



Digitalização de documentos e processos

Superintendência Regional de Goiânia / Grupo de Trabalho de
Implantação do Sistema Eletrônico de Informações—SEI

Goiânia / Rio de Janeiro
2017

Sumário

Apresentação	2
Escopo da Cartilha de Digitalização de Documentos e Processos	3
Não Escopo.....	3
Parâmetros para a Digitalização	3
Orientações Gerais para Digitalização	4
Preparação.....	5
Captura da Imagem.....	5
Conferência	6
Indexação.....	7
Nomenclatura dos Arquivos Digitais.....	7
Finalização	8
Recomendação Final	9
APÊNDICE 1 - Procedimento de digitalização utilizando o escâner Scanjet 5590 HP	10
APÊNDICE 2 - Procedimento de conversão de arquivos PDF para PDF com OCR	18
APÊNDICE 3 - Procedimento de armazenamento dos arquivos em PDF com OCR	21
ANEXO 1 - Ficha de controle de digitalização	25
Referências	27

Apresentação

Essa cartilha tem como objetivo orientar e prover **recomendações mínimas necessárias** quanto às atividades relativas à digitalização de documentos e processos e à Ilha de digitalização, a fim de garantir que os arquivos gerados sejam representantes digitais com qualidade arquivística, fidelidade ao documento original e capacidade de interoperabilidade.

Aqui, você encontrará informações para realização da digitalização dos **processos em tramitação** na CPRM, ou seja, o Manual define os procedimentos de digitalização a serem adotados na CPRM para os **processos ativos** que ainda permanecem em papel.

Todo o procedimento de uso do escâner e requisitos para digitalização do documento no formato recomendável para ser anexado no SEI – CPRM está disponível neste manual.

Quando houver a necessidade de tramitação de algum processo em papel, ou a necessidade de inclusão de algum documento em um dos processos que ainda estão em papel, deverá ser feita, imediatamente, a conversão do processo para o formato eletrônico no SEI-CPRM.

Qualquer dúvida, entre em contato pelo e-mail: sei@cprm.gov.br.

Escopo da Cartilha de Digitalização de Documentos e Processos

- Orientar os setores da CPRM para a digitalização de processos e documentos;
- Preparação dos documentos para o procedimento de digitalização;
- Digitalização de documentos em **tamanho A4**;
- Digitalização com OCR (Optical Character Recognition¹);
- Digitalização até o penúltimo volume dos processos em tramitação (correntes).

Não Escopo

- Auditoria dos documentos constantes no processo, como a numeração das folhas e conformidade documental;
- Restauração de documentos deteriorados;
- Digitalização de documentos maiores que A4 ou fora do padrão de conservação para os modelos de escâner disponíveis;
- Indexação para upload no SEI;
- Digitalização do último volume de processos em tramitação ou de processos não finalizados com apenas 1 (um) volume. Estes processos terão a tramitação em papel até a data de efetiva implantação do SEI e, portanto, serão digitalizados e carregados para o sistema a partir de 02/05/2017 por equipes específicas, responsáveis por fazer o upload dos volumes dos processos completos, acrescentando às imagens recém-digitalizadas (do último volume) as imagens pré-existentes (até o penúltimo volume).

Parâmetros para a Digitalização

É necessária a verificação de algumas configurações antes de utilizar escâneres para o procedimento de digitalização. Abaixo, segue uma tabela de ajustes principais conforme o tipo de documento a ser digitalizado, de modo que o arquivo gerado após o processo de digitalização seja apropriado para uso no processo eletrônico (Tabela 1).

Esses ajustes dependem do aparelho de escâner e software utilizado para a captura das imagens, podendo ser realizados em um ou ambos dos casos, de acordo com a infraestrutura de tecnologia da informação (TI) disponível em cada unidade.

¹ PDF com reconhecimento óptico de caracteres (pesquisável). Através do OCR é possível digitalizar uma folha de texto impresso e obter um arquivo de texto onde se pode buscar por palavras.

Tabela 1 - Ajustes de configurações necessários antes da digitalização

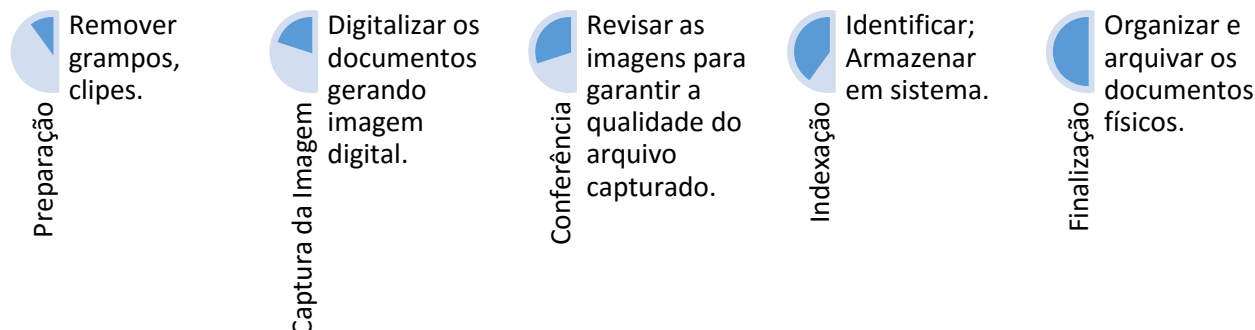
Resolução recomendada	Cor	Tipo original	Formato de arquivo	Reconhecimento de texto (OCR)
Tipo de documento: textos impressos, sem ilustração, em preto e branco				
300 dpi ou ppi	Monocromático (preto e branco)	Texto	PDF/A	Ativado
Tipo de documento: textos impressos, com ilustração em preto e branco				
300 dpi ou ppi	Escala de cinza	Texto/Foto	PDF/A	Ativado
Tipo de documento: textos impressos, com ilustração e cor				
300 dpi ou ppi	RGB (colorido)	Texto/Foto	PDF/A	Ativado
Tipo de documento: fotografias, mapas (com pouco ou nenhum texto)				
300 dpi ou ppi	RGB (colorido)	Foto	PNG	Desativado
Tipo de documento: plantas, mapas				
300 dpi ou ppi	Monocromático (preto e branco)	Texto/Foto	PNG	Desativado

Fonte: Adaptado de Ministério da Fazenda, 2017; CONSELHO NACIONAL DE ARQUIVOS, 2010.

Orientações Gerais para Digitalização

- Processos com até 30 páginas poderão ser digitalizados e salvos documento a documento, por tipo de documento.
- Processos com mais de 30 páginas poderão ser digitalizados e salvos em documento único contendo todos os documentos anexos ao processo.
- Processos com mais de um volume poderão ser digitalizados em arquivos separados, sendo um arquivo para cada volume ou um arquivo para cada aditivo, dependendo do teor do processo.
- Cabe ressaltar que, nos casos em que a digitalização for inviável, devido às características do suporte ou estado de conservação do documento, é possível a tramitação em meio físico, mediante registro e atualização do andamento do trâmite no SEI-CPRM.
- O servidor deverá se responsabilizar pela qualidade da imagem ao digitalizar a documentação. Dessa forma, é necessário que seja realizada conferência, após a digitalização, de todas as páginas a fim de constatar se foram realmente digitalizadas e estão legíveis.
- Deve ser preenchida a ficha “Controle de digitalização dos processos”, em anexo (Anexo 4).

PROCESSO DE DIGITALIZAÇÃO



Preparação

Todos os documentos e processos devem ser preparados previamente, a fim de torná-los aptos à digitalização:

- Desmontar processo;
- Remover elementos que prejudiquem a digitalização (grampos, clipes);
- Desamassar documentos dobrados ou amassados;
- Ordenar os documentos para digitalização (conferindo se todas as folhas estão na mesma direção).

Captura da Imagem

A captura consiste na conversão de um documento ou processo físico para imagem em formato digital.

Os modelos de escâneres podem ser utilizados para a captura da imagem em formato digital e que já possuem software conversor para OCR são:

- HP Scanjet 5590.

Os procedimentos de digitalização utilizando este modelo de escâner estão descritos no Anexo 1. Nem todos os escâneres possuem conversor para OCR embutido em seus aplicativos. Caso não haja a possibilidade de utilização do modelo de escâner citado acima, o software OCR Server está instalado em rede para realizar a conversão do arquivo em PDF comum para PDF pesquisável (OCR). O Anexo 2 contém instruções de como utilizá-lo para converter documentos digitalizados. Solicite a configuração da pasta “ocr-server” no seu computador através do Service Desk para poder realizar esta ação.



A digitalização de documentos impressos em papel reciclado também exige maior cuidado, pois esse tipo de papel tem baixa qualidade, o que pode comprometer a legibilidade. Para esse tipo de documento, recomenda-se digitalizar em resolução de 300 dpi, cor monocromática (escala de cinza) e aumentar o brilho para tornar a imagem mais clara.



Documentos com textos escritos à mão, com letras de tamanho muito pequeno, em idiomas diferentes do especificado, com fundo colorido, sombras ou sobre ilustrações e gráficos, e com letras decoradas podem não ser reconhecidos corretamente no processamento para o OCR.

Com os ajustes corretos, o arquivo digitalizado terá o menor tamanho possível e a melhor qualidade de imagem e de processamento de OCR, o que facilitará a pesquisa de texto no documento e o seu armazenamento em meio eletrônico.



Conferência

O controle da qualidade deve ser realizado por quem efetuou a digitalização, a fim de verificar a fidelidade da imagem capturada em relação ao documento original, ou seja, se todas as páginas foram digitalizadas e estão legíveis.

- Verificar se os arquivos digitalizados estão de acordo com os padrões apresentados na Tabela 1;
- Se o escâner utilizado não possuir a opção de reconhecimento de texto (ORC), deve-se submeter os arquivos a software que realiza essa ação²;
- Caso a qualidade do arquivo digital não atenda a algum dos requisitos apresentados, a digitalização deve ser realizada novamente, ajustando-se as configurações, até que atinja o padrão esperado.

² Abrir chamado “Helpdesk” (informática) para configuração do software OCR Server no seu computador.

Indexação

Os arquivos digitalizados necessitam ser devidamente organizados para, então serem inseridos no Sistema Eletrônico de Informações – SEI.

Nomenclatura dos Arquivos Digitais

Os arquivos digitais devem receber nomes que os identifiquem e que facilitem a sua localização e pesquisa. Para isso, é fundamental que essa identificação seja de tamanho pequeno, única e estruturada. Deve-se utilizar um padrão de nomenclatura que não permita existir no acervo dois arquivos com nomes iguais.

Para documentos de até 30 páginas:

Recomenda-se a utilização da seguinte estrutura padrão de nomenclatura dos arquivos digitais:

- Parte 1: espécie documental (sigla, nomenclatura do documento);
- Parte 2: número do documento ou processo;
- Parte 3: número do volume ou sequencial.

Exemplos:

Parte 1	Parte 2	Parte 3
memo003	6500572015	v1
Nome do documento e numeração	Nº Processo	Nº Volume

- Memo 003

memo003_6500572015_v1

- Nota Técnica 033

nt033_6500572015_v2



Não devem ser utilizados caracteres especiais nem letras maiúsculas na nomenclatura desses arquivos, tais como acentos, cedilha, asterisco, parênteses, colchetes, etc. O uso desses caracteres dificulta a interoperabilidade dos arquivos digitais.

Para processos com mais de 30 páginas em documentos:

Recomenda-se a utilização da seguinte estrutura padrão de nomenclatura dos arquivos digitais com mais de 30 páginas:

- Parte 1: número do documento ou processo;
- Parte 2: número do volume ou sequencial.

Exemplos:

Parte 1	Parte 2
6500572015	v1
Nº Processo	Nº Volume

- Processo 6500572015 com 4 volumes

6500572015_v1
6500572015_v2
6500572015_v3
6500572015_v4

Digitalize os volumes em separado e os nomeie em arquivos diferentes

Finalização

Nesta etapa, a documentação digitalizada deve ter seu trâmite físico encerrado e ser reorganizada e arquivada.

O encerramento do processo em papel a abertura do correspondente processo eletrônico deve ser realizado por meio do [Termo de Encerramento de Trâmite Físico de Processo](#), que deve conter, no mínimo:

- Identificação da seção/setor que realizou a digitalização;
- Quantidade de volumes e folhas do processo original em papel;
- Responsável e assinatura.

Exemplo:

CPRM

TERMO DE ENCERRAMENTO DE TRÂMITE FÍSICO DE PROCESSO

Processo nº: _____

O referido processo foi devidamente convertido do suporte físico para eletrônico e registrado no Sistema Eletrônico de Informação - SEI, mantendo o mesmo número do processo físico (NUP) e mesmo(s) interessado(s). Fica encerrada a tramitação do processo em suporte físico.

O processo físico ficará armazenado nesta unidade para cumprimento do prazo de guarda em fase corrente e, posteriormente, será encaminhado para o Arquivo.

Para fins de registro, o processo originalmente em suporte físico era composto de:

- Quantidade de Folhas: _____
- Quantidade de Volumes: _____
- Quantidade de Anexos: _____

O processo eletrônico resultante da presente conversão ficou composto da seguinte forma:

- Quantidade de Arquivos: _____
- Observações/Ocorrências: _____

Os arquivos PDF oriundos da digitalização da documentação em suporte físico (papel) foram devidamente submetidos a procedimento de conferência e autenticação.

Seção/Setor responsável pela conversão: _____
Data de conclusão do procedimento de conversão: _____
Responsável e Assinatura: _____

O **Termo de Encerramento de Trâmite Físico de Processo** deve ser inserido no SEI ao final do processo digitalizado (cópia do impresso, assinado eletronicamente), assim como no processo em papel (documento original, que deve ser armazenado corretamente).

Recomendação Final

Ao finalizar a digitalização de todos os documentos físicos, caso ocorra a tramitação de novos documentos os mesmos devem ser digitalizados conforme procedimentos já mencionados até a definitiva implantação do SEI.



Para esclarecimentos sobre o conteúdo deste informe, entre em contato com o Grupo de trabalho SEI da sua unidade e/ou envie um e-mail para sei@cprm.gov.br.

APÊNDICE 1

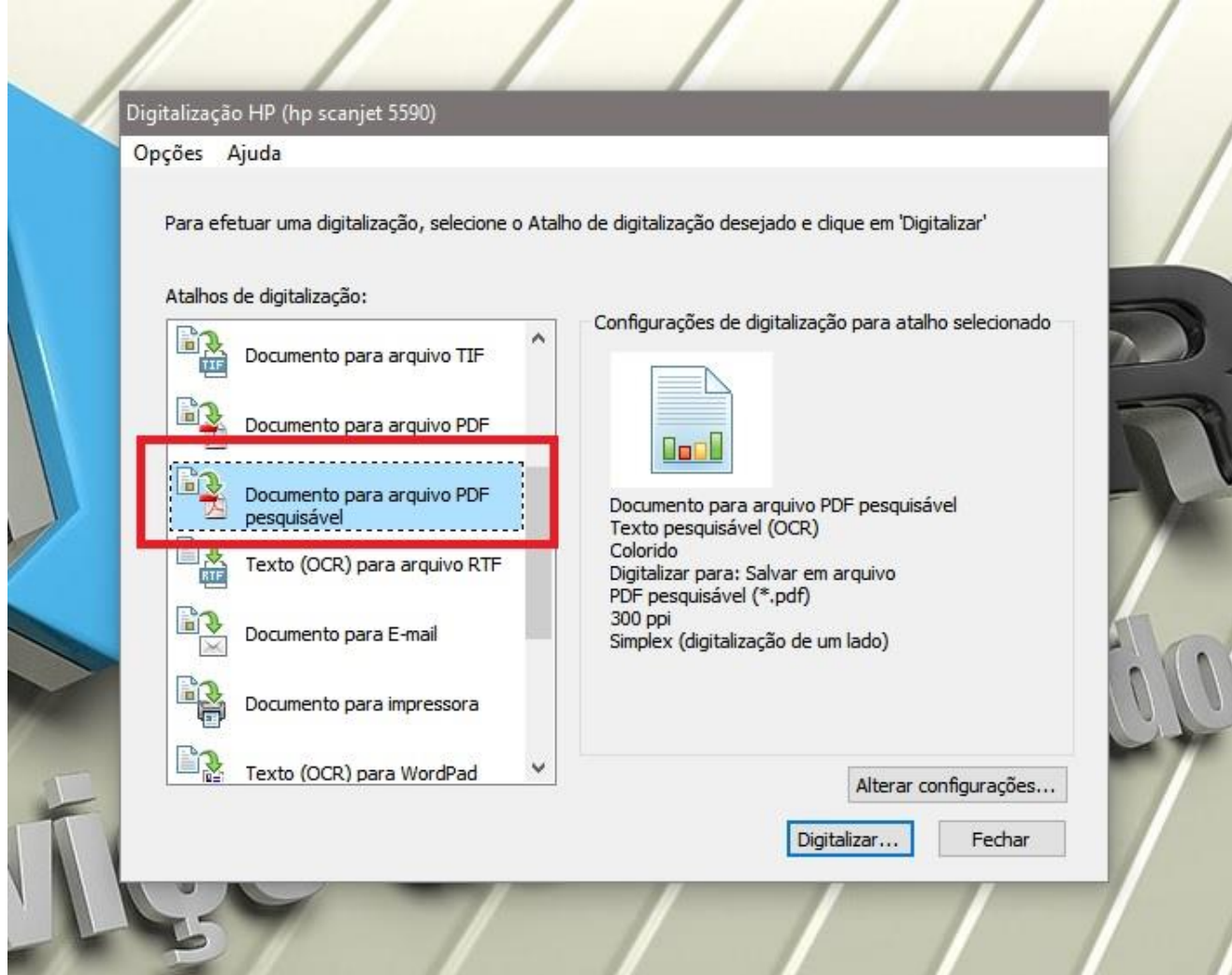
Procedimento de digitalização utilizando o escâner Scanjet 5590 HP

Abaixo, segue o passo a passo de digitalização de documentos a serem incluídos no SEI utilizando o escâner de mesa HP modelo Scanjet 5590.

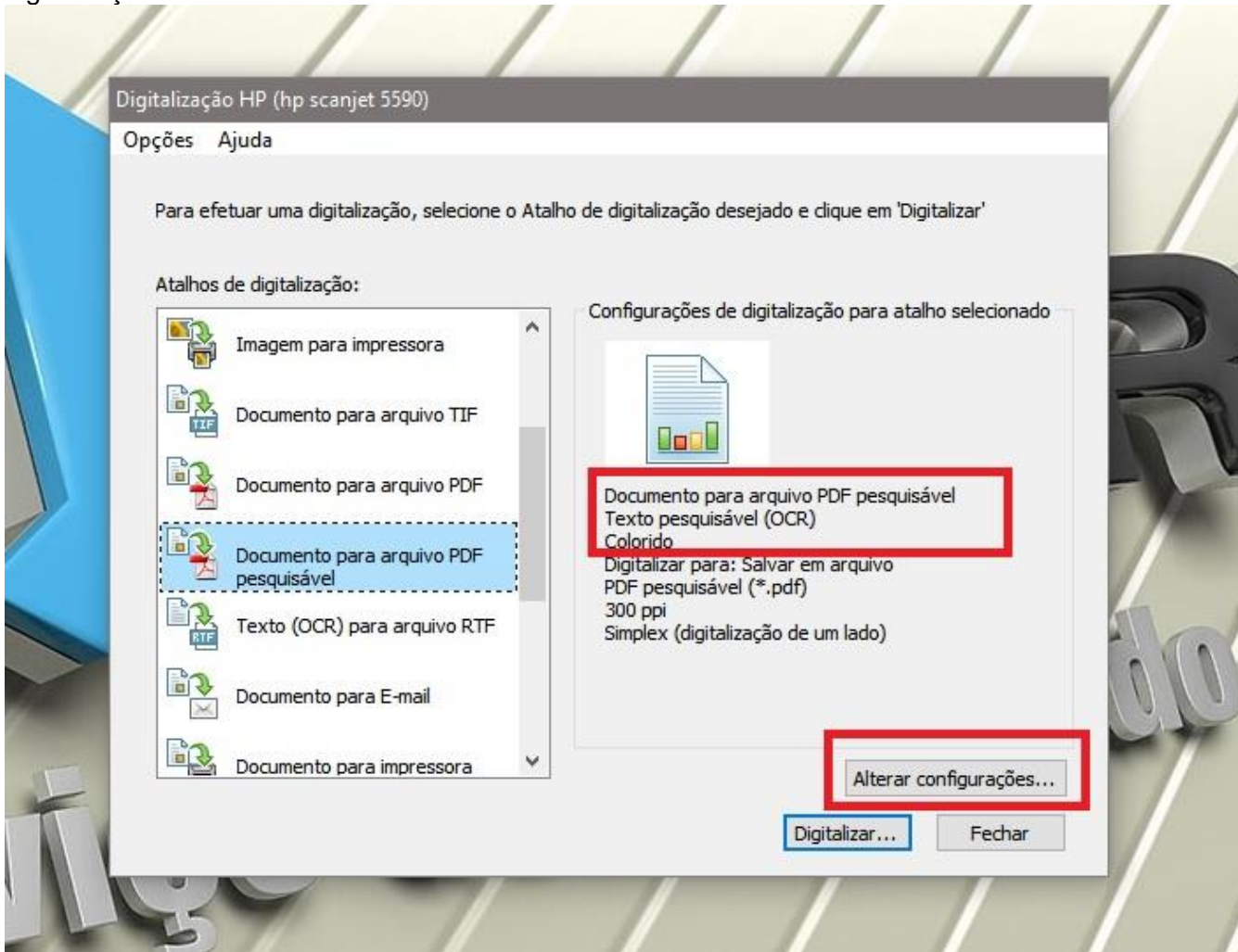
1. Na área de trabalho de seu computador, abra o software “Digitalização HP (5590)”, como mostrado na imagem abaixo.



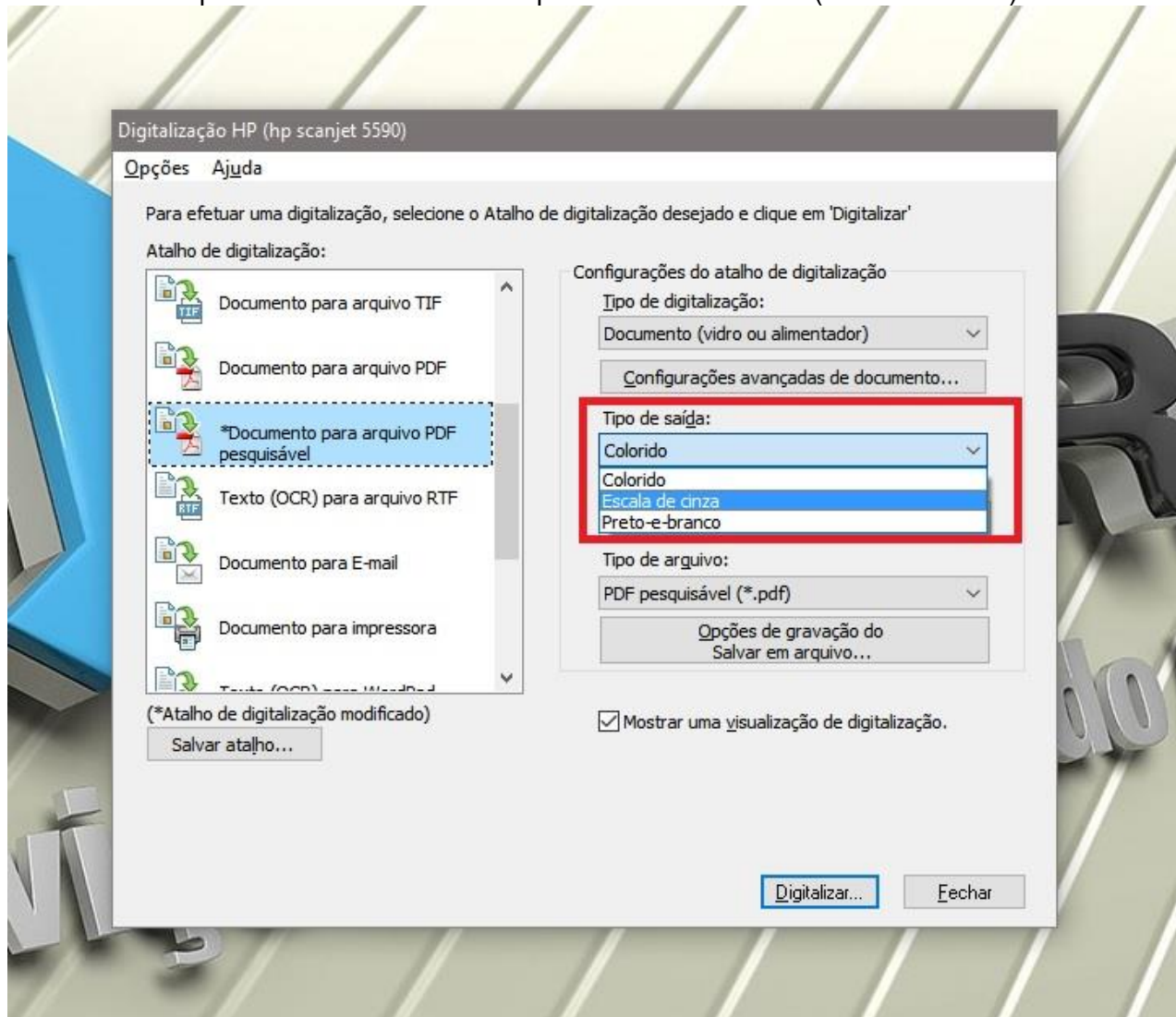
2. A janela inicial do aplicativo se abrirá. Nessa janela, note que há diversas opções de digitalização que podem ser selecionadas. A opção de digitalização para PDF com OCR e que deve ser selecionada é “Documento para arquivo PDF pesquisável”.



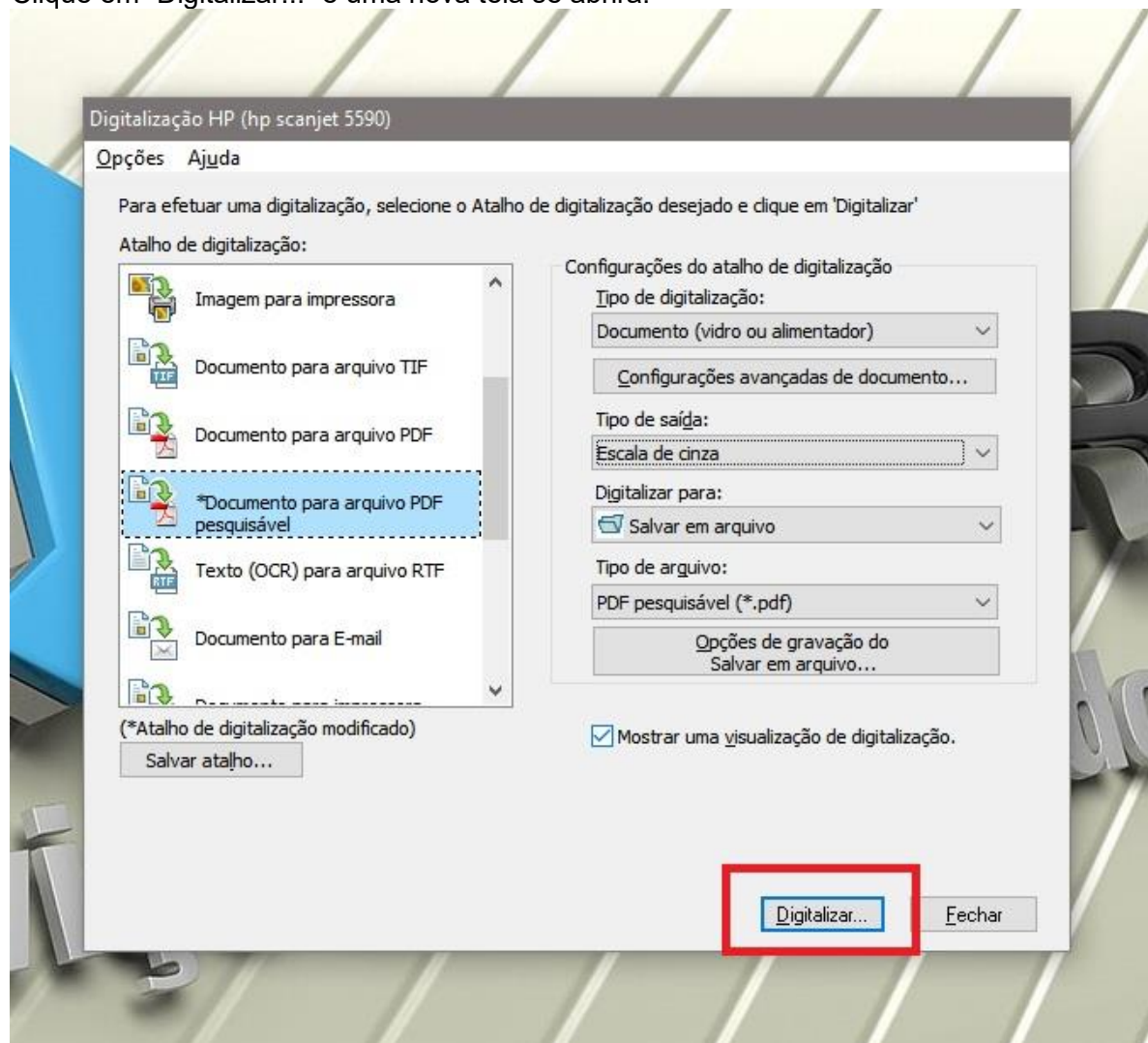
3. Note, na imagem abaixo, que aparecem as configurações do documento que será salvo, tais como Texto pesquisável (OCR), Colorido, 300 ppi (ou dpi), etc. Clique em “Alterar configurações...” para ajustar as configurações do arquivo que será criado a partir da digitalização.



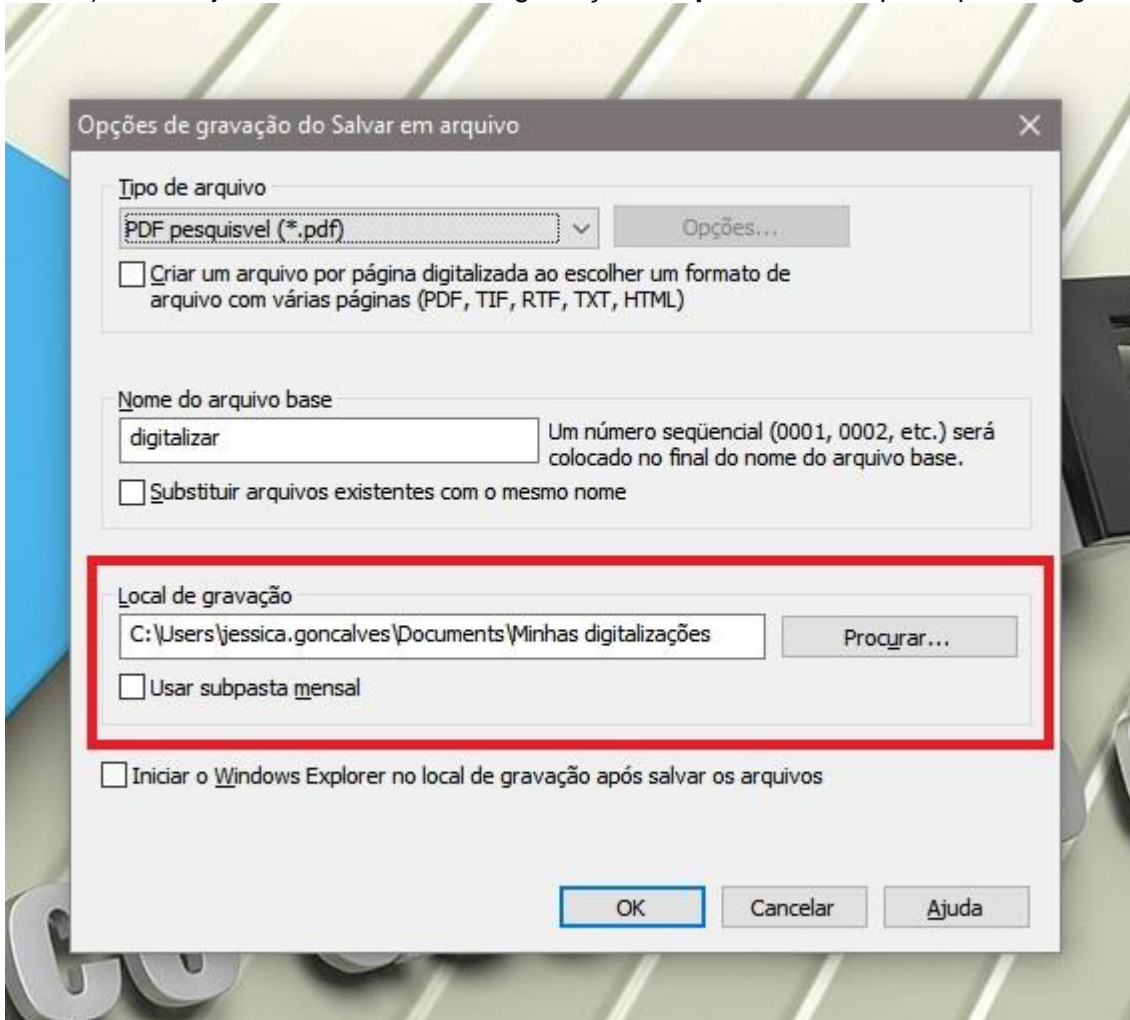
4. Selecionar o “Tipo de saída” do documento para “Escala de cinza” (monocromático) ao invés de “Colorido”, como na imagem abaixo.



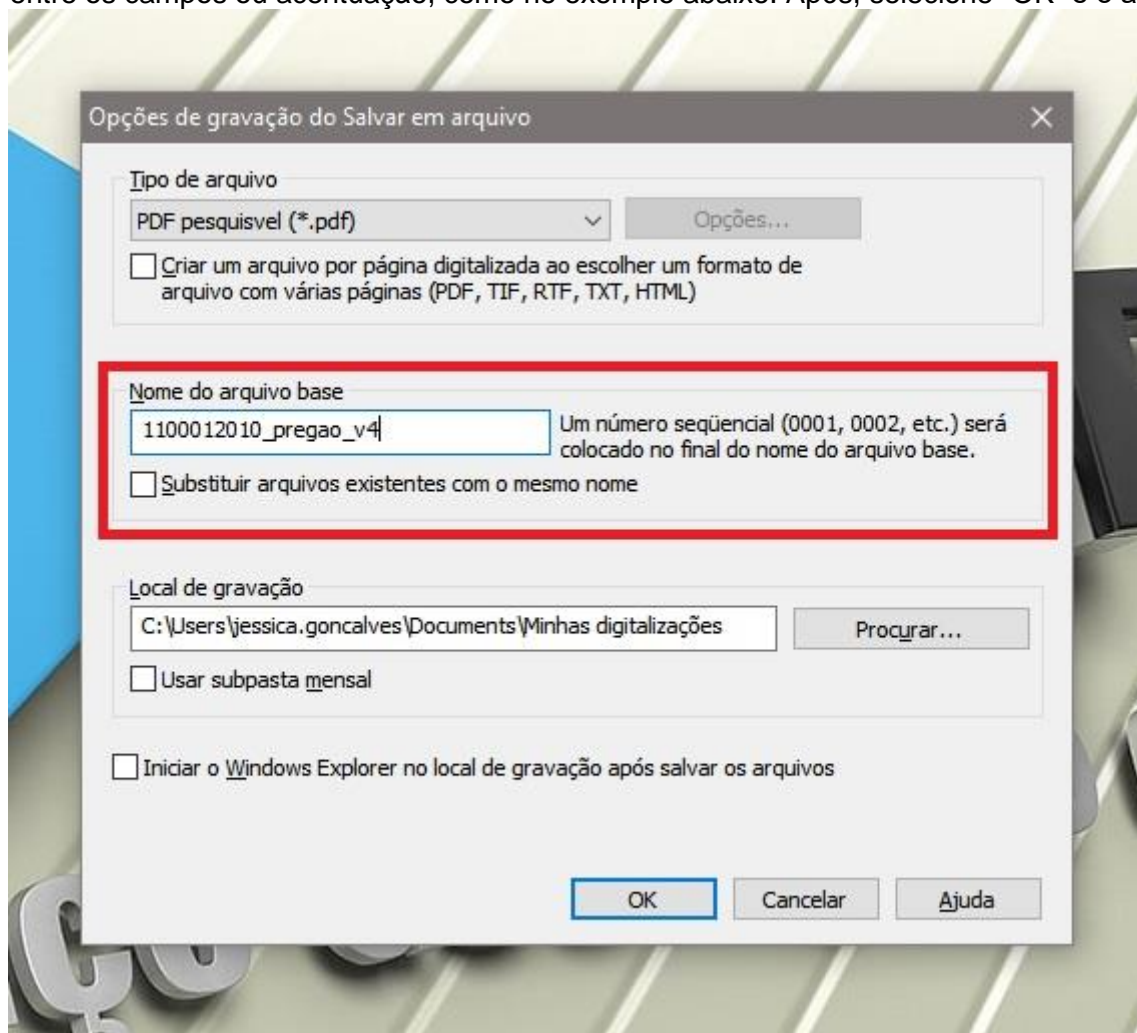
5. Clique em “Digitalizar...” e uma nova tela se abrirá.



6. Na nova tela que se abriu, note a área “Local de gravação”. Observe que o local para gravação automática do arquivo digitalizado é a pasta com final “\Documents\Minhas digitalizações”. Selecione o botão “Procurar...” na área “Local de gravação” (em detalhe na figura abaixo) se desejar modificar o local de gravação **temporária** do arquivo que será gerado após a digitalização.



7. Modifique o nome do arquivo com base na seguinte fórmula “numero_do_processo_tipo_de_processo_volume_ou_pasta”, sem espaço entre os campos ou acentuação, como no exemplo abaixo. Após, selecione “OK” e o arquivo será automaticamente digitalizado.

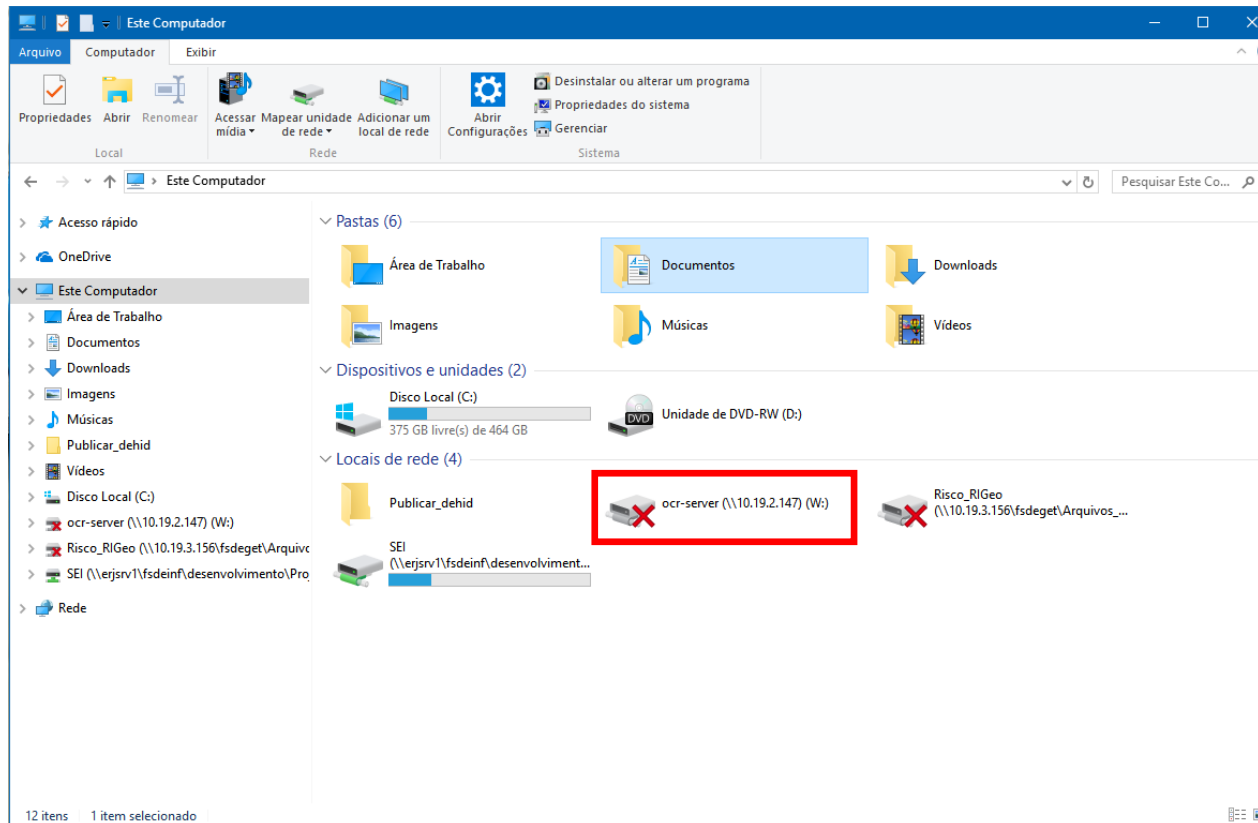


APÊNDICE 2

Procedimento de conversão de arquivos PDF para PDF com OCR

Procedimentos para conversão de PDF simples para PDF com texto pesquisável (OCR)

No caso de impossibilidade de utilização do modelo de escâner mostrado anteriormente, é possível a utilização de escâner de mesa ou de produção que não realize a conversão automática para formato PDF com OCR. Caso essa situação se mostre necessária, há um aplicativo instalado na rede da CPRM chamado OCR Server. O acesso a ele pode ser solicitado ao Service Desk pelos funcionários responsáveis pela conversão dos arquivos digitalizados e em PDF simples para arquivos em PDF pesquisável. Após a configuração do acesso, aparecerá uma pasta como essa em sua pasta “Este Computador”:



O X em vermelho, ao lado dela, indica que o acesso a esta pasta não foi autorizado. Caso isso ocorra, entre em contato com o servisse Desk para que as devidas correções sejam realizadas. Dentro do diretório ou pasta “ocr-server”, existem três outras pastas:

- Entrada: pasta onde devem ser colocados os arquivos que se deseja converter para PDF com OCR;
- Originais_Digitalizados: pasta onde ficam guardados os arquivos originais que acabaram de ser convertidos; e
- Saída: pasta onde ficam armazenados os arquivos em PDF já com OCR.

Para realizar o procedimento de conversão para PDF com OCR ou ocerização, deve-se copiar e colar o(s) arquivo(s) em PDF comum que se deseja converter na pasta Entrada, dentro da pasta “ocr-server”.

A conversão dos arquivos colados na pasta Entrada começa automaticamente, após alguns segundos. Dois arquivos com os nomes do arquivo adicionado, mais um complemento, serão gerados. Por exemplo, caso se copie e cole o arquivo de nome “48000_002936_2016-33_114.pdf”, após alguns segundos aparecerão os arquivos “48000_002936_2016-33_114.pdf.10.19.2.147.processing” e “48000_002936_2016-33_114.pdf.10.19.2.147.tmp”. Este é um procedimento normal, que significa que o(s) arquivo(s) adicionado(s) está(ão) sendo convertido(s).

Ao terminar a conversão para OCR, o(s) arquivo(s) irá(ão) ser removido(s) da pasta Entrada e aparecerá(ão) já convertido(s) na pasta Saída com o mesmo nome anterior, no exemplo “48000_002936_2016-33_114.pdf”.

O(s) arquivo(s) original(is) em PDF comum fica(m) armazenado(s) no diretório Originais_Processados.

Ao final, recorte o(s) arquivo(s) com OCR da pasta Saída do “ocr-server” e salve em seu computador para o posterior armazenamento na pasta em rede “\\erjsrv1\digitalizacao_SEI\unidade” (em vermelho pois deve ser substituída pela respectiva regional) e acréscimo ao SEI. A pasta Originais_Processados também deve ser limpa regularmente.

APÊNDICE 3

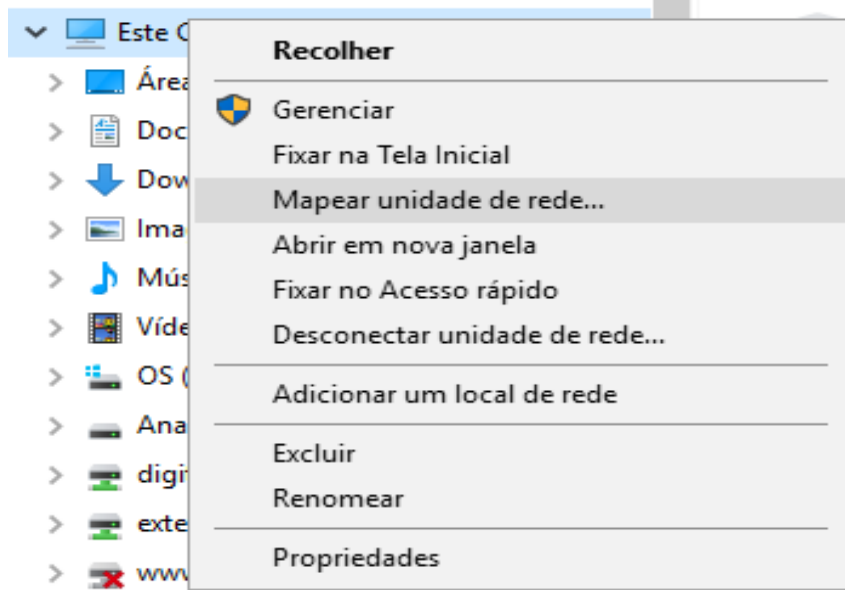
Procedimento de armazenamento dos arquivos em PDF com OCR

Procedimentos de armazenamento dos arquivos digitalizados em PDF pesquisável (OCR)

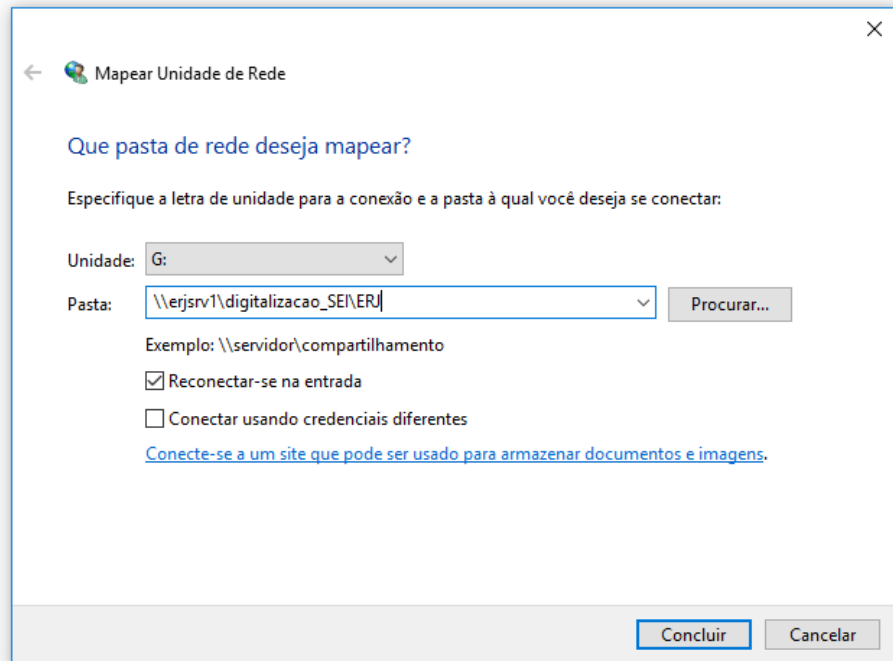
Foi criado um ambiente onde os arquivos digitalizados e ocerizados possam ser salvos para depois serem incluídos no sistema SEI - CPRM. Dessa forma, os responsáveis pela digitalização no ERJ e unidades regionais terão acesso a esse ambiente, sendo que cada usuário terá permissão de acesso à pasta que corresponde a sua unidade regional.

Para realizar o mapeamento da pasta na rede, siga as instruções abaixo:

1 - Vá em “Meu computador” ou “Este computador”, clique com o botão direito e escolha a opção “Mapear unidade de rede...”, como mostrado na figura abaixo.




2 – O mapeamento abaixo se abrirá. Escolha a unidade “G:” e insira o endereço do exemplo abaixo. No exemplo está ERJ, mas na unidade regional terá que alterar para a unidade correspondente, como nos exemplos “\\erjsrv1\digitalizacao_SEI\SUREG_BH”, “\\erjsrv1\digitalizacao_SEI\SUREG_SP”, etc.





Nome	Data de modificaç...	Tipo
ERJ	18/04/2017 10:34	Pasta de arquivos
REFO	18/04/2017 10:38	Pasta de arquivos
REPO	18/04/2017 10:37	Pasta de arquivos
RETE	18/04/2017 10:39	Pasta de arquivos
SEDE	18/04/2017 10:36	Pasta de arquivos
SUREG_BE	18/04/2017 10:37	Pasta de arquivos
SUREG_BH	18/04/2017 10:36	Pasta de arquivos
SUREG_GO	18/04/2017 10:38	Pasta de arquivos
SUREG_MA	18/04/2017 10:39	Pasta de arquivos
SUREG_PA	18/04/2017 10:39	Pasta de arquivos
SUREG_RE	18/04/2017 10:38	Pasta de arquivos
SUREG_SA	18/04/2017 10:38	Pasta de arquivos
SUREG_SP	18/04/2017 10:36	Pasta de arquivos

3 – Irá aparecer a pasta que foi mapeada e a subpasta com o “ANO” (no exemplo 2017) para que seja inserido o processo digitalizado e ocerizado.

Nome	Data de modificaç...	Tipo	Tamanho
 2017	18/04/2017 10:36	Pasta de arquivos	

4 – Todo o processo digitalizado e ocerizado, ao ser salvo na pasta, deverá seguir um padrão de nome como mostrado na figura abaixo. É importante seguir esse padrão já durante a conversão para OCR Server.

Caso o processo a ser digitalizado tenha mais de 1 pasta.

Nome	Data de modificaç...	Tipo	Tamanho
 nº processo + tipo_licitacao_contratacao_outros + volume_pasta.txt	18/04/2017 10:36	Documento de Te...	0 KB
 nº processo + tipo_licitacao_contratacao_outros.txt	18/04/2017 10:35	Documento de Te...	0 KB

Caso o processo a ser digitalizado tenha apenas 1 pasta

ANEXO 1
Ficha de controle de digitalização

Ficha Controle de Digitalização

Local:

Data:

Data	Hora		Matrícula	Nome do digitalizador	Nº do processo	Qntd. de pastas	Qntd. de folhas	Nome do validador	Status ¹
	Início	Fim							

¹ Corresponde a aprovação ou rejeição do documento. Deve ser preenchido com **A** de Aprova ou **R** de Rejeita.

Observações:

Referências

BRASIL. Casa Civil. **Procedimentos para a digitalização de processos**. [Brasília]: Casa Civil, [201-]. 8 f. Disponível em: <http://www.mda.gov.br/sei/sites/sei/files/user_29/Procedimentos%20para%20digitaliza%C3%A7%C3%A3o%20de%20processos%20SEI_final.pdf>. Acesso em: 31 out. 2017.

BRASIL. Ministério da Fazenda. Secretaria Executiva. Subsecretaria de Planejamento, Orçamento e Administração. Coordenação-Geral de Recursos Logísticos. **Procedimentos para digitalização de documentos e processos**: cartilha. Brasília: Coordenação-Geral de Recursos Logísticos/SPOA, 2017. 37 p. Disponível em: <<http://www.fazenda.gov.br/sei/publicacoes/cartilha-digitalizacao-de-documentos>>. Acesso em: 31 out. 2017.

BRASIL. Conselho Administrativo de Defesa Econômica. **Manual de digitalização de processo com OCR (PDF pesquisável)**: digitalização e envio de processo em OCR e (PDF) pesquisável para o sistema SEI. [Brasília]: CADE, 2014. 7 f. Disponível em: <<http://sempapel.cade.gov.br/wp-content/uploads/2014/12/DigitalizacaoOCR.pdf>>. Acesso em: 31 out. 2017.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). **Recomendações para digitalização de documentos arquivísticos permanentes**. [S.l.]: CONARQ, 2010. 28 f. Disponível em: <http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/Recomendacoes_digitalizacao_completa.pdf>. Acesso em: 27 dez. 2017.