).

Os resultados iniciais da extração de dados destacaram a necessidade de qualificação adicional para certos metadados. Campos como "COD_DOC_GESTAO", "COD_PROCESSO" e "COD_DOC_TRAMITAVEL" requerem descrição mais detalhada para serem totalmente úteis. Em contraste, campos como "NUM_DOC_FORMATADO", "NUM_TC_FORMATADO", "ASSUNTO" e "TIPO" foram considerados adequados às necessidades de preservação e gestão documental.

Para os testes iniciais de integração e extração de metadados, apontou-se duas abordagens possíveis: a primeira envolve a qualificação dos metadados existentes na planilha e a adição dos metadados do processo, permitindo a criação de pacotes de preservação em versão beta, evidenciando a estrutura de preservação em funcionamento; a segunda abordagem consiste na seleção de um conjunto mais

robusto de metadados da planilha geral para realização de uma extração mais estruturada, resultando em etapas mais longas de validação, mas com pacotes de preservação mais completos e definitivos.

3.1.2.3 Delimitar parâmetros de configuração de softwares de preservação e do RDC-Arq

Para a configuração do Archivematica, compreendeu-se ser necessário realizar capacitações de uso e operação da ferramenta, como descrito no Quadro 7. As capacitações, com carga horária de 8h, tiveram o objetivo de transferir conhecimento sobre o uso e o funcionamento do Archivematica de maneira síncrona, e a equipe de gestão de documentos como público-alvo. No Apêndice F é possível encontrar a apresentação realizada nesta capacitação.

Quadro 7 - Organização das capacitações

Data	Temáticas abordadas
18/06 10h-12h	O que é Archivematica? - o Dashboard - o Storage Service - o Menus Arquitetura e Componentes do Archivematica - o Microsserviços - o Ferramentas externas Plano de preservação - o Política de Identificação de Formatos - o Configuração de novos formatos/grupos de formatos - o Caracterização
20/06 8h30-10h30	Plano de preservação - o Extração - o Normalização - o Transcrição - o Validação Administração - o Geral - o Processing configuration - o PREMIS Agent - o Storage Locations - o Processing storage usage

24/06 8h30-10h30	Administração - o Transferir - o Admissão - o BackLog - o Archival Storage - o Arranjo
27/06 10h/12h	Introdução ao Storage Service

Fonte: Elaborado pelos autores (2024).

Após a realização do *Workshop*, no dia 09 de julho de 2024, efetuou-se o preenchimento do Formulário de *Processing Configuration*, que delimita parâmetros de configuração de processamento, disponível no Apêndice G.

3.1.2.4 Delimitar parâmetros de configuração de softwares de acesso

Considerando que o *software* de acesso aos documentos arquivísticos do TCU preservados será o AtoM (*Access do Memory*) da Artefactual, foram apontadas algumas ações necessárias para que essa plataforma esteja configurada adequadamente para receber a documentação recolhida.

Essas ações foram apresentadas à equipe do TCU por meio de reunião virtual, realizada no dia 11 de junho de 2024. São elas:

- Estabelecimento do quadro de arranjo a ser disponibilizado na plataforma: esse quadro de arranjo corresponde ao esquema de classificação dos documentos e é necessário para indicar o local de envio de metadados e objetos digitais no decorrer do processo de processamento dos pacotes de informação, em especial o DIP (*Dissemination Information Package*).
- Definição das taxonomias: o AtoM trabalha com entidades de informação chamadas taxonomias, uma espécie de grupo de termos de uma mesma categoria que apresentam algumas funções na representação dos conteúdos documentais e informacionais a serem disponibilizados na plataforma. As

principais taxonomias, nesse primeiro momento, seriam a de Níveis de Descrição, Assuntos e Locais.

- Levantamento das autoridades arquivísticas: os atores envolvidos na dinâmica documental são representados pela perspectiva arquivística com base na norma internacional do Conselho Internacional de Arquivos (ICA) ISAAR-CPF (Norma Internacional de Registro de Autoridade Arquivística – Entidade Coletiva, Pessoa e Família). Na plataforma, as autoridades arquivísticas constituem uma entidade que se relaciona diretamente com a entidade Descrição Arquivística, principalmente no que tange ao produtor arquivístico e sua respectiva história administrativa/biografia.
- Determinação dos aspectos de restrição de acesso: no conjunto documental do TCU há documentos com restrições de acesso em diferentes contextos e camadas. Com base nessas características documentais, é indispensável, considerando os recursos disponíveis ou não no AtoM, o estabelecimento de fluxos e configurações necessárias à melhor representação e disponibilização dos conjuntos documentais.

A equipe do TCU não apresentava familiaridade com o uso e a gestão do AtoM em ambiente institucional. Como uma das primeiras ações, foi realizado um repasse técnico de conhecimento em usabilidade do AtoM, de modo a permitir que a equipe pudesse saber quais são as funcionalidades, os recursos e os limites da plataforma. Isso permite que determinadas ações e decisões a respeito do acesso aos documentos aconteçam de modo adequado e de acordo com as perspectivas institucionais do TCU.

Seguem, abaixo, as informações sobre o repasse compartilhado com a equipe do TCU:

Público-Alvo: Equipe de gestão de documentos e de arquivos

Carga Horária: 12h

Objetivo: Compartilhar conhecimento sobre o uso do AtoM

Recursos necessários: AtoM instalado em ambiente de teste ou homologação.

Perfis de usuário administrador

Modalidade: Remota síncrona

Plataforma: Microsoft Teams

Cronograma de atividades:

Quadro 7 - Cronograma de atividades

Data e horário	Número de participantes	Conteúdo
02/07 08h30 às 11h30	24 pessoas	Introdução ao AtoM; Cadastro de usuários; Criação de grupos de usuários; Gestão de permissões; Configurações gerais.
03/07 08h30 às 11h30	21 pessoas	Gerenciamento de Taxonomias; Registros de autoridades.
04/07 08h30 às 11h30	17 pessoas	Entidade Custodiadora; Descrições arquivísticas.
05/07 08h30 às 11h30	14 pessoas	Publicações de descrição; Importação csv; Elementos visíveis.

Fonte: Elaborado pelos autores (2024).

Os encontros foram gravados pela plataforma Teams do Ibict e disponibilizados à equipe do TCU após a sua realização.

O *Workshop* aconteceu de modo prático, com o uso do AtoM do próprio do TCU como plataforma e não houve a confecção de uma apresentação em *Power Point*. O registro das atividades e práticas realizadas durante o repasse técnico pode ser acessado por meio do link: <https://atom.tcu.gov.br/index.php/>, disponível no ambiente do TCU.

CONSIDERAÇÕES FINAIS

O projeto de implementação do RDC-Arq no TCU foi pautado por dois objetivos centrais: diagnosticar aspectos relevantes para a proposição e implementação do RDC-Arq e estruturar a aplicação do Modelo Hipátia. As ações desenvolvidas para alcançar esses objetivos envolveram uma análise abrangente dos processos existentes e a adaptação de diretrizes internacionais às necessidades específicas do Tribunal.

Inicialmente, o diagnóstico detalhado das práticas de preservação digital do TCU foi fundamental para identificar pontos críticos e áreas de melhoria. Esta etapa incluiu a avaliação da infraestrutura tecnológica, das políticas institucionais vigentes e da capacidade técnica da equipe. A partir desse diagnóstico, foi possível propor soluções específicas para a implementação de um RDC-Arq eficaz, alinhado às melhores práticas internacionais.

Em seguida, foi estruturada a aplicação do Modelo Hipátia. Este modelo foi adaptado para atender às particularidades do TCU, garantindo a integridade, autenticidade e acessibilidade dos documentos digitais a longo prazo. A implementação envolveu a definição de diretrizes claras e a adoção de procedimentos específicos que asseguram a sustentabilidade da preservação digital na instituição.

A capacitação contínua da equipe do TCU foi uma parte crucial deste processo. Promoveu-se treinamentos e *workshops* para garantir que todos os envolvidos estivessem aptos a operar as tecnologias, fortalecendo o compromisso com a preservação digital e assegurando a adesão aos novos procedimentos.

Em conclusão, o trabalho realizado permitiu diagnosticar e propor soluções adequadas para a implementação de um RDC-Arq no TCU, além de estruturar uma base sólida para a preservação digital por meio do modelo Hipátia. Estes esforços conjuntos asseguram a proteção e a acessibilidade contínua dos documentos

digitais da instituição, contribuindo para a eficiência, transparência e responsabilidade das atividades do Tribunal.

REFERÊNCIAS

BRASIL. Conarq. **Resolução nº 43, de 04 de setembro de 2015**. Altera a redação da Resolução do Conarq nº 39, de 29 de abril de 2014, que estabelece diretrizes para a implementação de repositórios digitais. Rio de Janeiro: Conarq, 2015.

APÊNDICE A - Apresentações relacionadas ao Modelo Hipátia

1. Projeto de pesquisa: estudo para implementação do Modelo de Preservação Hipátia no Tribunal de Contas da União

PROJETO DE PESQUISA: ESTUDO PARA IMPLEMENTAÇÃO DO MODELO DE PRESERVAÇÃO HIPÁTIA NO TRIBUNAL DE CONTAS DA UNIÃO

Alexandre Oliveira • Cássio Pires • Marcito Campos • Marcelo Brondani • Marilete Pereira



Prazer, Ibict!



“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.”

HIPÁTIA

Matemática e filósofa que viveu em Alexandria em torno séculos IV e V d.C. Teve seu legado apagado dos registros históricos.



PARCEIROS QUEM IMPLEMENTARAM O HIPÁTIA



DJe e PJe



ARQUIVO NACIONAL

SEI



JUSTIÇA DO TRABALHO DO RIO GRANDE DO SUL

PJe



DJe e SEI



Tribunal de Justiça do
Estado de Minas Gerais

PROJUDI e DJe

OBJETIVOS DO PROJETO

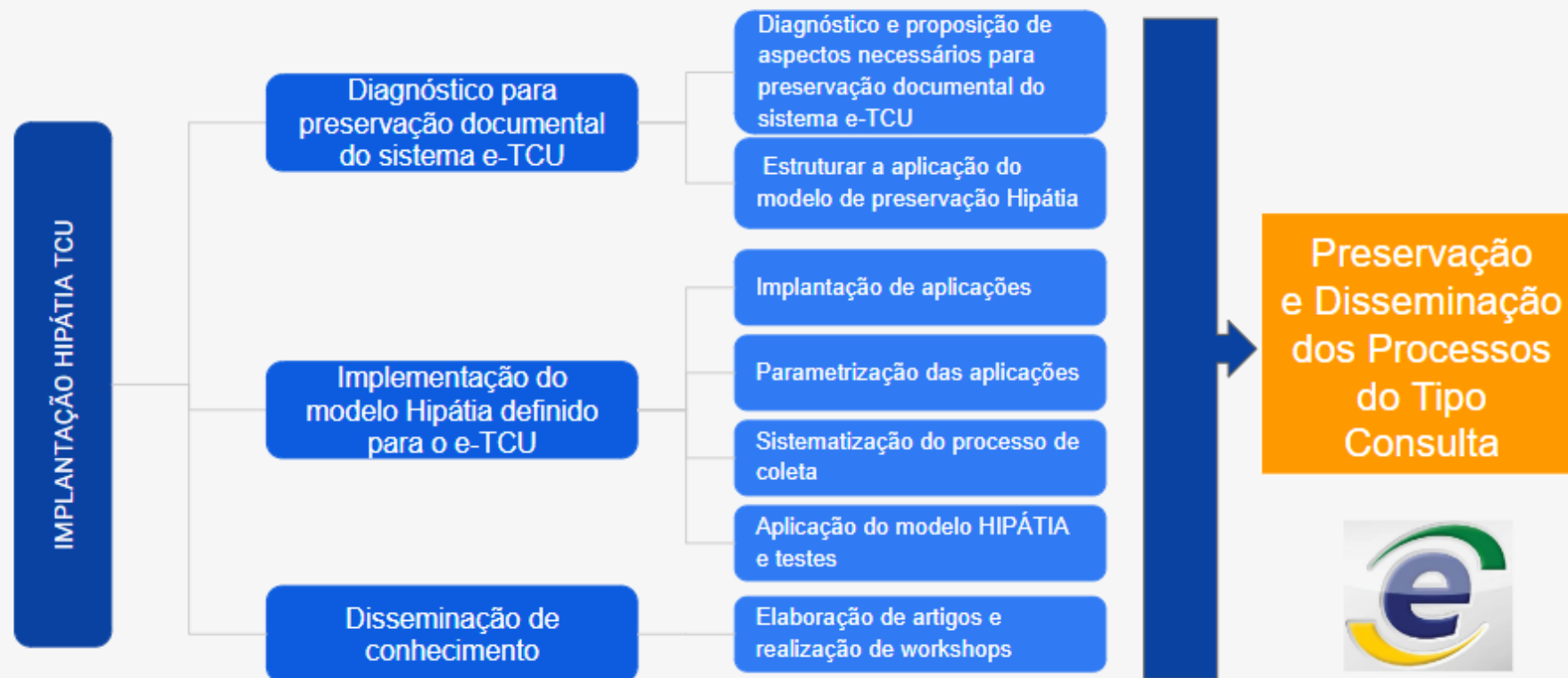
OBJETIVO GERAL

Implementar o RDC-Arq no Tribunal de Contas da União (TCU) adotando como piloto o sistema e-TCU controle externo para preservação baseado no modelo Hipátia em parceria técnica-científica com o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

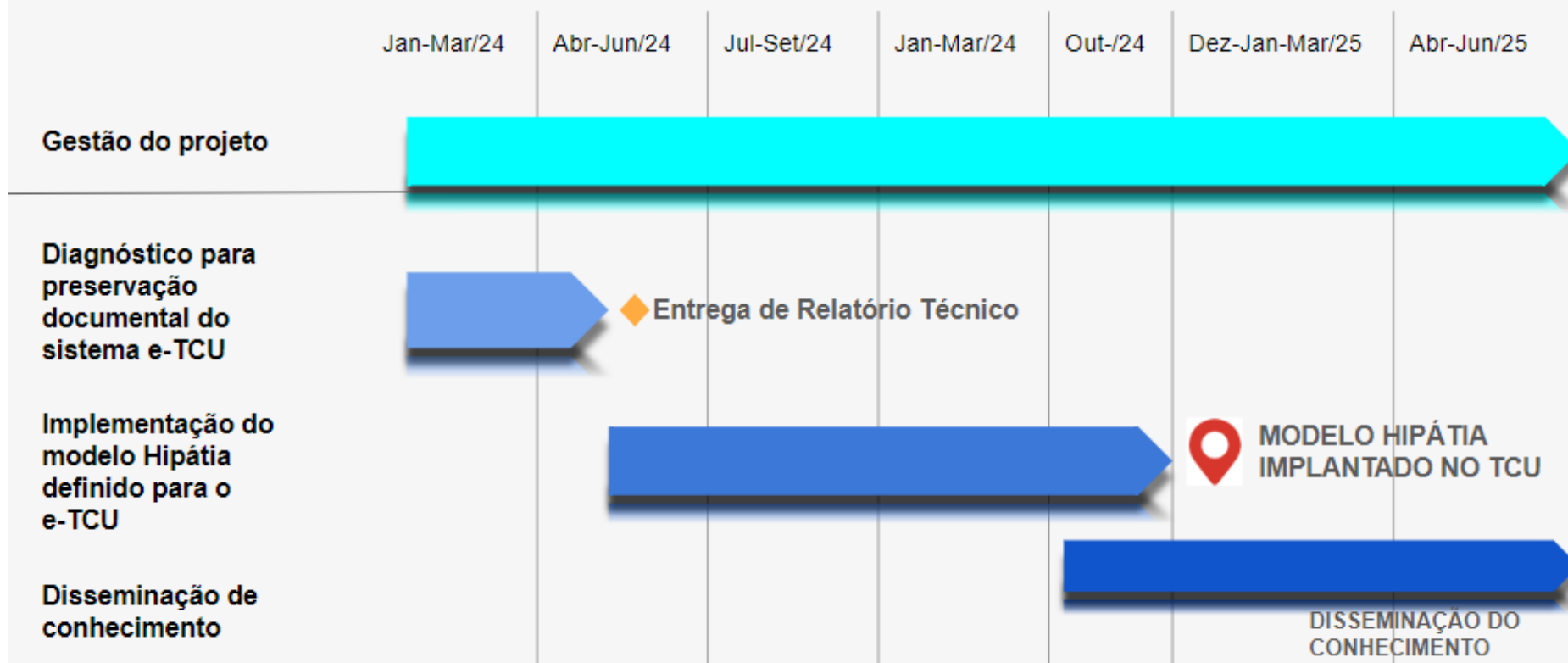
OBJETIVOS ESPECÍFICOS

- OE1 Diagnosticar aspectos relevantes para a proposição e implementação de um RDC-Arq no TCU;
- OE2 Estruturar a aplicação do modelo de preservação Hipátia;
- OE3 Implementar o modelo de preservação Hipátia definido para o sistema e-TCU controle externo;
- OE4 Disseminar o conhecimento construído durante a pesquisa;

ESTRUTURA ANALÍTICA DA PESQUISA



LINHA DO TEMPO



PREMISSAS

- ❖ Haverá apoio e disponibilidade da equipe do TCU para atender a equipe do projeto de pesquisa
- ❖ Não deve haver dependência externa ou de outros projetos para a realização da pesquisa
- ❖ As equipes trabalharão em conjunto para definições de critérios e parâmetros
- ❖ É um projeto de pesquisa e, dessa forma, é caracterizado como um processo de construção do conhecimento e tem como meta gerar novo conhecimento. É um processo de aprendizagem tanto da equipe de pesquisa como da equipe do Tribunal



Backup é
preservação?





Nome: Dorothy Counts

Ano: 1957

Local: Charlotte's Harry Harding High School

País: EUA

Ocasão: Primeira mulher negra a ingressar em uma escola exclusiva para pessoas brancas.

Descrição da cena: Alunos provocando a sra. Dorothy Counts durante evento escolar.

O QUE É PRESERVAÇÃO DIGITAL?

“Conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo o acesso e a interpretação dos documentos digitais pelo tempo que for necessário”.

e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão
Arquivística de Documentos, 2022.

POR QUE REALIZAR A PRESERVAÇÃO DIGITAL?

1. Garantir acesso contínuo à informação no longo prazo;
2. Aprimorar requisitos de segurança da informação;
3. Apoiar a preservação da documentação;
4. Promover a memória institucional;
5. Fomentar a pesquisa e o desenvolvimento científico;
7. Promover a sustentabilidade ambiental;
8. Elevar grau de conformidade regulatória.

RDC-ARQ: REPOSITÓRIO ARQUIVÍSTICO DIGITAL CONFIÁVEL

Fundamentos normativos

- Modelo de referência OAIS (*Open Archival Information System*): ISO 14721:2012;
- ACTDR (*Audit and Certification of Trustworthy Digital Repositories*): ISO 16363:2012;
- Diretrizes para a implementação de Repositórios Arquivísticos Digitais Confiáveis - RDC-Arq, **versão 2**: Resolução CONARQ nº 51, de 25 de agosto de 2023 (revoga Resolução nº 43, de 04 de setembro de 2015).

CONFORMIDADE REGULATÓRIA

EXEMPLOS DO PODER JUDICIÁRIO

- Resolução CNJ nº 324/2020, que institui diretrizes e normas de Gestão de Memória e de Gestão Documental e dispõe sobre o Programa Nacional de Gestão Documental e Memória do Poder Judiciário – Proname;
- Resolução CNJ nº 408/2021, que dispõe sobre o recebimento, o armazenamento e o acesso a documentos digitais relativos a autos de processos administrativos e judiciais;

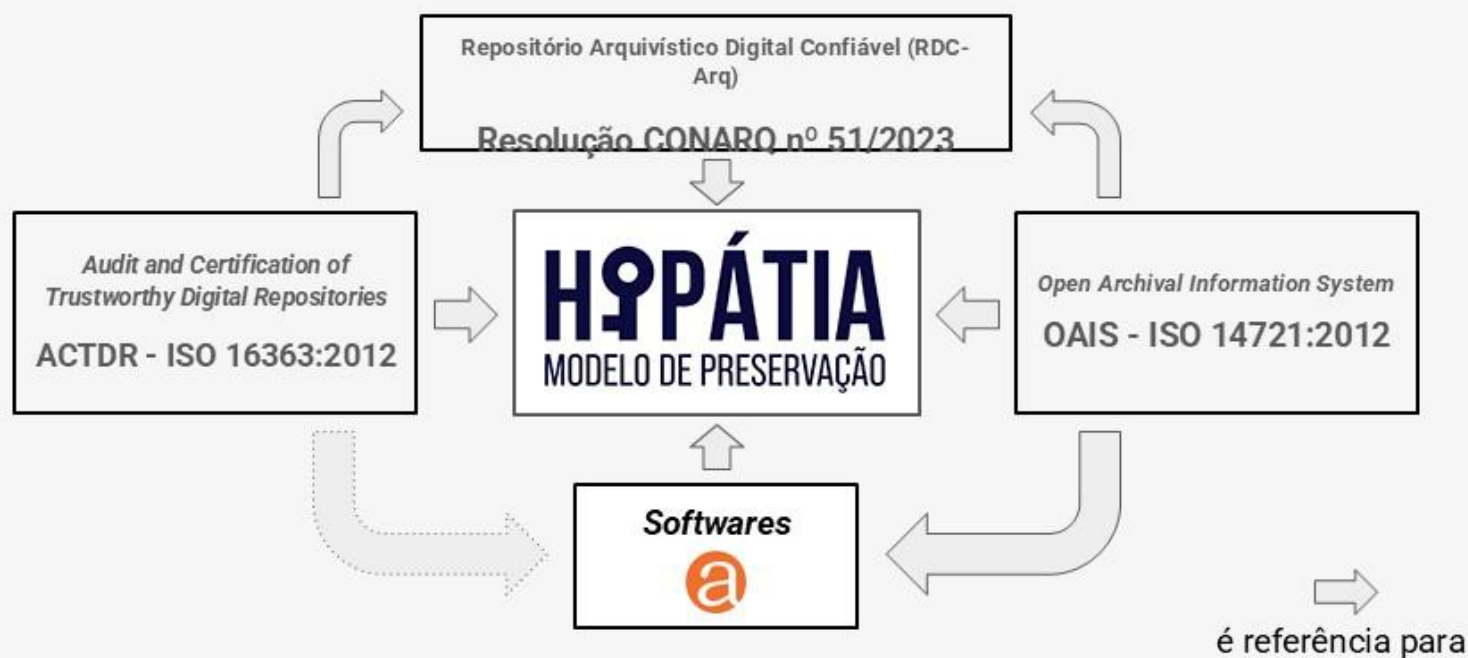
[...]

DIRETRIZ DO TRIBUNAL DE CONTAS DA UNIÃO

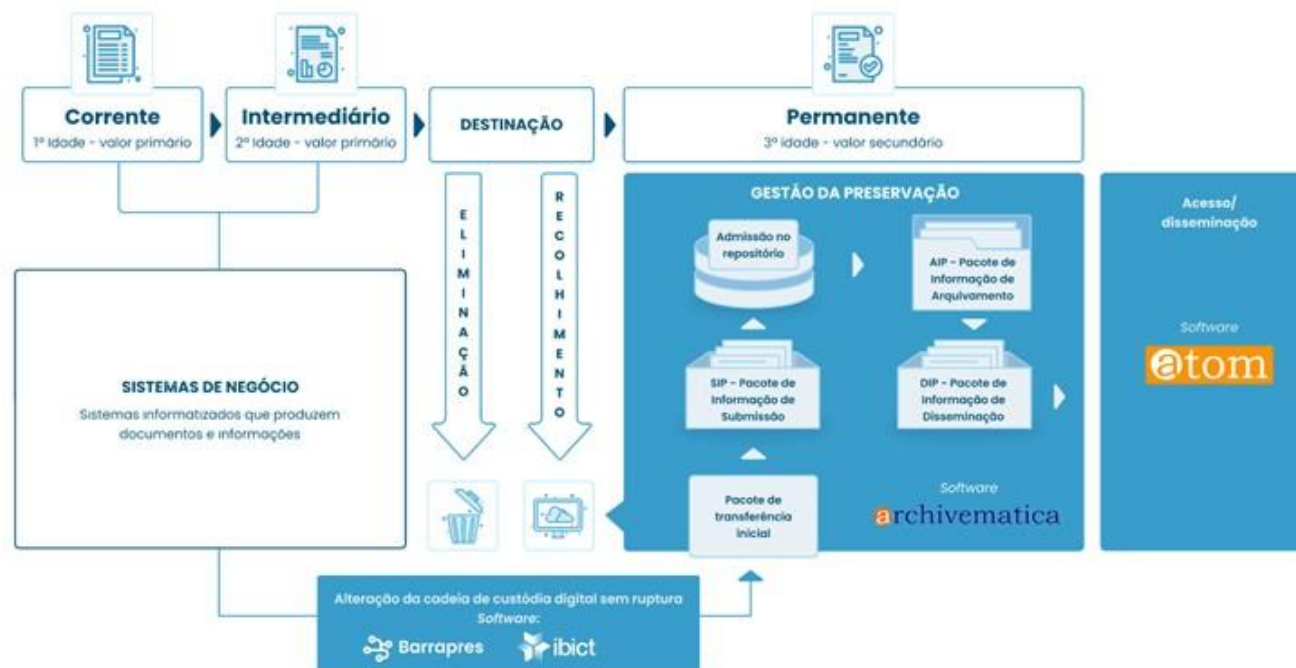
- Portaria TCU nº 85/2023, dispõe sobre o Plano de Preservação de Documentos Eletrônicos do Tribunal de Contas da União (PPDETCU);

[...]

RDC-ARQ: REPOSITÓRIO ARQUIVÍSTICO DIGITAL CONFIÁVEL



RDC-ARQ: REPOSITÓRIO ARQUIVÍSTICO DIGITAL CONFIÁVEL



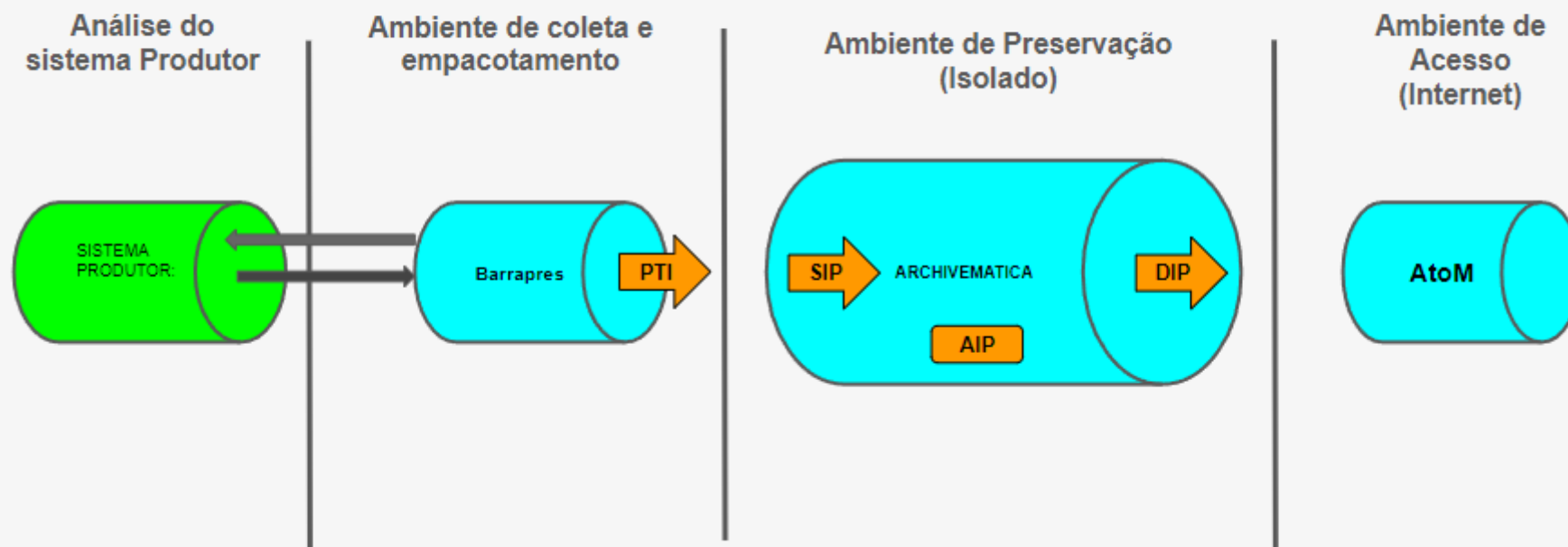
TECNOLOGIAS DO MODELO HIPÁTIA DE PRESERVAÇÃO DIGITAL



 **archivematica**[®]

 **atom**
access to memory

AMBIENTES - IMPLEMENTAÇÃO DO MODELO



AMBIENTE DE PRESERVAÇÃO (INTRANET) - ARCHIVEMATICA

archivematica. Transfer Ingest Preservation planning Access Administration marcelo ▾

Transfer type Accession no. Access system ID ☒ Approve automatically

- Hipatia
- 20189000001.zip (921.7 KB)
 - 20189000002.zip (321.3 KB)
 - 20189000003.zip (956.9 KB)
 - 20189000005.zip (1.4 MB)**
 - 20189000006.zip (1021.2 KB)
 - 20189000008.zip (1.5 MB)
 - 20189000009.zip (747.2 KB)
 - 20189000010.zip (443.1 KB)
 - 20189000650.zip (2.4 MB)
 - 20189001431.zip (61.4 MB)
 - 20189001535.zip (54.3 MB)
 - SF0694200604.zip (134.8 KB)

Transfer UUID Transfer start

archivematica. Transfer Ingest Preservation planning Access Administration marcelo ▾

Transfer type Transfer name Accession no. Access system ID ☒ Approve automatically

Transfer	UUID	Transfer start time	
20189000005	9E14d66ac-57b0-4a91-9761-04E5cc614993	2024-02-06 09:51	
+ Microservice: Create SIP from Transfer			
Job: Create SIP from transfer objects		Completed successfully	
Job: Serialize Dublin Core metadata to disk		Completed successfully	
Job: Move to processing directory		Completed successfully	
Job: Create SIP(s) [7]		Completed successfully	
Job: Load options to create SIPs		Completed successfully	
Job: Check transfer directory for objects		Completed successfully	
+ Microservice: Complete transfer			
+ Microservice: Generate METS.xml document			
Job: Generate METS.xml document		Completed successfully	
+ Microservice: Examine contents			
+ Microservice: Validation			
+ Microservice: Parse external files			
+ Microservice: Characterize and extract metadata			
Job: Load PREMIS events from metadata/premis.xml		Completed successfully	

AMBIENTE DE ACESSO (INTERNET) - AtoM

The screenshot displays the AtoM digital archive interface. At the top, there is a navigation bar with the 'HOPÁLIA' logo, a search bar, and a 'Navegar' dropdown menu. Below the navigation bar, a horizontal menu contains links for 'Início', 'Sobre', 'Política de Privacidade', and 'Ajuda'. The main content area is titled 'Mostrando 5744 resultados' (Showing 5744 results). On the left, there are two filter panels: 'Produtor' (Producer) and 'Nome' (Name). The 'Produtor' panel shows filters for 'Tribunal Regional do Trabalho de 4ª Região' (5737 results) and 'Servio Marco Castañeda Cebal, 1 resultado' (1 result). The 'Nome' panel shows filters for 'Reclamatória trabalhista, 4907 resultados' (4907 results), 'Anelú Lúcia Martin - Julia Vogel, 2389 resultados' (2389 results), and 'Conciliato, 2957 resultados' (2957 results). The main results area displays a grid of document thumbnails, each with a file icon, a file number (e.g., 001/79 - JCI), and the name of the institution (Montenegro). A 'Visualizar impressão' (View print) button is visible at the bottom left of the results area.

AtoM EM NÚMEROS

- 280 no mundo (36 países);
- 74 no Brasil;
- 17 no Judiciário brasileiro

Fonte: Observatório de Documentos Digitais, 2023.

INSTITUIÇÕES DO JUDICIÁRIO BRASILEIRO

Superior Tribunal de Justiça

Superior Tribunal Militar

Tribunal de Justiça de Pernambuco (TJPE) – PE

Tribunal de Justiça do Distrito Federal e Territórios (TJDFT) – DF

Tribunal de Justiça do Estado de Minas Gerais (TJMG) - MG

Tribunal de Justiça do Estado do Amazonas - AM

Tribunal de Justiça do Maranhão - MA

Tribunal de Justiça do Trabalho de Pernambuco (TRT6) - PE

Tribunal Regional Eleitoral de Minas Gerais (TRE-MG) - MG

Tribunal Regional Eleitoral do Ceará (TRE-CE) - CE

Tribunal Regional Eleitoral de Minas Gerais (TRE-MG) - MG

Tribunal Regional Eleitoral do Mato Grosso do Sul (TRE-MS) - MS

Tribunal Regional Federal da 4ª Região - Justiça Federal de Primeiro Grau do Rio Grande do Sul - RS

Tribunal Regional do Trabalho (TRT2) - Justiça do Trabalho da 2ª Região - SP

Tribunal Regional do Trabalho (TRT13) - Justiça do Trabalho da 13ª Região - PB

Tribunal Regional do Trabalho da 4ª Região (TRT4) - RS

Tribunal Regional do Trabalho (TRT21) - Justiça do Trabalho da 21ª Região - RN

Tribunal Superior do Trabalho (TST) – DF

Fonte: Observatório de Documentos Digitais, 2023.

Obrigado!

Alexandre Oliveira

Coordenador de Governança em Tecnologias para Informação e Comunicação (COTIC)

alexandreoliveira@ibict.br
hipatia.ibict.br

2. Estudo para implementação do Modelo de Preservação Hipátia no Tribunal de Contas da União: Seminário de Alinhamento

**ESTUDO PARA IMPLEMENTAÇÃO DO MODELO DE PRESERVAÇÃO
HIPÁTIA NO TRIBUNAL DE CONTAS DA UNIÃO**

SEMINÁRIO DE ALINHAMENTO



Prazer, Ibict!



“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.”



Ecosistema de Preservação



O que é Preservação Digital?

“Conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo o acesso e a interpretação dos documentos digitais pelo tempo que for necessário”.

E-ARQ BRASIL, 2020.

Qual a importância de realizar a Preservação Digital?

- 1- Acesso contínuo à informação;
- 2 - Proteção contra perda;
- 3 - Conservação do patrimônio cultural;
- 4 - Pesquisa e desenvolvimento contínuos;
- 5 - Conformidade regulatória;
- 6 - Memória histórica e documentação;
- 7 - Sustentabilidade ambiental;
- 8 - Aprendizado e educação.

Notícias

 **jornal da unesp**  [Artigos](#) [Acontece](#)

Reportagens

Muito além do backup: bem-vindo à era da preservação digital

Avanço rápido da tecnologia enseja desafio de preservar acessibilidade de arquivos apesar da obsolescência de ferramentas de hardware e software. Iniciativas como a Política de preservação Digital da Unesp buscam mudar cultura da universidade no manejo de objetos digitais.

[f](#) [t](#) [g](#) [p](#)

O GLOBO [Política](#)

Política

Toffoli não vê espaço para rever decisão sobre provas da Odebrecht após novos dados do Ministério da Justiça

A manipulação das provas sem respeitar a cadeia de custódia é um dos principais argumentos utilizados pelo ministro na decisão em que anulou os elementos obtidos a partir da leniência da Odebrecht

Por Jennifer Gualarte — Brasília
15/09/2023 14h19 - Atualizado há 5 meses

[f](#) [x](#) [s](#)

 **STJ SUPERIOR**
TRIBUNAL DE JUSTIÇA

[Institucional](#) [Processos](#) [Jurisprudência](#) [Precedentes](#) [Comunicação](#) [Leis e normas](#)

Você está em: [Início](#) | [Comunicação](#) | [Notícias](#)

Notícias

ESPECIAL

23/04/2023 06:50

A cadeia de custódia no processo penal: do Pacote Anticrime à jurisprudência do STJ

Em 2019, o **Pacote Anticrime (Lei 13.964)** regulamentou a **cadeia de custódia** no Código de Processo Penal (CPP), estabelecendo: "considera-se cadeia de custódia o conjunto de todos os procedimentos utilizados para manter e documentar a história cronológica do vestígio coletado em locais ou em vítimas de crimes, para rastrear sua posse e manuseio a partir de seu reconhecimento até o descarte".

A partir da definição jurídica da cadeia de custódia, o Pacote Anticrime instituiu a regulamentação sobre uma das questões mais sensíveis do processo penal: a guarda dos vestígios do delito. Afinal, caso não haja o recolhimento correto dos vestígios logo após o crime, a sua preservação durante as fases policial e judicial e o seu acondicionamento até a decisão final no processo, a chamada quebra da cadeia de custódia pode comprometer a apuração da verdade.

 **HOPÁTIA**
MODELO DE PRESERVAÇÃO

 **ibict**

Backup é
preservação?



**HIPÁTIA: modelo para
implementação de
repositórios
arquivísticos digitais
confiáveis**



Nome: Dorothy Counts

Ano: 1957

Local: Charlotte's Harry Harding High School

País: EUA

O caso: Primeira mulher negra a ingressar em uma escola exclusiva para pessoas brancas.

Descrição da cena: Alunos provocando a sra. Dorothy Counts durante evento escolar.

Modelo Hipátia

Hipátia

Homenagem à filósofa e matemática da Antiguidade, que viveu em Alexandria (no atual Egito) entre os séculos IV e V.

Início do projeto

Ano: 2018

Instituições: Ibict e TJDFT

Objetivo: apoiar a implementação de um Repositório Arquivístico Digital Confiável (RDC-Arq) por meio de tecnologias *softwares*, em particular, do desenvolvimento de uma camada de barramento interoperável.



Atriz representa Hipátia em peça teatral (1900).
Fonte: Wikipedia. Domínio Público.

Modelo Hipátia



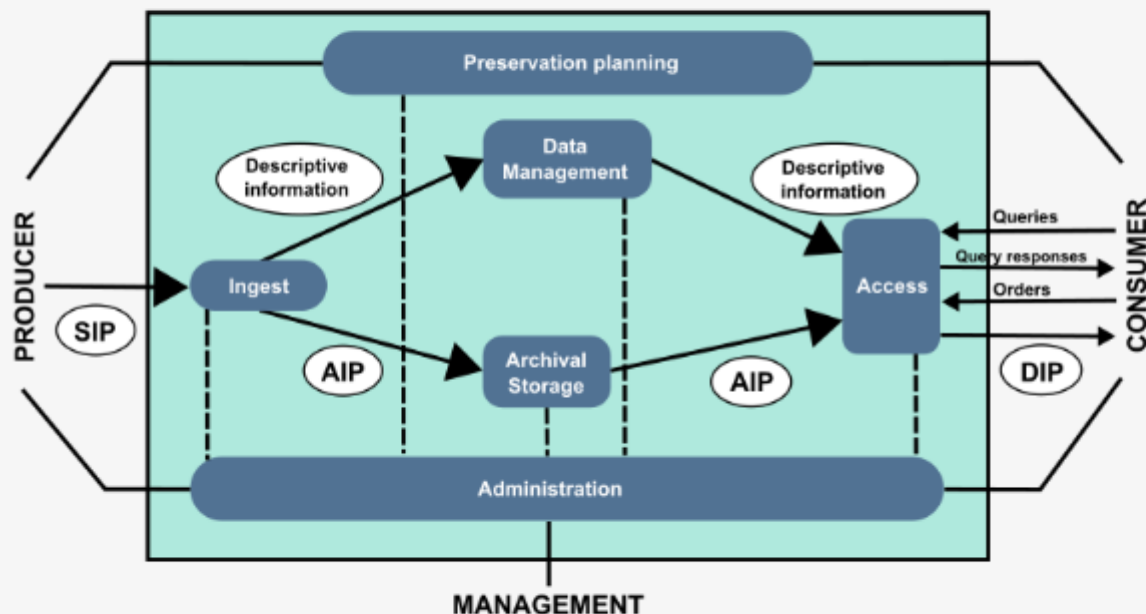
Modelo Hipátia

Modelo de implementação de
Repositórios Arquivísticos Digitais Confiáveis - RDC-Arqs (Conselho Nacional de Arquivos, 2023), que têm como referências:

- Modelo de referência OAIS (ISO 14721:2012) e atributos de um RDC
- Requisitos da TRAC (fundamentos da prática recomendada ACTDR ISO 16363:2012)
- Modelo de implementação de pacote de informação: padrão BagIt
- Padrões de metadados
 - Preservação: PREMIS, METS
 - Gestão: e-Arq Brasil
 - Descrição: ISAD(G), NOBRADE, EAD
- Protocolo para compartilhamento de metadados: OAI-PMH

Modelo de referência OAIS

Perspectiva das entidades funcionais do modelo



Legenda dos pacotes

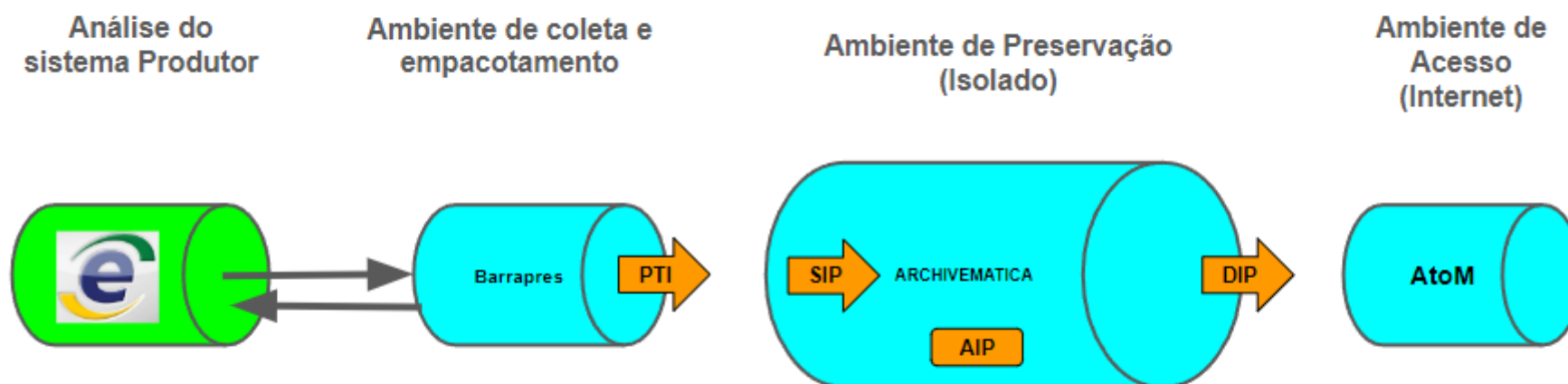
1. SIP
Pacote de Informação de Submissão
1. AIP
Pacote de Informação de Arquivamento
1. DIP
Pacote de Informação de Disseminação

Fonte: Wikipedia. Autores: Mathieu Alex Hache; Mess. CC BY-SA.

Modelo Hipátia

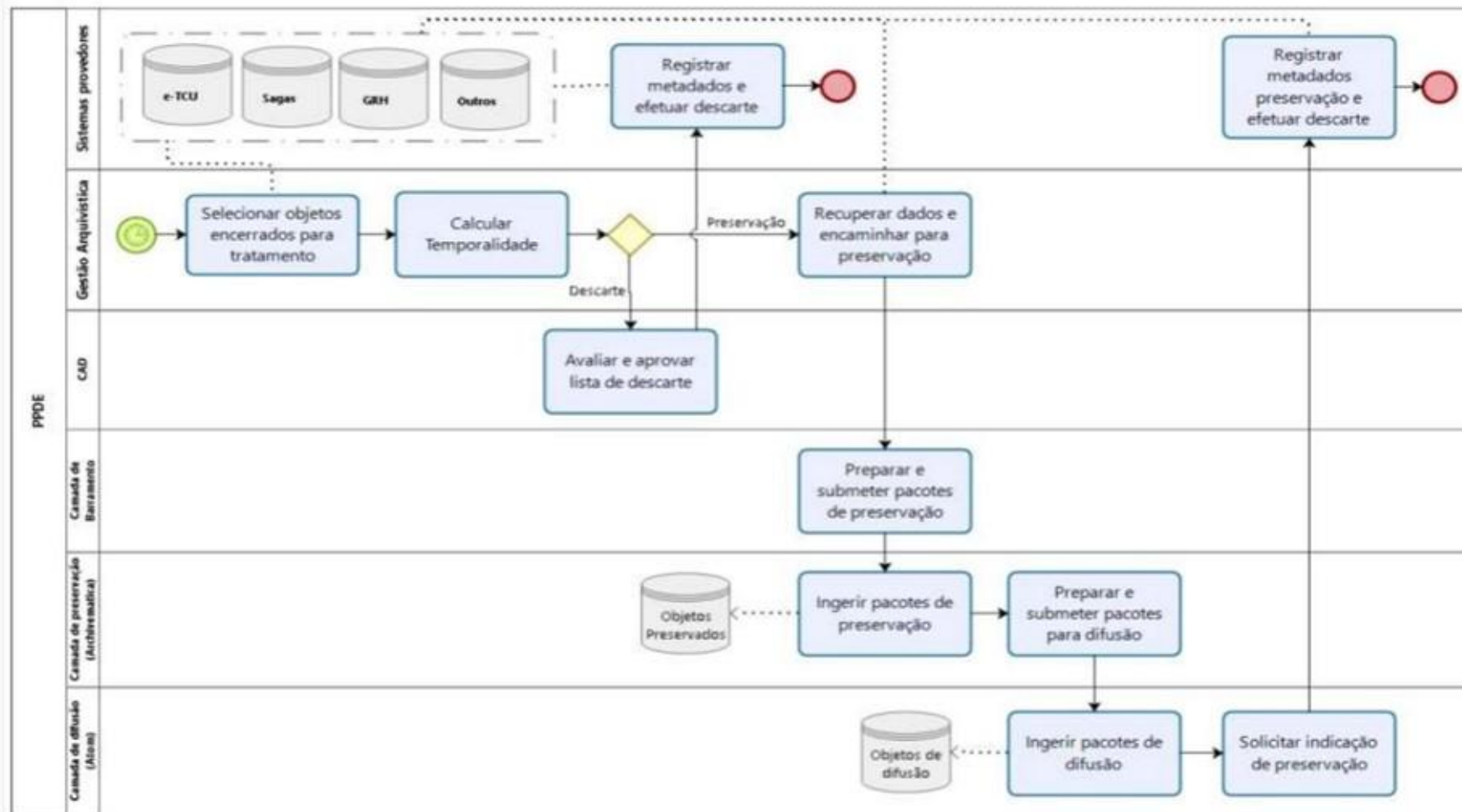


Ambientes - Implementação do Modelo



Tecnologias





Obrigado!
hipatia.ibict.br

3. Estudo para implementação do Modelo de Preservação Hipátia no Tribunal de Contas da União: Softwares do Modelo Hipátia

ESTUDO PARA IMPLEMENTAÇÃO DO MODELO DE PRESERVAÇÃO HIPÁTIA NO TRIBUNAL DE CONTAS DA UNIÃO

Softwares do Modelo Hipátia



Prazer, Ibict!



“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.”



Ecosistema de Preservação



O que é Preservação Digital?

“Conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo o acesso e a interpretação dos documentos digitais pelo tempo que for necessário”.

e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos, 2022.

Qual a importância de realizar a Preservação Digital?

1. Garantir acesso contínuo à informação no longo prazo;
2. Aprimorar requisitos de segurança da informação;
3. Apoiar a preservação da documentação;
4. Promover a memória institucional;
5. Fomentar a pesquisa e o desenvolvimento científico;
6. Promover a sustentabilidade ambiental;
7. Elevar grau de conformidade regulatória.

Backup é
preservação?





Nome: Dorothy Counts

Ano: 1957

Local: Charlotte's Harry Harding High School

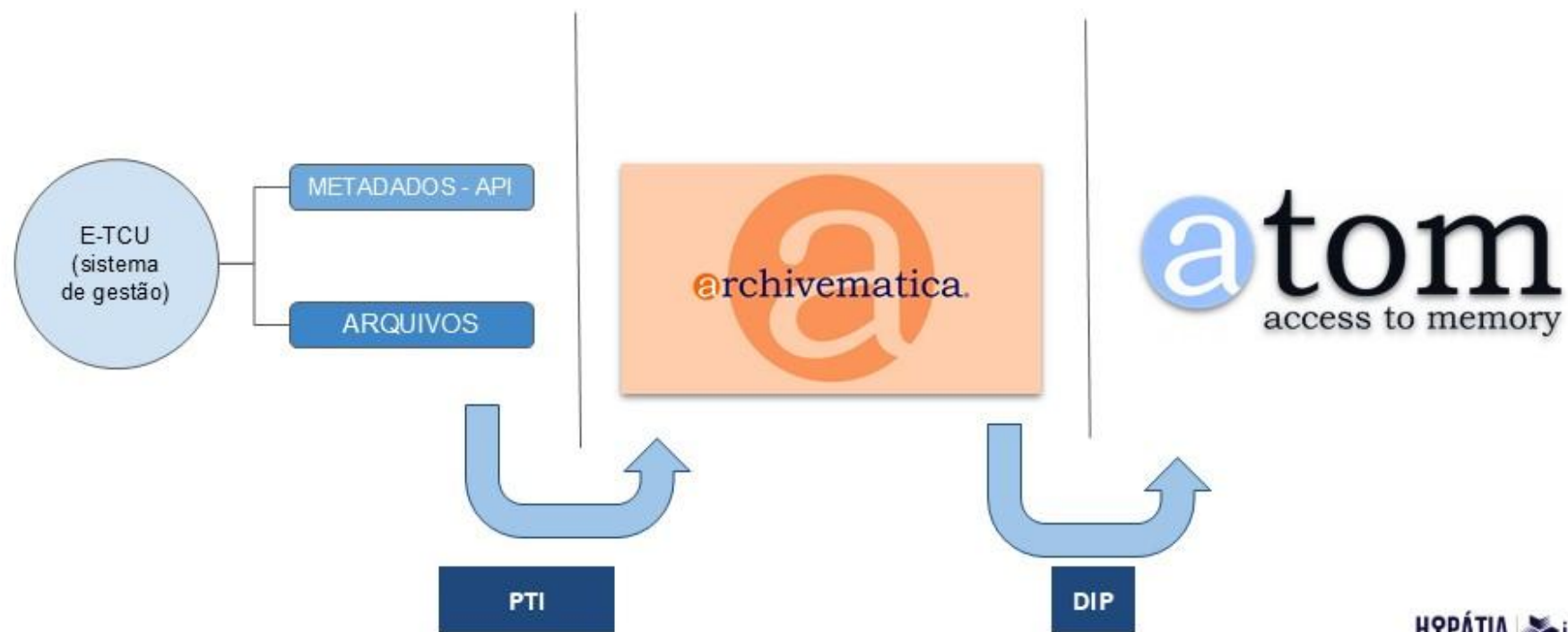
País: EUA

Ocasão: Primeira mulher negra a ingressar em uma escola exclusiva para pessoas brancas.

Descrição da cena: Alunos provocando a sra. Dorothy Counts durante evento escolar.

Modelo de Preservação

Softwares



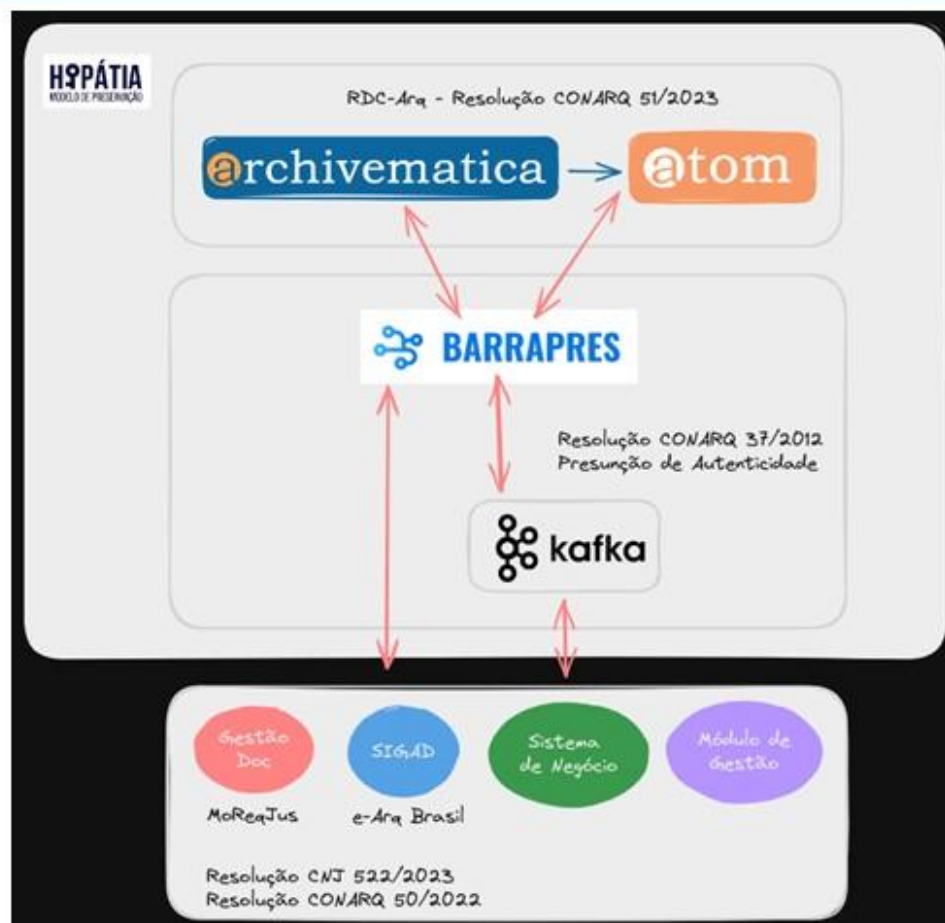


BarraPres

Em função da solução arquivística, em uma cadeia de custódia arquivística, para uma preservação digital sistêmica, faz-se necessária a instalação de um '*Barramento*', que tem como função, a integração e extração de metadados entre os sistemas, que irão compor o ambiente de preservação RDC-Arq.

- Arquivos Gerais
 - Init.py
 - Imports.py
 - Main.py
 - Arquivos Adicionais
 - Bagit.py
- Controllers
 - Adicionais.py
 - AtualizaStatus.py
 - Buscador.py
 - ClasseGeraL.py
 - CriarDiretorios.py
 - CriarStructMap.py
 - EnvioArquivematica.py
 - ObtemDados.py
 - VerificaStatus.py
- Models
 - Archivematica.py
 - Banco_sistema.py
 - Enviados.py
 - Lista_processos_manuais.py
 - Models.py
 - MySQLModel.py
 - Parametros.py
 - Sistemas.py
 - Tentativas.py
 - WebserviceSistemas.py
- Banco de Dados
 - Comando SQL: sqlCreateTables.sql
 - Comando SQL: sqlInsertData.sql





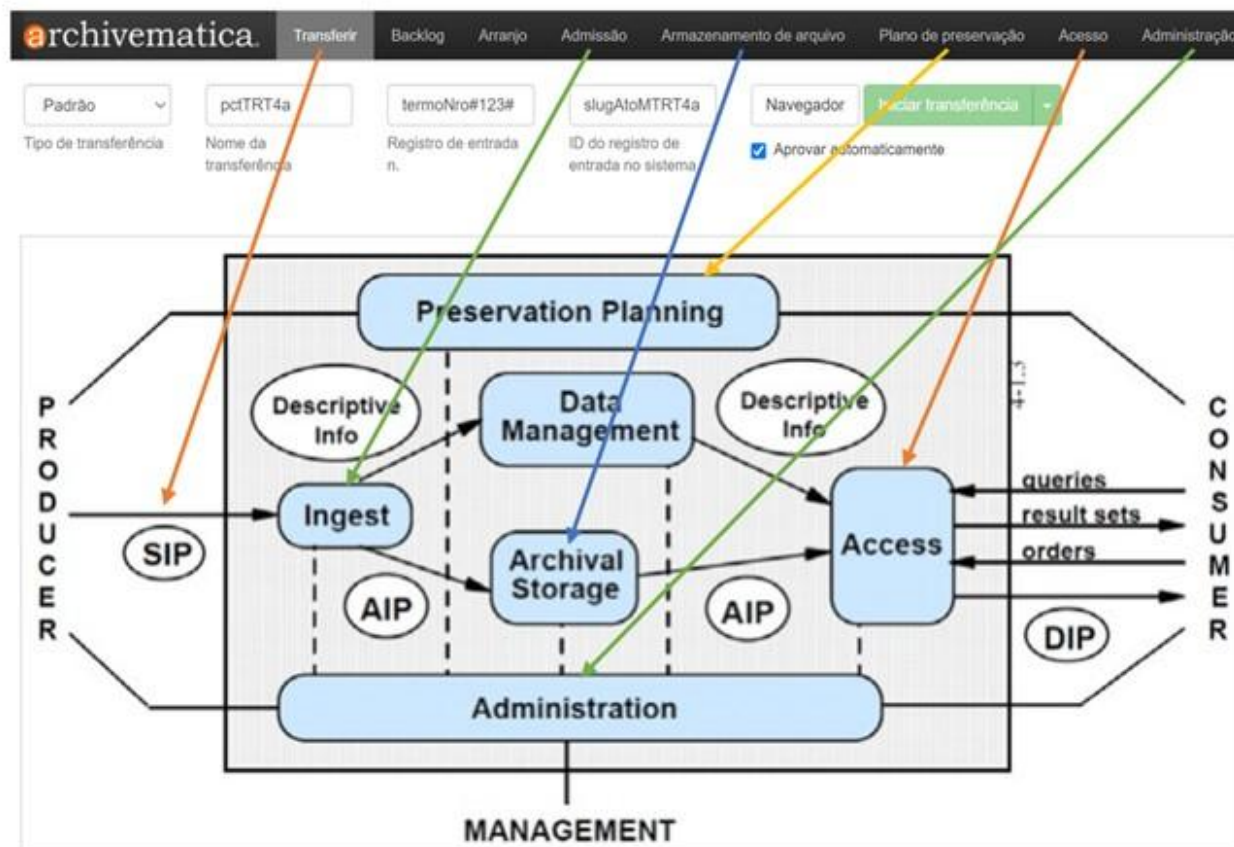
archivematica®

Composição

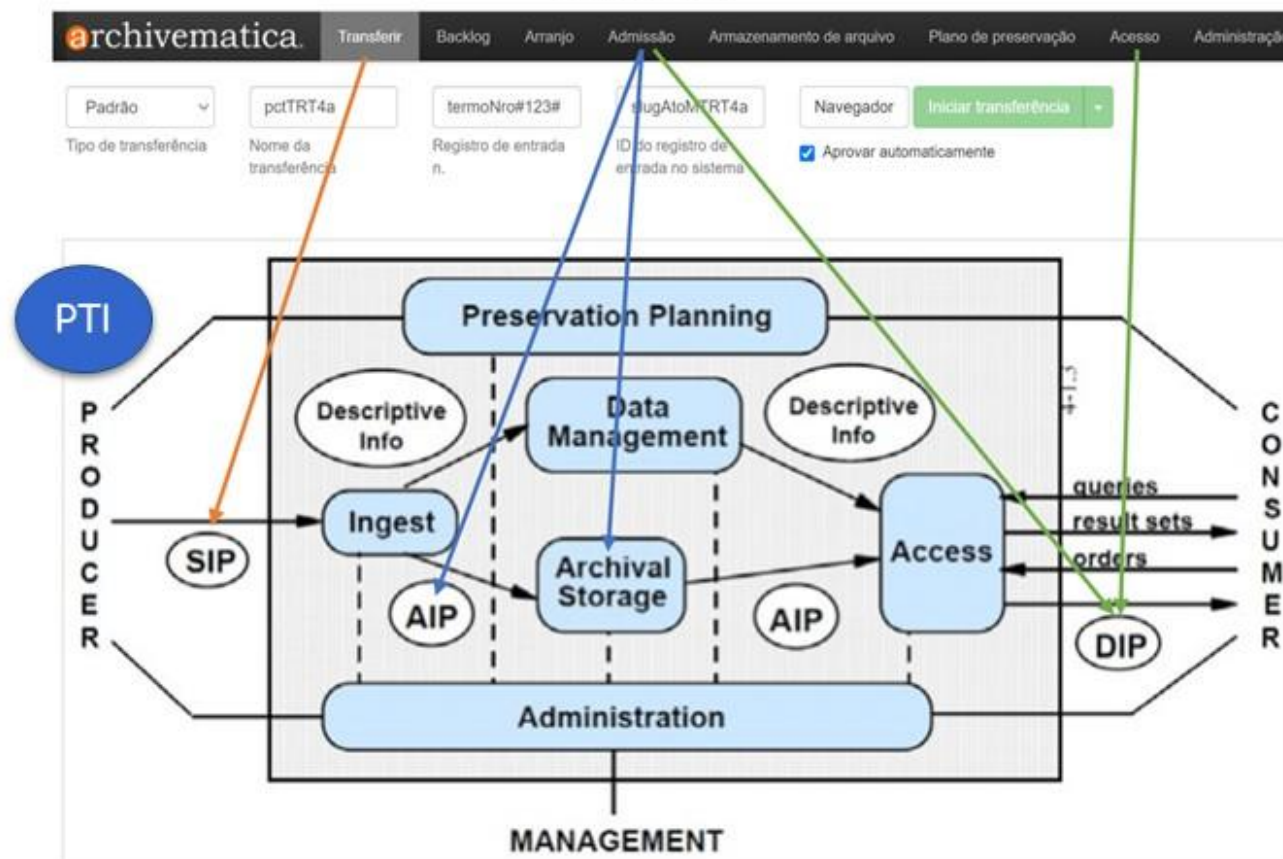
Sistema de dupla checagem



OAIS



OAIS



Bagit

Objetos digitais:

admitidos, derivadas de preservação/acesso

+

Informações de conteúdo do Pacote

metadados (gestão, descritivos, preservação)
salvos em METS XML

```
<base directory>/  
|   bagit.txt  
|   manifest-<algorithm>.txt  
|   [optional additional tag files]  
\--- data/  
    |   [payload files]  
\--- [optional tag directories]/  
    |   [optional tag files]
```

<https://tools.ietf.org/html/draft-kunze-bagit-10>

Dublin Core

Title

Part of AIC

Optional: leave blank if unsure

Creator

Subject

Description

Publisher

Contributor

Date

Use ISO 8601 (YYYY-MM-DD or YYYY-MM-DD/YYYY-MM-DD)

Format

Identifier

Source

Relation

Language

Use ISO 639

Coverage

Rights

atom

access to memory

Sobre o AtoM

- Plataforma de acesso
 - Desenhado para web.
 - Código aberto.
 - Baseado em padrões descritivos internacionais.
 - Importação/ exportação amigáveis.
 - Multilíngue.
 - Multi-repositório.

Fonte: Site da Artefactual - <https://www.accesstomemory.org/pt-br/> Acesso em 11/10/2022

Aspectos Arquivísticos AtoM

Aspectos arquivísticos

Contempla normas de descrição arquivística internacionais



2000



2004



2007



2008

Aspectos arquivísticos

- Contempla outros padrões descritivos
 - RAD – Rules for Archival Description
 - DACS – Describing Archives: A Content Standard
 - **DC – Dublin Core Metadata Element Set**
 - MODS – Metadata Object Description Schema

CANADIAN ARCHIVAL STANDARD

RULES FOR ARCHIVAL
DESCRIPTION



 **DublinCore**

MODS Metadata Object Description Schema
Official Web Site

Aspectos arquivísticos

- Explicita a estrutura de classificação/ arranjo dos documentos

Fundo TJDF - Fundo TJDF



Fonte: AtoM do TJDF - <https://memoria.tjdft.jus.br/index.php/> Acesso em 11/10/2022

AtoM em números

- 280 no mundo (36 países)
- 74 no Brasil;
- 17 no Judiciário brasileiro

Fonte: Observatório de Documentos Digitais, 2024.

Instituições do Judiciário Brasileiro

- Superior Tribunal de Justiça
- Superior Tribunal Militar
- Tribunal de Justiça de Pernambuco (TJPE) – PE
- Tribunal de Justiça do Distrito Federal e Territórios (TJDFT) – DF
- Tribunal de Justiça do Estado de Minas Gerais (TJMG) - MG
- Tribunal de Justiça do Estado do Amazonas - AM
- Tribunal de Justiça do Maranhão - MA
- Tribunal de Justiça do Trabalho de Pernambuco (TRT6) - PE
- Tribunal Regional Eleitoral de Minas Gerais (TRE-MG) - MG
- Tribunal Regional Eleitoral do Ceará (TRE-CE) - CE
- Tribunal Regional Eleitoral do Mato Grosso do Sul (TRE-MS) - MS
- Tribunal Regional Federal da 4ª Região - Justiça Federal de Primeiro Grau do Rio Grande do Sul - RS
- Tribunal Regional do Trabalho (TRT2) - Justiça do Trabalho da 2ª Região - SP
- Tribunal Regional do Trabalho (TRT13) - Justiça do Trabalho da 13ª Região - PB
- Tribunal Regional do Trabalho da 4ª Região (TRT4) - RS
- Tribunal Regional do Trabalho (TRT21) - Justiça do Trabalho da 21ª Região - RN
- Tribunal Superior do Trabalho (TST) – DF

Fonte: Observatório de Documentos Digitais, 2024.

Obrigado!
hipatia.ibict.br

4. Estudo para implementação do Modelo de Preservação Hipátia no Tribunal de Contas da União: Definições de infraestrutura computacional

ESTUDO PARA IMPLEMENTAÇÃO DO MODELO DE PRESERVAÇÃO HIPÁTIA NO TRIBUNAL DE CONTAS DA UNIÃO

DEFINIÇÕES DE INFRAESTRUTURA COMPUTACIONAL

Marcelo Moro Brondani

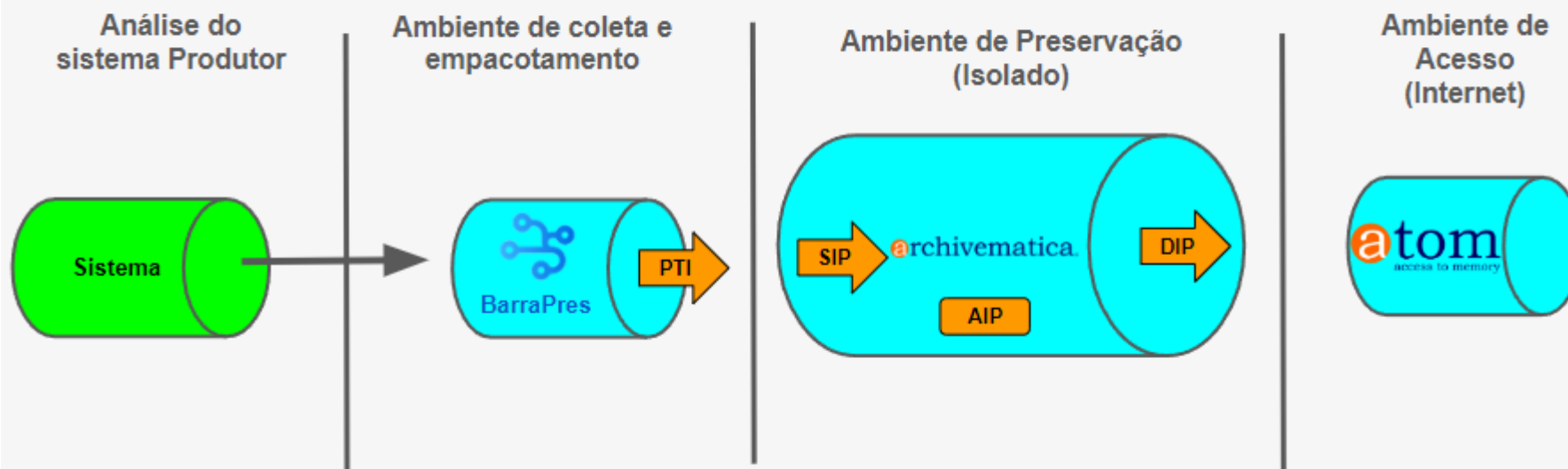


Prazer, Ibict!

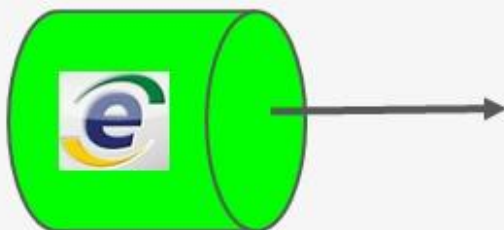


“Promover a competência, o desenvolvimento de **recursos e a infraestrutura de informação** em **ciência e tecnologia** para a produção, socialização e integração do conhecimento científico e tecnológico.”

Implementação do modelo

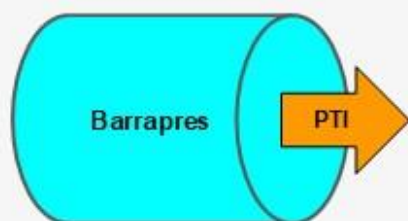


Análise do Sistema Produtor



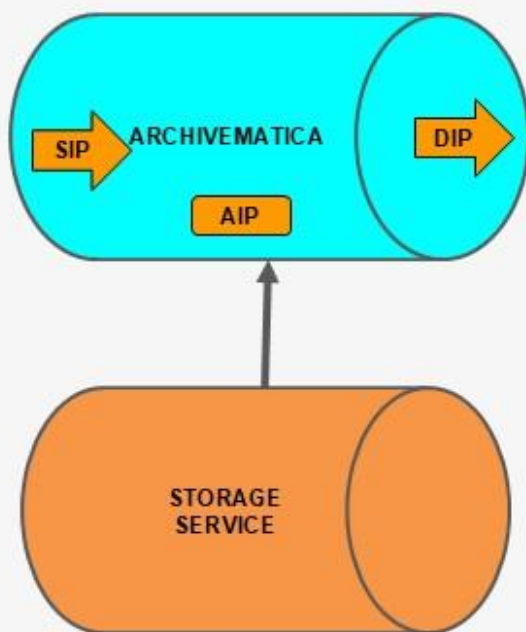
- Análise dos sistemas
- Construção de dicionários de dados
- Definição das conexões para coleta (via API, WEBSERVICE, BANCO DE DADOS, Coleta de diretórios)
- Definição dos parâmetros de coleta
- Construção dos parâmetros do PTI

Ambiente de coleta e empacotamento



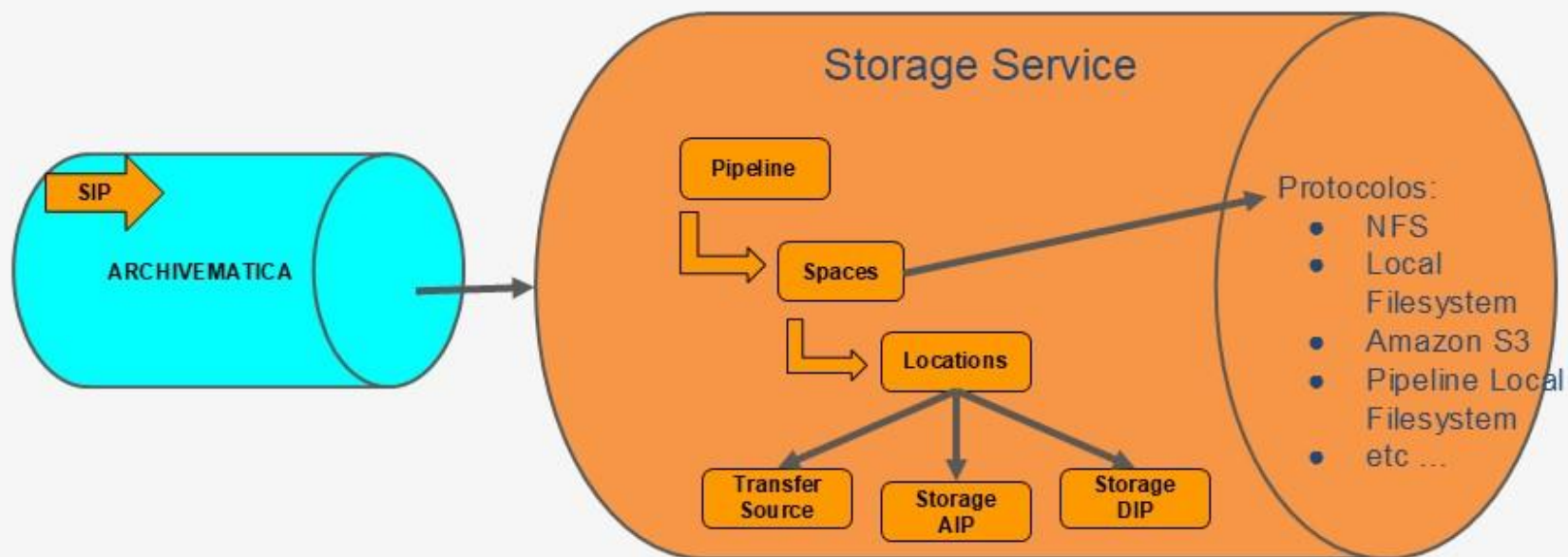
- BarraPres - Python, Mysql
 - Recursos:
 - Mapeamento dos metadados oriundos do produtor
 - Extração de metadado do objeto digital
 - Integração ao RDC-Arq
- ValidaPres - Captura de informações dos documentos digitalizados

Ambiente de Preservação

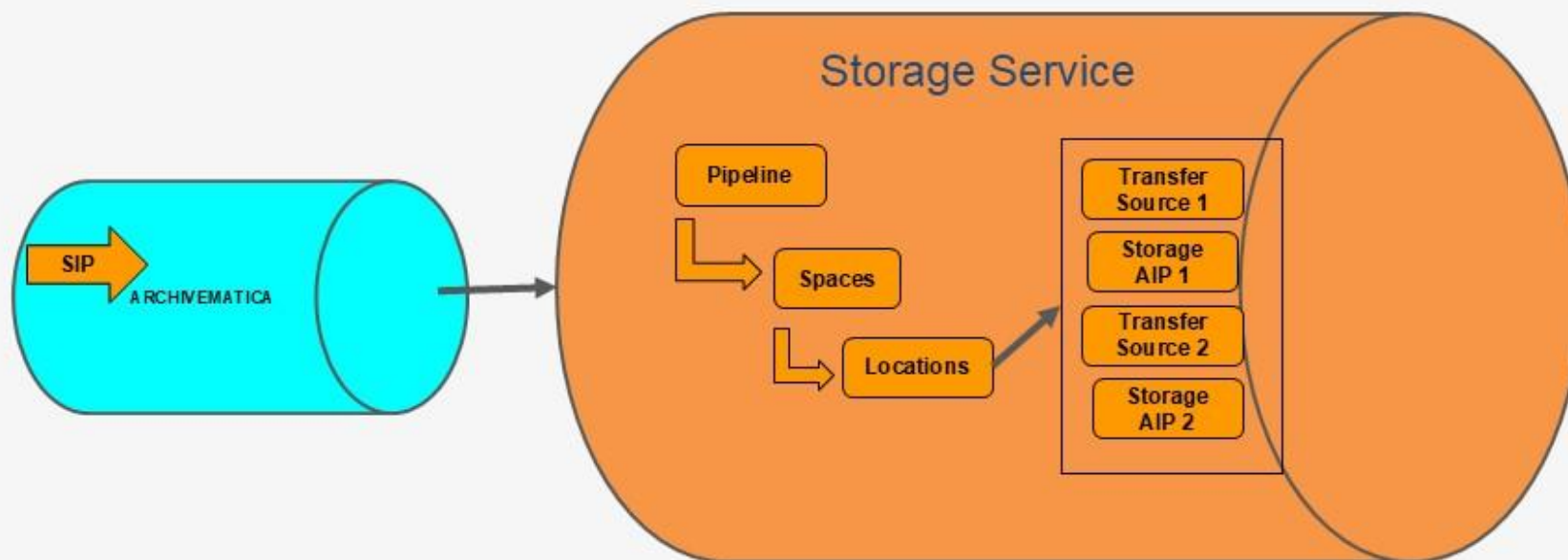


- Versão Atual
 - 1.15.0 - Archivematica (Dashboard, Server, Client)
 - 0.21.0 - Storage Service
- Requisitos de Hardware
 - 8 CPU Cores
 - 16 GB Memória RAM
 - 200GB de disco para Processamento
- Requisitos de Software
 - Ubuntu Server 22.04 LTS
 - Rocky Linux 9/Red Hat
 - MySQL 8
 - Elasticsearch 6.x (opcional)

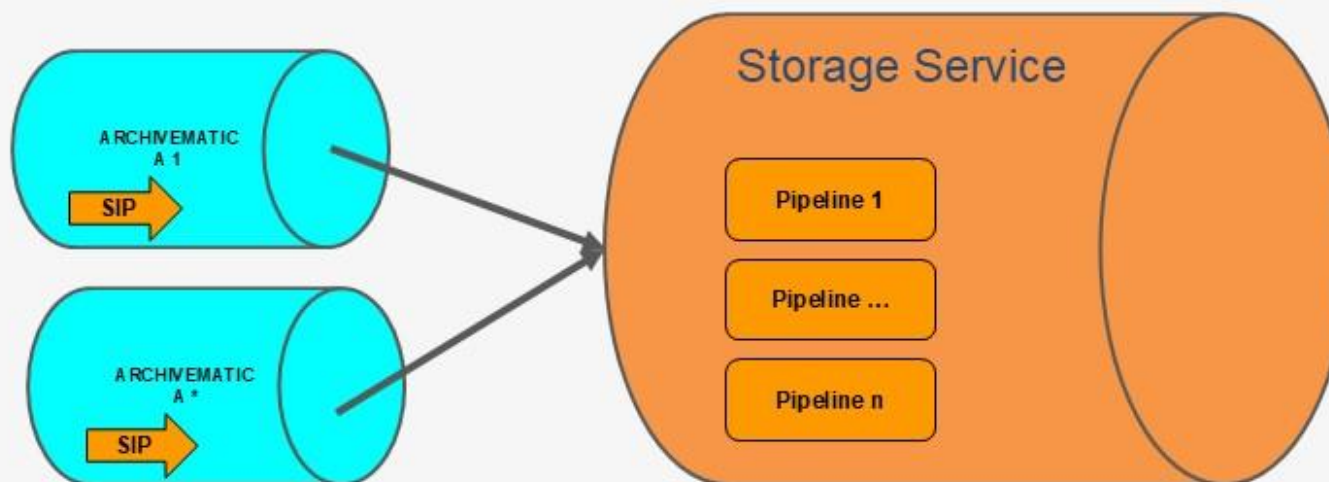
Arquiteturas - Padrão



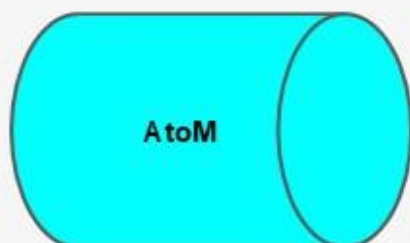
Arquitetura Centralizada - Multi Storage Location



Arquitetura Distribuída - Multi Pipelines



Ambiente de Acesso (Internet)



- Versão Atual
 - 2.7.3 - AtoM
- Requisitos de Hardware (Mínimo)
 - 2 CPU Cores
 - 7 GB Memória RAM
 - 50GB de disco para processamento
- Requisitos de Software
 - Ubuntu Server 20.04 LTS
 - CentOS 7
 - Java 8
 - MySQL 8
 - Elasticsearch 5.x
 - PHP 7.4

Obrigado!
hipatia.ibict.br

APÊNDICE B - Modelo de Maturidade em Preservação Digital

1. Apresentação - Modelo de Maturidade em preservação digital



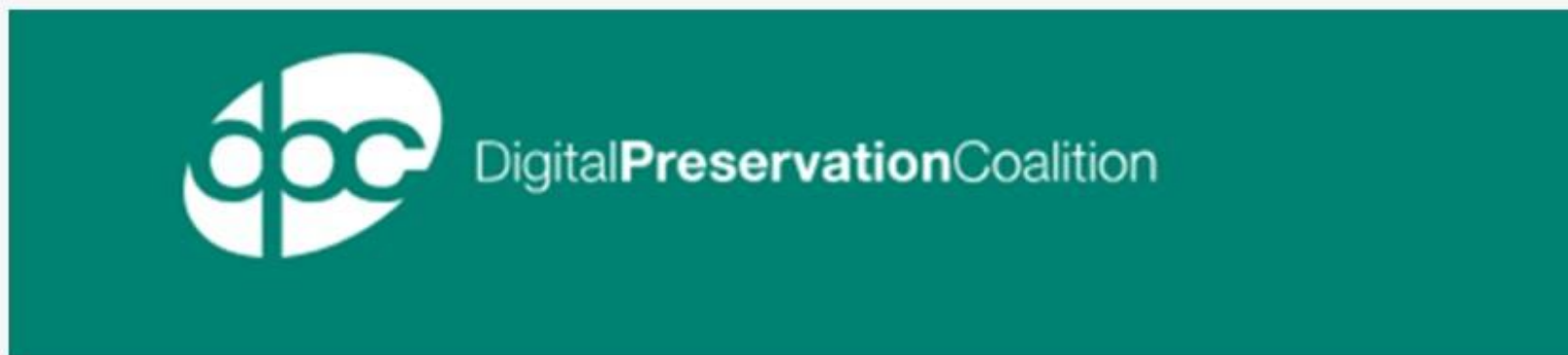
Modelo de Maturidade em Preservação Digital

DPC-RAM Rapid Assessment



Sobre o Modelo

11 seções cada uma com 5 níveis (Consciência Mínima, Consciência, Básico, Gerência, Otimizado)



Capacidades organizacionais

Viabilidade Organizacional

Política e estratégia

Base Legal

Capacidade de TI

Aperfeiçoamento Contínuo

Comunidade

Capacidades de serviços

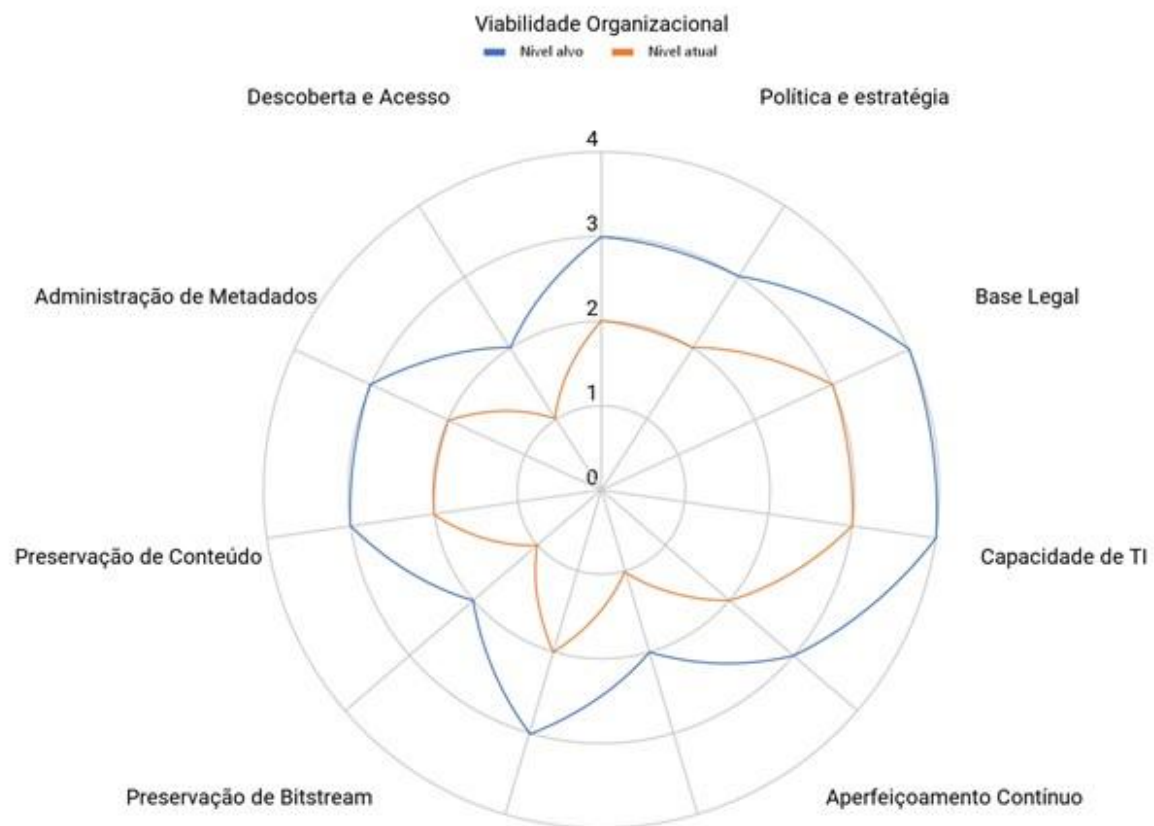
Aquisição, transferência e incorporação

Preservação de Bitstream

Preservação de Conteúdo

Administração de Metadados

Descoberta e Acesso



A. Viabilidade Organizacional: Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.	2 - Básico	Minuta de Política de Preservação
B. Políticas e Estratégia: Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.	2 - Básico	Estudos, capacitações
C. Base Legal: Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.	3 - Gerenciado	Auditoria de RDC-Arq e implementação do modelo OAIS
D. Capacidade de TI: Recursos de tecnologia da informação para apoiar atividades de preservação digital.	3 - Gerenciado	Instalação e transferência de conhecimento
E. Aperfeiçoamento Contínuo: Processos para avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.	2 - Básico	Capacitações e publicações
F. Comunidade: Engajamento e contribuição para a comunidade mais ampla de preservação digital.	1 - Conscientização	Eventos e capacitações

G. Aquisição, Transferência e Incorporação: Processos para adquirir ou transferir conteúdo e incorporá-lo a um arquivo digital.	2 - Básico	Implementação da Cadeia de Custódia
H. Preservação de Bitstream: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.	1 - Conscientização	Processamento de microsserviços do AM e agregador de padrões (PREMIS...) e armazenamento em RDC-Arq
I. Preservação de Conteúdo: Processos para preservar o sentido ou funcionalidade do conteúdo digital e garantir seu acesso e usabilidade contínuos no futuro.	2 - Básico	
J. Gerenciamento de Metadados: Processos para criar e manter metadados suficiente para dar suporte a preservação, gerenciamento e uso de conteúdo digital preservado.	2 - Básico	Estudo de metadados e extração deles (PTI)
K. Descoberta e Acesso: Processos para permitir a descoberta de conteúdo digital e prover acesso aos usuários.	1 - Conscientização	Implementação do ATOM



2. Digital Preservation Coalition Rapid Assessment Model (DPC RAM) - Modelo de Maturidade

Digital Preservation Coalition Rapid Assessment Model (DPC RAM)			
Assessment Worksheet			
Organização:			
Avaliação completa até:			
Avaliação completa em:			
Notas sobre o escopo da avaliação (tipo do conteúdo ou departamento):			
	Born digital archives / Digitised document collections / Digitised AV collections / Email / Ebooks / Etheses / Research data / Websites / Other (Delete as appropriate)		
Prazo utilizado para níveis almeçados (ex: 1/3/5/10 anos)			
Assessment completed by:			
Assessment complete on:			

ORGANIZATIONAL CAPABILITIES

	Nível Atual	Por que esse nível foi selecionado?	Nível Almejado	O que precisa ser implementado para chegar lá?
A. Viabilidade Organizacional: Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.				
B. Políticas e Estratégia: Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.				
C. Base Legal: Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.				
D. Capacidade de TI: Recursos de tecnologia da informação para apoiar atividades de preservação digital.				
E. Aperfeiçoamento Contínuo: Processos para avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.				
F. Comunidade: Engajamento e contribuição para a comunidade mais ampla de preservação digital.				

CAPACIDADES DE SERVIÇO

	Nível Atual	Por que esse nível foi selecionado?	Nível Almejado	O que precisa ser implementado para chegar lá?
G. Aquisição, Transferência e Incorporação: Processos para adquirir ou transferir conteúdo e incorporá-lo a um arquivo digital.				
H. Preservação de Bitstream: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.				
I. Preservação de Conteúdo: Processos para preservar o sentido ou funcionalidade do conteúdo digital e garantir seu acesso e usabilidade contínuos no futuro.				
J. Gerenciamento de Metadados: Processos para criar e manter metadados suficiente para dar suporte a preservação, gerenciamento e uso de conteúdo digital preservado.				
K. Descoberta e Acesso: Processos para permitir a descoberta de conteúdo digital e prover acesso aos usuários.				

3. Seções do DPC RAM

O Modelo

Existem 11 seções do DPC RAM, cobrindo diferentes elementos da capacidade de preservação digital, agrupados em duas partes. Os recursos organizacionais são definidos em um nível organizacional ou outro alto nível de granularidade apropriado. Os recursos de serviço se referem a níveis operacionais que podem ser considerados em um nível mais baixo de granularidade, possivelmente específico para um fluxo de conteúdo específico.

Capacidades Organizacionais		
A	Viabilidade Organizacional	Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.
B	Política e estratégia	Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.
C	Base Legal	Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.
D	Capacidade de TI	Recursos de tecnologia da informação para apoiar atividades de preservação digital.
E	Aperfeiçoamento Contínuo	Processos para a avaliação das capacidades de preservação digital atuais, a definição de metas e o monitoramento do progresso.
F	Comunidade	Engajamento e contribuição para a comunidade mais ampla de preservação digital.

Capacidades de Serviço		
G	Aquisição, transferência e incorporação	Processos para adquirir ou transferir conteúdo e incorporá-lo em um arquivo digital.
H	Preservação de Bitstream	Processos para garantir o armazenamento e a integridade do conteúdo digital a ser preservado.
I	Preservação de Conteúdo	Processos para preservar o significado ou a funcionalidade do conteúdo digital e garantir sua acessibilidade e usabilidade contínuas ao longo do tempo.
J	Administração de Metadados	Processos para criar e manter metadados suficientes para oferecer suporte à preservação, descobrimento e uso de conteúdo digital preservado.
K	Descoberta e Acesso	Processos para permitir a descoberta de conteúdo digital e fornecer acesso aos usuários.

Capacidades Organizacionais

A – Viabilidade Organizacional	
Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.	
0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de apoiar atividades de preservação digital.
1 – Consciência	A organização tem conhecimento da necessidade de apoiar atividades de preservação digital.
2 – Básico	As atividades de preservação digital são apoiadas e recebem recursos em um nível básico dentro da organização, por exemplo: • Há algum envolvimento da gerência sênior. • A equipe atribuiu responsabilidades e o tempo para assumi-las. • Um orçamento para preservação digital foi alocado (ainda que seja por tempo limitado). • Os requisitos de desenvolvimento da equipe foram identificados.

3 – Gerência	As atividades de preservação digital são geridas e apoiadas dentro da organização, por exemplo: • Há comprometimento da alta gerência. • A responsabilidade pela preservação digital é claramente assumida. • A equipe possui as habilidades necessárias para realizar atividades de preservação digital e acesso a conhecimentos relevantes, quando necessário. • Um orçamento principal dedicado à preservação digital foi alocado. • Orçamentos, funções da equipe e necessidades de desenvolvimento são avaliados regularmente. • Estatísticas e relatórios podem ser gerados sobre o arquivo digital para ajudar a informar os relatórios, o planejamento e o gerenciamento. • Os requisitos de desenvolvimento pessoal foram atendidos. • A Preservação Digital foi identificada como uma prioridade estratégica.
4 – Otimizado	As atividades de preservação digital são geridas de forma proativa, aprimoradas e desenvolvidas dentro da organização, por exemplo: • Os benefícios da preservação digital são reconhecidos, defendidos e incorporados em toda a organização. • Foi estabelecido um conselho administrativo inter-departamental de preservação digital. • Um ou mais funcionários são considerados especialistas em seu campo. • Orçamentos, funções da equipe e necessidades de desenvolvimento são avaliados proativamente em antecipação a mudanças futuras. • Estatísticas e relatórios sobre o arquivo digital são combinados com projeções de necessidades futuras para informar proativamente os relatórios, o planejamento e o gerenciamento. • A eficácia do desenvolvimento da equipe é monitorada regularmente. • Planos de continuidade e sucessão estão em vigor para garantir a preservação contínua das aquisições caso a organização não possa mais realizar essas atividades.

B – Política e Estratégia Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.	
0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de uma estrutura de políticas para preservação digital.
1 – Consciência	A organização está ciente da necessidade de desenvolver uma estrutura de políticas e pode ter algumas políticas relevantes, mas não existe nenhuma política ou estratégia de preservação digital.
2 – Básico	A organização possui uma estrutura de políticas básicas, por exemplo: • Existe uma política ou estratégia de preservação digital de alto nível. • Outras políticas relacionadas à preservação digital podem existir, mas existem lacunas na cobertura. • Alguns procedimentos para gerenciar e prover acesso ao conteúdo digital estão em vigor e podem ser documentados. • O escopo da coleção é definido e compreendido (por exemplo: política de desenvolvimento de coleções, cronograma de retenção). • O desenvolvimento de políticas e procedimentos é informado por um entendimento básico das necessidades do usuário.
3 – Gerência	A organização possui um conjunto abrangente e gerenciado de políticas, estratégias e procedimentos, por exemplo: • A política/estratégia de preservação digital está alinhada com outras políticas organizacionais e é revisada de acordo com um cronograma acordado. • A política e o procedimento levam em consideração quaisquer questões éticas relevantes. • Existe um conjunto de processos e procedimentos documentados para gerenciar e prover acesso ao conteúdo do arquivo digital. • Todo o pessoal relevante está ciente das políticas, estratégias e procedimentos de preservação digital. • O conhecimento de casos de uso atuais e futuros para conteúdo informa a política e o procedimento (por exemplo, coleta, abordagens de preservação, metadados e acesso).
4 – Otimizado	A organização gerencia proativamente suas políticas, estratégias e procedimentos e tem um compromisso com a melhoria contínua do processo, por exemplo: • Existe um conjunto completo de políticas, estratégias e procedimentos relacionados à preservação e acesso ao conteúdo digital • Política e estratégia são totalmente implementadas e a equipe se envolve ativamente. • Política, estratégia e procedimento são proativamente monitorados e atualizados para refletir mudanças internas, mudanças em outras políticas, necessidades de usuários ou outros fatores externos.

C – Base Legal

Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.

0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de gerenciar direitos e responsabilidades legais, de licenciamento e outros direitos e responsabilidades legais ou de princípios básicos para sua aplicação.
1 – Consciência	A organização está ciente da necessidade de gerenciar direitos e responsabilidades legais, de licenciamento e outros direitos legais e responsabilidades e possui um entendimento dos princípios básicos de como fazê-lo.
2 – Básico	• O gerenciamento básico de direitos e responsabilidades legais com relação ao conteúdo digital é realizado, por exemplo: • Os principais direitos e responsabilidades legais, juntamente com seus proprietários, foram identificados e documentados. • Existem modelos para acordos e licenças legais necessários. • Códigos de conduta relevantes relacionados à ética profissional são respeitados.
3 – Gerência	• Os direitos e responsabilidades legais relacionados ao conteúdo digital são gerenciados, por exemplo: • Informações relacionadas a licenciamento, direitos legais e contratos podem ser facilmente localizadas e acessadas quando necessário. • Questões e riscos legais são gerenciados e revisados regularmente. • As funções e responsabilidades de gerenciamento de questões e riscos jurídicos são claramente atribuídas. • Aconselhamento especializado pode ser acessado quando necessário (por exemplo, de especialistas jurídicos, de compras, gerenciamento de contratos ou conformidade de informações). • Ações realizadas devido a direitos e responsabilidades legais são documentadas. • Diferentes fluxos de trabalho de preservação ou acesso estão em vigor para conteúdo com diferentes requisitos legais ou regulamentares. • As responsabilidades de acessibilidade são cumpridas, de acordo com os requisitos da legislação local ou nacional relevante.
4 – Otimizado	• Os direitos e responsabilidades legais relacionados ao conteúdo digital são gerenciados proativamente, por exemplo: • Questões e riscos legais são proativamente monitorados e mitigados. • A organização se envolve e contribui para processos legais e judiciais que criam regulamentação.

D – Capacidade de TI

Recursos de tecnologia da informação para apoiar atividades de preservação digital.

0 – Consciência mínima	A organização tem um conhecimento mínimo seja da capacidade de TI necessária para suporte do arquivo digital ou dos princípios básicos para sua aplicação.
1 – Consciência	A organização está ciente da capacidade de TI necessária para suporte do arquivo digital e possui um entendimento dos princípios básicos.
2 – Básico	A organização tem acesso a instalações básicas de TI, incluindo infraestrutura e suporte técnico, por exemplo: • Suporte básico de TI está disponível para o arquivo digital. • A equipe com responsabilidades de TI tem uma compreensão básica de sua função no suporte à preservação digital. • Os sistemas de TI são documentados em um nível básico.
3 – Gerência	A organização tem acesso a instalações de TI gerenciadas de maneira abrangente, incluindo infraestrutura e suporte técnico, por exemplo: • Suporte adequado de TI está disponível para o arquivo digital. • As funções e responsabilidades de TI relacionadas à preservação digital são documentadas e revisadas regularmente. • Os sistemas de TI são regularmente corrigidos e atualizados. • Novas ferramentas e sistemas são implantados quando necessário. • Os sistemas de TI são amplamente documentados. • Contratos e serviços com provedores de serviços terceirizados (por exemplo, fornecedores de nuvem) são bem gerenciados e documentados.
4 – Otimizado	A organização tem acesso a instalações de TI gerenciadas proativamente que estão em constante evolução e aprimoramento, por exemplo: • Um nível aprimorado de suporte de TI está disponível para o arquivo digital • A TI demonstra boa compreensão e envolvimento com problemas de preservação digital. • Os requisitos de preservação digital são levados em consideração ao adquirir novos sistemas de TI. • Existe um roteiro detalhado para o desenvolvimento futuro de sistemas de TI. • Novas ferramentas e sistemas em potencial são proativamente identificados e testados.

E – Aperfeiçoamento Contínuo

Processos para a avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.

0 – Consciência mínima	A organização tem um conhecimento mínimo da posição ou dos objetivos atuais.
1 – Consciência	A organização está ciente da necessidade de entender sua posição atual e definir metas.
2 – Básico	A organização possui um entendimento básico dos recursos atuais de capacidade de preservação digital e das áreas de melhoria, por exemplo: • Um exercício inicial de benchmarking foi realizado. • Lacunas na capacidade de preservação digital foram identificadas. • Existe um entendimento de onde a organização é relativa aos pares.
3 – Gerência	A organização possui um processo gerenciado para fazer avaliação comparativa e estabelecer metas, por exemplo: • Metas foram estabelecidas e acordadas com os gerentes seniores. • O roteiro está em vigor para atingir as metas. • O exercício de avaliação comparativa é repetido periodicamente.
4 – Otimizado	A organização realiza a melhoria contínua do processo, com gerenciamento proativo, por exemplo: • A certificação/revisão externa foi alcançada e é mantida conforme apropriado • Recomendações para melhoria foram adotadas. • Metas e roteiro são revisados periodicamente.

F - Comunidade

Engajamento e contribuição para a comunidade mais ampla de preservação digital.

0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de se envolver com a comunidade mais ampla de preservação digital.
1 – Consciência	A organização está ciente dos benefícios da colaboração com a comunidade mais ampla de preservação digital.

2 – Básico	A organização se envolve com a comunidade mais ampla de preservação digital em um nível básico, por exemplo: • Uma ou mais redes de contatos relevantes foram estabelecidas. • Eventos relevantes da comunidade podem ser acessados. • Existe o compromisso de aprender com as experiências dos outros.
3 – Gerência	Há suporte e gerenciamento para o envolvimento com a comunidade mais ampla de preservação digital, por exemplo: • Redes e comunidades relevantes foram unidas. • Um papel ativo é desempenhado na comunidade de preservação digital. • A consultoria especializada em preservação digital pode ser acessada conforme apropriado. • Os sucessos e lições aprendidas do próprio trabalho são compartilhados com a comunidade.
4 – Otimizado	• A organização assume um papel de liderança na comunidade de preservação digital e gerencia proativamente esses compromissos, por exemplo: • É assumido um papel proativo no estabelecimento ou organização de redes comunitárias, atividades ou eventos colaborativos. • Contribuições são feitas a grupos de especialistas, comitês ou forças-tarefa.

Capacidades do Serviço

G - Aquisição, Transferência e Incorporação

Processos para adquirir ou transferir conteúdo e incorporá-lo em um arquivo digital.

0 – Consciência mínima	A organização tem um conhecimento mínimo seja da necessidade de adquirir ou transferir conteúdo para um arquivo digital ou dos princípios básicos para fazê-lo.
1 – Consciência	A organização está ciente da necessidade de adquirir ou transferir conteúdo para um arquivo digital e entende os princípios básicos para fazê-lo.

2 – Básico	A organização implementou um processo básico para aquisição, transferência e incorporação, por exemplo: • Existe um processo documentado de aquisição e incorporação. • Orientações básicas para doadores, depositantes e criadores de registros estão disponíveis quando apropriado. • Documentação e metadados são recebidos ou capturados por vezes como parte do processo de aquisição ou transferência. • Existe um processo documentado para selecionar e capturar conteúdo digital quando apropriado (por exemplo, arquivos da web, arquivos de email, conteúdo digitalizado, registros em EDRMS) • Parte do conteúdo é avaliado como parte de um processo manual, de acordo com as políticas relevantes. • Uma área de trabalho (física ou virtual) está disponível para atividades de anteriores a e de aquisição e incorporação (por exemplo, para realizar verificação de vírus e identificação de arquivos).
3 – Gerência	A organização implementou um processo gerenciado abrangente para transferência, aquisição e incorporação, por exemplo: • Os relacionamentos com doadores, depositantes e criadores de registros são gerenciados por meio de comunicação, orientação e suporte contínuos, quando necessário. • A avaliação é uma parte padrão do fluxo de trabalho de entrada. • Os fluxos de trabalho são eficientes e adequados à finalidade. • Partes do processo de aquisição e incorporação são automatizadas. • A transferência bem-sucedida do conteúdo é verificada pela verificação de integridade.
4 – Otimizado	A organização gerencia e aprimora proativamente o processo de transferência, aquisição e incorporação, por exemplo: • A organização coopera com possíveis doadores, depositantes e criadores de registros para apoiar o gerenciamento do ciclo de vida das melhores práticas. • Sistemas de TI internos que criam e mantêm conteúdo digital que será transferido para o arquivo, são adquiridos e configurados com uma consciência dos requisitos para preservação futura. • O processo de aquisição e incorporação é automatizado onde for benéfico fazê-lo, com a capacidade de fazer intervenções manuais quando necessário. • As ferramentas e sistemas em uso foram totalmente integrados. • Ferramentas de software são aplicadas para automatizar e aprimorar o processo, por exemplo, destacando informações confidenciais ou informando decisões de avaliação. • A reavaliação é realizada periodicamente, levando em consideração fatores como valor do conteúdo, métricas de uso e custos de preservação (financeiros e ambientais).

H – Preservação de Bitstream

Processos para garantir o armazenamento e a integridade do conteúdo digital a ser preservado.

0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de preservação de Bitstream ou dos princípios básicos para sua aplicação.
1 – Consciência	A organização está ciente da necessidade de preservação de Bitstream e possui um entendimento dos princípios básicos.
2 – Básico	A organização implementou um processo básico para preservação de Bitstream, por exemplo: • O armazenamento dedicado a isso está disponível para atender às necessidades atuais de preservação. • Os funcionários sabem onde o conteúdo está armazenado. • A replicação é baseada em regimes simples de backup. • As somas de verificação são geradas para todo o conteúdo. • Existe um entendimento de que funcionários devem estar autorizados a acessar o conteúdo.
3 – Gerência	A organização armazena o conteúdo de maneira gerenciada, consistente com as boas práticas de preservação para replicação e verificação de integridade. Por exemplo: • O conteúdo é gerenciado com uma combinação de verificação de integridade e replicação de conteúdo em um ou mais locais. • As decisões sobre a frequência da verificação de integridade e o número de cópias mantidas levam em consideração riscos, o valor do conteúdo e custos (financeiros e ambientais). • O conteúdo que falha na verificação de • A integridade é reparada. • As autorizações de funcionários para acessar o conteúdo são aplicadas e documentadas. • Testes são rotineiramente realizados para verificar a eficácia dos backups, replicação e verificação de integridade.

4 – Otimizado	A organização aplica um regime de armazenamento altamente gerenciado com gerenciamento proativo de riscos, por exemplo: • Cópias separadas geograficamente são mantidas para minimizar o risco de perda devido a desastres. • Diferentes tecnologias ou serviços de armazenamento estão em uso. • As necessidades futuras de armazenamento são previstas e atualizadas regularmente e a capacidade de armazenamento é monitorada e revisada de acordo. • A integridade do conteúdo e os processos para verificar a integridade são revisados independentemente • Todo o acesso ao conteúdo é registrado e revisado para uso não autorizado e/ou alterações feitas: qual conteúdo, quando e por quem.
---------------	---

I – Preservação de Conteúdo

Processos para preservar o significado ou a funcionalidade do conteúdo digital e garantir sua acessibilidade e usabilidade contínuas ao longo do tempo.

0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de preservação de conteúdo ou dos princípios básicos para sua aplicação.
1 – Consciência	A organização está ciente da necessidade de preservação de conteúdo e entende os princípios básicos para sua aplicação.
2 – Básico	A organização implementou um processo básico para entender o conteúdo que eles possuem, por exemplo: • Formatos de arquivo são identificados. • O conteúdo é caracterizado e avaliado por questões de preservação e qualidade, como conteúdo criptografado, quebrado ou incompleto e arquivos inválidos. • Há um entendimento básico dos usuários atuais e futuros e casos de uso para o conteúdo.

3 – Gerência	A organização implementou um processo gerenciado para monitorar e planejar a acessibilidade do conteúdo ao longo do tempo, por exemplo: • As atividades de observação tecnológica são realizadas e o conteúdo "em risco" é identificado. • Dependências técnicas são detectadas e documentadas. • Ocasionalmente, são realizadas ações para garantir a preservação e a qualidade do conteúdo, como migração, emulação ou modificação da criação ou captura de fluxos de trabalho. • As ações de preservação ocorrem com a compreensão das propriedades do objeto digital que devem ser retidas para dar suporte aos casos de uso atuais e futuros. • Todas as alterações no conteúdo digital são gravadas, incluindo detalhes de quando, o que, como, por que e quem.
4 – Otimizado	A organização adota uma abordagem proativa para priorizar e mitigar os riscos de preservação para garantir que o conteúdo seja acessível ao longo do tempo, por exemplo: • Riscos para formatos de arquivo específicos ou tipos de conteúdo mantidos são bem compreendidos. • Um processo rigoroso de planejamento de preservação identifica ações de preservação apropriadas para mitigação de riscos. • As decisões sobre a execução de ações de preservação levam em consideração riscos, valor do conteúdo, custos (financeiros e ambientais) e casos de uso. • Migrações de formato, normalizações, emulação e outras ações de preservação são implementadas de acordo com os planos de preservação. • Existe um controle de qualidade para avaliar (e registrar) o resultado das ações de preservação garantindo que o significado e / ou funcionalidade do conteúdo foi mantido conforme necessário. • Conteúdo digital e metadados são controlados por versão, quando apropriado.

J – Administração de Metadados

Processos para criar e manter metadados suficientes para oferecer suporte à preservação, descoberta e uso de conteúdo digital preservado.

0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de gerenciar metadados ou dos princípios básicos para sua implementação.
1 – Consciência	A organização está ciente da necessidade de gerenciar metadados e possui um entendimento dos princípios básicos.
2 – Básico	A organização cria e mantém metadados para preservação, descoberta e uso em um nível básico, por exemplo: • O conteúdo é descrito no nível da coleção em um registro de ativos digitais. • Existe um requisito descritivo mínimo adequado para metadados. • Metadados e a documentação adquirida com o conteúdo são mantidos e preservados. • Metadados básicos de preservação são capturados no nível do item.
3 – Gerência	A organização implementou um processo gerenciado para criar e manter metadados de preservação, descoberta e uso: • Padrões apropriados de metadados são identificados. • Orientação interna e vocabulários controlados estão em vigor para garantir a consistência da entrada de metadados. • Identificadores exclusivos persistentes são atribuídos e mantidos para conteúdo digital. • Os relacionamentos estruturais entre os elementos de dados e metadados que formam um objeto digital específico são mantidos.
4 – Otimizado	A organização realiza um gerenciamento proativo de metadados para preservação, descoberta e uso, e procura maneiras de desenvolver e aprimorar processos, por exemplo: • Existe uma rica quantidade de metadados para conteúdo digital, onde apropriado. • Padrões apropriados de metadados são aplicados. • A escolha dos padrões de metadados é revisitada e revisada periodicamente. • Metadados e documentação podem ser aprimorados ao longo da vida útil do conteúdo. • Metadados permite uma experiência mais avançada de renderização/reutilização para o usuário. • Metadados são colhíveis e reutilizáveis. • A estratégia de saída gerenciada é facilitada pelos procedimentos padronizados de empacotamento de conteúdo e metadados.

K – Descoberta e Acesso	
Processos para permitir a descoberta de conteúdo digital e fornecer acesso aos usuários.	
0 – Consciência mínima	A organização tem um conhecimento mínimo da necessidade de permitir à comunidade de usuários a descoberta e o acesso ao arquivo digital ou dos princípios básicos para realizar isso.
1 – Consciência	A organização está ciente da necessidade de permitir a descoberta e o acesso à sua comunidade de usuários e possui um entendimento dos princípios básicos.
2 – Básico	A organização implementou um mecanismo básico de descoberta e acesso (onde os direitos de acesso permitem), por exemplo: • Existe uma ferramenta básica de busca para alguns conteúdos digitais. • Os usuários podem visualizar ou acessar conteúdo e metadados digitais, remotamente ou no local. • O acesso dos usuários ao conteúdo digital é registrado. • Informações sobre a acessibilidade do conteúdo digital são fornecidas aos usuários.
3 – Gerência	A organização implementou um processo abrangente e gerenciado de descoberta e acesso (onde os direitos de acesso permitem), por exemplo: • A busca básica de recursos existe para todo o conteúdo digital. • A pesquisa de termos no texto completo está disponível para alguns conteúdos digitais. • As informações de direitos são exibidas e o acesso gerenciado pelo sistema sempre que possível. • Relatórios podem ser gerados sobre o acesso do usuário ao conteúdo digital. • Os sistemas de acesso são atualizados para refletir os comentários da comunidade de usuários. • As informações de descoberta de recursos estão disponíveis para os usuários em formatos acessíveis. • Há um caso de uso de acesso estabelecido para extração em massa de todo o conteúdo digital durante a invocação de uma estratégia de saída. As informações de descoberta de recursos estão disponíveis para os usuários em formatos acessíveis. • Há um caso de uso de acesso estabelecido para extração em massa de todo o conteúdo digital durante a invocação de uma estratégia de saída.

4 – Otimizado	A organização implementou um mecanismo avançado de descoberta e acesso (onde os direitos de acesso permitem) que é proativamente aprimorado e aprimorado, por exemplo: • Ferramentas avançadas de descoberta e acesso a recursos são fornecidas, como pesquisa facetada, visualização de dados ou acesso personalizado via APIs. • Diferentes opções estão disponíveis para acesso, renderização ou reutilização, como conteúdo migrado, emulado e visualizado. • Os direitos são totalmente gerenciados pelos sistemas de acesso, incluindo a emissão de contratos para reutilização. • O suporte ao acesso é fornecido aos usuários pela organização. • A comunidade de usuários é consultada proativamente para estabelecer e antecipar necessidades expectativas. • As informações coletadas sobre a descoberta e o acesso ao conteúdo digital são usadas para melhorar e aprimorar a experiência do usuário. • O conteúdo digital está disponível para usuários em formatos acessíveis. • Os mecanismos de acesso são compatíveis ou integram ferramentas de acessibilidade comuns.
---------------	--

**APÊNDICE C - Relatório Modelo de Maturidade em
Preservação Digital**

Modelo de Maturidade em Preservação Digital

Implementação do Modelo
Hipátia de Preservação Digital no Tribunal
de Contas da União (TCU)

PRESIDÊNCIA DA REPÚBLICA

Luiz Inácio Lula da Silva
Presidente da República

Geraldo José Rodrigues Alckmin Filho
Vice-Presidente da República

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Luciana Santos
Ministra da Ciência, Tecnologia e Inovação

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Tiago Emmanuel Nunes Braga
Diretoria

Carlos André Amaral de Freitas
Coordenação de Administração - COADM

Ricardo Medeiros Pimenta
Coordenação de Ensino e Pesquisa em Informação para a Ciência e Tecnologia - COEPI

Henrique Denes Hilgenberg Fernandes
Coordenação de Planejamento, Acompanhamento e Avaliação - COPAV

Cecília Leite Oliveira
Coordenação-Geral de Informação Tecnológica e Informação para a Sociedade - CGIT

Washington Luís Ribeiro de Carvalho Segundo
Coordenação-Geral de Informação Científica e Técnica - CGIC

Hugo Valadares Siqueira
Coordenação-Geral de Tecnologias de Informação e Informática - CGTI

Milton Shintaku
Coordenação de Tecnologias para Informação - COTEC

Alexandre Faria de Oliveira
Coordenação de Governança em Tecnologias para Informação e Comunicação - COTIC

©2024 INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA

Este Relatório é licenciado sob uma licença [Creative Commons Attribution 4.0](https://creativecommons.org/licenses/by/4.0/), sendo permitida a reprodução parcial ou total desde que mencionada a fonte.

EQUIPE TÉCNICA

Diretor do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)
Tiago Emmanuel Nunes Braga

Coordenador-Geral de Tecnologias de Informação e Informática (CGTI)
Hugo Valadares Siqueira

Coordenador do Projeto
Alexandre Faria de Oliveira

Autores
Alexandre Faria de Oliveira
Ívina Flores Melo
Marilete da Silva Pereira

Normalização
Marilete da Silva Pereira

Diagramação
Alisson Eugênio Costa

Este Relatório Técnico é um produto do projeto Implementação de modelo de preservação Hipátia no Tribunal de Contas da União (TCU).

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia e Inovação.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Sumário

1 INTRODUÇÃO	5
2 MODELOS DE MATURIDADE: GÊNESE DA PROPOSTA	5
2.1 Modelos de Maturidade em Preservação Digital: Ponderações Iniciais	7
3 RAM DPC: DINÂMICA DE FUNCIONAMENTO E APLICAÇÃO NO TCU	10
3.1 Capacidades Organizacionais	16
3.2 Capacidades de Serviço	17
4 CONSIDERAÇÕES FINAIS	19
BIBLIOGRAFIA CONSULTADA	21

1 INTRODUÇÃO

O presente relatório é apresentado como resultado da parceria estabelecida entre o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e o Tribunal de Contas da União (TCU) para a implementação de um Repositório Digital Confiável de Arquivos (RDC-Arq). O objetivo desta parceria é viabilizar a preservação digital dos documentos arquivísticos armazenados no sistema e-TCU.

Especificamente, este relatório aborda a aplicação do Modelo de Maturidade com o intuito de mensurar a evolução da preservação digital no TCU, identificando os pontos fortes e áreas que necessitam de melhoria. Para isso, foram realizados diagnósticos e análises, considerando as práticas atuais de preservação digital e comparando-as com os padrões internacionais estabelecidos. A implementação do Modelo de Maturidade proporciona uma visão detalhada do estágio de desenvolvimento do TCU em termos de preservação digital, oferecendo subsídios para o planejamento estratégico e a tomada de decisões informadas para o aprimoramento contínuo deste processo.

2 MODELOS DE MATURIDADE: GÊNESE DA PROPOSTA

Os modelos de maturidade são uma prática evolutiva na gestão da qualidade, inicialmente concebidos para a indústria de software. Phil Crosby introduziu a ideia em seu livro *Quality is Free* (1979). A aplicação moderna desses modelos começou com a criação do *Process Maturity Framework* no *Software Engineering Institute* no final da década de 1980, liderado por Watts Humphrey, baseado em ideias desenvolvidas por Crosby e seus colegas da IBM (Zapata, 2023).

Segundo o autor, em 2009, a ISO 9004 incorporou a autoavaliação como uma ferramenta crucial para revisar o nível de maturidade de uma instituição, abrangendo liderança, estratégia, gestão de sistemas, recursos e processos. Zapata (2023) destaca que os modelos de maturidade são regidos pelos seguintes princípios:

- **Avaliação de Processos:** Os atributos de um processo podem ser avaliados para determinar sua capacidade de contribuir para os objetivos da instituição.
- **Sustentabilidade dos Processos:** Processos podem não perdurar se a instituição não estiver madura o suficiente para sustentá-los.
- **Melhoria Contínua:** A melhoria de processos deve ser abordada como um programa de mudança organizacional, orquestrando melhorias para alcançar estados de capacidade organizacional sucessivamente mais previsíveis.
- **Níveis de Maturidade:** Cada etapa ou nível de maturidade estabelece padrões necessários sobre os quais melhorias futuras podem ser construídas.

A adaptabilidade no uso de modelos de maturidade permite que uma instituição os utilize para definir um ponto de partida, diagnosticar programas, processos ou serviços, ou definir planos específicos de melhoria. Idealmente, os modelos de maturidade devem ser aplicados de forma integral e contínua para abordar todo o ciclo evolutivo dos processos ou serviços.

As formas de avaliação de um modelo de maturidade organizam-se quatro possibilidades. A **avaliação inicial** é uma análise leve e econômica que dura alguns dias. Nesta etapa, as evidências não são profundamente revisadas e as entrevistas são limitadas, com foco na coleta de dados quantitativos. Já a **avaliação de progresso** envolve uma investigação detalhada de todas as áreas e práticas do processo no nível de maturidade, requerendo tempo para coletar e comparar dados quantitativos com resultados de entrevistas e revisão de artefatos. Esse tipo de avaliação visa estabelecer o progresso em direção ao próximo nível de maturidade ou antecipar os resultados de uma avaliação futura. (Zapata, 2023)

A **avaliação do provedor** é similar à de progresso, mas a equipe de avaliação não inclui membros da instituição avaliada. Dados quantitativos são coletados para verificar reclamações feitas nas propostas e estabelecer níveis contratuais de

desempenho ou melhoria. As descobertas desta avaliação podem ser usadas para desenvolver compromissos contratuais de melhorias, que são verificáveis durante o período de cumprimento do contrato através de avaliações de progresso, fornecedor ou confirmação. Por fim, a **avaliação confirmatória** é uma investigação minuciosa de todas as áreas e práticas do processo no nível de maturidade. Este tipo de avaliação envolve a coleta de cinco tipos de evidências para garantir que a equipe possa avaliar a amplitude da conformidade. Dados quantitativos são coletados para investigar o desempenho e os resultados das práticas. A força-tarefa avalia se cada prática foi implementada conforme sua intenção e se os objetivos de cada área de processo foram satisfeitos. (Zapata, 2023)

A definição de níveis de maturidade estabelece uma linha de base a partir da qual as entidades podem avançar gradualmente em direção a objetivos específicos de melhoria. Quanto maiores os níveis de maturidade dos sistemas de medição, melhor é o controle dos resultados, permitindo a correção de ações ou a revisão de metas de desempenho. Esses níveis são compostos por critérios que, coletivamente, permitem que a capacidade seja alcançada em cada nível. Cada área de atuação é projetada para atingir objetivos específicos de nível, indicando o que precisa ser feito, mas não necessariamente como deve ser feito.

Essas avaliações são adaptáveis e podem ser usadas para definir um ponto de partida para melhorar processos ou serviços, realizar diagnósticos de programas e processos, ou mesmo para elaborar planos específicos de melhoria. Idealmente, os modelos de maturidade devem ser aplicados de forma integral e contínua, abordando todo o ciclo evolutivo dos processos ou serviços ao longo do tempo. Isso permite que as organizações implementem melhorias de forma estruturada e sustentável, atingindo níveis ótimos de desempenho e eficiência.

2.1 Modelos de Maturidade em Preservação Digital: Ponderações Iniciais

Os modelos de maturidade em preservação digital fornecem ferramentas para que as instituições avaliem suas práticas de preservação, identifiquem lacunas e desenvolvam estratégias para aprimorar suas capacidades de preservação digital. Alguns exemplos de modelo de maturidade em Preservação Digital:

- **National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation:** O NDSA criou um modelo que define quatro níveis de maturidade em preservação digital, que cobrem cinco áreas: armazenamento e geografia, integridade de dados, controle de segurança, metadados e conteúdo.
- **Digital Repository Audit Method Based on Risk Assessment (DRAMBORA):** Desenvolvido pelo *Digital Curation Centre* (DCC) e pelo *Digital Preservation Europe* (DPE), o DRAMBORA fornece uma estrutura para avaliar os riscos e a maturidade dos repositórios digitais através de uma metodologia de autoavaliação.
- **Digital Preservation Capability Maturity Model (DPCMM):** Este modelo, desenvolvido pelo *Digital Preservation Capability Maturity Model*, avalia a maturidade de uma organização em 15 áreas de capacidade de preservação digital, ajudando a identificar pontos fortes e áreas que precisam de melhoria.
- **Rapid Assessment Model (RAM DPC):** um modelo desenvolvido pelo *Digital Preservation Coalition* (DPC) para ajudar organizações a avaliar e melhorar suas capacidades de preservação digital. O modelo proporciona uma estrutura para identificar riscos e implementar estratégias de mitigação, garantindo a preservação a longo prazo de registros digitais.

Cada modelo possui uma abordagem distinta, mas todos compartilham o objetivo comum de garantir a longevidade e integridade dos objetos digitais. A escolha do modelo adequado ou combinação deles depende dos objetivos específicos de preservação digital e das características da organização. Em termos de estrutura temos elencado no quadro 1:

Quadro 1- Modelos de Maturidade em Preservação Digital

VANTAGENS	DESVANTAGENS
<i>DSA Levels of Digital Preservation</i>	
Estrutura Simples: Fácil de entender e aplicar. Flexível: Adaptável a diferentes tipos de instituições. Incremental: Permite melhorias graduais em cinco áreas essenciais.	Simplificação Excessiva: Pode não abordar todas as complexidades de grandes instituições. Generalidade: Pode não ser específico o suficiente para casos particulares.
DRAMBORA	
Avaliação de Risco: Enfatiza a gestão de riscos. Autoavaliação: Facilita a compreensão interna das capacidades e desafios. Metodologia Detalhada: Aborda aspectos organizacionais e tecnológicos.	Foco em Autoavaliação: Pode resultar em avaliações tendenciosas ou subjetivas. Complexidade: Pode ser difícil de implementar sem experiência em gestão de riscos.
DPCMM	
Modelo Holístico: Avalia 15 áreas de capacidade de preservação. Identificação de Lacunas: Ajuda a identificar e priorizar áreas de melhoria. Foco em Melhoria Contínua: Promove o desenvolvimento contínuo das capacidades de preservação digital.	Abrangência: A avaliação de 15 áreas pode ser trabalhosa e consumir recursos significativos. Recurso Intensivo: Requer comprometimento contínuo para melhorias, o que pode ser um desafio para instituições com recursos limitados.
RAM- DPC	
Estrutura Detalhada: O RAM DPC fornece uma estrutura detalhada para avaliação e mitigação de riscos. Flexibilidade: Adaptável a diferentes tipos de organizações e contextos. Foco na Mitigação de Riscos: Ajuda as organizações a identificar e tratar potenciais ameaças à preservação digital.	Complexidade: Pode ser difícil de implementar sem conhecimento especializado em gestão de riscos. Recurso Intensivo: Requer comprometimento significativo de tempo e recursos para uma avaliação abrangente.

Fonte: elaboração própria, (2024).

O *Digital Preservation Coalition* (DPC) é uma comunidade sem fins lucrativos dedicada a estudos em preservação digital sediada no Reino Unido e tem como membros especialistas de todo o mundo. O DPC desenvolve modelos, materiais, guias de melhores práticas em preservação digital e os disponibiliza de maneira gratuita. Dentre os modelos desenvolvidos, está o *Rapid Assessment Model - Digital Preservation Coalition* (modelo rápido de avaliação - RAM-DPC), um modelo de avaliação que tem por objetivo mensurar os níveis de maturidade e desenvolvimento de instituições em torno da preservação digital.

O RAM-DPC, em nossa análise, em detrimento aos demais modelos, possui abordagem rápida e flexível e se destaca pela sua ênfase na mitigação de riscos, fornecendo uma estrutura detalhada para identificar e tratar ameaças potenciais à preservação digital. Comparado ao DRAMBORA, que também enfoca a gestão de riscos, o RAM DPC é mais direto e acessível, facilitando sua aplicação por equipes com diferentes níveis de experiência em gestão de riscos, razões pelas quais o modelo foi apropriado no Modelo Hipátia e aplicado aos seus parceiros de pesquisa.

O modelo Hipátia optou pela aplicação RAM-DPC que se demonstra, para fins de mapeamento das pesquisas aplicadas em preservação de documentos de arquivo, mais adequado e aderente à mensuração de andamento dos projetos. Nota-se que o RAM-DCP, por meio de seus indicadores, é condizente com o modelo OAIS e a cadeia de custódia bem como a legislação arquivística vigente.

3 RAM DPC: DINÂMICA DE FUNCIONAMENTO E APLICAÇÃO NO TCU

O RAM-DPC possui 11 seções temáticas (indicadores) que são organizadas em dois blocos de análise: capacidade organizacional e capacidade de serviço, como demonstrado abaixo:

Quadro 2 - Modelo de avaliação RAM-DPC

CAPACIDADES ORGANIZACIONAIS²	
A. Viabilidade Organizacional: Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.	
B. Políticas e Estratégia: Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.	
C. Base Legal: Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.	
D. Capacidade de TI: Recursos de tecnologia da informação para apoiar atividades de preservação digital.	
E. Aperfeiçoamento Contínuo: Processos para avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.	
F. Comunidade: Engajamento e contribuição para a comunidade mais ampla de preservação digital.	
CAPACIDADES DE SERVIÇO	
G. Aquisição, Transferência e Incorporação: Processos para adquirir ou transferir conteúdo e incorporá-lo a um arquivo digital.	
H. Preservação de Bitstream: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.	
I. Preservação de Conteúdo: Processos para preservar o sentido ou funcionalidade do conteúdo digital e garantir seu acesso e usabilidade contínuos no futuro.	
J. Gerenciamento de Metadados: Processos para criar e manter metadados suficiente para dar suporte a preservação, gerenciamento e uso de conteúdo digital preservado.	
K. Descoberta e Acesso: Processos para permitir a descoberta de conteúdo digital e prover acesso aos usuários.	

Fonte: Elaborado pelos autores, (2024).

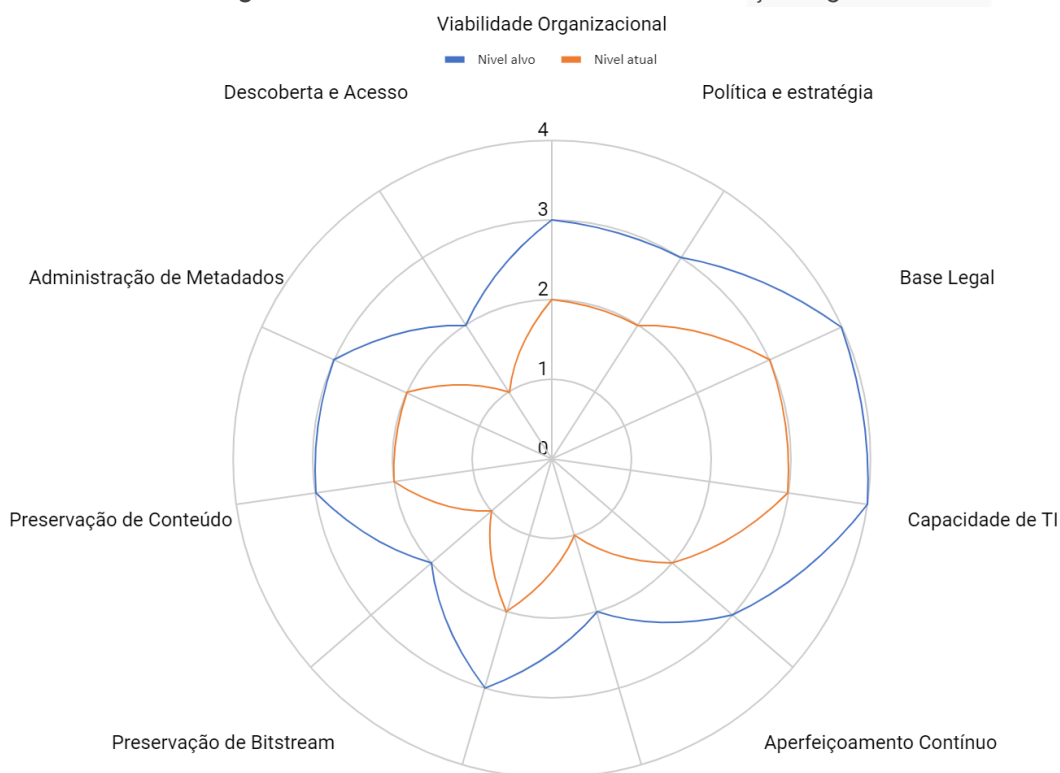
Para cada seção é possível atribuir avaliações de 0 a 4 (Consciência Mínima, Consciência, Básico, Gerência, Otimizado). Cada nível de cada seção apresenta uma descrição detalhada dos requisitos que são necessários para superar aquele nível e

² O Hipátia utiliza a tradução oficial do modelo em português (PT-PR).

somente se avança ao próximo nível se todos os requisitos do nível avaliado estejam cumpridos.

O RAM-DPC foi aplicado no TCU, tendo por escopo de observação e prospecção o período de março de 2024 a dezembro de 2024. A aplicação e a observação deram-se por meio de preenchimento de planilha contendo as seções e os níveis. A observação foi realizada pelos pesquisadores Hipátia a partir das reuniões realizadas, estudos de ambientes feitos, instalações realizadas e documentos disponibilizados pelo Tribunal. Após o preenchimento inicial, a planilha foi disponibilizada à equipe do TCU que a revisou resultando na seguinte matriz de maturidade.

Figura 1 – Matriz de Maturidade em Preservação Digital



Fonte: Elaborado pelos autores, (2024).

Para determinar o nível médio de maturidade da instituição conforme o gráfico fornecido pelo RAM DPC, podemos calcular a média dos níveis atuais para cada área avaliada.

Os níveis atuais são:

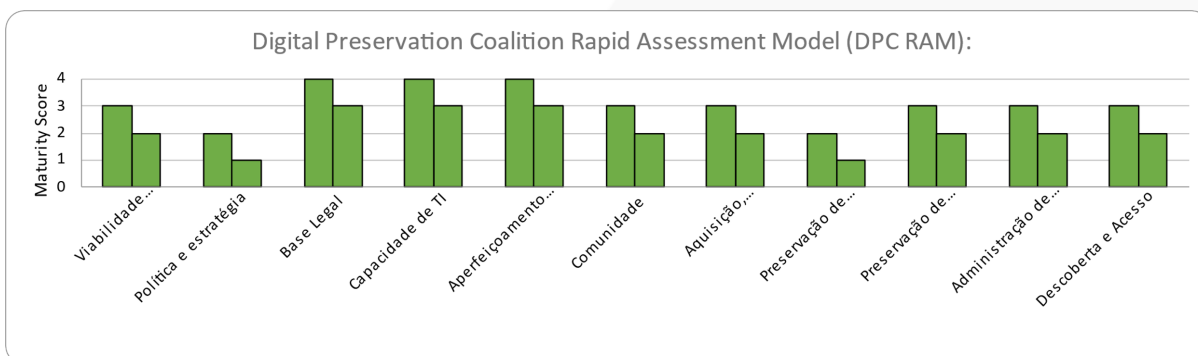
- Viabilidade Organizacional: 2
- Política e Estratégia: 2
- Base Legal: 1.5
- Capacidade de TI: 2
- Aperfeiçoamento Contínuo: 1
- Comunidade: 1
- Aquisição, Transferência e Incorporação: 1
- Preservação de Bitstream: 0
- Preservação de Conteúdo: 1
- Administração de Metadados: 2
- Descoberta e Acesso: 2

Somando esses valores: $2 + 2 + 1 + 2 + 1 + 1 + 1 + 0 + 1 + 2 + 2 = 15$

Dividindo pelo número de áreas (11): $15 / 11 \approx 1.36$

Portanto, o nível médio de maturidade da instituição é aproximadamente 1.36. Os resultados da análise de maturidade em preservação digital usando o RAM DPC mostram uma discrepância significativa entre os níveis atuais e os níveis alvo em várias áreas críticas.

Gráfico 1 – Níveis de maturidade



Fonte: Elaborado pelos autores (2024).

A Viabilidade Organizacional, Política e Estratégia, Capacidade de TI e Administração de Metadados estão moderadamente desenvolvidas, mas precisam de melhorias adicionais para atingir os níveis desejados. Áreas como Aperfeiçoamento Contínuo, Comunidade, e Aquisição, Transferência e Incorporação estão subdesenvolvidas e requerem atenção imediata. A Preservação de Bitstream é especialmente crítica, necessitando de esforços urgentes para melhorar. Melhorias nessas áreas ajudarão a instituição a alcançar os níveis desejados de maturidade em preservação digital.

Como pontos forte, identifica-se:

1. **Equipe engajada no projeto:** há equipes dedicadas ao projeto com conhecimento intermediário em preservação digital.
2. **Comprometimento da Alta Gerência:** A alta gerência do TCU está comprometida com a preservação digital, assumindo responsabilidade e alocando recursos financeiros .
3. **Base Legal Sólida:** A gestão dos direitos e responsabilidades legais é proativa.
4. **Capacidade de TI Gerenciada:** A infraestrutura de TI está bem desenvolvida, com suporte aprimorado disponível para atividades de preservação digital.

A partir da observação, foram atribuídos os níveis tais como descritos abaixo. Destaca-se que a observação foi realizada de maneira a testar a aplicação do RAM-DPC em um caso prático.

Quadro 3 - Modelo RAM-DPC TCU (março 2024-dez 2024)

CAPACIDADE ORGANIZACIONAL		
	Nível Atual (março/2024)	Nível Almejado (dez/ 2024)
A. Viabilidade Organizacional: Governança, estrutura organizacional, pessoal e recursos de atividades de preservação digital.	2 - Básico	3 - Gerenciado
B. Políticas e Estratégia: Políticas, estratégias e procedimentos que governam a operação e o gerenciamento do arquivo digital.	1 - Conscientização	2 - Básico
C. Base Legal: Gestão dos direitos e responsabilidades legais, cumprimento da regulamentação pertinente e adesão aos códigos éticos relativos à aquisição, preservação e disponibilização de acesso a conteúdos digitais.	3 - Gerenciado	4 - Otimizado
D. Capacidade de TI: Recursos de tecnologia da informação para apoiar atividades de preservação digital.	3 - Gerenciado	4 - Otimizado
E. Aperfeiçoamento Contínuo: Processos para avaliação das atuais capacidades de preservação digital, definição de metas e monitoramento do progresso.	3 - Gerenciado	4 - Otimizado
F. Comunidade: Engajamento e contribuição para a comunidade mais ampla de preservação digital.	2 - Básico	3 - Gerenciado
CAPACIDADES DE SERVIÇO		
	Nível Atual	Nível Almejado
G. Aquisição, Transferência e Incorporação: Processos para adquirir ou transferir conteúdo e incorporá-lo a um arquivo digital.	2 - Básico	3 - Gerenciado
H. Preservação de Bitstream: Processos para garantir o armazenamento e integridade do conteúdo digital a ser preservado.	1 - Conscientização	2 - Básico
I. Preservação de Conteúdo: Processos para preservar o sentido ou funcionalidade do conteúdo digital e garantir seu acesso e usabilidade contínuos no futuro.	2 - Básico	3 - Gerenciado
J. Gerenciamento de Metadados: Processos para criar e manter metadados suficiente para dar suporte a preservação, gerenciamento e uso de conteúdo digital preservado.	2 - Básico	3 - Gerenciado

K. Descoberta e Acesso: Processos para permitir a descoberta de conteúdo digital e prover acesso aos usuários.	2 - Básico	3 - Gerenciado
---	------------	----------------

Fonte: Elaborado pelos autores (2024).

O preenchimento da planilha, possibilitou a produção da seguinte visão, que pode ser interpretada como se segue. Primeiramente, cabe destacar que para fins de formação do RDC-Arq, relacionamos quando da análise RAM DPC os padrões *Trustworthy Repositories Audit & Certification* (TRAC) e ISO 16363 devido ao seu foco compartilhado na avaliação e aprimoramento da preservação digital e por ambas as normas serem as normas referência das resoluções do Conselho Nacional de Arquivos. Ademais, futuramente, quando da possibilidade de certificação do RDC-arq serão elas a serem analisadas.

3.1 Capacidades Organizacionais

A viabilidade organizacional do TCU está em um nível básico. O TCU possui equipe e recursos financeiros destinados ao projeto de preservação digital, mas ainda há necessidade de fortalecer a governança e a alocação de pessoal e recursos. Embora haja comprometimento da alta gerência, ainda há questões relacionadas ao orçamento, relatórios gerenciais e habilidades da equipe que precisam ser resolvidas antes de avançar para um nível mais alto de maturidade.

As políticas e estratégias estão no nível de conscientização, indicando que o TCU reconhece a necessidade de políticas formais. A Portaria 85/2023 aborda a preservação de documentos eletrônicos, mas carece de detalhes sobre estratégias de longo prazo, formatos e responsabilidades específicas.

A base legal do TCU está no nível gerenciado, indicando um cumprimento robusto das regulamentações e responsabilidades legais. O TCU tem bem definidas as regras de segurança da informação e um cumprimento robusto das regulamentações legais. No entanto, é necessário revisar as classificações e

metodologias para adequar-se completamente às necessidades de preservação digital.

A capacidade de TI do TCU está em um nível gerenciado, com recursos tecnológicos adequados para suportar as atividades de preservação digital. Investimentos contínuos em atualização tecnológica e infraestrutura são importantes para manter e melhorar essa capacidade. A infraestrutura de TI do TCU é suficiente para suportar as atividades de preservação digital, com processos gerenciados para avaliação comparativa e estabelecimento de metas. A organização realiza estudos internos e colabora com o IBICT para identificar e implementar as melhores práticas.

O aperfeiçoamento contínuo está no nível gerenciado, indicando que o TCU possui processos estabelecidos para avaliar e melhorar continuamente suas capacidades de preservação digital. Está em constante análise de experiências de outras instituições e tem projetos em parceria com o IBICT para avançar suas práticas de preservação digital.

O engajamento comunitário do TCU está no nível básico, mostrando que há participação inicial, mas é necessário fortalecer a colaboração e a contribuição para a comunidade de preservação digital, por exemplo com a participação em eventos e divulgação interna e externa do projeto. A parceria com o IBICT e a participação em eventos como palestras e congressos do poder Judiciário refletem o compromisso da organização com o compartilhamento de conhecimento e aprendizado contínuo.

O TCU demonstra progresso significativo em áreas críticas como base legal, capacidade de TI e aperfeiçoamento contínuo. No entanto, áreas como políticas e estratégia e engajamento comunitário ainda necessitam de desenvolvimento. Fortalecer a governança, implementar políticas claras e aumentar o engajamento comunitário são pontos de atenção para avançar a maturidade geral das capacidades de preservação digital do TCU.

3.2 Capacidades de Serviço

A gestão de documentos no TCU acontece, embora não se aplique ainda aos documentos natos digitais. Existe um manual de gestão de documentos e fluxos de transferência e recolhimento. No entanto, não se identificam estudos de metadados e aderência do e-TCU ao e-ARQ Brasil. Há uma unidade de arquivo no organograma, mas não há a implementação de uma cadeia de custódia.

O TCU possui processos iniciais para adquirir e incorporar conteúdos digitais ISO 16363 especifica a necessidade de políticas para a aceitação e verificação de novos conteúdos, assegurando sua autenticidade e integridade. Ainda destaca-se a importância de ter políticas claras para a aquisição e incorporação, garantindo que os conteúdos sejam adequadamente validados e protegidos desde o momento de sua admissão no repositório.

A preservação de bitstream refere-se à prática de garantir que os bits que compõem um arquivo digital permaneçam inalterados e acessíveis ao longo do tempo. Isso envolve a implementação de processos e tecnologias que asseguram a integridade e autenticidade dos dados armazenados, protegendo-os contra corrupção, perda e obsolescência tecnológica. A análise dos componentes da Preservação de Bitstream são:

1. Verificações de Integridade: Utilização de algoritmos de hash (soma de verificação) para monitorar a integridade dos arquivos e detectar qualquer alteração ou corrupção de dados.
2. Backup e Redundância: Implementação de estratégias de backup e armazenamento redundante para garantir que cópias dos dados estejam disponíveis em diferentes locais e sistemas.
3. Monitoramento Contínuo: Monitoramento regular dos arquivos para identificar e corrigir problemas de integridade e acessibilidade.

4. Migração de Dados: Planejamento e execução de migrações de dados para novos formatos ou plataformas conforme necessários para evitar a obsolescência tecnológica.

O nível atual indica uma conscientização inicial, mas a implementação de práticas básicas para proteger os dados contra corrupção e perda é necessária. A ISO 16363 enfatiza a necessidade de práticas robustas para a integridade dos bitstreams, incluindo verificações de integridade regulares, mecanismos de detecção de erros e procedimentos de recuperação.

No nível 1, também conhecido como "Conscientização", há entendimento da necessidade de preservação dos ativos digitais, mas não há implementação de soluções de processamento para preservação. O TCU conhece os formatos de produção, mas ainda não estuda ou processa componentes de integridade e credibilidade dos documentos.

A preservação de conteúdo está em estágio básico. Embora haja conhecimento dos formatos de produção, não há processamento de componentes de integridade em formato e padrões que adicionam segurança e completude. Os metadados e a documentação adquirida com o conteúdo ainda não são mantidos e preservados de forma eficaz. Metadados básicos de preservação não são capturados no nível do item.

O TCU implementou um mecanismo básico de descoberta e acesso, incluindo uma ferramenta básica de busca para alguns conteúdos digitais. Os usuários podem visualizar ou acessar conteúdo e metadados digitais, e o acesso é registrado. No entanto, não há plataformas de acesso integradas aos ambientes de preservação e acesso.

4 CONSIDERAÇÕES FINAIS

O nível médio de maturidade de 1.5 indica que o TCU está no início do desenvolvimento de suas capacidades de preservação digital. Para alcançar os níveis almejados, será necessário investir em processos estruturados e integrados, alinhados aos padrões internacionais como o TRAC e a ISO 16363. Focar em áreas críticas, como a preservação de bitstream e o gerenciamento de metadados, serão os pontos de atenção para fortalecer a infraestrutura de preservação digital e garantir a longevidade e integridade dos documentos.

Para os níveis almejados até dezembro de 2024, recomenda-se, por capacidade:

- A. Garantir que as habilidades da equipe sejam adequadas, desenvolver relatórios gerenciais e alocar um orçamento principal dedicado à preservação digital.
- B. Atualizar a estrutura da política atual e detalhar procedimentos, incluindo a definição do escopo da coleção e a gestão proativa dos direitos e responsabilidades legais.
- C. O TCU gerencia proativamente questões e riscos legais, contribuindo para processos legais e judiciais que criam regulamentações. Manter esta abordagem e envolver-se ativamente em questões legais continuará a mitigar riscos e assegurar conformidade.
- D. A infraestrutura de TI deve evoluir constantemente, com suporte aprimorado disponível para arquivos digitais e um entendimento claro das necessidades de preservação digital. A aquisição/ instalação de novos sistemas de TI deve considerar requisitos de preservação digital, e novas ferramentas e sistemas devem ser identificados e testados proativamente.
- E. O TCU deve manter a certificação/revisão externa e adotar recomendações para melhorias. Revisões periódicas de metas e roteiros são essenciais para garantir a melhoria contínua dos processos.

- F. A organização deve desempenhar um papel ativo na comunidade de preservação digital, compartilhando sucessos e lições aprendidas.
- G. O TCU deve gerenciar relacionamentos com os produtores dos documentos, automatizar partes do processo de aquisição e incorporação, e verificar a transferência de conteúdo por meio de verificações de integridade.
- H. A equipe deve saber onde o conteúdo está armazenado e garantir que a replicação e backup sejam realizados regularmente. Soma-se de verificação para todo o conteúdo que deve ser gerado.
- I. O conteúdo deve ser caracterizado e avaliado por questões de preservação e qualidade. Atividades de observação tecnológica devem identificar conteúdos "em risco", e ações de preservação devem ser realizadas para garantir a qualidade e integridade dos documentos.
- J. Os metadados e a documentação adquirida com o conteúdo devem ser mantidos e preservados. Padrões de metadados apropriados e vocabulários controlados devem estar em vigor para garantir a consistência.
- K. O TCU deve implementar uma busca básica para todo o conteúdo digital e permitir pesquisa de texto completo para alguns conteúdos. As informações de direitos devem ser exibidas e o acesso gerenciado pelo sistema. Relatórios sobre o acesso do usuário e sistemas de acesso atualizados com feedback da comunidade são necessários.

BIBLIOGRAFIA CONSULTADA

BRAGA, T. E. N. O modelo Hipátia: a proposta do IBICT para a preservação digital arquivística. In: **Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis**. Brasília: Ibict, 2022. Disponível em: <https://widat2022.ufes.br/wp-content/uploads/2022/11/st-4/st4-8-Modelo%20de%20>

0Preserva%C3%A7%C3%A3o%20Hipatia%20metodologia%20de%20estudo%20de%20metadados%20para%20extra%C3%A7%C3%A3o.pdf. Acesso em: 22 maio 2024.

DURANTI, L. **Diplomática: usos nuevos para una antigua ciencia**. Carmona (Sevilla): S&C ediciones, 1996.

INTERNATIONAL RESEARCH ON PERMANENT AUTHENTIC RECORDS IN ELETRONIC SYSTEMS (InterPARES 2). **Ontology C: Trustworthiness of a record**. 2007. Disponível em: http://www.interpares.org/ip2/display_file.cfm?doc=ip2_ontology.pdf. Acesso em: 15 maio 2024.

LIMA, R. de S.; OLIVEIRA, A. F.; D'AVILA, R. T.; CHAVES, E. P. da S. S. O modelo de referência OAIS e a preservação digital distribuída. **Ciência Da Informação**, v. 41, n. 1, 2014. Disponível em: <https://revista.ibict.br/ciinf/article/view/1352>. Acesso em: 16 jun. 2024.

MELO, Í. F.; LUZ, C. dos S. A aderência de sistemas informatizados de gestão arquivística ao e-ARQ Brasil: verificação de requisitos mínimos e obrigatórios. **Acervo**, [S. l.], v. 35, n. 1, p. 1–15, 2021. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1778>. Acesso em: 16 jun. 2024.

MELO, Í. F.; CANELHAS, T.; BRAGA, T. E. N. **Modelo de Preservação Hipátia: metodologia de estudo de metadados para extração**. In: V Workshop de Informação, Dados e Tecnologia, UFES, 2022. Disponível em: <https://widat2022.ufes.br/wp-content/uploads/2022/11/st-4/st4-8-Modelo%20de%20Preserva%C3%A7%C3%A3o%20Hipatia%20metodologia%20de%20estudo%20de%20metadados%20para%20extra%C3%A7%C3%A3o.pdf>. Acesso em: 14 jun. 2024.

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

ZAPATA CÁRDENAS, C. A. Los modelos de madurez en gestión documental y archivos: una herramienta eficaz de evaluación, planeación y mejora continua. **Métodos de Información**, v. 14, n. 26, p. 48-79, 2023. DOI: <https://dx.doi.org/10.5557/IIMEI14-N26-048079>.

APÊNDICE D - Modelo para Dicionário de Dados e-TCU

1. Documento Processo

<u>1º Envio</u>		EXTRAÇÃO OBRIGATÓRIA NO CASO eTCU Opções: Sim; Não; Não se aplica; Em revisão	JUSTIFICATIVA DA OBRIGATORIEDADE Opções: obrigatório eARQ (O); obrigatório, se aplicável eARQ (OA); facultativo eARQ (F); não se aplica eARQ (NA); metadado para acesso; Metadado Inserido pelo TCU (TCU) Metadados sugerido pelo IBICT (IBICT)	ORIGEM DO DADO Opções: webservice (API); não se aplica (NA)	DC	CAMINHO Opções: indicação do caminho; não se aplica (NA); não conhecido (NC)	EXEMPLO EXTRAÇÃO ATUAL Opções: indicação do valor; não há exemplo (NE)	Observações
1.a Documento	1.1 Identificador do documento	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		tcu.documento_gestao.cod_documento	16932223	Chave primária numérica dos documentos no GED do e-TCU (documento gestão)
	1.2. Número do documento	Sim	Processo/dossiê: NA Volume: NA Documento: OA	API	dc.identifier	tcu.documento_tramitavel.cod + calculadvd_documento(dt.cod)	58.085.668-4	Todo documento nato digital ou incluído como volume digitalizado no E-TCU recebe um numero de

								documento no formato 00.000.000-0. No E-TCU adm documentos tramitam soltos. No CE devem ser incluídos em processos.
	1.9. Status	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		tcu.arquivo_eletronico _catalogado.cod_ofici alizacao & tcu.vinculo_document o_lote.ind_tipo_vincul o	Alguns valores sugeridos: minuta, original, cópia	É possível calcular por meio de outros metadados
	1.10. Identificador de versão	Sim	Processo/dossiê: NA Volume: NA Documento: OA	API		tabela versao_arquivo_eletro nico	Data: "12/09/2017 15:02:14" Visibilidade: "TCU" Usuário de criação: "Fulano de Tal"	O campo "histórico" e "versões anteriores" do documento do e-TCU traz a data e hora de criação do documento, alteração de versões, responsável pelas alterações.

	1.11. Título	Sim	Processo/dossiê: F Volume: NA Documento: O	API	dc.title	NC	NE	Verificar o caminho e extração com a TI/TCU.
	1.12. Descrição	Não se aplica	Processo/dossiê: F Volume: NA Documento: F	API		NA	NE	
	1.13. Assunto	Sim	Processo/dossiê: F Volume: NA Documento: F	API	dc.subject	tcu.documento_tramitavel.texto_complemento_assunto	Despacho de encerramento - Processo 046.664/2012-2	Assunto do documento
	1.14. Autor	Sim	Processo/dossiê: O Volume: NA Documento: O	API		NC	Campo de texto livre	Nos documentos se refere ao campo "Autor" (mas, pode não ter sido informado)/ Nos processos, talvez a UT que autuou o processo seja melhor do que o campo "Interessados".
	1.15. Destinatário	Não	Processo/dossiê: F Volume: NA Documento: O	API	dc.contributor	NC	NE	
	1.16. Originador	Não se aplica	Processo/dossiê: NA Volume: NA Documento: OA	API		NA	NE	Não se aplica, pois o nome do originador é o mesmo do redator.

	1.17. Redator	Sim	Processo/dossiê: NA Volume: NA Documento: O	API	dc.publisher	tem caminho?	Unidade técnica responsável:SEGE CEX	Usuários que assinaram o documento
	1.18. Interessado (Parte Interessada)	Sim	Processo/dossiê: O Volume: NA Documento: NA	API		tem caminho?	Interessados: Senado Federal	Interessado(s) do Processo
	1.18. Interessado (Representante Legal)	Não se aplica	TCU					
	1.18. Interessado (Órgão Interessado)	Não se aplica	TCU					
	1.19. Identificador do componente digital	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API		NC	NE	Equivale ao item 5.1 - Identificador do componente digital?
	1.20. Gênero	Não	Processo/dossiê: NA Volume: NA Documento: F	API		NA	NE	Todos os documentos no e-TCU são textuais.
	1.21. Espécie	Não	Processo/dossiê: NA Volume: NA Documento: F	API		NA	NE	Campo Tipo, em dados gerais do documento. Mistura os conceitos de tipo e espécie
	1.22. Tipo	Sim	Processo/dossiê: NA Volume: NA Documento: F	API		tcu.tipo_documento	Despacho de expediente	Campo Tipo, em dados gerais do documento. Mistura os conceitos de tipo e espécie

	1.23. Idioma	Não se aplica	Processo/dossiê: NA Volume: NA Documento: F	API	dc.languag e	NA	NE	
	1.25. Numeração sequencial dos documentos	Sim	Processo/dossiê: NA Volume: NA Documento: OA	API		tcu.vinculo_document o_lote.num_peca_eletr onica	80	Quando juntado a processo, indica a numeração sequencial do número de peças. Quando avulso, indica a ordem de documentos anexados.
	1.26. Indicação de anexos	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		tcu.HIST_VINCULO_D OCUMENTO	tem exemplo?	Para documentos avulsos é possível ter anexos
	1.27. Indicação de anotação	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API		NA	NE	Comentários nos processos e documentos (Post it) Verificar se esses dados serão extraídos para preservação.
	1.28. Unidade responsável pela execução da ação	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		NC	NE	Campo histórico do documento, registro de todas as unidades/usuários

								responsável pelo evento/histórico.
	1.29. Relação com outros documentos	Sim	Processo/dossiê: OA Volume: NA Documento: OA	API		tcu.vinculo_documento_lote	000.000/0000-0	Em nível de processo, indica processos apensados Em nível de documento, indica processo relacionado
	1.30. Níveis de acesso	Sim	Processo/dossiê: O Volume: NA Documento: O	API		documento_gestao.co d_tipo_confidencialida de e arquivo_eletronico_ca talogado.ind_tipo_esc opo	Público, Restrito ou Sigiloso	Em nível de processo e documento, campo: Confidencialidade
	1.31. Previsão de desclassificação	Sim	Processo/dossiê: O Volume: NA Documento: O	API		Está ativo? Tem caminho?	NE	Evento de término da classificação (confidencialidade sigiloso)
	1.32. Data de produção	Sim	Processo/dossiê: O Volume: NA Documento: O	API	dc.date	versao_arquivo_eletronico.dthora_criacao		Data/Hora da criação da última versão do documento no GED
	1.33. Local de produção	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API	dc.coverage	NA	NE	Ver Ibict obrigatoriedade para documento.

								Não temos esse metadado no e-TCU.
	Metadados de assinatura/certificação (arquivo em P7s)	Não se aplica	IBICT	API		assinatura_versao_arq HIST_OFICIALIZA_AQ R_ELET_CAT		
Processo (NOVO)	1.3. Número do protocolo	Sim	Processo/dossiê: NA Volume: NA Documento: OA	API		tcu.documento_gestao.cod_registro_entrada + dv	57.977.215-5	Número de Registro de Entrada RE, documento nato digital não o possui
	1.4. Identificador do processo/dossiê	Sim	Processo/dossiê: O Volume: NA Documento: NA	API	dc.identifier	tcu.vinculo_documento_lote.cod_lote where ind_tipo_vinculo = 'O'	4344297	Chave primária numérica de identificação do processo no e-TCU
	1.5. Número de protocolo do processo	Sim	Processo/dossiê: O Volume: NA Documento: NA	API	dc.title	(num, ano, num_dv) tcu.processo_gestao	005.647/2009-6	nnn.nnn/aaaa-d: e-TCU, campo: Processo
	1.6. Identificador do volume	Não se aplica	Processo/dossiê: NA Volume: O Documento: NA	API		NA	NE	
	1.7. Número do volume	Não se aplica	Processo/dossiê: NA Volume: O Documento: NA	API		NA	NE	

	1.8. Tipo de meio	Sim	Processo/dossiê: F Volume: NA Documento: O	API		tcu.registro_entrada_p rotocolo.cod_tipo_orig em	"Eletrônico-conver tido" ou "Eletrônico"	É possível calcular por meio de outros metadados; inferir pelo meio de entrada (origem do RE), se protocolo físico ou serviço digital (Conecta, Protocolo Eletrônico, etc); se documento interno nasce eletrônico.
	1.11. Título	Sim	Processo/dossiê: F Volume: NA Documento: O	API		NC	NE	Verificar o caminho e extração com a TI/TCU.
	1.12. Descrição	Não se aplica	Processo/dossiê: F Volume: NA Documento: F	API		NA	NE	
	1.13. Assunto	Sim	Processo/dossiê: F Volume: NA Documento: F	API	dc.subject	tcu.documento_tramit avel.texto_compleme nto_assunto	Despacho de encerramento - Processo 046.664/2012-2	Assunto do documento
	1.14. Autor	Sim	Processo/dossiê: O Volume: NA Documento: O	API		NC	Campo de texto livre	Nos documentos se refere ao campo "Autor" (mas, pode não ter sido informado)/ Nos

								processos, talvez a UT que autou o processo seja melhor do que o campo "Interessados".
	1.16. Originador	Não se aplica	Processo/dossiê: NA Volume: NA Documento: OA	API	dc.publisher	NA	NE	Não se aplica, pois o nome do originador é o mesmo do redator.
	1.17. Redator	Sim	Processo/dossiê: NA Volume: NA Documento: O	API	dc.publisher	tem caminho?	Unidade técnica responsável: SEGE CEX	Usuários que assinaram o documento
	1.18. Interessado (Parte Interessada)	Sim	Processo/dossiê: O Volume: NA Documento: NA	API	dc.contributor	tem caminho?	Interessados: Senado Federal	Interessado(s) do Processo
	1.18. Interessado (Representante Legal)	Não se aplica	TCU	API	dc.contributor			
	1.18. Interessado (Órgão Interessado)	Não se aplica	TCU	API	dc.contributor			
	1.19. Identificador do componente digital	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API		NC	NE	Equivala ao item 5.1 - Identificador do componente digital?

	1.23. Idioma	Não se aplica	Processo/dossiê: NA Volume: NA Documento: F	API	dc.language	NA	NE	
	1.24. Quantidade de folhas	Sim	Processo/dossiê: OA Volume: OA Documento: F	API	dc.format	tcu.documento_gestao_num_paginas_digitalizadas	48	Se de origem externa o campo é preenchido
	1.26. Indicação de anexos	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		tcu.HIST_VINCULO_DOCUMENTO	tem exemplo?	Para documentos avulsos é possível ter anexos
	1.27. Indicação de anotação	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API		NA	NE	Comentários nos processos e documentos (Post it) Verificar se esses dados serão extraídos para preservação.
	1.28. Unidade responsável pela execução da ação	Sim	Processo/dossiê: NA Volume: NA Documento: O	API		NC	NE	Campo histórico do documento, registro de todas as unidades/usuários responsável pelo evento/histórico.
	1.30. Níveis de acesso	Sim	Processo/dossiê: O Volume: NA Documento: O	API		documento_gestao.co d_tipo_confidencialidade	Público, Restrito ou Sigiloso	Em nível de processo e documento, campo: Confidencialidade

						arquivo_eletronico_catalogado.ind_tipo_escopo		
	1.31. Previsão de desclassificação	Sim	Processo/dossiê: O Volume: NA Documento: O	API		Está ativo? Tem caminho?	NE	Evento de término da classificação (confidencialidade sigiloso)
	1.32. Data de produção	Sim	Processo/dossiê: O Volume: NA Documento: O	API	dc.date	versao_arquivo_eletronico.dthora_criacao		Data/Hora da criação da última versão do documento no GED
	1.33. Local de produção	Não se aplica	Processo/dossiê: NA Volume: NA Documento: O	API	dc.coverage	NA	NE	Ver Ibict obrigatoriedade para documento. Não temos esse metadado no e-TCU.
	1.34. Classe	Sim	Processo/dossiê: O Volume: NA Documento: O	API	ds.subject	tcu.classe_assunto	770.2 - Peça de Processo de Controle Externo	Em nível de processo e documento, campo: Classificação arquivística
	1.35. Destinação prevista	Sim	Processo/dossiê: O Volume: NA Documento: O	API		tem caminho?	G (=guarda) ou E(=eliminação)	
	1.36. Prazo de guarda	Sim	Processo/dossiê: O Volume: NA Documento: O	API		tem caminho?	99/99/9999	

	1.37. Localização	Sim	Processo/dossiê: OA Volume: F Documento: OA	API		ARMAZENAMENTO_P APEL_UNIDADE??	campo textual	Refere-se principalmente a parte física do processo, em qual local/unidade ele está arquivado. No eTCU temos a aba Localização quando for eletrônico-convertido. Para TC eletrônico o local de armazenamento (nuvem, servidor local, etc)
	Número de peças	Não se aplica	TCU	API				
	Relator	Não se aplica	TCU	API	dc.publisher			
	Representante do MPTCU	Não se aplica	TCU	API	dc.publisher			
	Unidade Jurisdicionada	Não se aplica	TCU	API	dc.contributor			Unidade que está sendo fiscalizada. Entidade, órgãos .

	Responsáveis	Não se aplica	TCU	API	dc.contributor			
	Legitimados (nome)	Não se aplica	TCU	API	dc.contributor			possibilidade de ter vistas ao processo; é um CPF (nome, qualificação...)
	Legitimados CPF	Não se aplica	TCU	API				
	Histórico	Não se aplica	TCU	API				buscar todas as unidades responsáveis pelo processo no período
	Situação	Não se aplica	TCU	API				aguardando providência; em instrução
	Deliberações	Não se aplica	TCU	API				
	Processos Apensados	Não se aplica	TCU	API				
	Processos apartados	Não se aplica	TCU	API				
	Processo de interesse do Congresso Nacional	Não se aplica	TCU	API				Sim/Não

	Urgência	Não se aplica	TCU	API				Sim/Não
	Repercussão na Mídia	Não se aplica	TCU	API				Sim/Não
	Selo histórico	Não se aplica	TCU	API	dc.subject			
	Metadados de assinatura/certificação	Não se aplica	IBICT	API				
2. Evento de gestão	2.1.1. Identificador do evento	Sim	O	API		tcu.evento_document o.cod	NUMBER(9,0)	o cod associado ao evento (histórico) no e-TCU
	2.1.2. Tipo de evento	Sim	O	API		tcu.evento_document o.cod_tipo_evento	Documento juntado ao processo 046.664/2012-2 por AudInovação (Walter Venson Filho).	Calculado a partir do histórico e situação do processo no e-TCU (cod_tipo_evento). Será necessário criar um campo no e-tcu para indicar o recolhimento para guarda permanente do processo e/ou possa ser eliminado.
	2.1.3. Identificador do processo/dossiê	Sim	OA	API		tcu.vinculo_document o_lote	nnn.nnn/aaaa-d: 046.664/2012-2	O código deve ser obtido no metadado 1.4 identificador do processo/dossiê.

								Número do TC formatado
2.1.4. Identificador do documento	Sim	O	API		tcu.documento_gestao_cod_documento	nnn.nnn.nnnn-d: 75.195.934-3		Obtido do metadado 1.1 Número do documento formatado
2.1.5. Identificador do lote	Sim	O	API		NA	NE		
2.1.6. Data e hora do evento	Sim	O	API		tcu.evento_documento	29/12/2023 15:55:27'		formato data padrão ISO 8601:2019, data do evento no e-tcu
2.1.7. Agente responsável pelo evento	Sim	O	API		tcu.evento_documento	WALTERVF		identificador de pessoa
2.1.8. Detalhes do evento	Sim	OA	API		TEXTO_FORMATADO + TEXTO_OBSERVACAO	Instrução iniciada por Walter Venson Filho		Texto combinado da descrição do evento.
2.2.1. Identificador do evento		O	API					Alguma coluna com chave primária?????
2.2.2. Tipo do evento	Sim	O	API			- Processo autuado em 01/08/2008 15:39:40; - Processo 031.344/2013-5		Histórico de eventos, coluna : Descrição do evento

							foi apensado a este processo; - Processo encerrado.	
	2.2.3. Identificador do processo	Sim	O	API				O Histórico é vinculado a um processo
	2.2.4. Identificador do volume	Não se aplica	OA	API				
	2.2.5. Data e hora do evento	Sim	O	API			4/7/14 8:45	Histórico de eventos, coluna : Data/Hora
	2.2.6. Agente responsável pelo evento	Sim	OA	API				Histórico de eventos, coluna : Descrição do evento
	2.2.7. Identificador do documento		OA	API				seria o mesmo do 1.1?
3. Classe	3.1. Identificador da classe	Sim	O	API		TCU.CLASSE_ASSUNTO.COD	956	
	3.2. Nome da classe	Sim	O	API		TCU.CLASSE_ASSUNTO.DESCR	Governança e gestão	eTCU, campo sem nome Sinergia, campo: Descrição da classe/assunto
	3.3. Código da classe	Sim	O	API		TCU.CLASSE_ASSUNTO.COD_CONARQ	002.2	eTCU, campo: TCU - Classe Assunto

								Sinergia, campo: Código CONARQ da classe/assunto
3.4. Subordinação da classe	Sim	O	API		TCU.CLASSE_ASSUN TO.COD_CLASSE_ASS UNTO_PAI	3		
3.5. Indicação de permissão de uso	Sim	O	API		NC	NE		
3.6. Indicação de classe ativa/inativa	Sim	O	API		TCU.CLASSE_ASSUN TO.DATA_EXCLUSAO	NE	Se nulo, ativo. Se preenchido, inativo.	
3.7. Prazo na idade corrente	Sim	O	API		TCU.HIST_ASSOC_RE GRA_FUNCAO_TEMP O.COD_REGRA_TEMP ORALIDADE + TCU.TIPO_REGRA_TE MPORALIDADE.QTD_ PRAZO_GUARDA	10 anos		
3.8. Evento de contagem na idade corrente	Sim	OA	API		TCU.HIST_ASSOC_RE GRA_FUNCAO_TEMP O.COD_REGRA_TEMP ORALIDADE + TCU.TIPO_REGRA_TE MPORALIDADE.QTD_ PRAZO_GUARDA	Encerramento de processo / arquivamento do documento. SITUACAO_DOCU MENTO: 6 = Documento arquivado. SITUACAO_PROC ESSO: 12 = Arquivado.	No Histórico há o evento: Ação - Encerramento de processo - ABERTO / ----- / ENCERRADO - Início em: 5/out/2010 - Término em: 5/out/2010	

	3.9. Prazo na idade intermediária	Sim	OA	API		TCU.HIST_ASSOC_RE GRA_FUNCAO_TEMP O.QTD_ANOS_FASE_I NTERMEDIARIA	25 anos	
	3.10. Evento de contagem na idade intermediária	Sim	OA	API		TCU.HIST_ASSOC_RE GRA_FUNCAO_TEMP O.COD_REGRA_TEMP ORALIDADE + TCU.TIPO_REGRA_TE MPORALIDADE.QTD_ PRAZO_GUARDA	NE	Conforme a Tabela de Temporalidade do TCU todos os prazos são em anos, então o evento de contagem na idade intermediária se dará automaticamente ao fim do prazo corrente.
	3.11. Destinação final	Sim	O	API		TCU.HIST_ASSOC_RE GRA_FUNCAO_TEMP O.IND_TIPO_DESTINA CAO_FINAL	G' = 'Guardar'. 'E' = 'Excluir'.	
	3.12. Sigilo associado à classe	Não se aplica	OA	API		NC	NE	
	3.13. Observação	Sim	F	API		TCU.CLASSE_ASSUN TO.TEXTTO_OBSERVA CAO	NE	
4. Eventos de gerenciamento de classe	4.1. Identificador do evento	Sim	O	API				Existe mas não foi confirmado

	4.2. Tipo do evento	Sim	O	API				Existe mas não foi confirmado
	4.3. Identificador da classe afetada	Sim	O	API				Existe mas não foi confirmado
	4.4. Data e hora do evento	Sim	O	API				Existe mas não foi confirmado
	4.5. Agente responsável pelo evento	Sim	O	API				Existe mas não foi confirmado
	4.6. Valor anterior do atributo	Sim	OA	API				Existe mas não foi confirmado
5. Componente digital	5.1. Identificador do componente digital	Sim	O	API		NC	999999999	number(9)
	5.2. Nome original	Sim	F	API		arquivo_eletronico.nome	xxxxxx	varchar(256)
	5.3. Tamanho	Sim	F	API		versao_arquivo_eletronico.valor_tamanho_arquivo	10000000000000000000	number(17)
	5.4. Software de criação	Não se aplica	F	API		NA	NE	
	5.5. Nível de composição	Sim	OA	API		valor será sempre 1	1	Pq sempre os arquivos serão zipados e alguns criptografados
	5.6. Inibidor	Sim	F	API		versao_arquivo_eletronico.num_chave_criptografia	9999	number(4)

	5.7. Formato de arquivo	Sim	O	API		versao_arquivo_eletronico.sigla_extensao	pdf docx doc tcu	varchar(20)
	5.8. Localização	Não	OA	API				Vai depender como será a integração
	5.9. Suporte	Sim	F	API			storage do data center	
	5.10. Dependência de software	Não	OA	API		NC	NE	
	5.11. Dependência de hardware	Não	OA	API		NC	NE	
	5.12. Outras dependências	Não	OA	API		NC	NE	
	5.13. Relação com outros componentes digitais	Não	OA	API		NC	NE	
	5.14. Fixidade	Sim	O	API		assinatura_versao_arq_eletron. texto_hash_criptografado	tem exemplo?	Varchar (4000)
	5.15. Assinatura digital	Sim	OA	API		assinatura_versao_arq_eletron. texto_hash_criptografado	tem exemplo?	Varchar (4000)
6. Evento de preservação	6.1. Identificador do evento	Não	O	API		NA	NE	Não existe, deve ser criado
	6.2. Tipo de evento	Não	O	API		NA	NE	Não existe, deve ser criado

	6.3. Identificador de componente digital	Não	O	API		NA	NE	Não existe, deve ser criado
	6.4. Data e hora do evento	Não	O	API		NA	NE	Não existe, deve ser criado
	6.5. Agente responsável pelo evento	Não	O	API		NA	NE	Não existe, deve ser criado
	6.6. Resultado do evento	Não	OA	API		NA	NE	Não existe, deve ser criado
	6.7. Detalhes do evento	Não	F	API		NA	NE	
7. Agente	7.1. Identificador do agente	Sim	O	API		tabela de usuários internos	WALTERVF	
	7.2. Nome do agente	Sim	O	API		tabela de usuários internos	Walter Venson Filho	
	7.1. Status do agente	Sim	O	API		tabela de usuários internos	Servidor Ativo	

2. Metadados a serem criados

		Proposta de Metadados	Sugestão IBICT	O que se espera?
8. Metadados do TCU				autor é o usuário responsável juridicamente falando pelo documento/ originador é a unidade/usuário autuadora/ redator é o usuário/ unidade que elabora o documento
	Processo	Relator	NOVO	
	Processo	Representante do MPTCU	NOVO	
		Parte Interessada	1.18	Documento herda do processo
		Representantes Legais	NOVO	
		Órgão Interessado	1.18	
	Processo	Unidade Jurisdicionada	NOVO	Unidade que está sendo fiscalizada. Entidade, órgãos .
	Processo	Responsaveis	NOVO	responsáveis pelas contas no processo ex tomada de contas especial
	Processo	Legitimadados	NOVO	possibilidade de ter vistas ao processo; é um CPF (nome,

				qualificação...)
	Processo	Se interesse Congresso Nacional	NOVO	campo marcado no sistema
	Processo	Se Urgente	NOVO	campo marcado no sistema
	Processo	Se Histórico	Selo Histórico (Lista de classe)	CPAD/ campo marcado no sistema
	Processo	Se Repercussão na Midia		CPAD/ campo marcado no sistema
	Documento/Processo	Data de autuação	1.32. Data de produção	
	Processo	Tipo	NOVO	forma legado de classificação dos processos anteriores ao PCD e TTD.
	Processo	Assunto do processo	NOVO	
	Processo	Número	NOVO	
	Processo	Unidade Responsável Técnica (histórico)	NOVO	buscar todas as unidades responsáveis pelo processo no período
	Processo	Processos apensados	NOVO	

	Processo	Processos apartados	NOVO	
	Documento/processo	Confidencialidade	1.30. Níveis de acesso	
	Processo	Estado	NOVO	encerrado/aberto/descartado
	Processo	Situação	NOVO	aguardando providência, em instrução...
	Processo	Deliberações	NOVO	decisão final (histórico)

3. Análise

		EXEMPLO EXTRAÇÃO ATUAL Opções: indicação do valor; não há exemplo (NE)	DUBLIN CORE	CAMINHO Opções: indicação do caminho; não se aplica (NA); não conhecido (NC)
1.a Documento	1.1 Identificador do documento			tcu.documento_gestao.cod_documento
	1.2. Número do documento		dc.identifier	tcu.documento_tramitavel.cod + calculadvdocumento(dt.cod)

	1.9. Status			tcu.arquivo_eletronico_catalogado.cod_oficializacao & tcu.vinculo_documento_lote.ind_tipo_vinculo
	1.10. Identificador de versão			tabela versao_arquivo_eletronico
	1.11. Título		dc.title	NC
	1.12. Descrição			NA
	1.13. Assunto		dc.subject	tcu.documento_tramitavel.texto_complemento_assunto
	1.14. Autor			NC
	1.15. Destinatário		dc.contributor	NC
	1.16. Originador			NA
	1.17. Redator		dc.publisher	tem caminho?
	1.18. Interessado (Parte Interessada)			tem caminho?
	1.18. Interessado (Representante Legal)			
	1.18. Interessado (Órgão Interessado)			
	1.19. Identificador do componente digital			NC
	1.20. Gênero			NA

1.21. Espécie			NA
1.22. Tipo			tcu.tipo_documento
1.23. Idioma		dc.language	NA
1.25. Numeração sequencial dos documentos			tcu.vinculo_documento_lote.num_peca_eletronica
1.26. Indicação de anexos			tcu.HIST_VINCULO_DOCUMENTO
1.27. Indicação de anotação			NA
1.28. Unidade responsável pela execução da ação			NC
1.29. Relação com outros documentos			tcu.vinculo_documento_lote
1.30. Níveis de acesso			documento_gestao.cod_tipo_confidencialidade e arquivo_eletronico_catalogado.ind_tipo_escopo
1.31. Previsão de desclassificação			Está ativo? Tem caminho?
1.32. Data de produção		dc.date	versao_arquivo_eletronico.dthora_criacao
1.33. Local de produção		dc.covergae	NA

	Metadados de assinatura/certificação (arquivo em P7s)			assinatura_versao_arq HIST_OFICIALIZA_AQR_ELET_CAT
Processo (NOVO)	1.3. Número do protocolo			tcu.documento_gestao.cod_registro_entrada + dv
	1.4. Identificador do processo/dossiê		dc.identifier	tcu.vinculo_documento_lote.cod_lote where ind_tipo_vinculo = 'O'
	1.5. Número de protocolo do processo		dc.title	(num, ano, num_dv) tcu.processo_gestao
	1.6. Identificador do volume			NA
	1.7. Número do volume			NA
	1.8. Tipo de meio			tcu.registro_entrada_protocolo.cod_tipo_origem
	1.11. Título			NC
	1.12. Descrição			NA
	1.13. Assunto		dc.subject	tcu.documento_tramitavel.texto_complemento_assunto
	1.14. Autor			NC
	1.16. Originador		dc.publisher	NA
	1.17. Redator		dc.publisher	tem caminho?
	1.18. Interessado (Parte Interessada)		dc.contributor	tem caminho?

1.18. Interessado (Representante Legal)		dc.contributor	
1.18. Interessado (Órgão Interessado)		dc.contributor	
1.19. Identificador do componente digital			NC
1.23. Idioma		dc.language	NA
1.24. Quantidade de folhas		dc.format	tcu.documento_gestao_num_paginas_digitalizadas
1.26. Indicação de anexos			tcu.HIST_VINCULO_DOCUMENTO
1.27. Indicação de anotação			NA
1.28. Unidade responsável pela execução da ação			NC
1.30. Níveis de acesso			documento_gestao.cod_tipo_confidencialidade e arquivo_eletronico_catalogado.ind_tipo_escopo
1.31. Previsão de desclassificação			Está ativo? Tem caminho?
1.32. Data de produção		dc.date	versao_arquivo_eletronico.dthora_criacao
1.33. Local de produção		dc.coverage	NA
1.34. Classe		ds.subject	tcu.classe_assunto

	1.35. Destinação prevista			tem caminho?
	1.36. Prazo de guarda			tem caminho?
	1.37. Localização			ARMAZENAMENTO_PAPEL_UNIDADE??
	Número de peças			
	Relator		dc.publisher	
	Representante do MPTCU		dc.publisher	
	Unidade Jurisdicionada		dc.contributor	
	Responsáveis		dc.contributor	
	Legitimados (nome)		dc.contributor	
	Legitimados CPF			
	Histórico			
	Situação			
	Deliberações			
	Processos Apensados			
	Processos apartados			
	Processo de interesse do Congresso Nacional			
	Urgência			
	Repercussão na Mídia			
	Selo histórico		dc.subject	

	Metadados de assinatura/certificação			
2. Evento de gestão	2.1.1. Identificador do evento			tcu.evento_documento.cod
	2.1.2. Tipo de evento			tcu.evento_documento.cod_tipo_evento
	2.1.3. Identificador do processo/dossiê			tcu.vinculo_documento_lote
	2.1.4. Identificador do documento			tcu.documento_gestao_cod_documento
	2.1.5. Identificador do lote			NA
	2.1.6. Data e hora do evento			tcu.evento_documento
	2.1.7. Agente responsável pelo evento			tcu.evento_documento
	2.1.8. Detalhes do evento			TEXTO_FORMATADO + TEXTO_OBSERVACAO
	2.2.1. Identificador do evento			
	2.2.2. Tipo do evento			
	2.2.3. Identificador do processo			

	2.2.4. Identificador do volume			
	2.2.5. Data e hora do evento			
	2.2.6. Agente responsável pelo evento			
	2.2.7. Identificador do documento			
3. Classe	3.1. Identificador da classe			TCU.CLASSE_ASSUNTO.COD
	3.2. Nome da classe			TCU.CLASSE_ASSUNTO.DESCR
	3.3. Código da classe			TCU.CLASSE_ASSUNTO.COD_CONARQ
	3.4. Subordinação da classe			TCU.CLASSE_ASSUNTO.COD_CLASSE_ASSUNTO_PAI
	3.5. Indicação de permissão de uso			NC
	3.6. Indicação de classe ativa/inativa			TCU.CLASSE_ASSUNTO.DATA_EXCLUSAO
	3.7. Prazo na idade corrente			TCU.HIST_ASSOC_REGRA_FUNCAO_TEMPO.COD_REGRA_TEMPORALIDADE + TCU.TIPO_REGRA_TEMPORALIDADE.QTD_PRAZO_GUARDA

	3.8. Evento de contagem na idade corrente			TCU.HIST_ASSOC_REGRA_FUNCAO_TEMPO.COD_REGRA_TEMPORALIDADE + TCU.TIPO_REGRA_TEMPORALIDADE.QTD_PRAZO_GUARDA
	3.9. Prazo na idade intermediária			TCU.HIST_ASSOC_REGRA_FUNCAO_TEMPO.QTD_ANOS_FASE_INTERMEDIARIA
	3.10. Evento de contagem na idade intermediária			TCU.HIST_ASSOC_REGRA_FUNCAO_TEMPO.COD_REGRA_TEMPORALIDADE + TCU.TIPO_REGRA_TEMPORALIDADE.QTD_PRAZO_GUARDA
	3.11. Destinação final			TCU.HIST_ASSOC_REGRA_FUNCAO_TEMPO.IND_TIPO_DESTINACAO_FINAL
	3.12. Sigilo associado à classe			NC
	3.13. Observação			TCU.CLASSE_ASSUNTO.TEXTTO_OBSERVACAO
4. Eventos de gerenciamento de classe	4.1. Identificador do evento			
	4.2. Tipo do evento			
	4.3. Identificador da classe afetada			
	4.4. Data e hora do evento			
	4.5. Agente responsável pelo evento			

	4.6. Valor anterior do atributo			
5. Componente digital	5.1. Identificador do componente digital			NC
	5.2. Nome original			arquivo_eletronico.nome
	5.3. Tamanho			versao_arquivo_eletronico.valor_tamanho_arquivo
	5.4. Software de criação			NA
	5.5. Nível de composição			valor será sempre 1
	5.6. Inibidor			versao_arquivo_eletronico.num_chave_criptografia
	5.7. Formato de arquivo			versao_arquivo_eletronico.sigla_extensao
	5.8. Localização			
	5.9. Suporte			
	5.10. Dependência de software			NC
	5.11. Dependência de hardware			NC
	5.12. Outras dependências			NC
	5.13. Relação com outros componentes digitais			NC
	5.14. Fixidade			assinatura_versao_arq_eletron. texto_hash_criptografado

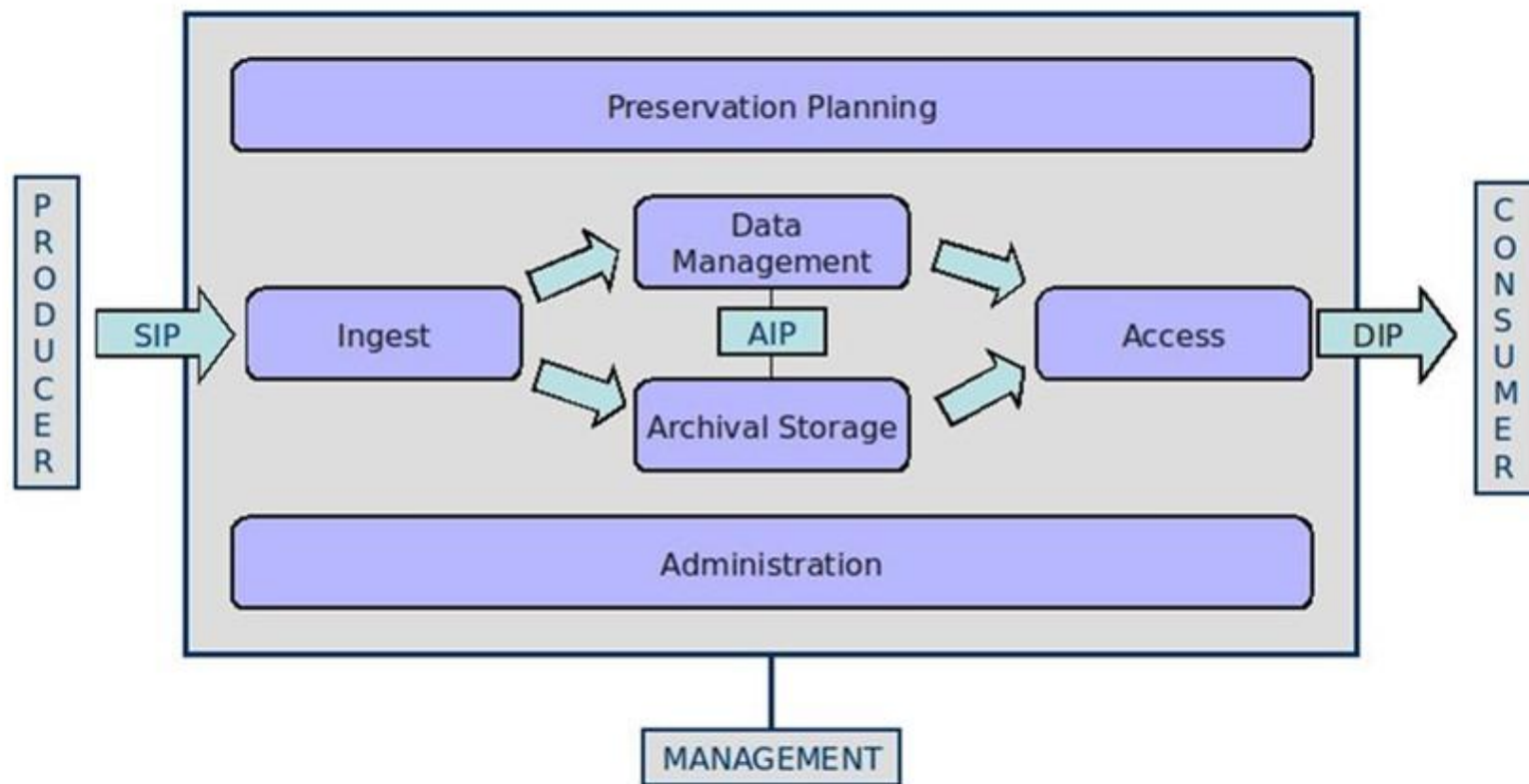
	5.15. Assinatura digital			assinatura_versao_arq_eletron. texto_hash_criptografado
6. Evento de preservação	6.1. Identificador do evento			NA
	6.2. Tipo de evento			NA
	6.3. Identificador de componente digital			NA
	6.4. Data e hora do evento			NA
	6.5. Agente responsável pelo evento			NA
	6.6. Resultado do evento			NA
	6.7. Detalhes do evento			NA
7. Agente	7.1. Identificador do agente			tabela de usuários internos
	7.2. Nome do agente			tabela de usuários internos
	7.1. Status do agente			tabela de usuários internos

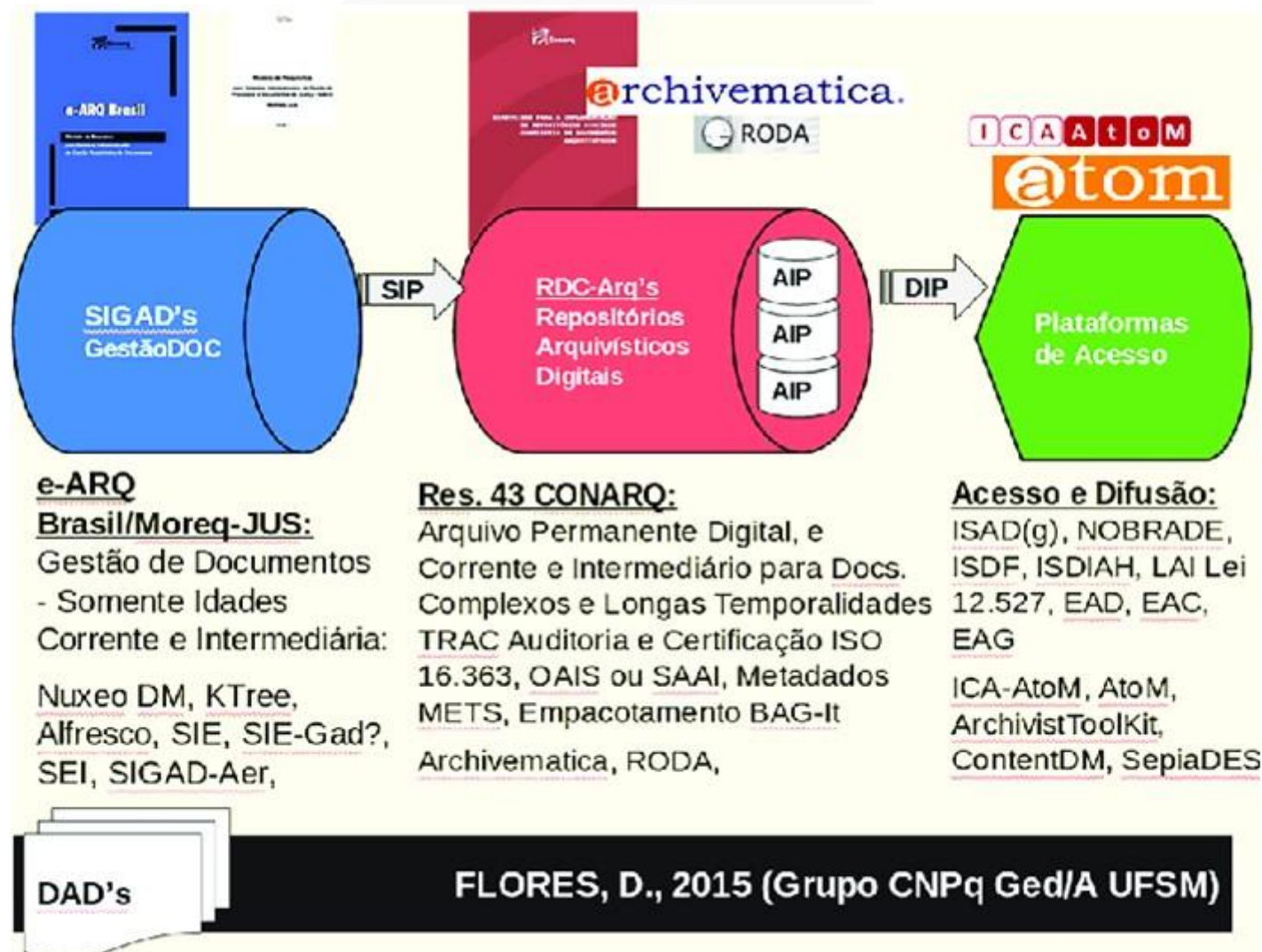
APÊNDICE E - Capacitação ao TCU

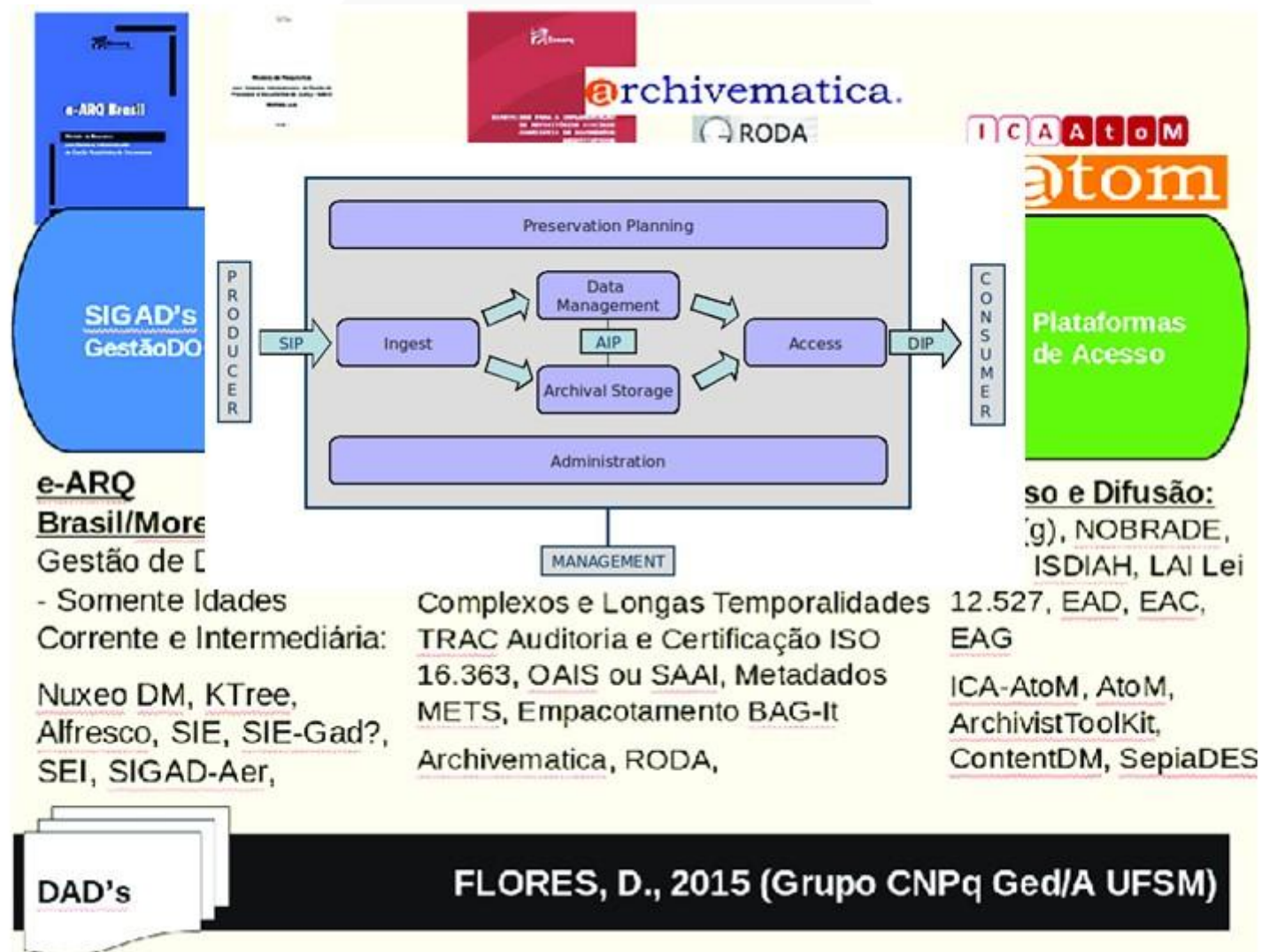
PROJETO DE PESQUISA: ESTUDO PARA IMPLEMENTAÇÃO DO MODELO DE PRESERVAÇÃO HIPÁTIA NO TRIBUNAL DE CONTAS DA UNIÃO

Ívina Flores Cássio Pires Marcelo Brondani Marilete Pereira







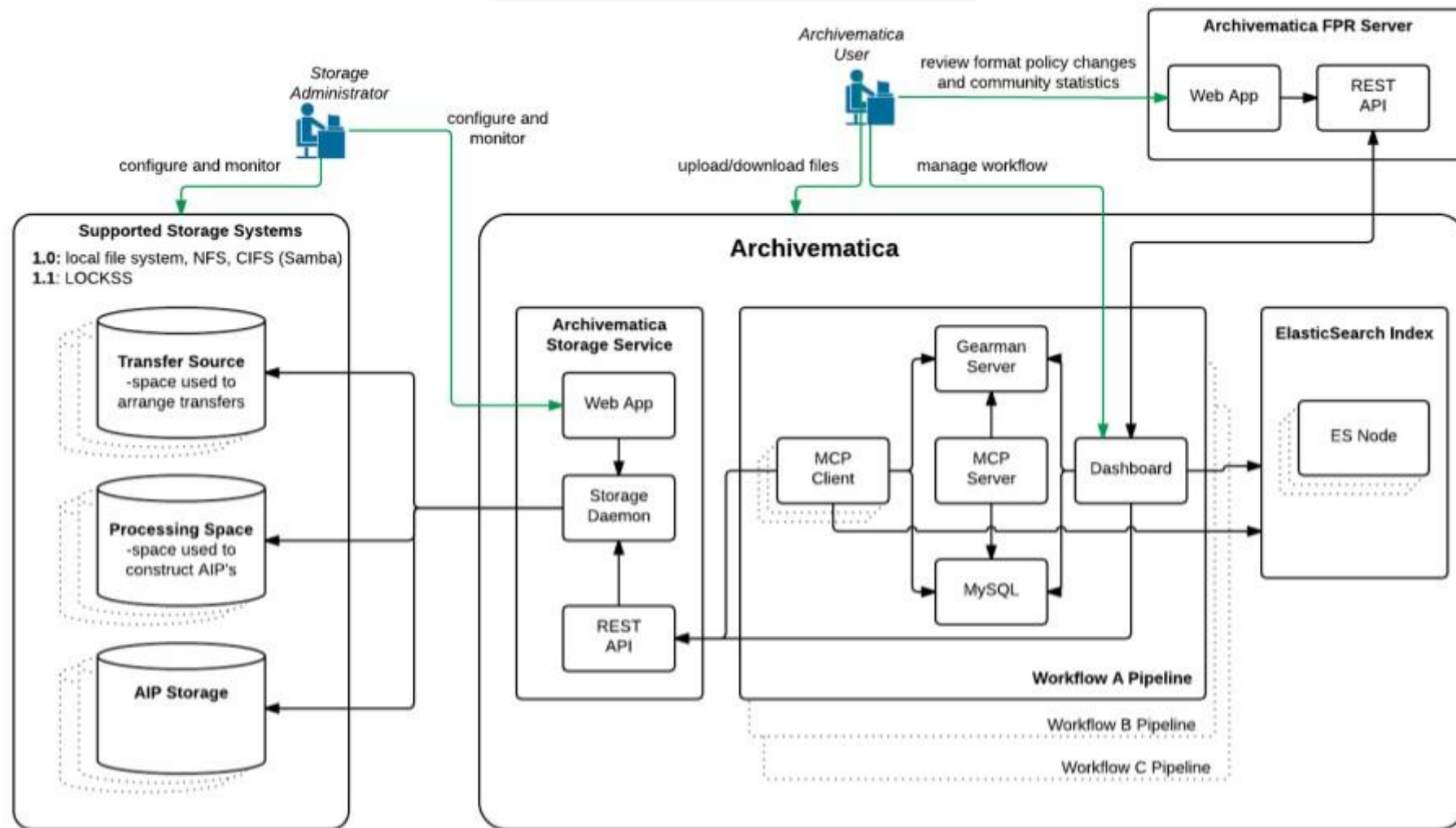


TECNOLOGIAS DO MODELO HIPÁTIA DE PRESERVAÇÃO DIGITAL



 archivematica[®]

 **atom**
access to memory



All Packages

Packages are Transfers, SIPs and AIPs uploaded to a Location managed by the storage service.

[View delete requests](#)

Show entries

Search:

UUID	Originating Pipeline	Current Location	Size	Type	Status
194303f9-f944-4f62-a6da-0080bb34e90e	Archivematica on archivematica1Dev02 (798afed6-d80b-46e1-b025-36dff21f29af)	/var/archivematica/sharedDirectory/www/AIPsStore/1943/03f9/f944/4f62/a6da/0080/bb34/e90e/test1-194303f9-f944-4f62-a6da-0080bb34e90e.7z	28.1 MB	AIP	Deleted
f5279c6d-52f1-4dbf-a579-3582877a91c2	Archivematica on archivematica1Dev02 (798afed6-d80b-46e1-b025-36dff21f29af)	/var/archivematica/sharedDirectory/www/AIPsStore/f527/9c6d/52f1/4dbf/a579/3582/877a/91c2/OD-f5279c6d-52f1-4dbf-a579-3582877a91c2.7z	844.4 KB	AIP	Uploaded

Showing 1 to 2 of 2 entries

[Previous](#) [Next](#)

Archivematica Dashboard - Tr... 

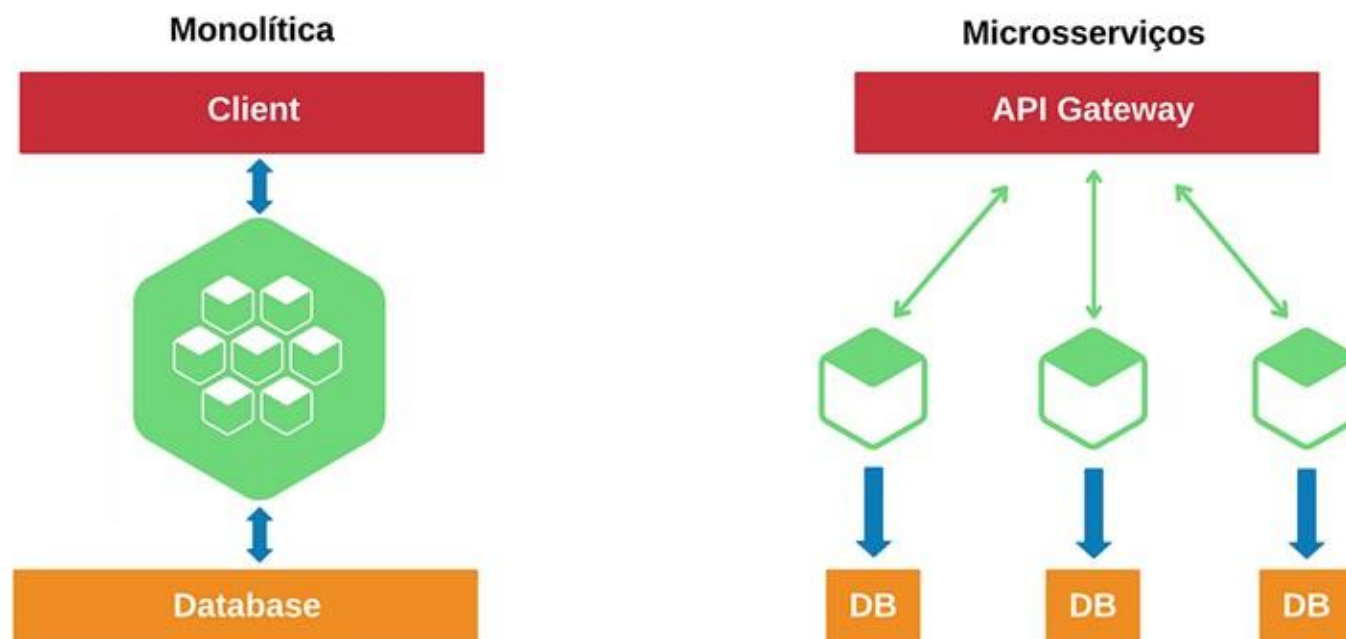
 sankofa.archivematica.org/transfer/    Google

archivematica [Transfer](#) [Ingest](#) [Archival storage](#) [Preservation planning](#) [Access](#) [Administration](#) [admin](#)

Standard /home/artefactual/archivematica-  

Type Transfer name Accession no.

Transfer	UUID	Transfer start time
 MultisIP ▶ Micro-service: Create SIP from Transfer ▶ Micro-service: Complete transfer ▼ Micro-service: Characterize and extract metadata Job: Check for specialized processing Job: Set file permissions Job: Load labels from metadata/file_labels.csv Job: Characterize and extract metadata ▼ Micro-service: Extract packages Job: Extract contents from compressed archives ▼ Micro-service: Identify file format Job: Identify file format Job: Determine which files to identify Job: Select file format identification command Job: Move to select file ID tool ▶ Micro-service: Clean up names ▶ Micro-service: Scan for viruses ▶ Micro-service: Quarantine ▶ Micro-service: Generate METS.xml document ▶ Micro-service: Verify transfer checksums	a7851af2-9033-4efc-97c5-d4b8d6330ef1	2014-02-06 11:41



Arquitetura de Microserviços

APÊNDICE F - Formulário de *Processing Configuration*

HYPÁTIA
MODELO DE PRESERVAÇÃO

**ibict**

Configuração de Processamento (CP) - TCU

A Configuração de Processamento define como algumas ações (“jobs”) que serão executadas pelo repositório ao receber objetos digitais (por exemplo, documentos) para serem preservados.

O objetivo deste formulário é levantar informações para a elaboração da Configuração de Processamento adotada pelo repositório do Tribunal de Contas da União na admissão de objetos digitais do e-TCU.

Este formulário utiliza trechos traduzidos da documentação do *software* Archivematica, disponível em: <https://www.archivematica.org/en/docs/> (inglês).

Denominação do CP *

No Archivematica, é possível atribuir um nome para cada Configuração de Processamento elaborada. Por exemplo: DJeCP, projudi-CP, documentosTextuais_CP, audiovisuais, cp-digitalizados etc.

Obs. Este campo do formulário não pode ser modificado depois de preenchido e salvo.

Sua resposta

Escanear vírus*

*

O Archivematica deve executar a busca por vírus no conteúdo das transferências?

☐ Sim

☐ Não

Atribuir UUIDs a diretórios *

*

O UUID é um identificador universal exclusivo, uma espécie de código único utilizado na computação, que possibilita a identificação de uma informação única em sistemas distribuídos.

O Archivematica atribui esses identificadores para os objetos digitais das transferências recebidas, mas ele também pode atribuir esse identificador aos diretórios (“pastas”). Na Configuração de Processamento, os diretórios processados também podem receber um UUID.

Esses identificadores devem ser registrados nos metadados do processamento?

- ☐ Sim - UUIDs são atribuídos.
- ☐ Não - UUIDs não são atribuídos.*

Gerar relatório de estrutura de transferência *



Os Pacotes de Transferência Inicial podem conter uma estrutura interna de diretórios e subdiretórios (“pastas e subpastas”). Na Configuração de Processamento, deve ser gerado um arquivo de texto mostrando a árvore de diretórios da transferência original?

OBS: A hierarquia mostrada é apenas do diretório que contém os objetos originais. Se produzido, este arquivo de texto será armazenado pelo repositório.

- ☐ Sim - o relatório de estrutura é criado.
- ☐ Não - relatório de estrutura não é criado.

Realizar a identificação do formato dos objetos na etapa de transferência *



O Archivematica pode utilizar programas auxiliares para identificar o formato dos arquivos da transferência recebida e registrar essa identificação nos metadados que preservará junto com esses arquivos. Por exemplo, o programa identificará se um arquivo “DJe.pdf” é realmente um arquivo de formato PDF e qual a sua versão. Esta Configuração de Processamento deve ativar esta ação?

- ☐ Sim - identifique o formato dos arquivos.
- ☐ Não - os formatos não serão identificados.

Extrair pacotes *



Pacotes de arquivos (por exemplo, arquivos .zip) enviados na transferência podem ser descompactados e extraídos em um diretório durante o processo. Esta ação deve ser realizada na presente Configuração de Processamento?

- ☐ Sim - o conteúdo do pacote é extraído.
- ☐ Não - o pacote é deixado como está.

Excluir pacotes após a extração *

*

Os pacotes de arquivos (por exemplo, arquivos .zip) que foram extraídos na etapa anterior podem ser excluídos após a realização da extração para evitar o armazenamento de conteúdo duplicado. Nesta Configuração de Processamento, esses pacotes devem ser automaticamente deletados após a extração?

- ☐ Sim - o pacote foi excluído.
- ☐ Não - o pacote é preservado junto com o conteúdo extraído.

Executar verificações de política nos objetos originais *

*

O Archivematica permite que os usuários utilizem um programa chamado [MediaConch](#) para criar “políticas” para arquivos. Essas “políticas” estabelecem as características técnicas que certos formatos de arquivo devem ter para serem validadas ou não pelo repositório. Se você criou políticas usando o [MediaConch](#), nesta Configuração de Processamento, o Archivematica deve verificar se os arquivos originais da transferência estão de acordo com essas “políticas”?

- ☐ Sim - a transferência é verificada em relação a quaisquer políticas.
- ☐ Não - as políticas não serão consideradas.

Examinar o conteúdo *

A opção de “examinar o conteúdo”, na Configuração de Processamento, executa uma análise dos documentos enviados para preservação. O Archivematica utiliza para isso a ferramenta [Bulk Extractor](#), que serve para reconhecer padrões em dados textuais, tais como números de cartão de crédito. Contudo, é preciso consultar a documentação da ferramenta para verificar se os padrões desejados serão efetivamente reconhecidos. Esses padrões que forem reconhecidos são registrados em um arquivo de *log*.

Para obter mais informações sobre como revisar *logs* do *Bulk Extractor*, consulte o [painel Análise](#) na guia de “Avaliação” do Archivematica.

Deseja que, nesta Configuração de Processamento, o Archivematica chame o programa *Bulk Extractor*?

- ☐ Pular examinar o conteúdo - Extrator em massa não é executado.
- ☐ Sim - o Bulk Extractor verifica o conteúdo e cria saídas de log de padrões reconhecidos para revisão.

Criar SIP(s) *

Após o processamento da transferência inicial, é preciso informar ao Archivematica o que fazer com o conteúdo processado até aquele ponto. A primeira alternativa é a produção imediata de um pacote de submissão de informação (o chamado "SIP"). Se esta opção for definida na Configuração de Processamento, o processamento inicial será, automaticamente, transformado em um SIP e o Archivematica dará início ao processo de "admissão" (*"ingest"*) do SIP no repositório.

A segunda alternativa é tornar como padrão que após o processamento de uma transferência inicial, esta sejam enviadas para armazenamento em uma lista de pendências, o chamado *"backlog"*.

Qual deve ser o padrão para esta Configuração de Processamento?

Obs: Caso seja enviado para *Backlog*, os outros *jobs* podem não acontecer de forma automatizada.

- ☐ Crie um SIP único e continue processando - a transferência torna-se um SIP e fica disponível para processamento adicional na guia de ingestão.
- ☐ Enviar para lista de pendências - a transferência é enviada para um espaço de armazenamento de lista de pendências para armazenamento temporário ou avaliação.

Realizar identificação do formato dos objetos do SIP *

*

O Archivematica pode identificar o formato dos arquivos que estão no pacote de informação ("SIP") com o apoio de outras ferramentas. Por padrão, na última versão do Archivematica, a ferramenta utilizada é o [Siegfried](#). Na aba de "planejamento de preservação" é possível customizar opções sobre identificação. Caso não sejam identificados os objetos do SIP, as informações de identificação da transferência inicial serão reutilizadas.

Executar a identificação de formatos do SIP nesta Configuração de Processamento?

- ☐ Sim - use o comando de identificação dos formatos do arquivo ativado.
- ☐ Não, use dados existentes - reutiliza dados de identificação dos formatos de arquivo da guia de transferência.

Normalizar *

A normalização é o processo de converter arquivos admitidos no repositório para formatos preferenciais. O Archivematica pode normalizar para a preservação e o acesso, apenas para a preservação, apenas para o acesso, ou, não normalizar (além de outras opções que ainda não utilizamos no modelo Hipátia). Quando é realizada a normalização para preservação, derivadas nos formatos preferenciais são criadas e incluídas no pacote de preservação (o "AIP"), que pode ser armazenado para o longo prazo. A normalização para acesso, por outro lado, cria derivadas de acesso utilizadas para produzir o pacote de disseminação (o "DIP"), que o usuário poderá enviar para uma plataforma de acesso (por exemplo, o AtoM). O "AIP" também será criado, mas não teremos a criação de derivadas de preservação, preservando apenas os originais. Na aba de "planejamento de preservação" do Archivematica é possível editar configurações de normalização.

Qual processo de normalização você prefere que seja executado nesta Configuração de Processamento?

- ☐ Normalizar para preservação e acesso - cria cópias de preservação dos objetos mais cópias de acesso que serão usadas para gerar o DIP.
- ☐ Normalizar para preservação - cria apenas cópias de preservação. Nenhuma cópia de acesso é criada e nenhum DIP será gerado.
- ☐ Normalizar para acesso - o AIP conterá apenas os originais. Nenhuma cópia de preservação será gerada. Serão criadas cópias de acesso que serão utilizadas para gerar o DIP.
- ☐ Não normalize - o AIP conterá apenas os originais. Nenhuma cópia de preservação ou acesso é gerada e nenhum DIP será gerado.

Aprovar normalização



O painel permite que os usuários revisem a saída de normalização e o relatório de normalização.

- ☐ Sim - pule a etapa de revisão e continue processando automaticamente.

Gerar miniaturas *

O processamento pode gerar automaticamente miniaturas (“*thumbnails*”) de arquivos para o pacote de preservação (“AIP”) e para o de disseminação (“DIP”). O Archivematica utiliza para isso as regras do *Format Policy Registry* (FPR), uma base de dados que permite que os usuários definam como o repositório deve proceder para cada formato de arquivo recebido. Esta Configuração de Processamento deve executar a ação de produção de miniaturas?

- ☐ Sim - as miniaturas serão geradas de acordo com as regras de formato do FPR. Os formatos que não possuem uma regra terão uma miniatura padrão gerada (ícone de documento cinza).
- ☐ Sim, sem ícones padrão - as miniaturas serão produzidas para qualquer formato que tenha uma regra de normalização para miniaturas no FPR. Formatos que não possuem regra não terão miniatura gerada.
- ☐ Não - as miniaturas não serão geradas.

Executar verificações de política em derivadas de preservação *



O Archivematica permite que os usuários utilizem um programa chamado [MediaConch](#) para criar “políticas”, ou conferencias de conformidades determinadas pela instituição. Essas “políticas” estabelecem características técnicas que certos formatos de arquivo devem ter para serem validados ou não pelo repositório. Se foram criadas “políticas” usando o [MediaConch](#), o Archivematica deve verificar se os arquivos **normalizados para a preservação** estão de acordo com essas “políticas” nesta Configuração de Processamento?

- ☐ Sim - os arquivos normalizados são verificados em relação a quaisquer políticas.
- ☐ Não - as políticas são ignoradas.

Executar verificações de política em derivadas de acesso *



O Archivematica permite que os usuários utilizem um programa chamado [MediaConch](#) para criar “políticas”. Essas “políticas” estabelecem características técnicas que certos formatos de arquivo devem ter para serem validados ou não pelo repositório. Se foram criadas “políticas” usando o [MediaConch](#), o Archivematica deve verificar se os arquivos **normalizados para o acesso** estão de acordo com essas “políticas” nesta Configuração de Processamento?

- ☐ Sim - os arquivos normalizados são verificados em relação a quaisquer políticas.
- ☐ Não - as políticas são ignoradas.

Vincular PIDs a objetos digitais *

*

PIDs são identificadores persistentes que o Archivematica pode atribuir aos objetos digitais, se estiver configurada uma integração do sistema com o serviço Handle.Net. Consulte [Vincular PIDs](#) para obter mais informações.

Nesta Configuração de Processamento, o Archivematica deve criar PIDs?

- ☐ Sim - os PIDs são criados e uma chamada de API envia os PIDs para o Handle Server.
- ☐ Não - PIDs não são criados.

Documentar existência de diretórios vazios *



Por padrão, o Archivematica remove diretórios transferidos que estejam vazios e não documenta sua existência. O processamento deve documentar em metadados a existência desses diretórios de acordo com esta Configuração?

- ☐ Sim - uma entrada para o diretório é criada no structmap.
- ☐ Não - o diretório não está documentado.

Adicionar metadados (lembrete) *

*

O Archivematica permite aos usuários adicionar metadados ou por meio de formulários preenchidos durante o processamento ou por meio de *upload* de CSV pela interface. Deseja que, nesta Configuração de Processamento, o sistema pule esta etapa e continue o processamento?

☐ Continuar - pule o lembrete e continue processando automaticamente.

Transcrever arquivos (OCR) *



Os usuários podem optar por executar o Tesseract, uma ferramenta de Reconhecimento Ótico de Caracteres (OCR) incluída no Archivematica, para produzir arquivos de texto contendo transcrições de arquivos simples (páginas únicas) de imagem. Nesta Configuração de Processamento, os caracteres em imagens devem ser reconhecidos e transcritos?

- ☐ Sim - o Tesseract é executado em todos os arquivos compatíveis com OCR.
- ☐ Não - o Tesseract não é executado.

Realizar a identificação do formato de documentação de submissão e de arquivos de metadados *

É possível identificar formatos de documentação de submissão e de arquivos de metadados que foram incluídos na transferência. A ferramenta que o Archivematica utilizará nesta identificação pode ser escolhida na aba de “Planejamento de preservação”. Para mais informações, consulte o item [“Identification”](#) da documentação.

O Archivematica deve executar, nesta Configuração de Processamento, o comando de identificação previamente configurado?

- ☐ Sim - use o comando de identificação de arquivo ativado.
- ☐ Não - os arquivos não serão identificados.

Selecionar o algoritmo para comprimir de pacotes de preservação (AIPs) *

★

Os pacotes de arquivamento (“AIPs”) para preservação criados pelo Archivematica podem ser armazenados como pacotes compactados ou descompactados, dependendo dos requisitos de armazenamento.

Como o Archivematica deve proceder em relação à compressão de AIPs nesta Configuração de Processamento?

- ☐ 7z usando bzip2 - um arquivo 7Zip é criado usando o algoritmo bzip2 .
- ☐ 7z usando LZMA - um arquivo 7Zip é criado usando o algoritmo LZMA .
- ☐ 7z sem compactação - um arquivo 7Zip é criado, mas o conteúdo não é compactado.
- ☐ Gzipped tar - o AIP é criado usando o algoritmo gzip .
- ☐ Parallel bzip2 - um arquivo 7Zip é criado usando o algoritmo Parallel bzip2 (pbzip2) .
- ☐ Descompactado - o AIP não é compactado.

Selecionar o nível de compressão do AIP *

Se foi definida uma opção de compactação do pacote de preservação ("AIP") para esta Configuração de Processamento (pergunta anterior), então, é possível escolher também o grau de compressão utilizado.

Selecionar um nível de compactação mais alto significa que o AIP resultante será menor, mas a compactação também levará mais tempo. Níveis de compactação mais baixos significam compactação mais rápida, por outro lado, o AIP será maior e ocupará mais volume em disco.

Se a opção para algoritmo de compactação escolhida na pergunta anterior for "Não executar algoritmo de compactação" ou "7z sem compactação", a seleção do nível de compressão do AIP não terá efeito prático na Configuração de Processamento.

Qual nível de compressão deve ser aplicado ao AIP nesta Configuração de Processamento?

- ☐ 1 - compactação mais rápida - o AIP será compactado o mais rápido possível.
- ☐ 3 - compactação rápida - um AIP maior que será compactado rapidamente.
- ☐ 5 - compactação normal - a ferramenta de compactação atingirá um equilíbrio entre velocidade e compactação para criar um AIP de tamanho moderado e compressão moderada.
- ☐ 7 - compressão máxima - um AIP menor que leva mais tempo para compactar.
- ☐ 9 - ultra compressão - o menor AIP possível.

Armazenar AIP *



A principal finalidade do Archivematica é produzir o pacote de preservação (“AIP”). Contudo, antes de armazená-lo, pode ser definida uma pausa no processamento para que o usuário revise o conteúdo do AIP.

Nesta Configuração de Processamento, armazenar o AIP sem fazer essa pausa?

☐ Sim - o AIP é marcado para armazenamento automaticamente.

Definir local de armazenamento do AIP (localização no servidor) *

*

O Serviço de Armazenamento (“*Storage Service*”) do Archivematica tem um local padrão, definido pelo administrador, para armazenar os pacotes de preservação (“AIPs”), produzidos como resultado da transferência e admissão de objetos digitais no repositório.

Na presente Configuração de Processamento, o local de armazenamento deve ser o padrão do *storage service* ou outro? Se for outro, especifique.

☐ Local padrão - o AIP é armazenado no local de armazenamento do AIP que foi definido como padrão no Serviço de Armazenamento.

☐ Outro: _____

Enviar DIP para sistema de acesso *

*

O processamento poderá resultar na produção de um pacote de disseminação (“DIP”) com a finalidade de promover o acesso à informação do repositório. Este pacote pode ser enviado automaticamente para um sistema de acesso que esteja integrado com o Archivematica, por exemplo, o Atom.

Para esta Configuração de Processamento, o “DIP” deve ser enviado para o Atom?

- ☐ Nenhum
- ☐ Carregar DIP para AtoM - consulte a documentação de upload do AtoM DIP.
- ☐ Não faça upload - o DIP não será carregado em um sistema de acesso.

Armazenar localização DIP *

*

Se a resposta à pergunta anterior, sobre armazenamento do DIP, for positiva, onde os DIPs devem ser armazenados?

Observe que, arquivos com configurações de processamento personalizadas, incluídos nas transferências para o Archivematica, poderão estabelecer um local de armazenamento diferente daquele definido pela Configuração de Processamento elaborada a partir das respostas deste formulário.

Local padrão - o DIP é armazenado no local de armazenamento do DIP que foi definido como padrão no Serviço de Armazenamento (*"storage service"*).

- ☐ Armazenar DIP - o DIP é encaminhado para armazenamento automaticamente.
- ☐ Não armazene - o DIP é descartado.

Armazenar localização DIP *

Se a etapa anterior e esta etapa estiverem configuradas, todos os DIPs serão enviados para o local de armazenamento selecionado (a menos que você tenha incluído uma configuração de processamento personalizada com a transferência que define outro local).

☐ Local padrão - o DIP é armazenado no local de armazenamento do DIP que foi definido como padrão no Serviço de Armazenamento.

☐ Outro: _____