



TRATAMENTO DE DADOS EM CARTÓRIO EXTRAJUDICIAL COM A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL

Marcos Rafael Martin¹

Palavras-chave: Tratamento de dados; Inteligência artificial; Ocerização – OCR; Sistema informatizado de automação; Cartório Extrajudicial.

INTRODUÇÃO

O presente trabalho visa apresentar o uso da inteligência artificial no tratamento de dados de um cartório extrajudicial de registro civil das pessoas naturais.

Em decorrência da necessidade de modernização dos cartórios e em cumprimento das normas do Conselho Nacional de Justiça foi necessário que os dados contidos nos livros físicos do cartório fossem compilados e inseridos no sistema informatizado de automação interno para posteriormente serem compartilhados para as centrais informatizadas dos órgãos públicos.

De regra geral os livros não pode ser deslocados do cartório e também não é conveniente que se faça isso, pois em caso de pedido de uma certidão é necessário ler o livro para a elaboração do documento. Bem como, também é necessário que o livro esteja no local do cartório para que nele sejam feitos outros serviços de atualização de dados como, por exemplo, averbações de divórcio, anotações de casamento, óbito etc.

Soma-se ainda o fato de que muitos dados estão manuscritos, de difícil reconhecimento de caracteres, tanto por humano quanto por máquina ou software.

O objetivo visa automatizar e ganhar escala na alimentação de dados, otimizando o espaço físico, a mão de obra, os equipamentos, os custo financeiro e o tempo, consdiderando, especialmente, as limitações físicas e qualidade dos dados do caso concreto, ou seja, digitalizar 100% do acervo físico de livros de registro, incluindo livros manuscritos e datilografados até dezembro de 2028, considerando a complexidade e o volume, mas assegurando a integridade dos dados, a conformidade legal (LGPD, Provimento nº 40/2021 CGJ/GO, etc.) e a facilidade de acesso e consulta futura. Sendo objetivo específico obter pelo menos 90% de assertividade na obtenção dos caracteres das imagens digitalizadas, diminuindo o tempo de conferência humana, dando efetividade ao trabalho realizado.

O método experimental com variáveis dependentes foi o utilizado para o andamento do procedimento.

¹ Mestre em Direito pela Atitus e Unisinos. E-mail: marcos@cartoriomartin.com.br.



1 FUNCIONAMENTO

Para enfrentar os desafios e atingir o objetivo proposto, foram adotadas dois procedimentos distintos, um para livros datilografados e outro para livros manuscritos, ambos com foco em uma assertividade superior a 90% na captação dos dados.

1.1. Digitalização de Livros Datilografados (Estimativa: 110.800 (cento e dez mil e oitocentos) registros).

Esta fase visa a captura eficiente dos dados contidos nos livros datilografados, mas que ainda não estão integrados ao sistema informatizado de automação.

Os recursos humanos e as metas foram progressivamente definidas:

Fase 1 (Até Agosto de 2025 - 1 a 2 funcionário com dedicação exclusiva): 100 registros por dia útil \times 82 dias = 8.200 registros.

Fase 2 (A partir de Setembro de 2025 - 2 a 4 funcionários com dedicação exclusiva): 200 registros por dia útil \times 122 dias = 24.800 registros.

Projeção para Conclusão (Livros Datilografados): Para digitalizar os 110.800 registros datilografados, mantendo a média de 200 registros/dia, seriam necessários aproximadamente 554 dias úteis (110.800 / 200). Considerando uma média de 250 dias úteis por ano, esta etapa levaria cerca de 4 semestres e meio.

1.1.1 Ferramentas e Tecnologias:

Uso de câmeras de alta resolução (podendo ser de smartphones modernos com boa qualidade óptica); scanners apropriados para documentos encadernados (para minimizar danos e garantir a qualidade da imagem); computadores com capacidade de processamento adequada; Inteligência Artificial Adapta One (para otimizar o reconhecimento e extração de dados); sistema Engagraph para a importação e gestão dos registros.

1.1.2 Procedimento Detalhado:

1.1.2.1 Identificação e Preparação: Localizar e preparar os livros datilografados que necessitam de digitalização ou migração para o sistema Engagraph.

1.1.2.2 Captura de Imagens: Realizar a captura das imagens dos registros. Para livros que permitam, o uso de scanners pode ser mais eficiente. Para volumes mais delicados ou de difícil manuseio no scanner, utiliza-se a fotografia com câmera de alta resolução.

1.1.2.3 Conversão Imagem-Texto (OCR e IA): As imagens capturadas serão processadas utilizando software de Reconhecimento Óptico de Caracteres (OCR) e a Inteligência Artificial



Adapta One. Esta combinação visa maximizar a precisão na conversão das imagens em texto editável e estruturado.

1.1.2.4 Formatação e Salvamento: O texto extraído será revisado e conferido por humano e salvo em formato JSON. Este formato é escolhido pela sua flexibilidade e facilidade de integração com sistemas.

1.1.2.5 Importação para o Sistema Engograph: Os arquivos JSON serão importados para o sistema Engograph, onde os dados serão validados e integrados ao banco de dados do cartório e imediatamente compartilhados para as Centrais dos órgãos públicos.

1.2. Digitalização de Livros Manuscritos (Estimativa: 184.400 registros):

Finalizada a etapa descrita no tópico 1.1, dar-se-á o início a esta etapa, que é a mais desafiadora devido à natureza dos documentos: manuscritos. Os recursos humanos e metas foram progressivamente definidas.

Estimativa com 3 funcionários com dedicação exclusiva: A estimativa de 30 registros por dia por colaborador. Com 3 funcionários dedicados a esta tarefa na nova estrutura do cartório (após a mudança de endereço ou reorganização do espaço), a meta diária seria de 90 registros.

Projeção Anual (Livros Manuscritos): Considerando 250 dias úteis em um ano, seriam digitalizados 22.500 registros manuscritos por ano. Se, por exemplo, o acervo manuscrito for de 184.500 registros levaria aproximadamente 16 semestres.

1.2.1 Ferramentas e Tecnologias: Serão utilizados computadores e o sistema Engograph. A principal ferramenta aqui é o recurso humano qualificado para a interpretação da caligrafia antiga e sua transcrição para o meio digital.

1.2.2 Procedimento Detalhado.

1.2.2.1 Identificação e Preparação: Localizar e organizar os livros manuscritos que não estão no sistema. Avaliando o estado de conservação e a legibilidade.

1.2.2.2 Digitação Direta no Sistema: Os dados dos livros manuscritos serão digitados diretamente no sistema Engograph por digitadores experientes. Recomenda-se um processo de dupla conferência por um segundo profissional para garantir a acurácia, dado o objetivo de assertividade superior a 90%.



RESULTADOS PRELIMINARES / CONSIDERAÇÕES FINAIS.

Os resultados obtidos pelo uso de inteligência artificial foram mais que satisfatório, dispensando, inclusive, revisão humana, eis que atingido o objetivo específico, ou seja, obtenção de 90% de assertividade na obtenção dos caracteres das imagens digitalizadas, mesmo na imagens em que há textos manuscritos.

Contudo, identificamos que a inteligência artificial “Chat GPT” utilizada como teste durante o processo, obteve maior assertividade em relação a “Adapta One”, pois seu aprendizado em conjunto com o “contexto” adquirido (Robo Notarial) foi superior.