



**Relatório Técnico:  
Preservação digital e gestão  
arquivística apoiada no  
aprimoramento da implantação  
do modelo RDC-Arq**

---

## **PRESIDÊNCIA DA REPÚBLICA**

Jair Messias Bolsonaro

Presidente da República

Hamilton Mourão

Vice-Presidente da República

## **MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES**

Marcos Cesar Pontes

Ministro da Ciência, Tecnologia e Inovações.

## **INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA**

Cecília Leite Oliveira

Diretora

Reginaldo de Araújo Silva

Coordenação de Administração – COADM

Gustavo Saldanha

Coordenação de Ensino e Pesquisa, Ciência e Tecnologia da Informação – COEPPE

José Luis dos Santos Nascimento

Coordenação de Planejamento, Acompanhamento e Avaliação – COPAV

Anderson Itaborahy

Coordenador-Geral de Pesquisa e Desenvolvimento de Novos Produtos - CGNP

Bianca Amaro de Melo

Coordenadora-Geral de Pesquisa e Manutenção de Produtos Consolidados - CGPC

Tiago Emmanuel Nunes Braga

Coordenador-Geral de Tecnologias de Informação e Informática – CGTI

Alexandre Faria de Oliveira

Coordenador de Governança em Tecnologias para Informação e Comunicação - COTIC



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES  
Instituto Brasileiro de Informação em Ciência e Tecnologia

**Relatório Técnico:**

**Preservação digital e gestão arquivística  
apoiada no aprimoramento da implantação do  
modelo RDC-Arq**

Brasília  
2021



---

© 2020 Instituto Brasileiro de Informação em Ciência e Tecnologia

Esta obra é licenciada sob uma licença Creative Commons - Atribuição CC BY 4.0, sendo permitida a reprodução parcial ou total desde que mencionada a fonte.

## EQUIPE TÉCNICA

Diretora do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)  
Cecília Leite Oliveira

Coordenador-Geral de Tecnologias de Informação e Informática (CGTI)  
Tiago Emmanuel Nunes Braga

Coordenador do Projeto  
Alexandre Faria de Oliveira

### Autores

Alexandre Faria de Oliveira

Antonio Luiz de Aguiar Alves Henrique

Claudio Gonçalves Silva Gama

Henrique Oliveira Costa

Jéssica Bilac Gaspareto

Lara Luiza da Silva Aquino

Marcello Silva Coutinho

Marília Cristina da Silva

Marcos Sigismundo da Silva

Maxwell Borges Bezerra

Raul Neres Carlos

Tatiana Canelhas Pignataro

## Editoração

Normalização

Ingrid Torres Schiessl

Projeto gráfico e capa

Juliana Colem Rodrigues

Este Relatório de Técnico é um produto do projeto Preservação digital e gestão arquivística apoiada no aprimoramento da implantação do modelo RDCArq.

Ref. TJDFT - Processo SEI no 01302.000067/2021-45 e 01302.000389/2020-11

Ref. FUNDEP 28209

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.



Setor de Autarquias Sul Quadra 05 Lote 06, Bloco H – 5º andar

Cep:70.070-912 – Brasília, DF

Telefones: 55 (61) 3217-6360/55 /(61)3217-6350

www.ibict.br

---

# Sumário

<b>1. INTRODUÇÃO.....</b>	<b>10</b>
<b>2. OBJETIVOS .....</b>	<b>10</b>
2.1 Objetivo Geral.....	10
2.2 Objetivos Específicos.....	10
<b>3. CONTEXTUALIZAÇÃO E RESULTADOS.....</b>	<b>11</b>
<b>3.1 Relatórios e interação das equipes .....</b>	<b>11</b>
3.1.1 Reuniões técnicas.....	11
3.1.2 Reuniões de alinhamento .....	12
<b>3.2 META 1: Estudos de infraestrutura para solução RDC-Arq .....</b>	<b>14</b>
3.2.1 Estudos de infraestrutura para solução RDCArq .....	14
3.2.2 Processo de implantação do sistema no Openshift do TJDFT .....	16
3.2.2.1 CRIAÇÃO DOS VOLUMES (PVC) .....	16
3.2.2.2 CRIAÇÃO E DEPLOYMENT DOS SERVIÇOS E APLICAÇÕES .....	16
3.2.2.3 MYSQL .....	17
3.2.2.4 ELASTICSEARCH .....	17
3.2.2.5 REDIS .....	17
3.2.2.6 GEARMAND .....	17
3.2.2.7 FITS.....	18
3.2.2.8 CLAMAV .....	18
3.2.2.9 NGINX.....	18
3.2.2.10 MCPSERVER .....	19
3.2.2.11 MCPCLIENT .....	20
3.2.2.12 DASHBOARD .....	20
3.2.2.13 STORAGE SERVICE .....	20
3.2.2.14 INICIALIZAÇÃO E PREPARAÇÃO DOS SERVIÇOS .....	21
3.2.2.15 CRIAÇÃO DOS BANCOS DE DADOS .....	21
3.2.2.16 CRIAÇÃO DAS TABELAS .....	21
3.2.2.17 CRIAÇÃO DAS ROTAS .....	22
3.3 Definições para alinhamento com as políticas de Segurança da Informação do Tribunal .....	23
3.4 Estudo de múltiplos Pipelines nos ambientes do Archivematica para preservação de documentos de guarda longa e guarda permanente do TJDFT.....	24

---

<b>3.5 META 2: Análise de padrão de design de código para solução RDC-Arq</b> .....	<b>25</b>
3.5.1 Análise de padrão de design de código para solução RDCArq .....	26
3.5.2 Tipos de design patterns.....	26
<b>3.6 META 3: Estudo para implantação de barramento RDC-Arq</b> .....	<b>27</b>
3.6.1 Estudo para implantação de barramento RDC-Arq orientado a eventos .....	27
3.6.1.1 Estudo para integração com o serviço de mensageria Kafka .....	30
3.6.1.1.1 Integração com o Hipátia .....	30
3.6.2 Estudos para implantação de banco de dados no barramento Hipátia para gerenciamento dos dados de acesso aos sistemas .....	31
3.6.2.1 Definição das tabelas do Hipátia (MYSQL VERSÃO 5.7 COM SSL) .....	31
3.6.3 Estudos dos metadados RDC-Arq do PJE.....	33
3.6.3.1 TIPOS DE METADADOS .....	33
<b>4. Transferência de conhecimento</b> .....	<b>34</b>
4.1 Website.....	34
4.2 Fórum de discussão .....	36
<b>5. Considerações finais</b> .....	<b>36</b>
<b>Referências Bibliográficas</b> .....	<b>38</b>
<b>Dados do Processo</b> .....	<b>39</b>
Número do Processo .....	39
Classe Processual - Código .....	39
Classe Processual - Nome .....	39
Código da Localidade da Jurisdição .....	39
Nome da Localidade .....	40
Identificador da Competência.....	40
Nome da Competência .....	40
Processo Vinculado (São campos que se repetem) .....	40
Modalidade de vinculação de processo (São campos que se repetem).....	41
Prioridade.....	41
Valor da causa .....	41
Nível de Sigilo .....	42
Data da autuação .....	42
Pedido de liminar.....	42
Número SISPL.....	42
Processo de Origem .....	43
Incidental.....	43

---

---

Objeto do processo .....	43
Justiça 100 Digital .....	43
Processo Digitalizado.....	43
Órgão Julgador - Código.....	44
Órgão Julgador - Nome.....	44
Plano Classificação e Tabela de Temporalidade Utilizada.....	44
Prazo de guarda .....	45
Destinação prevista.....	45
<b>Submission Documentation.....</b>	<b>45</b>
Identificador do documento.....	45
Identificador do documento principal.....	45
Tipo de documento.....	46
Tipo de documento - Nome.....	46
Data e hora da criação .....	46
Descrição .....	46
Hash .....	47
Documento Sigiloso .....	47
Mimetype.....	47
Assinatura .....	47
Usuário da assinatura.....	48
Algoritmo hash.....	48
Cadeia do certificado .....	48
Codificação do certificado .....	48
* Data da assinatura.....	49
* Signatário login .....	49
Identificação do tipo de meio do documento .....	49
Gênero Documental .....	49
<b>Partes .....</b>	<b>50</b>
*Polo.....	50
*Nome da parte .....	50
*Sexo da parte.....	50
*Nome do genitor.....	50
*Nome da genitora.....	51
(Referenciado na tabela) Data de nascimento da parte .....	51
(Referenciado na tabela) Data de óbito da parte .....	51
*Número do documento principal.....	51
*Tipo de pessoa .....	52

---

(Referenciado na tabela) Nome Fantasia .....	52
(Referenciado na tabela) Pessoa vinculada .....	52
(Referenciado na tabela) Cidade natural .....	52
(Referenciado na tabela) Estado natural .....	53
(Referenciado na tabela) Nacionalidade da parte .....	53
*Código do documento de identificação da parte.....	53
*Emissor do documento de identificação da parte .....	53
*Tipo do documento de identificação da parte .....	54
Nome no documento de identificação.....	54
*Outro nome da parte .....	54
*CEP endereço da parte.....	55
*Logradouro endereço da parte.....	55
*Número endereço da parte .....	55
*Complemento endereço da parte .....	55
*Bairro endereço da parte.....	56
*Cidade endereço da parte .....	56
*Estado endereço da parte.....	56
*País endereço da parte.....	56
<b>Dados dos representantes processuais.....</b>	<b>57</b>
(Referenciado na tabela) Nome.....	57
(Referenciado na tabela) Inscrição .....	57
(Referenciado na tabela) Número do documento principal .....	57
(Referenciado na tabela) Tipo de representante.....	57
<b>Dados dos assuntos processuais.....</b>	<b>58</b>
*Principal.....	58
*Código Nacional.....	58
Código Pai Nacional.....	58
*Descrição igual ao nome.....	58
<b>Movimentação Processual .....</b>	<b>59</b>
*Código da movimentação processual .....	59
*Nome da movimentação processual.....	59
Descrição completa da movimentação processual .....	59
*Data da movimentação processual .....	59
<b>Estudo dos campos do PJE com os campos do MOREQ-JUS .....</b>	<b>60</b>
Nível de Sigilo .....	60
Usuário do cadastro do processo .....	60

---

Nome do usuário do protocolo .....	60
Processo Digitalizado .....	60
Hash .....	61
Mimetype .....	61
Assinatura .....	61
Usuário da assinatura .....	61
Algoritmo hash.....	61
Cadeia de certificado.....	62
Codificação de certificado .....	62
Data assinatura .....	62
Signatário login .....	62
Assinatura válida .....	62
<b>Apêndice B - Metadados Descritivos Dublin Core.....</b>	<b>63</b>
<b>Apêndice C - Metadados Descritivos ISAD(G) .....</b>	<b>66</b>
<b>Apêndice D - Outros Metadados ISAD(G).....</b>	<b>69</b>
<b>Apêndice E - Avaliação das políticas de segurança da informação do Tribunal e o modelo Hipátia .....</b>	<b>73</b>
Política Corporativa de Segurança da Informação .....	73
RESOLUÇÃO 21 DE 8 DE NOVEMBRO DE 2016 .....	73
Avaliação em relação ao Guia de Boas Práticas.....	73
Itens em conformidade com a política de segurança .....	79
Itens em estudo para aplicação das conformidades .....	80
<b>Considerações finais .....</b>	<b>81</b>

---

# 1. INTRODUÇÃO

O Tribunal de Justiça do Distrito Federal e Territórios (TJDFT) estabeleceu diretrizes internas para orientar a gestão documental. Assim, o órgão avançou ao estabelecer diretrizes internas capazes de orientar a gestão e preservação de documentos no formato digital em consonância com as tendências nacionais e internacionais da área. A resolução 16, de 25 de agosto de 2016, permitiu a gestão documental dos processos em meio analógico.

O sucesso dessa resolução demanda avanços no processo de modernização da gestão documental já regulamentado no âmbito da política de gestão documental. Nesse contexto, documentos em formato digital são incorporados ao processo e ampliam o alcance da gestão documental. O tema está no âmbito da política de gestão documental do Poder Judiciário, fruto de discussões ocorridas em meio ao Programa Nacional de Gestão Documental e Memória do Poder Judiciário (PRONAME). Ao reconhecer esses procedimentos técnicos preconizados pela Ciência da Arquivologia, o Conselho Nacional de Justiça (CNJ) editou a Recomendação 37/2011 veiculando os requisitos e instrumentos necessários à gestão documental, independente do suporte em que se encontre a informação.

A parceria do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e TJDFT promove soluções viáveis de implantação de um Repositório Arquivístico Digital Confiável (RDC-Arq) e identifica a necessidade de aprimoramento do modelo de RDC-Arq com a incorporação de outros sistemas. Esse projeto ainda está em curso e incorpora diversos avanços. Todavia, dada a natureza da pesquisa e o ineditismo desse projeto, novas ações foram identificadas durante a evolução da pesquisa. Essas necessidades estão expressas neste projeto de pesquisa que aborda a Preservação digital e gestão arquivística apoiada no aprimoramento da implantação do modelo RDC-Arq.

Este relatório tem o objetivo de apresentar os principais resultados e entregas de 2021 do projeto de pesquisa para a implantação do RDC-Arq no contexto de preservação de processos judiciais eletrônicos.

Este relatório está estruturado em 5 partes. Um breve histórico da necessidade e consequente definição da elaboração do projeto está descrito na **Introdução**. Os objetivos geral e específicos são abordados em **Objetivos**. **Contextualização e Resultados** mostra as tratativas entre equipes técnicas em **Relatórios e interação das equipes**. Nesta mesma seção, os estudos que fundamentam a escolha de ferramentas e de procedimentos metodológicos estão em Metas 1, 2 e 3. O processo sistemático para disseminação e apropriação do conhecimento é tratado em **Transferência de conhecimento**. Por fim, a avaliação da condução do projeto destacando as entregas e apontando os próximos passos estão em **Considerações Finais**.

## 2. OBJETIVOS

### 2.1 Objetivo Geral

Aprimorar a implementação do modelo RDC-Arq na estrutura do TJDFT.

### 2.2 Objetivos Específicos

- Melhorar o desempenho e infraestrutura do barramento RDC-Arq;
- Disponibilizar novas formas de acesso ao RDC-Arq;
- Possibilitar integração com o Processo Judicial Eletrônico - PJe;
- Desenvolver módulo de interface para o AtoM;
- Provar modelos propostos; e
- Disseminar conhecimento.

## 3. CONTEXTUALIZAÇÃO E RESULTADOS

### 3.1 Relatórios e interação das equipes

#### 3.1.1 Reuniões técnicas

As reuniões técnicas nortearam o desenvolvimento das atividades durante o ano de 2021.

DATA	ATIVIDADES
10/02/2021	Reunião Técnica - Oficina de alinhamento RDC-Arq PJe (Lean Inception)
24/02/2021	Reunião Técnica - Definições de infraestrutura e arquitetura da solução RDC-ARQ PJE
03/03/2021	Reunião Técnica - Apresentação dos ambientes de desenvolvimento do TJDFT
14/05/2021	Reunião Técnica - RDC-ARQ - preservação de processos arquivados digitalizados que não estão no PJE
17/05/2021	Reunião Técnica - Revisão final dos metadados do PJE
20/05/2021	Reunião Técnica - Interface PJE Arq - Barramento Hipátia (RDC - Arq)
21/05/2021	Reunião Técnica - Finalização metadados PJE RDC-ARQ
17/06/2021	Reunião Técnica - Estudos sobre os metadados para o AtoM
21/06/2021	Reunião Técnica - Arquitetura de integração PJE-Arq e Hipátia
30/07/2021	Reunião Técnica - Infraestrutura Kafka para o RDC-ARQ
23/09/2021	Reunião Técnica - Integração PJE - RDC
14/10/2021	Reunião Técnica - Teste piloto da integração PJE - RDC - Arquitetura baseada em eventos + API + MNI
04/11/2021	Reunião Técnica - Alinhar sobre a abordagem de preservação digital e validação de metadados
04/11/2021	Reunião Técnica - Apresentação dos estudos realizados pelo IBICT - Ambientes Archivematica (guarda longa e permanente)
08/11/2021	Reunião Técnica - Acesso à API do PJE para extração de metadados
09/12/2021	Apresentação dos resultados de 2021 do projeto RDC-ARQ PJe
14/12/2021	Reunião Técnica - Alinhamento técnico para teste de integração do RDC-Arq (PJE-Arq, Hipátia e PJE)

### 3.1.2 Reuniões de alinhamento

Em fevereiro de 2021, foi realizado um workshop colaborativo (*Lean Inception*) onde conseguimos dividir as atividades de trabalho em partes menores e fazer estimativas de entregas, otimizar o trabalho e ordenar os itens em uma lista de atividades (*backlog*). Utilizando conceitos da metodologia Ágil, a dinâmica de interação e atualização das informações do projeto entre as equipes se deu em reuniões de ponto de controle, que possibilita o acompanhamento do projeto por meio de artefato conhecido como *roadmap*. Essa ferramenta permite monitorar os MVPs (Produto mínimo viável) ou etapas do projeto e é revisada periodicamente para adequação às necessidades do projeto. A execução, cadência, priorização, organização e transparência das atividades são gerenciadas pelo sistema/conceito chamado Kanban.

As reuniões de ponto de controle de acompanhamento das atividades do projeto de pesquisa ocorrem a cada 15 dias com a participação das equipes técnicas e de negócio do TJDFT e IBICT.

DATA	TIPO	ATIVIDADES
29/01/2021	Ponto de controle RDC-Arq PJE	<ul style="list-style-type: none"><li>• Reunião para refinar e priorizar o backlog;</li><li>• Definir meta;</li><li>• Identificar dependências;</li></ul>
10/02/2021	Lean Inception: Reunião para estruturar os “entregáveis”	<ul style="list-style-type: none"><li>• Estruturar os entregáveis (MVP) em um roadmap, priorizados pelo maior valor</li><li>• Estruturar o fluxo do projeto</li></ul>
08/03/2021	Reunião de planejamento do próximo ciclo - 8/3 a 26/3	<ul style="list-style-type: none"><li>• Reunião para refinar e priorizar o backlog;</li><li>• Definir meta;</li><li>• Organizar o quadro kanban;</li><li>• Definir atribuições;</li><li>• Identificar dependências;</li><li>• Kanban meeting; WIP;</li><li>• Atualização do roadmap ao final do ciclo. (verificar tarefas necessárias para entrega de valor mínimo esperado em outubro de 2021 - meta 4)</li></ul>
29/03/2021	Reunião de planejamento do próximo ciclo - 29/3 a 16/4	<ul style="list-style-type: none"><li>• Reunião para refinar e priorizar o backlog;</li><li>• Definir meta do ciclo;</li><li>• Organizar o quadro kanban;</li><li>• Definir atribuições;</li><li>• Identificar dependências e impedimentos</li></ul>
14/04/2021	Ponto de Controle RDC-ARQ PJE	<ul style="list-style-type: none"><li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li><li>• Revisão do roadmap.</li></ul>
30/04/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"><li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li><li>• Revisão do roadmap.</li><li>• Avaliação dos avanços com o Hipátia, definição dos metadados e implantação da solução em container.</li></ul>

10/05/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> <li>• Revisão do roadmap com a possibilidade de paralelismo implantação do projeto em container e adaptação do Hipátia para on-demand.</li> </ul>
24/05/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo. Revisão do roadmap com a possibilidade de paralelismo implantação do projeto em container e adaptação do Hipátia para on-demand.</li> </ul>
11/06/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
09/07/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> <li>• Revisão do roadmap com a possibilidade de paralelismo implantação do projeto em container e adaptação do Hipátia para on-demand</li> </ul>
23/07/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
06/08/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
20/08/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
03/09/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> <li>• Revisão do roadmap;</li> </ul>
01/10/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
14/10/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> <li>• Atualização do roadmap.</li> </ul>
28/10/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> <li>• Atualização do roadmap.</li> <li>• Reunião de Retrospectiva</li> </ul>
12/11/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
26/11/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>
10/12/2021	Ponto de controle RDC ARQ PJE	<ul style="list-style-type: none"> <li>• Avaliação do progresso do último ciclo e planejamento para o próximo.</li> </ul>

---

## 3.2 META 1: Estudos de infraestrutura para solução RDC-Arq

A meta 1 é dedicada ao levantamento bibliográfico de propostas de modelos de infraestruturas associados a projetos similares, bem como o estudo de infraestruturas implementadas em instituições públicas nacionais e internacionais. Além disso, é destinada à análise de requisitos de desempenho e confiabilidade necessários ao ambiente para guarda de informações relacionadas a processos judiciais e ao estudo para estabelecer a infraestrutura de preservação, regras e políticas de gerenciamento do acervo para a preservação em fase de guarda intermediária, como continuidade do tratamento arquivístico dos processos judiciais eletrônicos. Todo esse processo investigativo resulta na proposição do modelo de arquitetura computacional ideal para projetos de barramento RDC-Arq que considera aspectos de segregação de ambientes.

### 3.2.1 Estudos de infraestrutura para solução RDCArq

O ambiente computacional do TJDFT utiliza a abordagem de DevOps (junção das palavras *development* e *operations*), que encurta o ciclo de vida de desenvolvimento de sistemas e fornece a entrega contínua com alta qualidade de software. Facilita a integração entre os times de desenvolvimento e infraestrutura, a gerência de ambientes e sistemas, além de promover maior aderência ao ambiente de execução das aplicações desde o início do desenvolvimento.

O TJDFT utiliza primariamente duas ferramentas para implementação do modelo DevOps: *Openshift* e *GitLab*, que, por sua vez, empregam Kubernetes e *Docker*. Os relacionamentos entre as ferramentas utilizadas nesse modelo DevOps seguem as seguintes etapas:

1. Desenvolvedor realiza um push no repositório no *GitLab* do TJDFT;
2. O pipeline de CI/CD é ativado no *GitLab*;
3. A leitura das diferentes aplicações descritas nos DockerFiles é realizada;
4. As imagens *Dockers* são construídas;
5. As imagens são enviadas para o *Registry* do TJDFT;
6. O *Openshift* busca as imagens no *Registry* e as utiliza no Kubernetes para criar as aplicações.

#### GITLAB

O *GitLab* é um gerenciador de repositório de software baseado em *git*, que permite o versionamento de código, com suporte a Wiki, gerenciamento de tarefas e *Continuous Integration* (CI) /*Continuous Delivery* (CD), muito comumente chamadas apenas de CI/CD, que são ferramentas de integração e entrega contínua, essenciais ao modelo DevOps. No TJDFT, o *Openshift* não faz o build (geração do código executável das aplicações) de projetos por questões de segurança, repassando esse trabalho para a ferramenta de entrega contínua do *GitLab*, a partir do código dos projetos versionados na plataforma.

#### DOCKER

O *Docker* é uma plataforma *open source* que facilita a criação e administração de ambientes isolados. Com ele, é possível empacotar uma aplicação ou ambiente dentro de um container, como se fossem máquinas virtuais modulares e extremamente leves. Em conjunto com o Kubernetes, o *Docker* ainda realiza as mesmas tarefas do seu objetivo original, a diferença é que o sistema automatizado solicita que o *Docker* realize essas tarefas, em vez do administrador fazer essas solicitações manualmente.

#### KUBERNETES

O Kubernetes é uma plataforma *open source* que automatiza as operações dos containers Linux. Ele possibilita a criação e gerenciamento de um cluster de containers em clouds privadas ou públicas, eliminando grande parte dos processos manuais necessários para implantar e escalar as aplicações em containers. Dessa forma, o gerenciamento dos clusters é feito com facilidade e efetividade.

## OPENSIFT

O *Openshift* é uma plataforma de containers Kubernetes para empresas com operações automatizadas em todo o conjunto de sistemas. Com ele é possível criar, desenvolver e implantar aplicações de forma simples e rápida, em qualquer infraestrutura. A plataforma inclui um sistema operacional Linux para empresas, ambiente de execução de container, rede, monitoramento, registro e soluções de autorização e autenticação. Assim, a integração de arquitetura, processos, plataformas e serviços necessários para impulsionar o trabalho das equipes de desenvolvimento e de operações é feita de forma aliada e harmoniosa.

## O SOFTWARE ARCHIVEMATICA

O *Archivematica* é um sistema para a preservação digital gratuito e de código aberto. Ele foi projetado para manter os dados baseados em padrões de preservação digital e o acesso a longo prazo para coleções de objetos digitais. O software fornece um conjunto integrado de ferramentas livres que permite aos usuários processar objetos digitais para o armazenamento de arquivos em conformidade com o modelo funcional ISO-OAIS. O código do *Archivematica* é liberado sob a licença GNU *Affero General Public License* e sua documentação é liberada sob uma licença *Creative Commons*.

O software *Archivematica* consiste na implantação de diversas aplicações que trabalham em conjunto na preservação dos arquivos. São eles:

- *MCPServer* (servidor): é um gerenciador de tarefas que usa o *gearman* e atribui tarefas ao serviço de *MCPCClient* e determina o fluxo de trabalho.
- *MCPCClient* (cliente): é um cliente que utiliza o *gearman* para despachar os trabalhos dos vários scripts.
- *Dashboard* (GUI): é uma interface gráfica web voltado para o usuário escrita em Django. Esse serviço lida com sistemas externos e configurações de fluxo de trabalho, status e decisões de processos de transferência e SIP e dá acesso ao backlog e armazenamento de arquivos.
- *Storage Service* (armazenamento): serviço responsável por mover arquivos para o *Archivematica* para processamento e depois da operação, movê-los de volta para o armazenamento.
- *Clamavd*: é um kit de ferramentas de software antivírus gratuito, de plataforma cruzada e de código aberto, capaz de detectar muitos tipos de software malicioso, incluindo vírus. Um de seus principais usos é em servidores de e-mail como verificador de vírus de e-mail no servidor.
- *Gearmand*: é uma estrutura de aplicativo de código aberto projetada para distribuir para outras máquinas ou processos que são mais adequados para fazer o trabalho. Ele permite que você trabalhe em paralelo, para balancear a carga de processamento e para chamar funções entre linguagens. ([gearman.org](http://gearman.org)).
- *Mysql*: é um sistema de gerenciamento de banco de dados relacional de código aberto.
- *Redis*: é um armazenamento de estrutura de dados na memória, usado como um banco de dados de chave-valor distribuído na memória, cache e agente de mensagens, com durabilidade opcional.
- *Nginx*: é um servidor web que também pode ser usado como *proxy* reverso, balanceador de carga, *proxy* de e-mail e cache HTTP. Gratuito e de código aberto, lançado sob os termos da licença BSD de 2 cláusulas, uma grande fração dos servidores web o usa frequentemente como um balanceador de carga.
- *Elasticsearch*: serviço que fornece um mecanismo de pesquisa de texto completo distribuído e com capacidade para vários locatários, com uma interface da web HTTP e documentos JSON sem esquema.
- *Fits*: é usado para caracterizar arquivos que não têm um formato de arquivo reconhecido.

*Archivematica* é um projeto ativo e dinâmico de código aberto com uma grande comunidade de usuários. Os desenvolvedores estão em constante atividade na melhora do sistema. Atualmente o software se encontra na versão v1.13.0.

Neste projeto foi utilizado a versão v1.12.0 por se tratar de uma versão mais estável e amplamente utilizada pelas principais organizações

### 3.2.2 Processo de implantação do sistema no Openshift do TJDFT

A implementação do software Archivematica em ambiente de contêineres começa com a criação de um repositório de versionamento de código no *GitLab*. Em seguida, é necessário a elaboração de uma imagem *docker*, representada por um arquivo de configuração chamado *DockerFile*. Neste arquivo são definidas as especificações da imagem, isto é, como ela deverá ser construída e comandos necessários para sua correta configuração.

O arquivo de configuração define as regras de implementação que o *Openshift* envia para o *Kubernetes*, que implementa a instância, denominada de *POD*, que executará o sistema em container. Um *POD* é a menor unidade de aplicação *Kubernetes*, podendo ser composto por um ou mais containers fortemente acoplados, de forma a compartilhar acessos.

A partir da implementação em *Docker Compose* descrita no *Github* oficial do software Archivematica, foi desenvolvida a containerização para o ambiente do TJDFT seguindo as orientações pré-definidas pelo Tribunal.

As etapas da implementação consistem na criação de volumes (*Persistent Volume Claim - PVC*) no *Openshift*, na criação dos arquivos necessários para integração contínua dos serviços utilizados pelo Archivematica, no *deployment* de cada serviço no *Openshift*, na inicialização e preparação dos serviços e, por fim, na criação das rotas no *Openshift*.

Todo o projeto encontra-se no repositório de códigos do Tribunal, acessível pela rede interna, disponível no endereço: <https://GitLab.tjdft.jus.br/docker-images/archivematica-ibict>. A seguir descrevem-se as etapas da implementação em detalhes.

#### 3.2.2.1 CRIAÇÃO DOS VOLUMES (PVC)

A primeira etapa a ser realizada para implementação do Archivematica é a criação de cinco volumes (*PVC*) no *Openshift* que são utilizados pelas aplicações: *Storage Service*, *Elasticsearch*, *Fits*, *Clamav*, *MCPServer*, *MCPClient*, *Dashboard* e *MySQL*. O Quadro 1 apresenta os volumes criados no projeto Archivematica - Stage no *Openshift* com seus respectivos nomes, capacidades e aplicações associadas.

Quadro 1 - Listagem dos volumes

Nome do PVC	Capacidade	Aplicação associada
<i>Archivematica-storage-service-location-data</i>	2GB	<i>Storage Service</i>
<i>Archivematica-storage-service-location-data</i>	2GB	<i>Storage Service</i>
<i>elasticsearch-data</i>	1GB	<i>Elasticsearch</i>
<i>Archivematica-pipeline-data</i>	2GB	<i>Fits, Clamav, MCPServer, MCPClient, Dashboard e Storage Service</i>
<i>mysql-data</i>	1GB	<i>MySQL</i>

Fonte: Elaboração dos autores (2021)

A classe de serviço utilizada para o volume ***elasticsearch-data*** foi a ***production-iscsi***, por orientação da equipe de infraestrutura do TJDFT. As demais utilizam a classe padrão *stage*.

#### 3.2.2.2 CRIAÇÃO E DEPLOYMENT DOS SERVIÇOS E APLICAÇÕES

Esta etapa consiste na criação dos conjuntos de dados necessários para integração contínua no *GitLab* que são: (i) *DockerFile* e respectivos arquivos da imagem; e (ii) a descrição da imagem no arquivo ***gitlab-ci.yml***. As subseções a seguir retratam os serviços e aplicações utilizadas.

---

### 3.2.2.3 MYSQL

Os arquivos dessa aplicação estão disponíveis no endereço: <https://gitlab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/mysql-am>. A descrição da imagem no `gitlab-ci.yml` foi a seguinte:

```
variables:
  MYSQL_IMAGE_NAME: archivematica-stage/mysql-am

deploy_mysql_desenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $SA_USER -p $SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -t $REGISTRY_ADDR/$MYSQL_IMAGE_NAME:$IMAGE_TAG
      ./mysql-am/
    - docker push $REGISTRY_ADDR/$MYSQL_IMAGE_NAME:$IMAGE_TAG
```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome “**mysql-am**” e associado a seu respectivo PVC.

### 3.2.2.4 ELASTICSEARCH

O *deployment* dessa aplicação foi realizado a partir da imagem no *DockerHub*, sem necessitar do registro de uma nova imagem no *Registry* do TJDFT. Dessa forma, a imagem utilizada foi:

**Elasticsearch 6.5.4:docker.elastic.co/elasticsearch/elasticsearch:6.5.4.**

Também foi necessária a configuração de limites de memória para essa aplicação. Na edição de limites de recurso da aplicação no *Openshift*, foi definido o mínimo garantido de 1Gb e um máximo de 2Gb.

Por fim, foi realizado o *deployment* da imagem no *Openshift* com o nome **elasticsearch-am** e associado a seu respectivo PVC.

### 3.2.2.5 REDIS

O **deployment** dessa aplicação foi realizado a partir da imagem no *DockerHub*, sem necessitar do registro de uma nova imagem no *Registry* do TJDFT. Dessa forma, a imagem utilizada foi a **Redis 3.2: redis:3.2-alpine**. Foi realizado o **deployment** da imagem no *Openshift* com o nome “**redis-am**”. Nenhuma configuração adicional foi necessária.

### 3.2.2.6 GEARMAND

Os arquivos dessa aplicação estão disponíveis no endereço:

<https://GitLab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/gearmand-a m>.

A descrição da imagem no `Gitlab-ci.yml` foi a seguinte:

```

variables:
  GEARMAND_IMAGE_NAME: archivematica-stage/gearmand-am

deploy_gearmand_desenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $SA_USER -p $SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -t
    $REGISTRY_ADDR/$GEARMAND_IMAGE_NAME:$IMAGE_TAG ./gearmand-am/
    - docker push $REGISTRY_ADDR/$GEARMAND_IMAGE_NAME:$IMAGE_TAG

```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome “**gearmand-am**”. Nenhuma configuração adicional foi necessária.

### 3.2.2.7 FITS

O **deployment** dessa aplicação foi realizado a partir da imagem no *DockerHub*, sem necessitar do registro de uma nova imagem no *Registry* do TJDFT. Dessa forma, a imagem utilizada foi a **Fits 0.8.4** da **Artefactual: artefactual/fits-ngserver:0.8.4**.

Foi realizado o **deployment** da imagem no *Openshift* com o nome “**fits-ngserver-am**” e associado a seu respectivo PVC.

### 3.2.2.8 CLAMAV

O **deployment** dessa aplicação foi realizado a partir da imagem no *DockerHub*, sem necessitar do registro de uma nova imagem no *Registry* do TJDFT. Dessa forma, a imagem utilizada foi *Clamav* mais atual da **Artefactual: artefactual/clamav:latest**. Ademais, foram definidas algumas variáveis de ambiente descritas no Quadro 2.

Quadro 2 - Variáveis de ambiente do Clamav

Nome da variável de ambiente	Valor
CLAMAVA_MAX_FILE_SIZE	42
CLAMAV_MAX_SCAN_SIZE	42
CLAMAV_MAX_STREAM_LENGTH	100

Fonte: Elaboração dos autores (2021)

Por fim, foi realizado o *deployment* da imagem no *Openshift* com o nome “**clamav-am**” e associado a seu respectivo PVC.

### 3.2.2.9 NGINX

Os arquivos dessa aplicação estão disponíveis no endereço: <https://gitlab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/nginx-am>. A descrição da imagem no **gitlab-ci.yml** foi a seguinte:

```

    NGINX_IMAGE_NAME: archivematica-stage/nginx-am

deploynginxdesenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $SA_USER -p $SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -t $REGISTRY_ADDR/$NGINX_IMAGE_NAME:$IMAGE_TAG
      ./nginx-am/
    - docker push $REGISTRY_ADDR/$NGINX_IMAGE_NAME:$IMAGE_TAG

```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome “**nginx-am**”. Nenhuma configuração adicional foi necessária.

**Observação:** o *Openshift* do TJDFT já possui um balanceador de carga para suas aplicações, de forma que não é necessário a utilização do serviço do *nginx* no projeto. Contudo, este produto descreveu a implementação da imagem do serviço seguindo a documentação da implementação em *Docker Compose* descrita no *github* oficial.

### 3.2.2.10 MCP SERVER

Os arquivos dessa aplicação estão disponíveis no endereço:

<https://gitlab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/archivematica-mcp-server>

A descrição da imagem no ***Gitlab-ci.yml*** foi a seguinte:

```

variables:
  AM_MCP_SERVER_IMAGE_NAME: archivematica-stage/archivematica-mcp-server

deployarchivematicamcpserverdesenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $SA_USER -p $SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -f archivematica-mcp-server/Dockerfile -t
      $REGISTRY_ADDR/$AM_MCP_SERVER_IMAGE_NAME:$IMAGE_TAG .
    - docker push $REGISTRY_ADDR/$AM_MCP_SERVER_IMAGE_NAME:$IMAGE_TAG

```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome **archivematica-mcp-server** e associado a seu respectivo PVC.

Atenção: as variáveis de ambiente dessa aplicação foram definidas no *DockerFile*.

---

### 3.2.2.11 MCPCLIENT

Os arquivos dessa aplicação estão disponíveis no endereço: <https://GitLab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/archivematica-mcp-client>. A descrição da imagem no `gitlab-ci.yml` foi a seguinte:

```
variables:
  AM_MCP_CLIENT_IMAGE_NAME: archivematica-stage/archivematica-mcp-client

deploy_archivematica-mcp-client_desenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $$SA_USER -p $$SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -f archivematica-mcp-client/Dockerfile -t
      $REGISTRY_ADDR/$AM_MCP_CLIENT_IMAGE_NAME:$IMAGE_TAG .
    - docker push $REGISTRY_ADDR/$AM_MCP_CLIENT_IMAGE_NAME:$IMAGE_TAG
```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome “**archivematica-mcp-client**” e associado a seu respectivo PVC.

Atenção: as variáveis de ambiente dessa aplicação foram definidas no *DockerFile*.

### 3.2.2.12 DASHBOARD

Os arquivos dessa aplicação estão disponíveis no endereço: <https://GitLab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/archivematica-dash-board>. A descrição da imagem no `gitlab-ci.yml` foi a seguinte:

```
deploy_archivematica-dash-board_desenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $$SA_USER -p $$SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -f archivematica-dash-board/Dockerfile -t
      $REGISTRY_ADDR/$AM_DASHBOARD_IMAGE_NAME:$IMAGE_TAG .
    - docker push $REGISTRY_ADDR/$AM_DASHBOARD_IMAGE_NAME:$IMAGE_TAG
```

Em seguida, foi realizado o **deployment** da imagem no *Openshift* com o nome “**archivematica-dash-am**” e associado a seu respectivo PVC.

Atenção: as variáveis de ambiente dessa aplicação foram definidas no *DockerFile*.

### 3.2.2.13 STORAGE SERVICE

Os arquivos dessa aplicação estão disponíveis no endereço: <https://gitlab.tjdft.jus.br/docker-images/archivematica-ibict/-/tree/master/archivematica-storage-service>.

A descrição da imagem no **gitlab-ci.yml** foi a seguinte:

```
variables:
  AM_STORAGE_IMAGE_NAME: archivematica-stage/archivematica-storage-service

deploy_archivematica-storage-service_desenv:
  stage: deploy
  image: docker:latest
  only:
    - master
  variables:
    IMAGE_TAG: desenv
  script:
    - docker login -u $SA_USER -p $SA_USER_TOKEN $REGISTRY_ADDR
    - docker build --pull -t
    $REGISTRY_ADDR/$AM_STORAGE_IMAGE_NAME:$IMAGE_TAG
    ./archivematica-storage-service/
    - docker push $REGISTRY_ADDR/$AM_STORAGE_IMAGE_NAME:$IMAGE_TAG
```

Em seguida, foi realizado o *deployment* da imagem no *Openshift* com o nome **archivematica-storage-se** e associado aos seus respectivos PVCs.

Atenção: as variáveis de ambiente dessa aplicação foram definidas no *DockerFile*.

### 3.2.2.14 INICIALIZAÇÃO E PREPARAÇÃO DOS SERVIÇOS

Nessa etapa foi realizada a criação do banco de dados (*database*) e a criação das tabelas na aplicação *MySQL*. Em seguida, deve-se reinicializar as aplicações *MCPServer*, *MCPClient*, *Dashboard* e *Storage Service*.

### 3.2.2.15 CRIAÇÃO DOS BANCOS DE DADOS

Para criar os bancos de dados, deve-se entrar na aplicação *MySQL* e executar os seguintes comandos:

```
sh-4.2$ bash
bash-4.2$ mysql -u root -p12345

mysql> DROP DATABASE IF EXISTS SS;
mysql> CREATE DATABASE SS;
mysql> GRANT ALL ON SS.* TO 'archivematica'@'%' IDENTIFIED BY 'demo';"

mysql> DROP DATABASE IF EXISTS MCP;
mysql> CREATE DATABASE MCP;
mysql> GRANT ALL ON MCP.* TO 'archivematica'@'%' IDENTIFIED BY 'demo';"
```

### 3.2.2.16 CRIAÇÃO DAS TABELAS

Para criar as tabelas nos respectivos bancos de dados, deve-se entrar nas aplicações *Dashboard* e *Storage Service* (armazenamento) e executar um *script python*.

Contudo, para executar o *script python*, foi necessário modificar o *entrypoint* de cada um dos serviços (*Dashboard* e *Storage Service*) para que o *POD* inicie sem erros no *Openshift*. Dessa forma, o *entrypoint* padrão foi comentado no *DockerFile* e em seu lugar, foi usado o seguinte *entrypoint*:

```
CMD ["tail", "-f", "/dev/null"]
```

Uma vez modificados os *entrypoints* e feito o *redploy* das aplicações, é possível acessar cada aplicação e executar os *scripts python* necessários para a criação das tabelas nos respectivos bancos de dados.

---

Para a aplicação do *Storage Service*, são executados os seguintes comandos:

```
$ /src/storage_service/manage.py migrate --noinput
$ /src/storage_service/manage.py create_user --username="test"
--password="test" --email="test@test.com" --api-key="test" --superuser
```

Para a aplicação do *Dashboard*, são executados os seguintes comandos:

```
$ /src/storage_service/manage.py migrate --noinput
$ /src/dashboard/src/manage.py install --username="test"
--password="test" --email="test@test.com" --org-name="test"
--org-id="test" --api-key="test"
--ss-url="http://archivematica-storage-se:8000" --ss-user="test"
--ss-api-key="test" --site-url="http://archivematica-dash-am:8000"
$ cd /src/dashboard/frontend
$ npm install --unsafe-perm
```

Após a execução dos comandos, redefine-se cada *entrypoint* para o original e é realizado o *redeploy* da aplicação. Neste momento, o software do Archivematica e todos seus serviços estão operando como o esperado.

### 3.2.2.17 CRIAÇÃO DAS ROTAS

Para disponibilizar o sistema na rede interna do TJDFT, é necessário configurar as rotas dentro do *Openshift*. As orientações da equipe de infraestrutura do Tribunal estão no seguinte endereço:

<https://gitlab.tjdft.jus.br/infraestrutura/openshift/integracao/dns>.

Deve-se criar uma rota segura para a aplicação, que usa TLS/HTTPS. Uma rota é uma referência ao serviço do POD, que deve expor uma determinada porta. Nas configurações do *Openshift* os tipos de *Termination* de uma rota atualmente são: *Edge*, *Reencrypt* ou *Passthrough*. A recomendação da equipe de infraestrutura do Tribunal foi para usar a terminação *Edge*.

Uma rota segura é publicada com HTTPS (porta 443) ou possui redirecionamento das requisições de HTTP (porta 80) para HTTPS (porta 443), como mostra o vídeo no endereço (acessado na rede interna):

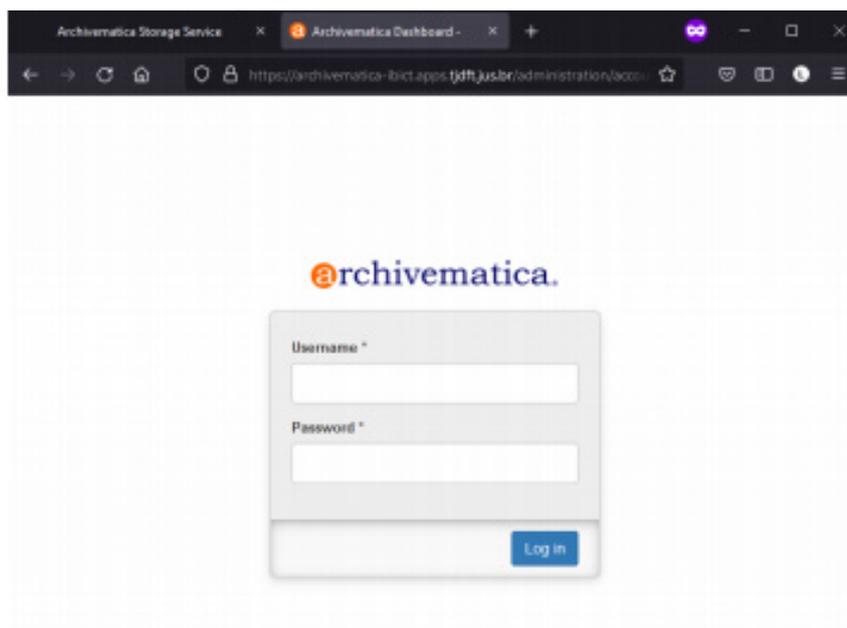
<https://drive.tjdft.jus.br/index.php/s/D7dYX65pNi3ggAs>.

Dessa forma, foram criadas duas rotas para os serviços:

- *Dashboard*: <https://archivematica-ibict.apps.tjdft.jus.br>
- *Storage*: <https://archivematica-storage-ibict.apps.tjdft.jus.br>

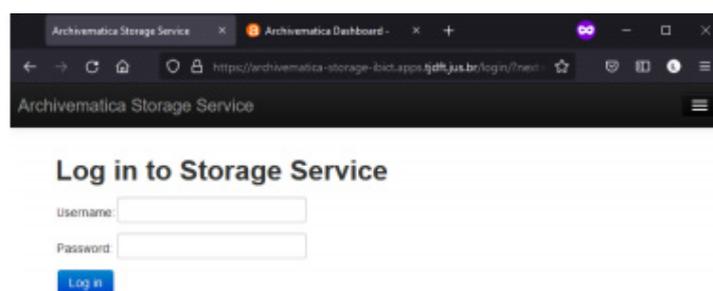
As Figuras 1 e 2 apresentam respectivamente os serviços do *Dashboard* e do *Storage* sendo executados com sucesso.

Figura 1 - Serviço de Dashboard



Fonte: Captura de tela da página web do Archivematica (2021).

Figura 2 - Serviço de Storage



Fonte: Captura de tela da página web do Archivematica (2021)

### 3.3 Definições para alinhamento com as políticas de Segurança da Informação do Tribunal

Foi realizado estudo com o objetivo de avaliar o atual estágio de desenvolvimento do barramento tecnológico, integrante do modelo Hipátia de preservação digital, frente ao Guia de Boas Práticas de Desenvolvimento, Infraestrutura e Segurança das Aplicações do TJDFE e a Política Corporativa de Segurança da Informação baseada na RESOLUÇÃO 21 DE 8 DE NOVEMBRO DE 2016. O estudo completo foi anexado no Apêndice E deste relatório.

A premissa é seguir as recomendações e melhores práticas definidas pelo Guia em relação aos requisitos de segurança da informação, arquitetura, experiência dos usuários, ferramentas, infraestrutura e operação, obedecendo as melhores práticas de mercado, antes e durante a implantação, o desenvolvimento e a manutenção das aplicações.

O Ibict considera de extrema importância padronizar os procedimentos de desenvolvimento da ferramenta em relação às recomendações e boas práticas do TJDFE para implantação, manutenção, sustentação e eventual desligamento, de forma a garantir isonomia nos princípios de segurança de sistemas durante todo o ciclo de vida da aplicação.

---

## 3.4 Estudo de múltiplos Pipelines nos ambientes do Archivematica para preservação de documentos de guarda longa e guarda permanente do TJDFT

Diante da necessidade do Tribunal de Justiça do Distrito Federal e Territórios – TJDFT em utilizar dois Archivematicas com regras de preservação distintas para preservar os documentos de guarda permanente e os documentos de guarda longa, foi realizado um estudo que analisou a possibilidade de se operar múltiplos pipelines ao invés de múltiplos Archivematicas no ambiente tecnológico.

Em um primeiro momento, após analisar casos aplicados em outros países, alguns problemas de atualização de versões, falta de documentação e de assistência da Artefactual, empresa responsável pelo software, foram identificados. Contudo, posteriormente, foi realizado um diálogo mais aprofundado com a comunidade no exterior, mostrando que existe a possibilidade desta configuração em ambiente de produção. As discussões foram apresentadas em uma reunião para a equipe do TJDFT.

Destaca-se assim que o Archivematica é composto por duas interfaces que são instaladas e interoperam concomitantemente. A interface Dashboard é a interface de gestão documental onde os pacotes de preservação são submetidos e processados. Já a interface *Storage Service* fica sob a responsabilidade dos gestores de preservação e de Tecnologia da Informação, por ser um local de acesso direto aos pacotes preservados e também sob a responsabilidade do profissional de tecnologia da informação, já que o *Storage Service* apresenta as configurações técnicas do sistema do Archivematica.

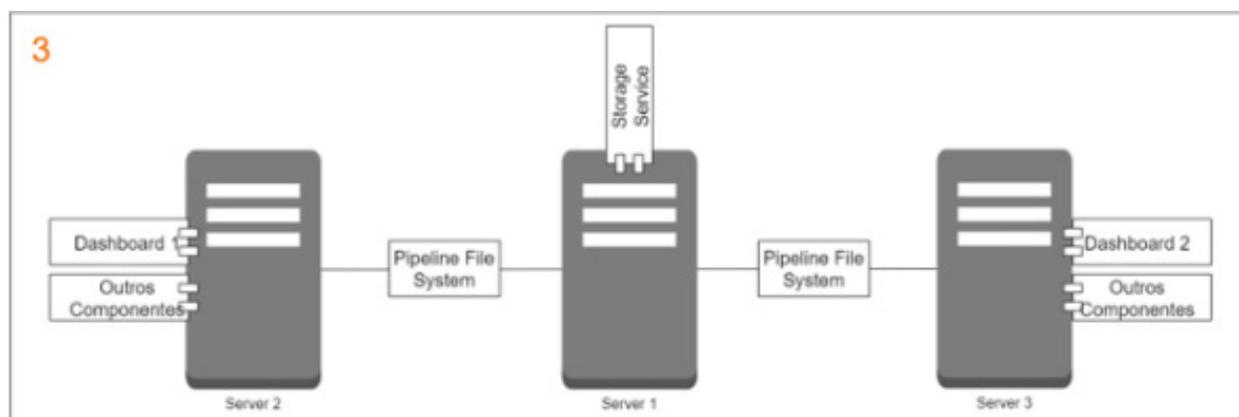
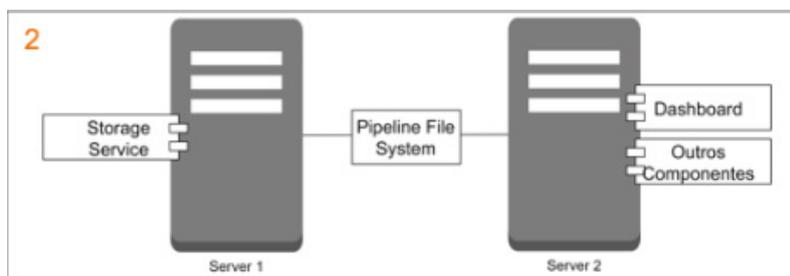
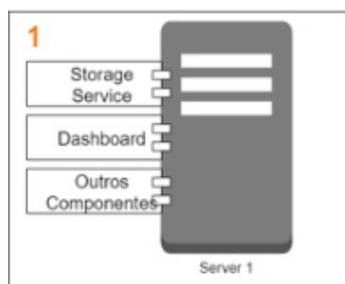
No *storage service* encontra-se as configurações dos *spaces*, que são os protocolos de conexão que serão usados entre o *dashboard* e um *storage*. Cada pipeline representa uma interface *dashboard*, que um *Storage Service* pode trabalhar e operar com múltiplos pipelines que vão ter protocolos específicos conforme sua funcionalidade.

De maneira geral, os múltiplos pipelines podem ser utilizados para restringir o acesso a dados sigilosos por pessoas que estariam utilizando o Dashboard, mas não tem permissão para visualizar tal conteúdo conforme seja necessário. Os múltiplos pipelines também podem ser utilizados quando há planos de preservação diferenciados para documentos que possuem destinações finais diferentes, por exemplo, como é o caso do TJDFT.

Desse modo, o acesso a um único *Storage Service* por administrador independentemente da quantidade de dashboards, e a possibilidade de ter uma única interface de configuração do *Storage Service*, facilitando o acesso e a administração do sistema são algumas das vantagens da utilização de múltiplos pipelines. Através dessa configuração, também é possível restringir acessos a dados sigilosos, mover os pacotes entre os diferentes pipelines sem a perda da rastreabilidade e operar tendo diferentes configurações de regras e comandos do plano de preservação.

Por conseguinte, a desvantagem a princípio envolveria apenas os estudos, as implicações e a aplicabilidade dos múltiplos pipelines em ambiente de *Docker*, já que é o tipo de sistema computacional utilizado pelo TJDFT. Na instalação em *Virtual Machine*, foi necessário separar o componente *Storage Service* dos demais componentes do Archivematica e de maneira prática, o *Storage Service* por SSH captura os documentos do dashboard e salva nos respectivos locais do *Storage Service*. Diferentemente da instalação comum, para conseguir operar o Archivematica com diferentes pipelines é necessário criar um protocolo de acesso no *Storage Service* chamado de “*Pipeline Local File System*” que vai realizar a comunicação entre o *storage Service* e seus respectivos pipelines.

As imagens abaixo demonstram as diferenças entre os diferentes tipos de instalação. A primeira imagem representa uma instalação comum na qual o *Storage Service*, o *Dashboard* e os outros componentes ficam na mesma máquina. A segunda imagem demonstra um primeiro teste realizado, onde o dashboard e os outros componentes foram separados do *Storage Service* e conectados através do protocolo *Pipeline File System*. Já a terceira imagem representa a aplicação de múltiplos dashboards geridos por apenas um *Storage Service*.



O estudo está em fase final das análises, sendo a integração com o Openshift, o principal foco das atenções, pois todos os parâmetros antes estudados, mudam com essa nova abordagem da solução tecnológica.

## PLANO DE PRESERVAÇÃO DIGITAL DO PJE – PPD-PJE – **VERSÃO PRELIMINAR**

**O plano visa apoiar as operações de preservação digital dos processos judiciais digitais, por meio da integração de um repositório arquivístico digital confiável - RDC-Arq, com o sistema PJe e a solução de gerenciamento documental do PJe (PJe-Arq), com foco na preservação e disseminação das informações preservadas.**

Pretende-se, dessa forma, que o PPD-PJe seja o documento de referência na resolução de questões e promoções de melhorias / atualizações das práticas e tecnologias que visam a implementação da Política de Preservação Digital do TJDFT no âmbito do sistema PJe.

### 3.5 META 2: Análise de padrão de design de código para solução RDC-Arq

Aqui são realizados os levantamentos de padrões de design de código e análise da aplicação dos padrões abordados na estrutura do barramento RDC-Arq. Por consequência, a partir da relação de requisitos de negócio e aplicação das validações automatizadas nos códigos, elabora-se o modelo para estruturação de códigos para barramento RDC-Arq e o aplica à estrutura do TJDFT. Essa meta se finaliza com a apresentação em reunião presencial/virtual e entrega de relatório com detalhamento da etapa.

---

### 3.5.1 Análise de padrão de design de código para solução RDCArq

O desenvolvimento de softwares deve seguir padrões modernos de codificação. Ele deve ser pensado para a solução de problemas de forma específica, mas também ser flexível o suficiente para ser reutilizável e atender novos requisitos.

Para (GAMMA, 2000), os padrões de projeto (*Design Patterns*) permitem melhor reutilizar projetos e arquiteturas, fornecendo alternativas que deixam mais clara a documentação e manutenção de sistemas. Padrões de projeto visam a criação de projetos de software flexíveis e extensíveis.

### 3.5.2 Tipos de design patterns

Os criadores do conceito de *design patterns* propuseram 23 padrões, divididos em três categorias

- **Criacionais:** padrões que propõem soluções flexíveis para criação de objetos.
  - *Abstract Factory*
  - *Factory Method*
  - *Singleton*
  - *Builder*
  - *Prototype*
- **Estruturais:** padrões que propõem soluções flexíveis para composição de classes e objetos.
  - *Proxy*
  - *Adapter*
  - *Facade*
  - *Decorator*
  - *Bridge*
  - *Composite*
  - *Flyweight*
- **Comportamentais:** padrões que propõem soluções flexíveis para interação e divisão de responsabilidades entre classes e objetos.
  - *Strategy*
  - *Observer*
  - *Template Method*
  - *Visitor*
  - *Chain of Responsibility*
  - *Command*
  - *Interpreter*
  - *Iterator*
  - *Mediator*
  - *Memento*
  - *State*.

A pesquisa com *design patterns* está compatibilizando os estudos com o método de programação orientada a eventos adotada no TJDFT. A programação orientada a eventos é um conceito que difere como as rotinas são executadas. Nesse contexto, as rotinas são acionadas por indicações de um serviço de mensagens (ou eventos) que indica aos sistemas as funções e respostas de acordo com os acontecimentos de cada um, tornando todas as comunicações entre sistemas bastante flexível e assíncronas.

---

A pesquisa para atualização do Hipátia está mesclando esses conceitos para aprimorar sua flexibilidade na comunicação com os sistemas geradores, permitindo a reutilização do código de acordo com a demanda e necessidade das instituições.

## 3.6 META 3: Estudo para implantação de barramento RDC-Arq

Inicialmente essa etapa previa o levantamento de requisitos informacionais para implantação do modelo *OnDemand*, propondo uma arquitetura informacional para suporte a chamadas *OnDemand*. Previa também o estudo e implementação da descrição da tipologia documental a partir de códigos disponibilizados pelo MNI, ou outro serviço de fornecimento de dados, aplicada ao PJe para endereçamento dos pacotes na árvore de arranjo do ambiente de acesso AtoM.

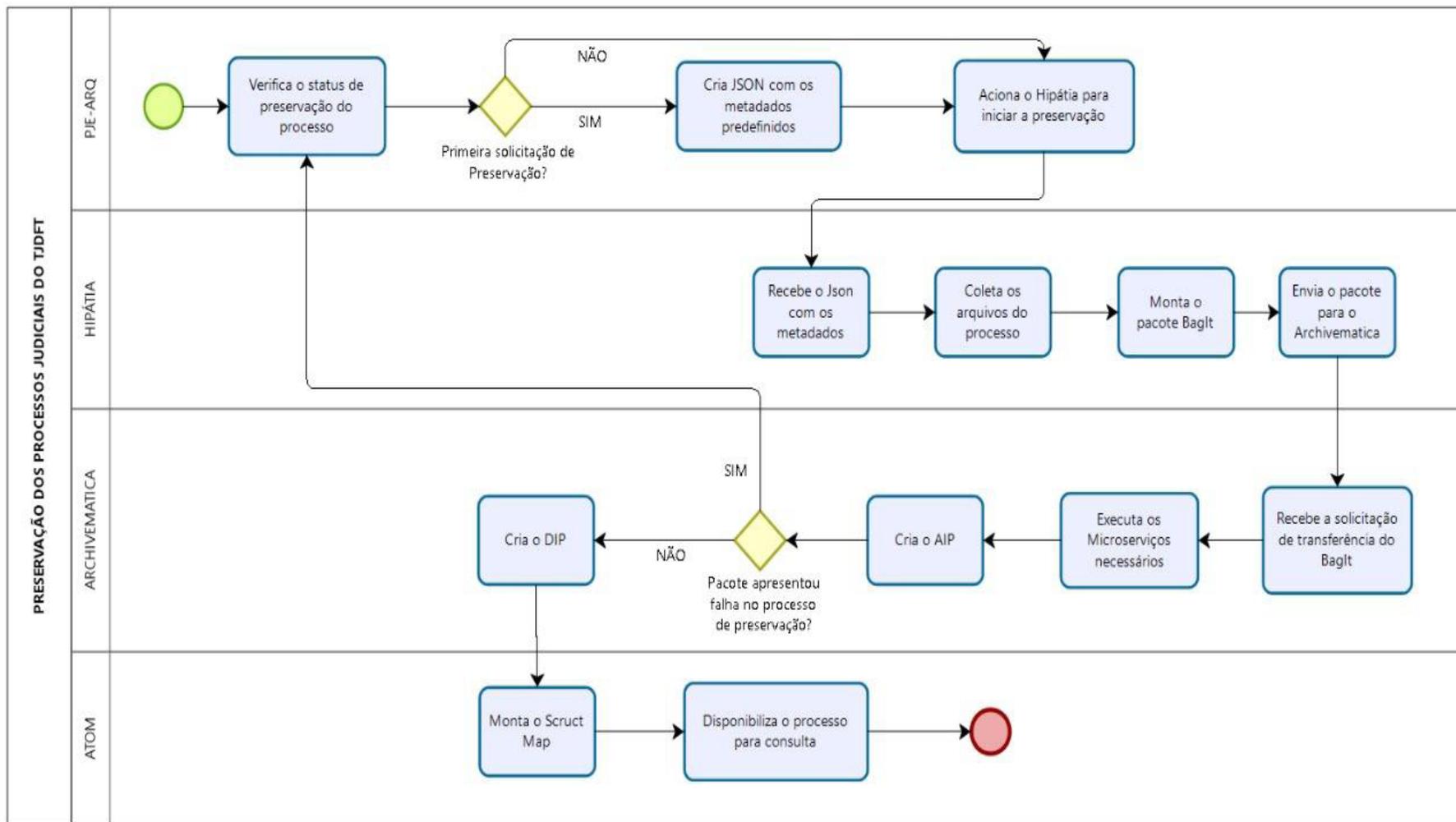
**O Termo Aditivo 02 do TED 001/2020 formalizou a alteração de escopo da meta 3 com o objetivo de tornar as aplicações da solução RDC-Arq desacopladas, em conformidade com as diretrizes da Plataforma Digital do Poder Judiciário – PDPJ-Br, e utilizar a arquitetura baseada em eventos para comunicação e troca de informações.**

### 3.6.1 Estudo para implantação de barramento RDC-Arq orientado a eventos

Havia o planejamento de uma solução para que o barramento Hipátia fosse *OnDemand*, isto é, acionado de acordo com a demanda de preservação por intermédio de algum comando dentro da aplicação PJE-Arq, conforme apresentado na Figura 3.

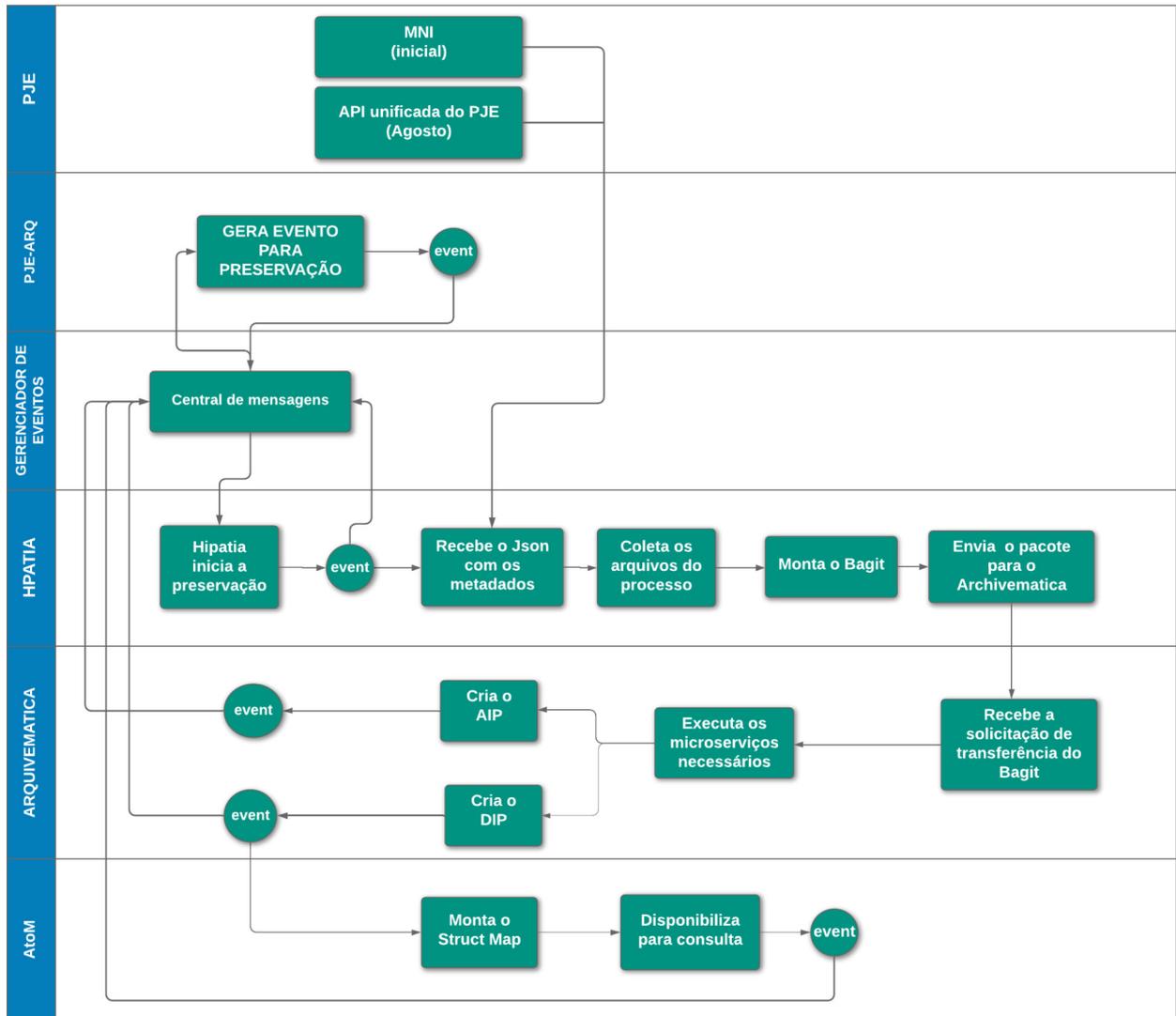
Entretanto, em reuniões para alinhamento da solução de integração PJE-Arq e Hipátia, foi apresentado o conceito de arquitetura baseada em eventos que o TJDFT vem adotando no desenvolvimento de suas aplicações. Nesse método, as aplicações não ficam acopladas aguardando os comandos, mas ligadas a um serviço de “mensageria” que alerta sobre eventos disponíveis nas diversas aplicações e inicia os processos específicos de cada um. Esse fluxo está apresentado na Figura 4 e mostra o novo conceito de execução baseada em eventos com o objetivo de alinhar as ações de cada equipe dentro da arquitetura da solução RDC-Arq.

Figura 3 - Modelo de fluxo de processos para Hipátia PJE com ênfase no acionamento do OnDemand



Fonte: Elaboração dos autores (2021)

Figura 4 - Modelo de fluxo de processos para Hipátia PJE com a comunicação entre o Hipátia e a mensageria



Fonte: Elaboração dos autores (2021).

Nesta versão ficam entendidas as comunicações com o sistema de mensageria e como o Hipátia recuperará as informações a partir dos eventos e quando ele irá gerar eventos para a infraestrutura.

Atividade executadas

- Reunião com a equipe para entendimento dos conceitos;
- Proposta inicial: Hipátia acionado on-demand;
- Sugestão para manter o Hipátia ativo com coleta agendada;
- Mudança no escopo: Arquitetura orientada a eventos;
  - Aprimoramento para estrutura de mensagens;
  - Foco na escalabilidade.

---

### 3.6.1.1 Estudo para integração com o serviço de mensageria Kafka

O Apache Kafka é um sistema de mensageria utilizado para processar informações oriundas de várias fontes, de maneira assíncrona, e disponibilizar esses dados para os diversos consumidores integrados a plataforma. De maneira simples, para que seja implantada uma estrutura onde aplicações se comunicam entre si através de trocas de mensagens ou, como a nomenclatura mais disseminada determina, eventos, é necessário a definição de três atores primordiais:

- Broker: Recebe os eventos dos produtores e as disponibiliza para os consumidores, funciona como uma espécie de central, pode ser configurado para atender diferentes necessidades nos mais variados casos, como por exemplo, garantir um armazenamento de eventos mais prolongado, ou até mesmo ser implantado em um ambiente “clusterizado”, visando uma melhor performance e garantindo a alta disponibilidade.
- Consumer: Consome os eventos armazenados no servidor do Kafka.
- Producer: Envia eventos para o servidor do Kafka.

Após a definição do papel de cada aplicação no decorrer do fluxo de execução do sistema é necessária a criação dos tópicos, que funcionam como agrupadores dos eventos, facilitando assim o consumo dos itens. Todo evento produzido é agrupado sob algum dos diversos tópicos que são criados no servidor do Kafka.

#### 3.6.1.1.1 Integração com o Hipátia

Conforme definido no fluxo do processo de integração entre o PJE-ARQ e o Hipátia, o segundo se comportará como consumidor no início do fluxo, quando o processo for disponibilizado pelo PJE, e então passará a exercer o papel de produtor de eventos, nas etapas subsequentes. Para facilitar o entendimento foi ilustrado como seria essa troca de mensagens nas imagens a seguir.

Exemplo de tópicos criados no servidor

```
paul@DESKTOP-IAGB5P3:/usr/local/kafka$ bin/kafka-topics.sh --list --bootstrap-server localhost:9092
HPT_AIP_CRIADO
HPT_START_EXECUCAO
HPT_STRUCT_MAP_CRIADO
PJE_NOVO_PROCESSO
```

Processo disponibilizado no PJE

```
paul@DESKTOP-IAGB5P3:/usr/local/kafka$ bin/kafka-console-consumer.sh --bootstrap-server localhost:9092
--topic PJE_NOVO_PROCESSO --from-beginning
processo0008227000026201923,13072021
```

Ex.: N° do processo e data do arquivamento

Início de uma execução do Hipátia

```
paul@DESKTOP-IAGB5P3:/usr/local/kafka$ bin/kafka-console-producer.sh --broker-list localhost:9092
--topic HPT_START_EXECUCAO
>0008227000026201923,534
```

Criação do AIP pelo Archivematica

```
paul@DESKTOP-IAGB5P3:/usr/local/kafka$ bin/kafka-console-producer.sh --broker-list localhost:9092
--topic HPT_AIP_CRIADO
>0008227000026201923,534,809e18eb-74a9-436b-b200-c07c54126693
>
```

Ex.: N° do processo, id do processo e hash do arquivo

Criação do Struct Map e disponibilização do processo no AtoM

```
paul@DESKTOP-IAGB5P3:/usr/local/kafka$ bin/kafka-console-producer.sh --broker-list localhost:9092
--topic HPT_STRUCT_MAP_CRIADO
>0008227000026201923,534
```

---

O formato e conteúdo das mensagens são utilizadas para exemplificação, assim, podem ter mais ou menos conteúdo. Assim, para que seja integrado o barramento com o sistema de mensageria, deve ser implementado no código do barramento o sistema de consumo e produção de eventos, bem como as configurações de conexão (URL do servidor de broker e os tópicos).

### 3.6.2 Estudos para implantação de banco de dados no barramento Hipátia para gerenciamento dos dados de acesso aos sistemas

A informação é um dos bens mais valiosos das instituições, ela pode representar uma vantagem competitiva, ou se mal utilizada, ou mal protegida, tornar-se fonte de vulnerabilidade.

Essa necessidade de proteção, fica ainda mais evidenciada quando se trata dos dados de acessos aos sistemas dentro de um RDC-Arq. Por esse motivo e atendendo as recomendações do Guia de Boas Práticas de Desenvolvimento do TJDF, foi iniciado o estudo para implantação de um banco de dados para gerenciar alguns dos dados mais importantes para o bom funcionamento da aplicação.

O SGBD (Sistema Gerenciador de Banco de Dados) escolhido é o MySQL, por já ser o banco de dados usado no Archivematica, facilitando assim o processo de definição, construção, manipulação e compartilhamento do banco de dados e as aplicações.

Abaixo estão descritas as definições das tabelas e campos que serão utilizados na construção da nova versão do barramento Hipátia.

#### 3.6.2.1 Definição das tabelas do Hipátia (MYSQL VERSÃO 5.7 COM SSL)

##### **TABELA - Sistemas (SEI, PJE, DJE, AtoM, etc.)**

- ID
- SIGLA\_SISTEMA
- NOME\_SISTEMA (DESCRIÇÃO)
- DRIVER\_BANCO
- URL\_ACESSO
- DATABASE
- USER
- PASSWORD a definir o tipo de criptografia)
- PORTA
- SISTEMA\_NEGOCIO (!- TRUE/FALSE)
- QTE\_ENVIO\_SEQUENCIAL

##### **TABELA - WEBSERVICES\_SISTEMAS**

- ID
- SIGLA\_SISTEMA
- DESCRICAO (MNI, PJE-ARQ, ATOM, ARCHIVEMATICA, SEI)
- URL\_WEBSERVICE

##### **TABELA - PARAMETROS**

- ID
- ID\_WEBSERVICE\_SISTEMAS
- CHAVE
- VALOR

---

## **TABELA - ENVIADOS**

- ID
- UUID (Archivematica)
- STATUS\_AIP: (0-sucesso, 1-processando, 2-erro)
- STATUS\_DIP: (a definir quais status)
- ID\_PROCESSO
- N\_PROCESSO
- ID\_SISTEMA (sistema produtor)
- DESTINACAO
- DESCRICAO
- TIPO\_PROCESSUAL
- FAILED/ERROS (ARCHIVEMATICA)
- TENTATIVAS (DE ENVIO)
- DATA\_ENVIO (date\_time)
- DATA\_CONCLUSAO (date\_time)

## **TABELA - LISTA\_PROCESSOS\_MANUAIS**

- ID
- ID\_SISTEMA
- ID\_PROCESSO
- N\_PROCESSO
- STATUS (0 - processado, 1-nao\_processado (default))
- DATA (date\_time)

## **TABELA - LOGS\_CONSULTAS**

- ID
- ID\_ENVIADOS
- ID\_SISTEMA
- ID\_PROCESSO
- N\_PROCESSO
- CONTEUDO (RETORNO DA CONSULTA)
- CONSULTA\_ENVIADA
- DATA (date\_time)

## **TABELA - LOGS\_GERAL**

- ID
- ID\_SISTEMA
- ID\_ENVIADOS
- CONTEUDO (payload-kafka (envio, recebimento), url de envio do Archivematica (pipeline), url de envio para o AtoM)
- TIPO (consumer\_kafka, producer\_kafka, envio\_Archivematica, envio\_AtoM)
- DATA (date\_time)

---

#### TABELA - Processing\_configuration (gera XML com regras por tipo de documento)

- ID
- ID\_SISTEMA
- CLASSIFICACAO
- ARQUIVO\_XML (? arquivo ou via api)
- STATUS (ativo ou inativo)

#### TABELA - SLUG\_ATOM

- ID
- ID\_SISTEMA
- CODIGO\_CLASSIFICACAO (\*solicitar ao sistema produtor)
- SLUG
- SLUG\_DEFAULT (sim ou não)
- STATUS (ativo ou inativo)

### 3.6.3 Estudos dos metadados RDC-Arq do PJE

Metadados são dados que carregam informações sobre outros dados, eles reúnem informações estruturadas que descrevem e identificam dados, auxiliando no gerenciamento, localização e preservação da informação classificada.

#### 3.6.3.1 TIPOS DE METADADOS

Existem 3 tipos principais de metadados (RILEY, 2017, p. 6, *tradução nossa*):

- **Metadados descritivos:** descrevem um recurso com o propósito, por exemplo, de descoberta ou identificação. Isso pode incluir elementos como título, resumo, autor e palavras-chave.
- **Metadados estruturais:** indicam como objetos compostos são colocados juntos, por exemplo, como é que páginas são ordenadas para formar capítulos.
- **Metadados administrativos:** fornecem informações para auxiliar no gerenciamento de um recurso, como por exemplo, quando e como ele foi criado, tipo de arquivo e outras informações técnicas, e sobre quem tem acesso a ele. Existem vários subconjuntos de dados administrativos; dois deles, às vezes, são listados separadamente como tipos metadados:
  - **Metadados para gerenciamento de direitos**, que tratam dos direitos de propriedade intelectual, e
  - **Metadados para preservação**, que contêm informações necessárias ao arquivamento e à preservação de um determinado recurso.

Para este projeto foram definidos e pesquisados os metadados necessários para a preservação dos processos judiciais no sistema PJE, bem como os metadados descritivos Dublin Core<sup>1</sup> e ISAD(G)<sup>2</sup>.

Os **Apêndices A, B, C, e D** apresentam os metadados levantados e validados pelas equipes de gestão da informação do Ibict e TJDFT em reuniões técnicas.

---

1 Disponível em: <https://dublincore.org/>

2 Disponível em: [https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/isad\\_g\\_2001.pdf](https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/isad_g_2001.pdf)

## 4. Transferência de conhecimento

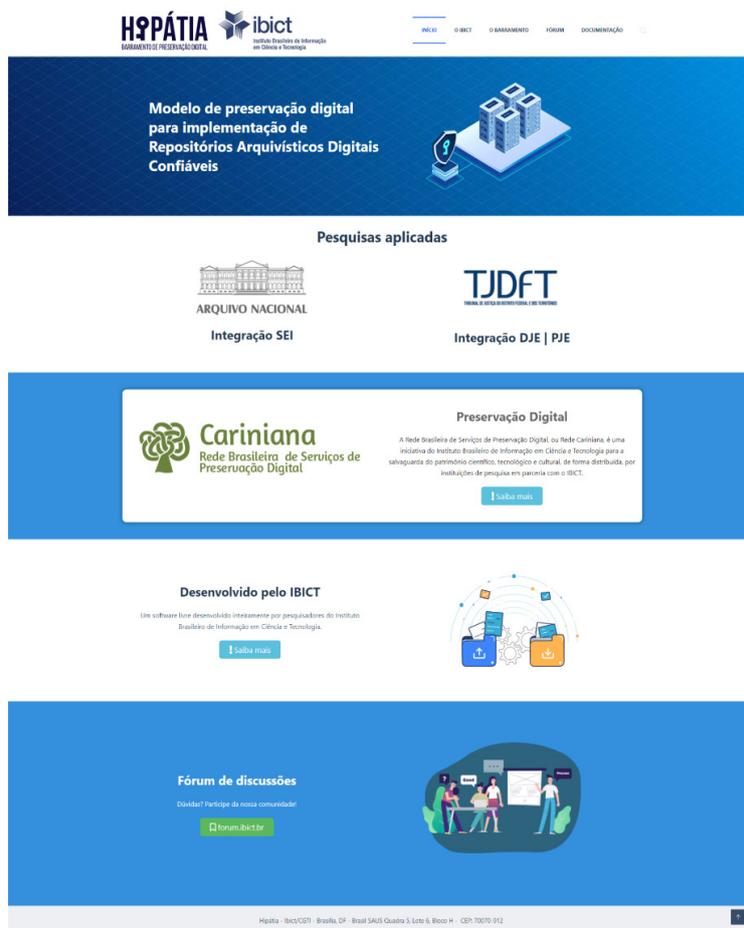
Os resultados da pesquisa foram amplamente disseminados em dois canais de comunicação com a comunidade de modo a repassar orientações sobre o uso do Hipátia.

A disseminação dos resultados da pesquisa é de suma importância em um projeto de pesquisa desenvolvido por um instituto público. Afinal, os atos da Administração Pública, como define a Constituição Federal em seu artigo 37, devem ser pautados pela publicidade. Além disso, a Lei de Acesso à Informação (Lei 12.527, de 2011) atribui aos órgãos de governo a responsabilidade de disseminação ativa de informações sobre sua gestão. Tais normativas e os princípios da Ciência Aberta definem a atuação do Ibict, que tem tradição na socialização do conhecimento científico e tecnológico.

### 4.1 Website

Foi criada uma página web, disponível em [hipatia.ibict.br](http://hipatia.ibict.br), que mostra informações relevantes do modelo de preservação digital para implementação de Repositórios Arquivísticos Digitais Confiáveis, conforme a Figura 5:

Figura 5 - Página inicial do portal Hipátia



Fonte: Captura de tela da página inicial do portal Hipátia (2021)

Destacam-se a parceria com o TJDFT e casos de uso do barramento. Para complementar a documentação do projeto, foram disponibilizadas três publicações em formatos de guia e manuais, elaboradas em outros projetos do IBICT que são:

### a) Manual de instalação Hipátia

O barramento Hipátia, é um programa que surgiu do projeto firmado com o TJDFT, com a finalidade de automatizar o envio dos arquivos por parte do agente Produtor para o Repositório Digital Confiável Arquivístico (RDC-ARQ), proposto pelo Modelo OAIS de preservação digital. O seu funcionamento segue o padrão de se conectar ao sistema Produtor por meio de Banco de Dados ou *WebService*, extrair metadados e arquivos a serem preservados e criar o Pacote de Submissão de Informação (SIP), enviando para o RDC-ARQ<sup>3</sup>.

### b) Guia Archivematica

O guia é resultado de um trabalho coletivo e compartilhado entre todos os autores identificados nesta obra, desenvolvido pela tradução, interpretação, ou adequação terminológica ou revisão ou mesmo adequação do referencial do Archivematica aos referenciais do Conselho Nacional de Arquivos no tocante aos Repositórios Arquivísticos Digitais Confiáveis, os RDC-Arq. É uma primeira aproximação na questão de adequar a terminologia aos referenciais introduzidos pelos Repositórios Digitais, na medida em que esse software foi selecionado para apoiar esse projeto. Igualmente apoia outras instituições que aderem ao uso do Archivematica para o desenvolvimento de repositórios arquivísticos digitais confiáveis, com documentação técnica em português.

**COSTA, Milene et al. Guia do usuário Archivematica. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2016. 136 p. DOI: 10.18225/978-85-7013-122-5 Disponível em: <http://livroaberto.ibict.br/handle/123456789/1063>.**

### c) Guia AtoM

O guia se baseia na documentação oficial do AtoM, versão 2.2 abrangendo não apenas a operacionalização da ferramenta, mas contando com informações voltadas à equipe de tecnologia. Com isso, procura atender a todos os envolvidos na instalação, manutenção, aperfeiçoamento, administração e operação, na medida em que a relação entre a equipe de informática e o arquivo requer estar afinada para o uso eficaz da ferramenta.

**ABREU, Jorge Phelipe Lira de; CASTRO, Priscila de Paiva; SHINTAKU, Milton; SANTAREM SEGUNDO, José Eduardo. Guia de usuário do AtoM. Brasília: Ibict, 2017. 164 p. Disponível em: <https://livroaberto.ibict.br/handle/123456789/1065>**

Como forma de divulgação, duas apresentações foram realizadas virtualmente:



**WEBINAR Preservação digital: o case do TJDFT na implementação do RDC-Arq. 2020. 1 vídeo (2 horas e 22 minutos). Publicado pelo canal TJDFToficial. Disponível em: <https://www.youtube.com/watch?v=kWKDlwQbk1k>.**



O PAPEL DA preservação em uma sociedade digital. 2020. 1 vídeo (1 hora e 52 minutos). Publicado pelo canal Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT. Disponível em: <https://www.youtube.com/watch?v=O0NJgqHzJbs>

<sup>3</sup> Disponível em: [https://git.ibict.br/cgti/barrapres\\_/barramento/-/wikis/Manual-Hip%C3%A1tia](https://git.ibict.br/cgti/barrapres_/barramento/-/wikis/Manual-Hip%C3%A1tia)

## 4.2 Fórum de discussão

No âmbito da Web, a colaboração de usuários é cada vez mais presente. Um local adequado para essa interatividade é o fórum de discussão que também está contemplado nessa etapa do projeto.

Figura 6 - Página inicial do Fórum do Ibict tópico Hipátia

Hipátia é um modelo de preservação digital que visa implantar os Repositórios Arquivísticos Digitais Confiáveis, garantindo toda a cadeia de custódia dos objetos digitais arquivísticos. Ele integra os sistemas barramento Hipátia, Archivematica e AtoM.

Tópico	Usuários	Respostas	Visualizações	Atividade
<b>Sobre a categoria HIPÁTIA</b> Hipátia é um modelo de preservação digital que visa implantar os Repositórios Arquivísticos Digitais Confiáveis, garantindo toda a cadeia de custódia dos objetos digitais arquivísticos. Ele integra os sistemas barramento... <a href="#">leia mais</a>		0	66	Jun 17
Manual de instalação Hipátia		1	98	Jun 30

Fonte: Captura de tela do tópico Hipátia no Fórum do Ibict (2021)

A Figura apresenta o Fórum de Discussão disponível no website que se destina a coletar insights, sugestões e críticas. Essa ferramenta está aberta a toda comunidade e se alinha às boas práticas do serviço público possibilitando a publicidade e a interação com os usuários.

## 5. Considerações finais

Este projeto elaborou um conjunto de propostas e vem realizando várias atividades que atendem a necessidade no TJDFT por melhorias e modernização na gestão documental aderente às melhores práticas nacionais e internacionais do mundo jurídico. Essas atividades foram pensadas e formuladas em conjunto com o Ibict que proveu sua expertise metodológica e científica na condução de projetos.

Nessa concepção, foram estabelecidos os objetivos específicos relacionados às entregas do projeto que se caracterizam como mapas que conduzem a equipe ao objetivo final e à entrega das atividades previstas. Assim, de forma a estabelecer explicitamente esses objetivos às entregas de trabalho, apresentamos o Quadro 3:

Quadro 3 - Objetivos Específicos X Entregas de Trabalho

OBJETIVOS ESPECÍFICOS	ENTREGAS DE TRABALHO
Melhorar o desempenho e infraestrutura do barramento RDC-Arq	Estudos de infraestrutura para solução RDC-Arq Estudos dos metadados RDC-Arq do PJE
Disponibilizar novas formas de acesso ao RDC-Arq	Estudo para implantação de barramento RDC-Arq usando o modelo de mensageria

Fonte: Elaboração dos autores (2021)

O Quadro 3 mostra expressamente o andamento do projeto dado que as entregas estão relacionadas com o propósito inicial do projeto em que foram estabelecidos os objetivos específicos.

Portanto, a implantação do software Archivematica de acordo com o modelo de infraestrutura *DevOps*, no qual o TJDFT se embasa, foi realizada com sucesso. Os volumes foram definidos, os arquivos de configuração *DockerFile* e demais arquivos necessários estão versionados no repositório *GitLab*, as imagens *Docker* estão descritas no **gitlab-ci.yml** e a aplicação está sendo executada sem erros conforme as especificações do Termo de Referência.

De tal modo, neste documento, descrevemos as atividades desenvolvidas pelas equipes das duas instituições para a operacionalização do projeto. Implementar o Archivematica no ambiente *OpenShift* figurou como o maior desafio para o Ibict. A pesquisa pioneira servirá de base para outras instituições.

Também foi dada ênfase a pesquisa inicial para implantação de um ambiente Archivematica com dois pipelines. Uma forma mais eficaz de se implantar o sistema com múltiplos fluxos de trabalhos com diferentes regras de preservação. Ao se usar uma única instalação para trabalhar diferentes formas de gestão, otimiza-se recursos informacionais e humanos.

A programação orientada a eventos foi outro grande salto evolutivo no barramento Hipátia, de modo a não deixá-lo vinculado e dependente dos outros sistemas. Integrar em um ambiente de mensageria, permite que ele continue a coleta e comunicação com os outros softwares do RDC-Arq, mesmo que o sistema de origem esteja temporariamente inoperante.

Todas as inovações pesquisadas e aplicadas neste ano de 2021, permitiram que as recomendações de segurança descritas no Guia de Boas Práticas do TJDFT, fossem avaliadas para que o ambiente Hipátia entre em conformidade com os itens apresentados e possa ser validado pela equipe responsável para entrar em produção.

O projeto evidencia a exitosa parceria entre o TJDFT e o IBICT. Ambas instituições tem colaborado com a elaboração de ferramentas metodológicas e científicas para que as barreiras para implantação de um RDC-Arq sejam transpostas. Dessa maneira está sendo construído um modelo de Repositório Digital Confiável que poderá ser replicado em todos os Tribunais de Justiça e outras instituições interessadas.

---

## Referências Bibliográficas

- ARCHIVEMATICA. Disponível em: <https://www.archivematica.org/>. Acesso em: 24 set. 2021.
- BACA, Murtha (ed.). Introduction to Metadata, Pathway to Digital Information. 2. ed. [S.l.]:[S.n.], 2008. Disponível em: <https://www.getty.edu/publications/resources/virtuallibrary/0892368969.pdf>. Acesso em: 24 set. 2021.
- BRASIL. Conselho Nacional de Justiça. **Modelo de Requisitos para Sistemas Informatizados de Gestão de Processos e Documentos do Judiciário brasileiro** (MoReq-Jus). Brasília: CNJ, 2009. 163 p. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2011/01/manualmoreq.pdf>. Acesso em: 24 set. 2021.
- BRASIL. Tribunal de Justiça do Distrito Federal e Territórios. **Relatório de gestão**. 2020. [https://www.tjdft.jus.br/transparencia/processo-de-contas-anual/rg-tcu-2020\\_final.pdf](https://www.tjdft.jus.br/transparencia/processo-de-contas-anual/rg-tcu-2020_final.pdf). Acesso em: 24 set. 2021.
- BRASIL. Tribunal de Justiça do Distrito Federal e Territórios. **Relatório de gestão**. 2019. Disponível em: [https://www.tjdft.jus.br/institucional/governanca/governanca-e-gestao-estrategica/relatorio-de-gestao\\_tjdft.pdf](https://www.tjdft.jus.br/institucional/governanca/governanca-e-gestao-estrategica/relatorio-de-gestao_tjdft.pdf). Acesso em: 24 set. 2021.
- CONSELHO INTERNACIONAL DE ARQUIVOS (Conarq). **ISAD(G)**: Norma geral internacional de descrição arquivística: segunda edição, adotada pelo Comitê de Normas de Descrição, Estocolmo, Suécia, 19-22 de setembro de 1999, versão final aprovada pelo CIA. 2. ed. Rio de Janeiro: Arquivo Nacional, 2000. 119 p. ISBN: 85-7009-032-3. Disponível em: [https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/isad\\_g\\_2001.pdf](https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/isad_g_2001.pdf). Acesso em: 24 set. 2021.
- COSTA, Milene *et al.* **Guia do usuário Archivematica**. Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2016. 136 p. DOI: 10.18225/978-85-7013-122-5 Disponível em: <http://livroaberto.ibict.br/handle/123456789/1063>. Acesso em: 24 set. 2021.
- DUBLIN CORE METADATA INITIATIVE. Disponível em: <https://dublincore.org/>. Acesso em: 24 set. 2021.
- FERRITER, MEGHAN. BagIt at the Library of Congress. **The Signal**, 4 apr. 2019. Disponível em: <https://blogs.loc.gov/thesignal/2019/04/bagit-at-the-library-of-congress/>. Acesso em: 24 set. 2021.
- GAMMA, Erich *et al.* **Padrões de projetos**: soluções reutilizáveis de software orientado a objetos. Porto Alegre: Bookman, 2000.
- GITLAB. GitLab CI/CD. [201-]. Disponível em: <https://docs.gitlab.com/ee/ci/>. Acesso em: 24 set. 2021.
- GOMES, Rafael. **Container versus máquina virtual**. Disponível em: 24 fev. 2016. <https://imasters.com.br/devsecops/container-versus-maquina-virtual>. Acesso em: 24 set. 2021.
- REDHAT. O que é DEVOps. [201-]. Disponível em: <https://www.redhat.com/pt-br/topics/devops>. Acesso em: 24 set. 2021.
- REDHAT. O que é Docker?. [201-]. Disponível em: <https://www.redhat.com/pt-br/topics/containers/what-is-docker>. Acesso em: 24 set. 2021.
- REDHAT. O que é Kubernetes?. [201-]. Disponível em: <https://www.redhat.com/pt-br/topics/containers/what-is-kubernetes>. Acesso em: 24 set. 2021.
- RILEY, Jenn. **Understanding Metadata**: what is metadata, and what is it for?: a primer. Bethesda, MD: NISO Press, 2017. ISBN: 978-1-937522-72-8. Disponível em: [https://groups.niso.org/apps/group\\_public/download.php/17446/Understanding%20Metadata.pdf](https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf). Acesso em 24 set. 2021.

# Apêndice A - Metadados de Gestão PJe (MoReq-Jus, e-Arq e Resolução nº 24 do conarq)

## DADOS DO PROCESSO

### Número do Processo

Descrição	Número do processo judicial, que deve estar de acordo com as regras definidas pela Resolução do CNJ nº 65, de 16 de dezembro de 2008.
Campo CSV	numeroProcesso
Campo WS	(nr_processo em core.tb_processo)
Exemplo	07196889620188070003

### Classe Processual - Código

Descrição	Número identificador da classe processual, segundo o padrão estabelecido no Sistema de Gestão de Tabelas Processuais Unificadas <sup>1</sup> . As classes estão disponíveis, de acordo com a configuração de classes para cada jurisdição.
Campo CSV	codigoClasseProcessual
Campo WS	(cd_classe_judicial em client.tb_classe_judicial)
Exemplo	6017

### Classe Processual - Nome

Descrição	Descrição da classe processual, segundo o padrão estabelecido no Sistema de Gestão de Tabelas Processuais Unificadas <sup>2</sup> . As classes estão disponíveis, de acordo com a configuração de classes para cada jurisdição.
Campo CSV	nomeClasseProcessual
Campo WS	(ds_classe_judicial em client.tb_classe_judicial)
Exemplo	Usucapião

### Código da Localidade da Jurisdição

Descrição	Código da Localidade da Jurisdição
Campo CSV	codigoDaLocalidade
Campo WS	(id_jurisdicao em client.tb_jurisdicao)
Exemplo	18

---

### Nome da Localidade

Descrição	Nome da Jurisdição.
Campo CSV	nomeDaLocalidade
Campo WS	(ds_jurisducao em client.tb_jurisducao)
Exemplo	Brasília - Fórum Des. Joaquim Sousa Neto

### Identificador da Competência

Descrição	Identificador da competência.
Campo CSV	idDaCompetencia
Campo WS	(id_competencia em client.tb_competencia)
Exemplo	4

### Nome da Competência

Descrição	Nome da competência.
Campo CSV	nomeDaCompetencia
Campo WS	(ds_competencia em client.tb_competencia)
Exemplo	Criminal

### Processo Vinculado (São campos que se repetem)

Descrição	Número do processo vinculado, no formato CNJ.
Campo CSV	processoVinculado
Campo WS	(tabela: client.tb_processo_trf_conexao)
Exemplo	0725943-26.2021.8.07.0016

## Modalidade de vinculação de processo (São campos que se repetem)

Descrição	Tipo de elemento destinado a permitir a identificação entre processos judiciais. As vinculações podem ser: 'AP' (apenso); 'AR' (ação rescisória); 'CD' (competência delegada); 'CT' (continência); 'CX' (conexão); 'DP' (dependência); 'OG' (originário); 'OR' (outro tipo de associação que não os anteriores); 'PU' (precedente de uniformização); 'RG' (repercussão geral); 'RR' (recurso repetitivo).
Campo CSV	modalidadeVinculacaoProcesso
Campo WS	( <a href="#">tp_tipo_conexao</a> em <a href="#">client.tb_processo_trf_conexao</a> )
Exemplo	Apenso

## Prioridade

Descrição	Elemento destinado a permitir a identificação da existência de prioridades processuais não óbvias, ou seja, aquelas que não são resultado direto da identificação da classe processual (ex.: habeas corpus ou mandado de segurança). Na versão 2.0, será texto livre, mas é recomendável utilizar os seguintes textos identificadores: "IDOSO" "RÉU PRESO" "PERECIMENTO" "MENOR"
Campo CSV	prioridade
Campo WS	( <a href="#">ds_prioridade</a> em <a href="#">client.tb_prioridade_processo</a> )
Exemplo	Idoso

## Valor da causa

Descrição	Valor da relação jurídica. Não usar caracteres especiais, apenas números.
Campo CSV	valorDaCausa
Campo WS	( <a href="#">vl_causa</a> em <a href="#">client.tb_processo_trf</a> )
Exemplo	5.000,00

## Nível de Sigilo

Descrição	Nível de sigilo a ser aplicado ao processo. Dever-se-á utilizar os seguintes níveis: - 0: públicos, acessíveis a todos os servidores do Judiciário e dos demais órgãos públicos de colaboração na administração da Justiça, assim como aos advogados e a qualquer cidadão - 1: segredo de justiça, acessíveis aos servidores do Judiciário, aos servidores dos órgãos públicos de colaboração na administração da Justiça e às partes do processo. - 2: sigilo mínimo, acessível aos servidores do Judiciário e aos demais órgãos públicos de colaboração na administração da Justiça - 3: sigilo médio, acessível aos servidores do órgão em que tramita o processo, à(s) parte(s) que provocou(ram) o incidente e àqueles que forem expressamente incluídos - 4: sigilo intenso, acessível a classes de servidores qualificados (magistrado, diretor de secretaria/escrivão, oficial de gabinete/assessor) do órgão em que tramita o processo, às partes que provocaram o incidente e àqueles que forem expressamente incluídos - 5: sigilo absoluto, acessível apenas ao magistrado do órgão em que tramita, aos servidores e demais usuários por ele indicado e às partes que provocaram o incidente.
Campo CSV	nivelDeSigilo
Campo WS	<a href="#">(cd_nivel_acesso em client.tb_processo_trf)</a>
Exemplo	Segredo de Justiça

## Data da autuação

Descrição	Este atributo corresponde a data de autuação do processo e sempre virá preenchido nos dadosBasicos do processo após uma consulta. Observação: Na entrega da manifestação não é preciso preencher este atributo.
Campo CSV	dataDaAutuacao
Campo WS	<a href="#">(dt_autuacao em client.tb_processo_trf)</a>
Exemplo	01/01/2021 14:35:00 Formato: 'DD/MM/AAAA HH:MM:SS'

## Pedido de liminar

Descrição	Indica se houve pedido de liminar no processo.
Campo CSV	pedidoDeLiminar
Campo WS	<a href="#">(in_tutela_liminar em client.tb_processo_trf)</a>
Exemplo	Sim / Não;

## Número SISPL

Descrição	Número do processo antigo do SISPL, para processos digitalizados.
Campo CSV	numeroSISPL
Campo WS	<a href="#">(numero_processo_sispl em client.tb_processo)</a>
Exemplo	2010.01.1.032950-9

### Processo de Origem

Descrição	Número do processo de origem no formato do CNJ.
Campo CSV	processoDeOrigem
Campo WS	(nr_processo_origem em client.tb_processo)
Exemplo	2010.01.1.032950-9

### Incidental

Descrição	Indica se o processo é incidental.
Campo CSV	incidental
Campo WS	(in_tutela_liminar em client.tb_processo_trf)
Exemplo	Sim / Não

### Objeto do processo

Descrição	Campo utilizado internamente para classificar o objeto do processo.
Campo CSV	objetoDoProcesso
Campo WS	(ds_objeto em client.tb_processo_trf)
Exemplo	SISTJ (indica que foi digitalizado do SISTJ para o PJE)

### Justiça 100 Digital

Descrição	Campo que indica se houve pedido de juízo 100% digital.
Campo CSV	justica100Digital
Campo WS	(in_solicitado_juizo_100_digital em client.tb_processo_trf)
Exemplo	sim/não

### Processo Digitalizado

Descrição	Campo que indica se o processo foi digitalizado.
Campo CSV	processoDigitalizado (boolean)
Campo WS	(ds_codigo em client.tb_tipo_sit_processual // id_tipo_situacao_processual em client.tb_situacao_processual )
Exemplo	Sim/não

### Órgão Julgador - Código

Descrição	Identificador do órgão julgador do processo.
Campo CSV	codigoOrgaoJulgador
Campo WS	( <a href="#">id_orgao_julgador</a> em <a href="#">client.tb_orgao_julgador</a> )
Exemplo	144

### Órgão Julgador - Nome

Descrição	Descrição do órgão julgador do processo.
Campo CSV	nomeOrgaoJulgador
Campo WS	( <a href="#">ds_orgao_julgador</a> em <a href="#">client.tb_orgao_julgador</a> )
Exemplo	2ª Vara Cível de Ceilândia

### Plano Classificação e Tabela de Temporalidade Utilizada

Descrição	Descrever qual a versão do plano de classificação e tabela de temporalidade foi utilizado para extrair os metadados deste documento.
Campo CSV	versaoPCTT
Onde buscar	No Sistema de Gestão de Tabelas Processuais Unificadas ( <a href="https://www.cnj.jus.br/sgt/versoes.php">https://www.cnj.jus.br/sgt/versoes.php</a> )  function getDataUltimaVersao() /** * Retorna uma String contendo a data da última versão * * @return String */ (Obs: data extraída no momento do cálculo da temporalidade no PJE-Arq)
Exemplo	Plano de classificação e tabela de temporalidade de documentos - PCTT-AD TJDFT (Minuta de Proposta do PCTT-TJDFT - Versão 1-47)

### Prazo de guarda

Descrição	Indicação do prazo estabelecido em tabela de temporalidade para o cumprimento da destinação.
Campo CSV	prazoGuardaClassificacao
Onde buscar	PJe-Arq
Exemplo	Um (01) ano de guarda em arquivo corrente e depois é recolhido para o arquivo permanente.

### Destinação prevista

Descrição	Indicação da próxima ação de destinação (transferência, eliminação ou recolhimento) prevista para o documento, em cumprimento à tabela de temporalidade.
Campo CSV	destinacaoPrevista
Onde buscar	PJe-Arq
Exemplo	Repositório Arquivístico Digital - Tribunal de Justiça do Distrito Federal e dos Territórios

## SUBMISSION DOCUMENTATION

Ficará no JSON os dados que se repetem e não são possíveis de elaborar um **csv** com 2 ou mais linhas. São os campos referentes a: documento; partes; movimentação processual;

### Identificador do documento

Descrição	Identificador do Documento.
Campo CSV	idDocumento
Campo WS	( <a href="#">id_processo_documento_trf</a> em <a href="#">client.tb_processo_documento_trf</a> )
Exemplo	123456

### Identificador do documento principal

Descrição	Identificador do documento principal caso este documento seja um anexo.
Campo CSV	idDocumentoVinculado
Campo WS	( <a href="#">id_doc_associado</a> em <a href="#">client.tb_proc_doc_associacao</a> )
Exemplo	123456

### Tipo de documento

Descrição	Código do tipo de documento disponível no serviço "ConsultaPJe" - Ex.: 58 - Petição Inicial, 118A - Informações. O código variará dependendo da instalação do PJe.
Campo CSV	codigoTipoDocumento
Campo WS	(id_tipo_processo_documento em core.tb_processo_documento)
Exemplo	123456

### Tipo de documento - Nome

Descrição	Descrição do tipo de documento disponível no serviço "ConsultaPJe" - Ex.: 58 - Petição Inicial, 118A - Informações. Obs.: A ser desenvolvido pela SEDES/TJDFT.
Campo CSV	nomeTipoDocumento
Campo WS	(ds_tipo_processo_documento em core.tb_tipo_processo_documento)
Exemplo	Petição Inicial, 118A - Informações.

### Data e hora da criação

Descrição	Sequência de ano+dia+mês/hora+minuto+segundo. Atributo destinado a permitir a informação do momento em que o documento foi recebido pelo Poder Judiciário. É opcional no caso de documento enviado para protocolo.
Campo CSV	dataHoraCriacao
Campo WS	(dt_inclusao em core.tb_processo_documento_bin)
Exemplo	12/12/2012 14:32:00

### Descrição

Descrição	Descrição do documento.
Campo CSV	descricao
Campo WS	(ds_documento em core.tb_documento)
Exemplo	Sentença; Carteira de motorista.

---

## Hash

Descrição	Resumo matemático do documento enviado pelo algoritmo SHA-256.
Campo CSV	hash
Campo WS	<a href="#">(ds_md5_documento em core.tb_processo_documento_bin)</a>
Exemplo	b0c020615d91c745a8cb54025cece674

## Documento Sigiloso

Descrição	Propriedade que indica se o documento é ou não sigiloso.
Campo CSV	documentoSigiloso
Campo WS	<a href="#">(in_documento_sigiloso em core.tb_tipo_processo_documento)</a>
Exemplo	Sim / não

## Mimetype

Descrição	Tipo do arquivo que será enviado, seguindo o seguinte domínio: application/pdf ou html/text. O PJe suporta os seguintes tipos de arquivos: pdf (máximo até 3Mb), mpeg (máximo até 5Mb), ogg (máximo até 10Mb), vorbis (máximo até 5Mb), png (máximo até 3Mb), mp4 (máximo até 10Mb) e quicktime (máximo até 10Mb).
Campo CSV	mimetype
Campo WS	<a href="#">(ds_extensao em core.tb_processo_documento_bin)</a>
Exemplo	video/mp4, application/pdf

## Assinatura

Descrição	Contém os dados em Base64 do arquivo de assinatura do documento. Para arquivos em PDF ou HTML, que não sejam P7S, é necessário informar os dados da assinatura.
Campo CSV	assinatura
Campo WS	<a href="#">(ds_signature em core.tb_processo_documento_bin)</a>
Exemplo	assinaturadeteste

## Usuário da assinatura

Descrição	Identificador do usuário que realizou a assinatura. O MNI salva o ID do usuário, mas pode-se gravar o CPF ou ambos.
Campo CSV	usuarioDaAssinatura
Campo WS	( <a href="#">id_usuario_ultimo_assinar</a> em <a href="#">core.tb_processo_documento_bin</a> )
Exemplo	123456

## Algoritmo hash

Descrição	Algoritmo utilizado para gerar a assinatura: MD5withRSA, SHA1withRSA, SHA256withRSA ou ASN1MD5withRSA. Atualmente, o mais utilizado é o SHA256withRSA. Atualmente, o mais utilizado é o SHA256withRSA
Campo CSV	algoritmoHash
Campo WS	campo fixo com o valor MD5
Exemplo	MD5withRSA

## Cadeia do certificado

Descrição	String do certificado. Caso seja do tipo PEM é necessário incluir inclusive as strings "BEGIN CERTIFICATE" e "END CERTIFICATE".
Campo CSV	cadeiaDeCertificado
Campo WS	( <a href="#">ds_cert_chain</a> em <a href="#">core.tb_processo_documento_bin</a> )
Exemplo	-----BEGIN CERTIFICATE----- certchaindeteste= -----END CERTIFICATE-----

## Codificação do certificado

Descrição	Codificação utilizada no certificado: "PEM" ou "PkiPath".
Campo CSV	codificacaoDoCertificado
Campo WS	( <a href="#">ds_md5_documento</a> em <a href="#">core.tb_processo_documento_bin</a> )
Exemplo	PkiPath

---

### \* Data da assinatura

Descrição	Data e Hora da assinatura, obedecendo a sequência de ano+dia+mês+minuto+segundo.
Campo CSV	dataAssinatura
Campo WS	(dt_primeira_assinatura em core.tb_processo_documento_bin)
Exemplo	20190725181340

### \* Signatário login

Descrição	Informação referente a signatário que fez uso de login e senha para assinar o documento.
Campo CSV	signatarioLogin
Campo WS	(?)
Exemplo	

### Identificação do tipo de meio do documento

Descrição	Documento/volume/processo/dossiê: digital, não digital (convencionais/analógicos) ou híbrido, para controlar as relações entre os meios e o monitoramento de preservação.
Campo CSV	meioDocumental
Onde buscar	Será um texto fixo.
Exemplo	Documento digital

### Gênero Documental

Descrição	É a configuração da informação no documento de acordo com o sistema de signos utilizado na comunicação do documento. Exemplos: Audiovisual; textual; cartográfico; iconográfico; multimídia.
Campo CSV	generoDocumental
Onde buscar	Será de acordo com o mime type do documento.
Exemplo	Textual

---

## PARTES

### \*Polo

Descrição	Sigla para identificação do tipo de polo. Para polo ativo, usar "AT"; Passivo, usar "PA"; "TC" = Terceiro; "FL" = Fiscal da lei diverso; "TJ" = Testemunha do juízo; "AD" = Assistente simples desinteressado (amicus curiae); "VI" = Vítima.
Campo CSV	polo
Campo WS	(ds_tipo_parte em client.tb_tipo_parte // client.tb_processo_parte)
Exemplo	Ativo

### \*Nome da parte

Descrição	Nome da parte que compõe o polo do processo.
Campo CSV	nome
Campo WS	(ds_nome em acl.tb_usuario_login)
Exemplo	Vanessa Rocha

### \*Sexo da parte

Descrição	Correspondente a pessoa parte, onde "M" masculino; "F" feminino; "D" desconhecido.
Campo CSV	sexo
Campo WS	(in_sexo em client.tb_pessoa_fisica)
Exemplo	Feminino

### \*Nome do genitor

Descrição	Nome completo do pai da pessoa parte.
Campo CSV	nomeDoGenitor1
Campo WS	(nm_genitora em client.tb_pessoa_fisica)
Exemplo	Francisco Rocha

---

**\*Nome da genitora**

Descrição	Nome completo da mãe da pessoa parte.
Campo CSV	nomeDoGenitor2
Campo WS	(nm_genitor em client. tb_pessoa_fisica)
Exemplo	Maria Rocha

**(Referenciado na tabela) Data de nascimento da parte**

Descrição	Data do nascimento no formato AAAAMMDD.
Campo CSV	dataDeNascimento
Campo WS	(dt_nascimento em tb_pessoa_fisica)
Exemplo	19900104

**(Referenciado na tabela) Data de óbito da parte**

Descrição	Data do falecimento no formato AAAAMMDD.
Campo CSV	dataDeObito
Campo WS	(dt_obito em client. tb_pessoa_fisica)
Exemplo	20100529

**\*Número do documento principal**

Descrição	Número do documento principal da pessoa parte, devendo ser utilizado o RIC ou o CPF para pessoas físicas, nessa ordem, ou o CNPJ para pessoas jurídicas. O atributo é opcional em razão da possibilidade de haver pessoas sem documentos ou cujos dados não estão disponíveis.
Campo CSV	numeroDocumentoPrincipal
Campo WS	(nr_documento_identificacao em client.tb_pess_doc_identificacao)
Exemplo	999999999999

### \*Tipo de pessoa

Descrição	Indicar qual o tipo de pessoa que será integrante do processo. Enviar um dos seguintes valores, sem acentos: "fisica";"juridica"; "autoridade"; "orgaorepresentacao".
Campo CSV	tipoDePessoa
Campo WS	(ds_tipo_pessoa em client.tb_tipo_pessoa // in_tipo_pessoa em client.tb_pessoa)
Exemplo	Física

### (Referenciado na tabela) Nome Fantasia

Descrição	Nome fantasia, para o caso da parte pessoa jurídica.
Campo CSV	nomeFantasia
Campo WS	(nm_fantasia em client.tb_pessoa_juridica)
Exemplo	

### (Referenciado na tabela) Pessoa vinculada

Descrição	Indicação da existência de um relacionamento entre uma autoridade e uma pessoa jurídica ou ente que ela representa. Exemplo: a escola é vinculada a seu diretor ou reitor. Sugestão: Salvar o nome da pessoa ou criar outro campo, se necessário, para salvar o CPF.
Campo CSV	pessoaVinculada
Campo WS	(nr_cpf_responsavel em client.tb_pessoa_juridica OU id_pessoa em client.tb_autoridade_publica)
Exemplo	Cristiano

### (Referenciado na tabela) Cidade natural

Descrição	Local de nascimento da pessoa parte.
Campo CSV	cidadeNatural
Campo WS	(id_municipio_nascimento em client.tb_pessoa_fisica)
Exemplo	Brasília

**(Referenciado na tabela) Estado natural**

Descrição	Estado de nascimento da pessoa parte. A implementação de codificação/decodificação deverá interpretar o valor do campo como se tratando da sigla de duas letras identificadoras das unidades federativas brasileiras.
Campo CSV	estadoNatural
Campo WS	( <a href="#">id_estado em core.tb_municipio</a> // <a href="#">id_municipio_nascimento em client.tb_pessoa_fisica</a> )
Exemplo	DF

**(Referenciado na tabela) Nacionalidade da parte**

Descrição	País de nascimento da pessoa parte. Exemplo: "BR"-Brasil. Dever-se-á utilizar, preferencialmente, o código ISO-3166-1-alpha-2 ( <a href="http://www.iso.org/iso/english_country_names_and_code_elements">http://www.iso.org/iso/english_country_names_and_code_elements</a> ) pertinente ao país respectivo.
Campo CSV	nacionalidade
Campo WS	( <a href="#">id_pais em client.tb_pessoa_fisica</a> )
Exemplo	BR

**\*Código do documento de identificação da parte**

Descrição	Número do documento identificador vinculado ao tipo de pessoa, podendo ser o número do CPF, CNPJ, RG, TIT, etc. São uma sequência de caracteres que tornam único o documento descrito sem os separadores característicos.
Campo CSV	codigoDoDocumentoDeIdentificacao
Campo WS	( <a href="#">nr_documento_identificacao em client.tb_pess_doc_identificacao</a> )
Exemplo	99999999999

**\*Emissor do documento de identificação da parte**

Descrição	Entidade emissora do documento. Exemplo: SSP, Cartórios, Justiça Eleitoral, SRFB.
Campo CSV	emissorDoDocumentoDeIdentificacao
Campo WS	( <a href="#">dr_orgao_expedidor em client.tb_pess_doc_identificacao</a> )
Exemplo	SSP

### \*Tipo do documento de identificação da parte

Descrição	Elemento destinado a permitir a identificação dos diversos tipos de documentos de identificação existentes. Trata-se de enumeração dos seguintes tipos: [Quebra da Disposição de Texto]- CI: carteira de identidade[Quebra da Disposição de Texto]- CNH: carteira nacional de habilitação[Quebra da Disposição de Texto]- TE: título de eleitor[Quebra da Disposição de Texto]- CN: certidão de nascimento[Quebra da Disposição de Texto]- CC: certidão de casamento[Quebra da Disposição de Texto]- PAS: passaporte[Quebra da Disposição de Texto]- CT: carteira de trabalho[Quebra da Disposição de Texto]- RIC: registro individual do cidadão[Quebra da Disposição de Texto]- CMF: cadastro no Ministério da Fazenda Brasileiro [Quebra da Disposição de Texto]- PIS_PASEP: número no programa de integração social [Quebra da Disposição de Texto]- CEI: cadastro específico do INSS [Quebra da Disposição de Texto]- NIT: número de identificação do trabalho [Quebra da Disposição de Texto]- CP: cadastro em conselhos profissionais [Quebra da Disposição de Texto]- IF: identidade funcional[Quebra da Disposição de Texto]- OAB: número de cadastro na Ordem dos Advogados do Brasil [Quebra da Disposição de Texto]- RJC: número de inscrição empresarial; [Quebra da Disposição de Texto]- RGE: registro de identificação do estrangeiro
Campo CSV	tipoDoDocumentoDeldocumentacao
Campo WS	( <a href="#">client.tb_pess_doc_identificacao</a> em <a href="#">client.tb_pess_doc_identificacao</a> )
Exemplo	CNH

### Nome no documento de identificação

Descrição	Nome da parte detentora do documento.
Campo CSV	nomeNoDocumentoDeldocumentacao
Campo WS	( <a href="#">ds_nome_pessoa</a> em <a href="#">client.tb_pess_doc_identificacao</a> )
Exemplo	

### \*Outro nome da parte

Descrição	Elemento destinado a permitir a inclusão de outros nomes da pessoa individualizada. Devem ser incluídos sob esse elemento os nomes que NÃO constam em documentos específicos. Em caso de nomes divergentes do principal existentes em documentos específicos, esses nomes devem ser indicados no atributo pertinente do documento.
Campo CSV	outroNome
Campo WS	( <a href="#">ds_pessoa_nome_alternativo</a> em <a href="#">client.tb_pessoa_nome_alternativo</a> )
Exemplo	

### \*CEP endereço da parte

Descrição	CEP da parte do processo. Atributo indicador do código de endereçamento postal do endereço no diretório nacional de endereços da ECT. O valor deverá ser uma sequência de 8 dígitos, sem qualquer separador. O atributo é opcional para permitir a apresentação de endereços desprovidos de CEP e de endereços internacionais.
Campo CSV	cep
Campo WS	(nr_cep em core.tb_cep // id_cep em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

### \*Logradouro endereço da parte

Descrição	Logradouro vinculado ao CEP, tais como as ruas, avenidas, praças, jardins, etc.
Campo CSV	logradouro
Campo WS	(nm_logradouro em core.tb_cep // id_cep em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

### \*Número endereço da parte

Descrição	Número vinculado ao CEP, podendo ser número da casa, apartamento, etc.
Campo CSV	numero
Campo WS	(nm_logradouro em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

### \*Complemento endereço da parte

Descrição	Dados adicionais do endereço.
Campo CSV	string
Campo WS	(ds_complemento em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

---

**\*Bairro endereço da parte**

Descrição	Parte que compõe a cidade, devendo ser vinculada ao CEP.
Campo CSV	bairro
Campo WS	(nm_bairro em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

**\*Cidade endereço da parte**

Descrição	Vinculada ao CEP, ou seja, nome da zona geográfica onde se agrupam as residências, comércios e indústrias.
Campo CSV	cidade
Campo WS	(nm_logradouro em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	

**\*Estado endereço da parte**

Descrição	Sigla do Estado brasileiro, composto por duas letras.
Campo CSV	estado
Campo WS	(id_municipio em core.tb_cep // id_cep em core.tb_endereco // id_endereco em tb_processo_parte_endereco)
Exemplo	DF

**\*País endereço da parte**

Descrição	Sigla referente ao país do Estado que compõe o endereço, composto por duas letras. Dever-se-á utilizar, preferencialmente, o código ISO-3166-1-alpha-2 ( <a href="http://www.iso.org/iso/english_country_names_and_code_elements">http://www.iso.org/iso/english_country_names_and_code_elements</a> ) pertinente ao país respectivo.
Campo CSV	país
Campo WS	(como o PJE não aceita endereço sem CEP ainda, o país referente aos endereços é sempre BR)
Exemplo	BR

## DADOS DOS REPRESENTANTES PROCESSUAIS

### (Referenciado na tabela) Nome

Descrição	Nome do advogado ou do escritório de advocacia ou órgão de representação processual.
Campo CSV	nome
Campo WS	(ds_nome em acl.tb_usuario_login)
Exemplo	"B & G Advogados" (Campo que se repete)

### (REFERENCIADO NA TABELA) INSCRIÇÃO

Descrição	Inscrição do advogado ou escritório de advocacia no cadastro da Ordem dos Advogados do Brasil. O campo é preenchido no formato CCDDDDDDDC, sendo os primeiros 'C' dois caracteres indicativos da unidade federativa brasileira em que há a inscrição, os sete dígitos 'D' seguintes o número da inscrição, devendo incluir zeros ('0') à esquerda caso o número tenha menos de 7 dígitos e o último caracter a letra identificadora do tipo de inscrição.
Campo CSV	inscricao
Campo WS	(nr_oab + ds_letra_oab + id_uf_oab em client.tb_dado_oab_pess_advogado)
Exemplo	1254DF

### (REFERENCIADO NA TABELA) NÚMERO DO DOCUMENTO PRINCIPAL

Descrição	O número RIC ou o CPF, para advogados, ou o CNPJ, para escritórios de advocacia e órgãos de representação processual (MP, advocacia pública e defensoria pública).
Campo CSV	numeroDoDocumentoPrincipal
Campo WS	(ds_login em acl.tb_usuario_login)
Exemplo	1111111

### (REFERENCIADO NA TABELA) TIPO DE REPRESENTANTE

Descrição	Atributo que permite indicar o tipo de representante processual.
Campo CSV	tipoDeRepresentante
Campo WS	(id_tipo_representante em client.tb_proc_parte_representante)
Exemplo	Advogado"; "Escritório de advocacia"; "Ministério público"; "Defensoria pública";

---

## DADOS DOS ASSUNTOS PROCESSUAIS

### \*Principal

Descrição	Indica se o assunto será o principal do processo.
Campo CSV	principal
Campo WS	(in_assunto_principal em client.tb_processo_assunto)
Exemplo	sim/não

### \*Código Nacional

Descrição	Código do assunto processual, segundo SGT.
Campo CSV	codigoNacional
Campo WS	(cd_assunto_trf em client.tb_assunto_trf)
Exemplo	1945 (descrever o nome e código).

### Código Pai Nacional

Descrição	Atributo destinado à entrada do código de assunto nacional de que o assunto local é filho.
Campo CSV	codigoPaiNacional
Campo WS	(id_assunto_trf_superior.cd_assunto_trf em client.tb_assunto_trf)
Exemplo	4563

### \*DESCRIÇÃO IGUAL AO NOME

Descrição	Texto descritivo do assunto processual.
Campo CSV	descricaoigualao nome
Campo WS	(ds_assunto_trf em client.tb_assunto_trf)
Exemplo	“Alimentos”

---

## MOVIMENTAÇÃO PROCESSUAL

### \*Código da movimentação processual

Descrição	Código da movimentação de acordo com a SGT.
Campo CSV	codigoDaMovimentacao
Campo WS	( <a href="#">cd_evento</a> em <a href="#">core.tb_evento</a> // <a href="#">core.tb_processo_evento</a> )
Exemplo	390

### \*Nome da movimentação processual

Descrição	Nome da movimentação processual, conforme SGT.
Campo CSV	nomeDaMovimentacao
Campo WS	( <a href="#">ds_evento</a> em <a href="#">core.tb_evento</a> // <a href="#">core.tb_processo_evento</a> )
Exemplo	Sentença Proferida

### Descrição completa da movimentação processual

Descrição	Descrição completa da movimentação, já com os complementos preenchidos.
Campo CSV	descricaoCompletaDaMovimentacao
Campo WS	( <a href="#">ds_texto_final_interno</a> em <a href="#">core.tb_processo_evento</a> )
Exemplo	Decorrido prazo de Vanessa Rocha em 19/03/2021 23:59:59

### \*Data da movimentação processual

Descrição	Data da movimentação em sequência de ano+dia+mês/hora+minuto+segundo.
Campo CSV	dataDaMovimentacao
Campo WS	( <a href="#">dt_atualizacao</a> em <a href="#">core.tb_processo_evento</a> )
Exemplo	2021/03/21 23:59:59

## ESTUDO DOS CAMPOS DO PJE COM OS CAMPOS DO MOREQ-JUS

### Nível de Sigilo

Descrição	Nível de sigilo a ser aplicado ao processo. Dever-se-á utilizar os seguintes níveis: - 0: públicos, acessíveis a todos os servidores do Judiciário e dos demais órgãos públicos de colaboração na administração da Justiça, assim como aos advogados e a qualquer cidadão - 1: segredo de justiça, acessíveis aos servidores do Judiciário, aos servidores dos órgãos públicos de colaboração na administração da Justiça e às partes do processo. - 2: sigilo mínimo, acessível aos servidores do Judiciário e aos demais órgãos públicos de colaboração na administração da Justiça - 3: sigilo médio, acessível aos servidores do órgão em que tramita o processo, à(s) parte(s) que provocou(ram) o incidente e àqueles que forem expressamente incluídos - 4: sigilo intenso, acessível a classes de servidores qualificados (magistrado, diretor de secretaria/escrivão, oficial de gabinete/assessor) do órgão em que tramita o processo, às partes que provocaram o incidente e àqueles que forem expressamente incluídos - 5: sigilo absoluto, acessível apenas ao magistrado do órgão em que tramita, aos servidores e demais usuários por ele indicado e às partes que provocaram o incidente.
Moreq	Metadados segurança; MSG1.1.20 Sigilo_recebido

### Usuário do cadastro do processo

Descrição	Identificador do usuário que protocolou a ação. Identificação do perfil de usuário para alteração de documento ou processo/dossiê/volume
Moreq	Metadados segurança; MSG1.1.16 Perfil_alteração_corpo

### Nome do usuário do protocolo

Descrição	Nome do usuário que efetuou o protocolo do processo. Identificação do perfil de usuário para alteração de metadado de documento ou processo/dossiê/volume.
Moreq	Metadados segurança; MSG1.1.17 Perfil_alteração_metadado

### Processo Digitalizado

Descrição	Campo que indica se o processo foi digitalizado.
Moreq	Metadados preservação; MPR1.3.6

## HASH

Descrição	Resumo matemático do documento enviado pelo algoritmo SHA-256.
Moreq	Metadados preservação; MPR1.3.24 - Checksum

## MIMETYPE

Descrição	Tipo do arquivo que será enviado, seguindo o seguinte domínio: application/pdf ou html/text. O PJe suporta os seguintes tipos de arquivos: pdf (máximo até 3Mb), mpeg (máximo até 5Mb), ogg (máximo até 10Mb), vorbis (máximo até 5Mb), png (máximo até 3Mb), mp4 (máximo até 10Mb) e quicktime (máximo até 10Mb).
Moreq	Metadados preservação; MPR1.3.19 - Formato_arquivo_digital

## ASSINATURA

Descrição	Contém os dados em Base64 do arquivo de assinatura do documento.
Moreq	Metadados segurança; MSG1.1.4 - Assinatura_digital_assinatura

## Usuário da assinatura

Descrição	Identificador do usuário que realizou a assinatura. (Nesse campo, o MNI salva o ID do usuário, mas no novo serviço podemos gravar o CPF, ou o nome, ou ambos, o que for mais atemporal)
Moreq	Metadados segurança; MSG1.1.5 - Assinatura_digital_validador

## Algoritmo hash

Descrição	Algoritmo utilizado para gerar a assinatura: MD5withRSA, SHA1withRSA, SHA256withRSA ou ASN1MD5withRSA. Atualmente, o mais utilizado é o SHA256withRSA.
Moreq	Metadados preservação; MPR1.3.23 - Algoritmo_checksum

---

### Cadeia de certificado

Descrição	String do certificado. Caso seja do tipo PEM é necessário incluir inclusive as strings "BEGIN CERTIFICATE" e "END CERTIFICATE".
Moreq	Metadados segurança; MSG1.1.3 - Certificado_digital

### Codificação de certificado

Descrição	Codificação utilizada no certificado: "PEM" ou "PkiPath". Registro de que a assinatura digital foi verificada com sucesso.
Moreq	Metadados segurança; MSG1.1.2 - Assinatura_digital_validação

### Data assinatura

Descrição	Data e Hora da assinatura, obedecendo a sequência de ano+dia+mês+minuto+segundo.
Moreq	metadados segurança; MSG1.1 - Assinatura_digital_data.1

### Signatário login

Descrição	Informação referente a signatário que fez uso de login e senha para assinar o documento. (Esse campo é para identificar o usuário que assinou o documento, para os casos de terem utilizado o PJE Token, que é um aplicativo em que servidores do TJ conseguem assinar os documentos com login e senha, utilizando o token do próprio TJDFE para garantir a assinatura.)
Moreq	metadados segurança; MSG1.1.5 - Assinatura_digital_validador

### Assinatura válida

Descrição	Indica se a assinatura foi verificada com sucesso.
Moreq	metadados segurança; MSG1.1.2 - Assinatura_digital_validação

## Apêndice B - Metadados Descritivos Dublin Core

CAMPO	CAMPOS DC CSV	EXEMPLO DE PREENCHIMENTO	ONDE BUSCAR?	PERMITE REPETIÇÃO NO ATOM
<b>Título</b>	dc.title	Processo no 07193054-20.2019.8.07.0003	Nro do Processo Não precisa estar com máscara no número do processo!	Não permite repetição.
<b>Criador (Produtor)</b>	dc.creator	Tribunal de Justiça do Distrito Federal e dos Territórios	Texto fixo	Permite repetição.
<b>Assunto (Ponto de acesso de assunto)</b>	dc.subject[Quebra da Disposição de Texto]	Campo de assunto da tabela Classe; Assuntos principais e secundários; Jurisdição; órgão julgador colegiado; órgão julgador; partes.	(Assuntos cobertos pelo conteúdo)	Permite repetição.
<b>Descrição</b> [Quebra da Disposição de Texto] <b>(Âmbito e conteúdo)</b>	dc.description	-	(Uma descrição resumida sobre o conteúdo)	Não permite repetição.
<b>Editor</b> [Quebra da Disposição de Texto] <b>(Sem correspondência)</b>	dc.publisher		(Pessoa que fez a descrição)	Permite repetição.
<b>Contribuidor / Colaborador</b> [Quebra da Disposição de Texto] <b>(Ponto de acesso de local)</b>	dc.contributor		(Pessoa que contribuiu com o conteúdo)	Permite repetição.

<b>Data</b> [Quebra da Disposição de Texto]	dc.date	2019-03-29	Campo vem dos metadados de gestão: Data de Ajuizamento <dataAjuizamento>	Permite repetição.
			Formato: AAAA-MM-DD ou AAAA-MM ou AAAA	
<b>Tipo do Recurso</b> [Quebra da Disposição de Texto] <b>(Tipo)</b>	dc.type	Texto	(Uma categoria para o tipo de conteúdo) Exemplos: Collection, Dataset, Event, Image, InteractiveResource, MovingImage, PhysicalObject, Service, Software, Sound, StillImage, Text. Campo fixo	Permite repetição.
<b>Formato</b>	dc.format	application/pdf	Mimetype Campo fixo	Não permite repetição.
<b>Identificador do Recurso</b>	dc.identifier	Processo no 07193054-20.2019.8.07.0003	Um identificador único para o seu conteúdo. Pode ser um texto ou um número gerado por um sistema de identificação formal - ou apenas uma URL	Não permite repetição.
<b>Fonte</b>	dc.source		Onde o conteúdo foi inicialmente publicado / originado. Um conteúdo que tem alguma relação intelectual com o conteúdo descrito. Se for conteúdo original, veiculado pela primeira vez e sem nenhuma relação intelectual com outros conteúdos, deve ser a URL para esse conteúdo. Campo fixo	Não permite repetição.

<b>Idioma</b>	dc.language	pt_BR	Em que linguagem o conteúdo foi escrito. Você precisa especificar o código correto da língua. Campo fixo	Permite repetição.
<b>Relação</b>	dc.relation	campo: processos vinculados	**Concatenação das informações.	Não permite repetição.
<b>Abrangência / Cobertura (Ponto de acesso de local)</b>	dc.coverage		Onde o conteúdo está localizado fisicamente. Normalmente inclui localização espacial (nome do lugar ou coordenadas geográficas), período temporal (data) ou jurisdição.	Permite repetição.
<b>Gerenciamento de Direitos Autorais</b>	dc.rights		Informação dos direitos autorais do conteúdo. Pode conter a declaração de Copyright ou outros tipos de licença. Sigilo absoluto, acessível apenas ao magistrado do órgão em que tramita, aos servidores e demais usuários por ele indicado e às partes que provocaram o incidente.	Não permite repetição.

## Apêndice C - Metadados Descritivos ISAD(G)

CAMPO	CAMPOS DC CSV	EXEMPLO DE PREENCHIMENTO	ONDE BUSCAR?	PERMITE REPETIÇÃO NO ATOM
Título	dc.title	Processo no 07193054-20.2019.8.07.0003	Nro do Processo Não precisa estar com máscara no número do processo!	Não permite repetição.
Criador (Produtor)	dc.creator	Tribunal de Justiça do Distrito Federal e dos Territórios	Texto fixo	Permite repetição.
Assunto (Ponto de acesso de assunto)	dc.subject[Quebra da Disposição de Texto]	Campo de assunto da tabela Classe; Assuntos principais e secundários; Jurisdição; órgão julgador colegiado; órgão julgador; partes.	(Assuntos cobertos pelo conteúdo)	Permite repetição.
Descrição[Quebra da Disposição de Texto] (Âmbito e conteúdo)	dc.description	-	(Uma descrição resumida sobre o conteúdo)	Não permite repetição.
Editor[Quebra da Disposição de Texto](Sem correspondência)	dc.publisher		(Pessoa que fez a descrição)	Permite repetição.

Contribuidor / Colaborador[Quebra da Disposição de Texto](Ponto de acesso de local)	dc.contributor		(Pessoa que contribuiu com o conteúdo)	Permite repetição.
Data[Quebra da Disposição de Texto]	dc.date	2019-03-29	Campo vem dos metadados de gestão: Data de Ajuizamento <dataAjuizamento>	Permite repetição.
			Formato: AAAA-MM-DD ou AAAA-MM ou AAAA	
Tipo do Recurso[Quebra da Disposição de Texto] (Tipo)	dc.type	Texto	(Uma categoria para o tipo de conteúdo) Exemplos: Collection, Dataset, Event, Image, InteractiveResource, MovingImage, PhysicalObject, Service, Software, Sound, StillImage, Text. Campo fixo	Permite repetição.
Formato	dc.format	application/pdf	Mimetype Campo fixo	Não permite repetição.
Identificador do Recurso	dc.identifier	Processo no 07193054-20.2019.8.07.0003	Um identificador único para o seu conteúdo. Pode ser um texto ou um número gerado por um sistema de identificação formal - ou apenas uma URL	Não permite repetição.
Fonte	dc.source		Onde o conteúdo foi inicialmente publicado / originado. Um conteúdo que tem alguma relação intelectual com o conteúdo descrito. Se for conteúdo original, veiculado pela primeira vez e sem nenhuma relação intelectual com outros conteúdos, deve ser a URL para esse conteúdo. Campo fixo	Não permite repetição.

---

Idioma	dc.language	pt_BR	Em que linguagem o conteúdo foi escrito. Você precisa especificar o código correto da língua. Campo fixo	Permite repetição.
Relação	dc.relation	campo: processos vinculados	**Concatenação das informações.	Não permite repetição.
Abrangência / Cobertura (Ponto de acesso de local)	dc.coverage		Onde o conteúdo está localizado fisicamente. Normalmente inclui localização espacial (nome do lugar ou coordenadas geográficas), período temporal (data) ou jurisdição.	Permite repetição.
Gerenciamento de Direitos Autorais	dc.rights		Informação dos direitos autorais do conteúdo. Pode conter a declaração de Copyright ou outros tipos de licença. Sigilo absoluto, acessível apenas ao magistrado do órgão em que tramita, aos servidores e demais usuários por ele indicado e às partes que provocaram o incidente.	Não permite repetição.

---

## Apêndice D - Outros Metadados ISAD(G)

Esses metadados foram extraídos na planilha oficial da *Artefactual*, para importação de metadados ISAD(G) à plataforma de difusão AtoM.

CAMPO	CAMPOS ISAD CSV	EXEMPLO DE PREENCHIMENTO	OBSERVAÇÃO
Locais	[isad.placeAccessPoints]	Brasília - DF	Campo igual ao dc.coverage. Onde o conteúdo está localizado fisicamente. Normalmente inclui localização espacial (nome do lugar ou coordenadas geográficas), período temporal (data) ou jurisdição Campo fixo.
Assuntos	[isad.subjectAccessPoints]	Campo de assunto da tabela Classe; Assuntos principais e secundários; Jurisdição; órgão julgador colegiado; órgão julgador; partes.	Campo igual ao dc.subject.
Pessoas / Entidades	[isad.nameAccessPoints]	Nomes das partes ( <a href="#">ds_nome em acl.tb_usuario_login</a> ) Nomes dos representantes processuais ( <a href="#">ds_nome em acl.tb_usuario_login</a> )	Lista de pessoas ou entidades relacionadas ao documento.
Gênero	[isad.genreAccessPoints]	Textual	Campo fixo.
Repositório	[isad.repository]	Tribunal de Justiça do Distrito Federal e dos Territórios	Campo igual ao dc.relation (isLocatedAt). Campo fixo.

Sistema de Escrita	[isad.script]	Português ou latim	Registrar a escrita do documento Campo fixo
Notas sobre o Idioma	[isad.languageNote]		Especificar qualquer tipo de alfabeto, escrita, sistema de símbolos ou abreviaturas utilizadas.
URI do Obj	[isad.digitalObjectURI]		URI enviada do Archivematica para o AtoM, preenchida pelo sistema.
Checksum	[isad.digitalObjectChecksum]	Esses metadados estão dentre aqueles que serão preservados do PJe. Sua extração deverá ser feita juntamente com os demais metadados.	Calculo matemático do documento.
Identificador da Descrição	[descriptionIdentifier]		Identificador único dado pelo usuário da descrição.
Identificador da Instituição	[isad.institutionIdentifier]	Tribunal de Justiça do Distrito Federal e dos Territórios	Identificação da instituição responsável pela descrição. Campo fixo.
Status da Descrição	[isad.descriptionStatus]	Draft	Registrar o estado da descrição, indicando se está em elaboração (rascunho), revisada ou final (concluída). Campo fixo
Nível de detalhamento da Descrição	[isad.levelOfDetail]	Minimal	Indicar o nível de detalhe da descrição (mínima, parcial ou completa) de acordo com as regras ou diretrizes nacionais e internacionais. Campo fixo.
Idioma(s) da Descrição	[isad.languageOfDescription]	pt_BR	Indicar o idioma(s) utilizado para criar a descrição do material arquivístico Campo fixo

Escrita(s) da descrição	[isad.scriptOfDescription]	Português ou Latim	Indicar a escrita utilizada para criar a descrição do material arquivístico. Campo fixo
Fontes	[isad.sources]	Site de pesquisa do PJe: <a href="https://pje.tjdft.jus.br/consultapublica/ConsultaPublica/listView.seam">https://pje.tjdft.jus.br/consultapublica/ConsultaPublica/listView.seam</a>	Site do PJe Campo fixo
Status da publicação	[isad.publicationStatus]	Draft	Preencher entre Rascunho ou Publicado Campo fixo
Nome do Objeto físico	[isad.physicalObjectName]	Existe a possibilidade para aqueles processos judiciais que possuem documentos físicos (por algum motivo, em razão da digitalização para inserção no PJe e os documentos em papel tiverem sido preservados)	Ex.: Caixa Nro 10
Endereçamento do objeto físico	[isad.physicalObjectLocation]		Ex.: Arquivo Judiciário, Módulo 05, Estante 03, Prateleira 01
Tipo do objeto físico	[isad.physicalObjectType]		Exemplos: caixa, ficha, pasta, prateleira, armário, etc.
Identificador alternativo	[isad.alternativeIdentifiers]		Introduza um código de referência anterior, identificador alternativo, ou qualquer outra sequência alfanumérica associada ao registro.
Etiqueta do Id. alternativo	[isad.alternativeIdentifierLabels]		Digite um nome para o campo identificador alternativo que indique sua finalidade e uso. Ex. código de barras, AIP

---

Tipo de data	[isad.eventTypes]	Data de Ajuizamento <dataAjuizamento>	Indique se são data(s) em que os documentos de arquivo foram acumulados pelo produtor no exercício das suas atividades, ou seja, integrados num sistema de arquivo, ou, se são data(s) em que os documentos foram produzidos, ou seja, que inclui a data de cópias, edições, versões, anexos ou originais de peças produzidas antes da sua acumulação ou integração no sistema de arquivo. Escolher entre Produção ou Acumulação.
Data Inicial	[isad.eventStartDates]	Data de Ajuizamento <dataAjuizamento>	Indique a data de início.
Data Final	[isad.eventEndDates]	Data do arquivamento - Para documentos permanentes.	Indique a data final, se houver.
Descrição do evento	[isad.eventDescriptions]	Data da preservação digital no RDC-Arq	
Data Tópica	[isad.eventPlaces]	Tribunal de Justiça do Distrito Federal e dos Territórios.	Local de produção de um documento. Campo fixo.
Cultura	[isad.culture]		Dado automático, preenchido pelo sistema.

---

# Apêndice E - Avaliação das políticas de segurança da informação do Tribunal e o modelo Hipátia

## Política Corporativa de Segurança da Informação

### Introdução

O presente estudo tem por objetivo avaliar o atual estágio de desenvolvimento do barramento tecnológico, integrante do modelo Hipátia de preservação digital, frente ao Guia de Boas Práticas Desenvolvimento, Infraestrutura e Segurança das Aplicações do TJDFT e a Política Corporativa de Segurança da Informação baseada na RESOLUÇÃO 21 DE 8 DE NOVEMBRO DE 2016.

A premissa é seguir as recomendações e melhores práticas definidas pelo Guia em relação aos requisitos de segurança da informação, arquitetura, experiência dos usuários, ferramentas, infraestrutura e operação, obedecendo as melhores práticas de mercado, antes e durante a implantação, o desenvolvimento e a manutenção das aplicações.

O Ibict considera de extrema importância padronizar os procedimentos de desenvolvimento da ferramenta em relação aos critérios do TJDFT para implantação, manutenção, sustentação e eventual desligamento, de forma a garantir isonomia nos princípios de segurança de sistemas durante todo o ciclo de vida da aplicação.

## RESOLUÇÃO 21 DE 8 DE NOVEMBRO DE 2016

Dispõe, no âmbito do Tribunal de Justiça do Distrito Federal e dos Territórios - TJDFT, sobre o Sistema de Gestão de Segurança da Informação - SGSI e a Política Corporativa de Segurança da Informação - PCSI.

Art. 1º O objetivo desta Resolução é estabelecer, implementar, manter e melhorar, continuamente, mecanismos e controles para promover a gestão da segurança da informação, de forma a garantir o direito de acesso previsto em Lei, a proteção de dados, informações e conhecimentos gerados e custodiados, a redução de riscos de ocorrência de perdas, alterações e acessos indevidos, preservando a disponibilidade, integridade, confiabilidade e autenticidade das informações, no âmbito do Tribunal de Justiça do Distrito Federal e Territórios - TJDFT.

Art. 2º As normas desta Resolução aplicam-se a todas as autoridades, servidores, colaboradores e quaisquer pessoas que tenham acesso a informações do TJDFT.

## Avaliação em relação ao Guia de Boas Práticas

### Repositório de Código e Versionamento

O código fonte da aplicação deve ser versionado no Gitlab institucional, preferencialmente. O controle das versões deve estar organizado com os branches e as tags adequadas. É necessário verificar quais usuários têm acesso ao projeto na ferramenta. É necessário verificar a visibilidade do projeto, sendo mais seguro defini-lo como PRIVATE.

Critérios	Situação	Obs
Versionamento do Git institucional	OK	
Nomenclatura de branches	NÃO OK	Revisando de acordo com política interna
Nomenclatura de tags	NÃO OK	Revisando de acordo com política interna
Visibilidades do projeto	OK	
Controle de acesso ao projeto	OK	

### Autenticação Única

A ferramenta de SSO possibilita o controle de acesso único a vários sistemas, sejam eles conectados ou independentes. Recomenda-se a interação entre as aplicações e o SSO institucional para fins de login.

Critérios	Situação	Obs
integrada ao SSO institucional	não se aplica	
Cadastros locais de usuários e senhas não devem ser usados	não se aplica	
Mecanismos legados	não se aplica	

### Linguagens de Programação e Tecnologias

As tecnologias e linguagens em que a aplicação será desenvolvida devem ser modernas e evoluídas por seus mantenedores, amplamente aceitas e validadas pela comunidade e deve-se observar o domínio da equipe técnica local.

Critérios	Situação	Obs
Front-Ends: Angular, Vue, React, Ember	não se aplica	
Back-Ends: Spring Boot, Laravel, Django, Nodejs, Wildfly,	não se aplica	<b>utiliza-se PYTHON</b>
Plataformas: Jboss EAP, <b>Kafka</b> , RabbitMQ, <b>Openshift</b>	OK	
Cache: Redis	OK	
Bancos de dados: PostgreSQL, Oracle, MongoDB	não se aplica	<b>utiliza-se MySQL (BD das soluções Arquivemática e Atom)</b>
Indexação: <b>ElasticSearch</b> , MeiliSearch	OK	
Autenticação: Keycloak	não se aplica	

## Revisão de código

A revisão de código visa garantir a qualidade por meio da conferência no código por outros desenvolvedores, de forma consciente e sistemática, para verificar se há erros ou falhas.

O código só estará pronto para ser implantado caso seja aprovado, segundo regras definidas, como por exemplo a necessidade de teste automatizado.

Critérios	Situação	Obs
Usar uma wiki, que é um documento formado por todas as atividades relacionadas ao projeto;	OK	utiliza-se a wiki e as issues do próprio Git do TJDFT
Usar checklists;	OK	Passo a passo de instalação e configurações
Estar atento às novas ameaças, para a segurança da aplicação;	OK	Por meio de grupos oficiais do Archivematica e AtoM
Usar a automação em conjunto com o trabalho manual, no intuito de identificar padrões e falhas recorrentes;	Não OK	reavaliando fluxo de trabalho
Usar ferramentas para otimizar o processo, como o Android Lint, o Findbugs e o Checkstyle.	Não se aplica	desenvolvimento em python

## Padrão de Arquitetura

Por padrão de arquitetura entende-se a adoção de uma solução anteriormente estudada, testada e documentada para um problema.

Critérios	Situação	Obs
Arquitetura Orientada a Mensagens (Eventos)	OK	
desacoplamento entre aplicações	OK	
Evitar a construção de sistemas monolíticos	OK	
Comunicações entre sistemas devem ser feitas via serviço	OK	
Aplicação deve ter seu próprio banco de dados	OK	

## Testes Automatizados

Os testes automatizados são códigos implementados especificamente para testar o funcionamento da aplicação. Por meio desses, os programas tornam-se auto-testáveis.

Critérios	Situação	Obs
Testes automatizados dos tipos unitários	não OK	A ser implementado até o final do projeto
Cobertura mínima de 70%	não OK	A ser implementado até o final do projeto
Pipeline deve estar integrado com a ferramenta de qualidade	não OK	A ser implementado até o final do projeto

## Testes de Carga e Testes de Stress

O teste de carga avalia o volume de transações, acessos simultâneos ou usuários que um servidor/software/sistema suporta. O teste de stress pretende descobrir em que ponto a aplicação ultrapassa a sua capacidade de funcionamento.

Critérios	Situação	Obs
Teste de carga	não OK	avaliando metodologia, previsão de entrega até o final do projeto
Teste de stress	não OK	avaliando metodologia, previsão de entrega até o final do projeto

## Implantação de Software Terceiro

Uma aplicação desenvolvida externamente - por contratação, cessão ou outro modelo - é necessária, primeiramente, a indicação formal de um indivíduo ou setor como responsável técnico. Esta designação deve compreender todo o ciclo de vida da solução - desde sua implantação, sustentação, eventuais evoluções e, fatalmente, seu desligamento. A necessidade de atualizações de versões, sempre que disponíveis, é medida crucial para a integridade e segurança do software. Da mesma forma, as tecnologias de suporte devem seguir compatíveis, com suas respectivas modernizações.

Critérios	Situação	Obs
Técnico responsável pelo acompanhamento	OK	

## Monitoramento das Aplicações

Percebe-se a necessidade de manter as aplicações e suas funcionalidades com alta disponibilidade. Para tanto, devem ser utilizados métodos para monitoramento da performance dos ambientes, bem como, sempre que possível, é recomendável a supervisão de funcionalidades críticas ou falhas frequentes conhecidas.

Critérios	Situação	Obs
Monitoramento de níveis 1, 2 e 3	em avaliação	previsão de entrega até o final do projeto
Ferramentas, tais como zabbix ou sentry	em avaliação	previsão de entrega até o final do projeto

## Documentação

É essencial para auxiliar usuários e programadores sobre as rotinas que estão contidas no software, facilitando o uso e o desenvolvimento de futuras evoluções.

Critérios	Situação	Obs
Documentação técnica no wiki do Gitlab	em avaliação	Atualizando wiki no Git do TJDFT

## Integração contínua e Entrega contínua

A aplicação da metodologia ágil de desenvolvimento, faz-se necessário adotar práticas alinhadas à aceleração de entregas de novas versões. O modelo de MVPs (menor produto viável) exige maior frequência de atualizações, com a adoção de um método de trabalho que seja flexível e inteligente.

As etapas devem estar separadas, contendo passos de desenvolvimento, segurança e operação (DEVSECOPs) e sugere-se as seguintes: build -> test -> quality/security -> external.

Critérios	Situação	Obs
Pipeline configurado com as etapas adequadas no Gitlab institucional	OK	
<b>test</b> são executados os testes automatizados	NÃO OK	em avaliação
<b>quality/security</b> a varredura do SonarQube	NÃO OK	em avaliação
<b>external</b> está a integração com a plataforma de gerenciamento de contêineres	OK	
<b>release</b> automático no repositório	OK	
<b>deploy</b> automático em produção	OK	

## Balanceamento de Carga

Importante uma boa distribuição do potencial tecnológico para garantir que todas as aplicações funcionem plenamente e atendam aos seus propósitos.

Critérios	Situação	Obs
Balanceamento de carga dos sistemas sobre múltiplas instâncias	em avaliação	<ul style="list-style-type: none"><li>• Estudo sobre o uso de pipelines no Archivematica.</li><li>• Não se aplica ao barramento hipátia.</li></ul>

## Contêineres

O ambiente de contêiner é uma tecnologia que armazena todas as informações necessárias para que uma aplicação possa ser executada ou disponibilizada de forma isolada.

Critérios	Situação	Obs
Utilizar imagens de origem segura e minimalistas (com o mínimo de recursos necessários);	OK	
Subir o serviço do contêiner sem o uso do usuário root (administrador);	OK	
Utilizar volumes para persistência de dados de aplicação;	OK	
Escolher protocolo de compartilhamento de arquivos adequado (NFS, ISCSI, Objeto), evitando o uso do CIFS (somente para exceções);	OK	

---

### Atualização de Tecnologias Utilizadas

Toda a estrutura de linguagens, servidores de aplicação etc. precisa ser constantemente atualizada, tanto em suas versões mais recentes quanto para novas ferramentas que podem surgir.

Critérios	Situação	Obs
Atualizar as tecnologias usadas nas aplicações em produção	OK	

### Identificação e Tratamento de Vulnerabilidades

É fundamental a existência de um fluxo sistematizado, e de ação contínua, com foco em medidas de segurança das aplicações e na gestão de riscos. Some-se a isso a necessidade do uso cotidiano de critérios rígidos para o desenvolvimento, com observação das melhores práticas.

Critérios	Situação	Obs
Pipeline de sua aplicação deve estar integrado à ferramenta de qualidade	não OK	avaliar em qual ferramenta

### Armazenamento das Credenciais

As credenciais, chaves de acessos, keys, passphrases, tokens, entre outros, devem ser protegidos e armazenados da forma mais segura possível, evitando exposição em códigos e arquivos de configurações.

Critérios	Situação	Obs
Utilizar em suas aplicações, sempre que possível, as soluções de gerenciamento de credenciais homologadas e em uso pela SETI	não OK	ENV para OPENSIFT

### Publicação

Publicar uma aplicação é torná-la acessível aos usuários, por meio de uma URL, na rede interna ou na internet.

Critérios	Situação	Obs
Seguir o fluxo interno de requisição para publicação	OK	

## Itens em conformidade com a política de segurança

ITENS	Obs
Repositório de Código e Versionamento	
Versionamento do Git institucional	
Visibilidades do projeto	
Controle de acesso ao projeto	
Linguagens de Programação e Tecnologias	
plataformas: Jboss EAP, <b>Kafka</b> , RabbitMQ, <b>Openshift</b>	
Cache: Redis	
Indexação: <b>ElasticSearch</b> , MeiliSearch	
Revisão de código	
Usar uma wiki, que é um documento formado por todas as atividades relacionadas ao projeto;	
Usar checklists;	
Estar atento às novas ameaças, para a segurança da aplicação;	
Padrão de Arquitetura	
Arquitetura Orientada a Mensagens (Eventos)	
desacoplamento entre aplicações	
Evitar a construção de sistemas monolíticos	
Comunicações entre sistemas devem ser feitas via serviço	
Aplicação deve ter seu próprio banco de dados	
Implantação de Software Terceiro	
Técnico responsável pelo acompanhamento	
Integração contínua e Entrega contínua	
Pipeline configurado com as etapas adequadas no Gitlab institucional	
<b>external</b> está a integração com a plataforma de gerenciamento de contêineres	
<b>release</b> automático no repositório	
<b>deploy</b> automático em produção	

Contêineres	
Utilizar imagens de origem segura e minimalistas (com o mínimo de recursos necessários);	
Subir o serviço do contêiner sem o uso do usuário root (administrador);	
Utilizar volumes para persistência de dados de aplicação;	
Escolher protocolo de compartilhamento de arquivos adequado (NFS, ISCSI, Objeto), evitando o uso do CIFS (somente para exceções);	
Atualização de Tecnologias Utilizadas	
Atualizar as tecnologias usadas nas aplicações em produção	
Publicação	
Seguir o fluxo interno de requisição para publicação	

## Itens em estudo para aplicação das conformidades

ITENS	Obs
Repositório de Código e Versionamento	Revisando de acordo com a política interna
Nomenclatura de branches	Revisando de acordo com política interna
Nomenclatura de tags	Revisando de acordo com política interna
Revisão de código	reavaliando fluxo de trabalho
Usar a automação em conjunto com o trabalho manual, no intuito de identificar padrões e falhas recorrentes;	reavaliando fluxo de trabalho
Testes Automatizados	em análise a capacitação da equipe
Testes automatizados dos tipos unitários	Em análise a capacitação da equipe
Cobertura mínima de 70%	Em análise a capacitação da equipe
Pipeline deve estar integrado com a ferramenta de qualidade	Em análise a capacitação da equipe
Testes de Carga e Testes de Stress	Avaliando metodologia
Teste de carga	avaliando metodologia, previsão de entrega até o final do projeto

Teste de stress	avaliando metodologia, previsão de entrega até o final do projeto
Monitoramento das Aplicações	previsão de entrega até o final do projeto
Monitoramento de níveis 1, 2 e 3	previsão de entrega até o final do projeto
Ferramentas, tais como zabbix ou sentry	previsão de entrega até o final do projeto
Documentação	Atualizando wiki no GIT do TJDFE
Documentação técnica no wiki do Gitlab	Atualizando wiki no git do TJDFE
Integração contínua e Entrega contínua	Em avaliação para implementação
<b>test</b> são executados os testes automatizados	em avaliação
<b>quality/security</b> a varredura do SonarQube	em avaliação
<b>Balanceamento de Carga</b>	Em avaliação para a produção
Balanceamento de carga sobre múltiplas instâncias	Estudo sobre o uso de pipelines no Archimatica. Não se aplica ao barramento hipátia.
Identificação e Tratamento de Vulnerabilidades	A definir a ferramenta
Pipeline de sua aplicação deve estar integrado à ferramenta de qualidade	avaliar em qual ferramenta
Armazenamento das Credenciais	Definição de uso do ENV para openshift
Utilizar em suas aplicações, sempre que possível, as soluções de gerenciamento de credenciais homologadas e em uso pela SETI	ENV para OPENSIFT

## Considerações finais

A análise do guia de boas práticas foi fundamental para avaliação do desenvolvimento do barramento Hipátia e das integrações com Archimatica e AtoM. A equipe de desenvolvimento do Ibict já iniciou a implementação de alguns dos itens e estuda a incorporação de metodologias para alinhar o código do sistema às conformidades apontadas no guia.