

# IARA: Aplicação de Inteligência Artificial para Análise Automatizada de Matrículas Imobiliárias

Carlos S. Portela<sup>1,2</sup>, Elton S. Siqueira<sup>2</sup>, Edinaldo H. Sousa<sup>1,2</sup>, Gleyciane B. Freitas<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Computação Aplicada (PPCA) – Universidade Federal do Pará (UFPA) - Tucuruí - PA - Brasil

<sup>2</sup>Faculdade de Sistemas de Informação (FASI) – Universidade Federal do Pará (UFPA) Cametá - PA - Brasil

{csp,eltonss}@ufpa.br, {edinaldohenriques02, gleycianefreitas04}@gmail.com.br

**Abstract.** *The analysis of real estate registrations in registry offices still relies heavily on manual processes, which impacts efficiency and consistency. This paper presents the IARA (Artificial Intelligence Applied to Results and Analysis), an information system based on Artificial Intelligence designed to support the automated extraction and analysis of property registration documents. The solution was developed through a proof of concept grounded in an OCR benchmark and validated with two real estate registry offices in the state of Bahia.*

**Resumo.** *A análise de registros imobiliários em cartórios ainda depende muito de processos manuais, o que impacta a eficiência e a consistência. Este artigo apresenta o IARA (Inteligência Artificial Aplicada a Resultados e Análises), um sistema de informação baseado em Inteligência Artificial projetado para auxiliar na extração e análise automatizada de documentos de registro de imóveis. A solução foi desenvolvida por meio de uma prova de conceito baseada em um benchmark de OCR e validada em dois cartórios na Bahia.*

## 1. Contexto

O registro de imóveis no Brasil é regido por princípios e normas estabelecidos pela Lei nº 6.015, de 31 de dezembro de 1973, que disciplina os registros públicos e fundamenta a segurança jurídica da propriedade imobiliária [Brasil 1973], sendo sustentado por princípios como publicidade, continuidade, especialidade e concentração [Costa e Silva 2024]. Apesar de sua relevância institucional, os processos internos dos cartórios de registro de imóveis ainda enfrentam desafios significativos relacionados à análise e à interpretação de documentos registrais [Filho, Falavigna e Lima 2019].

Na prática, grande parte desses documentos apresenta características que dificultam sua análise, como *layouts* heterogêneos, baixa qualidade de digitalização, uso de terminologia jurídica antiga e ausência ou inconsistência de informações essenciais. Soma-se a isso a necessidade de identificar sobreposições de áreas georreferenciadas, um problema recorrente em imóveis rurais e urbanos [Souza 2005].

Este cenário foi observado em dois cartórios de registro de imóveis do estado da Bahia, que atuam como clientes e parceiros no contexto deste trabalho. Nesses cartórios, a análise manual de matrículas representa um gargalo operacional, impactando prazos, aumentando o risco de inconsistências e limitando a escalabilidade dos serviços prestados. Nesses dois cartórios, o acervo documental ultrapassa 5.000 registros de

imóveis ativos, muitos deles compostos por matrículas antigas, digitalizadas em diferentes padrões e com elevado grau de heterogeneidade.

Diante dessas dificuldades, emerge a necessidade de Sistemas de Informação (SI) que incorporem técnicas de Inteligência Artificial (IA) [Davenport e Ronanki 2018] capazes de apoiar e qualificar o processo de análise registral, sem comprometer os requisitos legais, normativos e de proteção de dados que caracterizam o domínio cartorial. O objetivo deste trabalho é apresentar o sistema IARA (Inteligência Artificial aplicada a Resultados e Análises), um SI baseado em IA que automatiza a verificação de consistência registral e detecção de sobreposições em matrículas imobiliárias.

## 2. Processo Adotado

O desenvolvimento da solução seguiu uma abordagem incremental, iniciando-se por uma prova de conceito fundamentada em um *benchmark* de ferramentas de Reconhecimento Óptico de Caracteres (OCR) aplicadas ao domínio de registros imobiliários. Esse *benchmark* permitiu comparar a eficácia de ferramentas proprietárias e *open source* em diferentes hardwares (CPU/GPU), considerando critérios de acurácia, eficiência e adequação ao contexto de documentos legais complexos, em conformidade com Hegghammer (2022) e Vijayarani e Sakila (2015). Essa prova de conceito utilizou arquivos de matrículas de imóveis em formato PDF como dados de entrada, totalizando 219 páginas, abrangendo documentos de diferentes épocas e qualidades de digitalização. Os dados de saída consistem na extração de texto estruturado contendo dados do imóvel e coordenadas geográficas para verificação de sobreposições e consistência registral.

Os resultados do *benchmark* subsidiaram a escolha das tecnologias mais adequadas para o desenvolvimento do sistema, conforme Tabela 1.

**Tabela 1. Benchmark da Acurácia das Ferramentas OCR**

Método OCR	Similaridade Média	WER Médio	CER Médio	Velocidade (Chars/s)
AWS Textract	0.8631	0.2737	0.2624	420.5
Google Gemini	0.8570	0.3223	0.2592	155.8
Cloud Vision API	0.8467	0.3302	0.2826	448.3
Easyocr (GPU)	0.8319	0.4663	0.3084	226.5
Easyocr (CPU)	0.8314	0.4690	0.3089	47.0
Paddleocr (CPU)	0.8199	0.4522	0.3113	720.3
Paddleocr (GPU)	0.8199	0.4525	0.3116	2538.6
Pytesseract (Sequencial)	0.7588	0.9609	0.4131	1298.6
Pytesseract (Paralelizado)	0.7588	0.9609	0.4131	2303.6

Com base nesses resultados, optou-se por uma arquitetura híbrida, explorando técnicas de OCR da AWS Textract e algoritmos de IA em Python capazes de lidar com a heterogeneidade dos documentos registraes. Em seguida, foram elaborados protótipos de interface, contemplando os principais fluxos de uso do sistema. Esses protótipos foram validados iterativamente com os cartórios parceiros, garantindo aderência às práticas e necessidades reais do domínio.

As matrículas de propriedades rurais apresentam desafios adicionais quando comparadas às urbanas, uma vez que muitas delas foram elaboradas antes da adoção sistemática de padrões de georreferenciamento [Souza 2005]. Essas características dificultam a extração automática de informações, a verificação de consistência registral e a identificação de sobreposição de áreas.

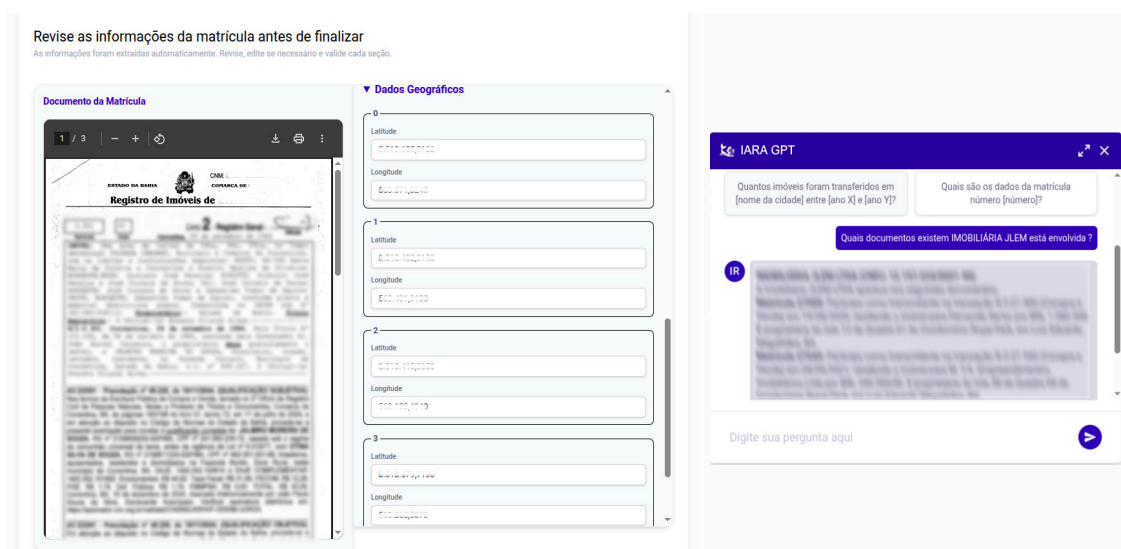
Paralelamente, foram definidos indicadores de avaliação, com foco na comparação entre o tempo médio necessário para análise manual de uma matrícula imobiliária e o tempo requerido pelo sistema com apoio de Inteligência Artificial. Esses indicadores permitiram mensurar os ganhos de eficiência proporcionados pela solução.

Todo o processo de desenvolvimento considerou requisitos legais relacionados à Lei Geral de Proteção de Dados (LGPD) [Brasil 2018] e ao Provimento nº 149/2023 [CNJ 2023], seguindo as seguintes etapas: Coleta: escaneamento de matrículas imobiliárias via upload de PDFs pelos usuários autorizados (tabeliães/cartórios); Retenção: armazenamento temporário em servidores com criptografia AES-256; Processamento: uso de IA para validação jurídica e georreferenciamento; Compartilhamento: restrito a autorização expressa do titular ou obrigação legal; e Eliminação: sob solicitação do titular.

### 3. Solução

O sistema denominado IARA faz referência à sereia do folclore brasileiro, associada às águas e à profundidade, evocando simbolicamente a capacidade do sistema de “mergulhar” em grandes volumes de documentos e emergir com informações relevantes para análise jurídica. Funcionalmente, o IARA permite a integração com bases documentais de escritórios jurídicos e cartórios, bem como a importação de matrículas de imóveis em formato PDF.

A Figura 1 apresenta as telas de extração da matrícula (A) e um chat que usa IA (B) para responder perguntas sobre essas matrículas.

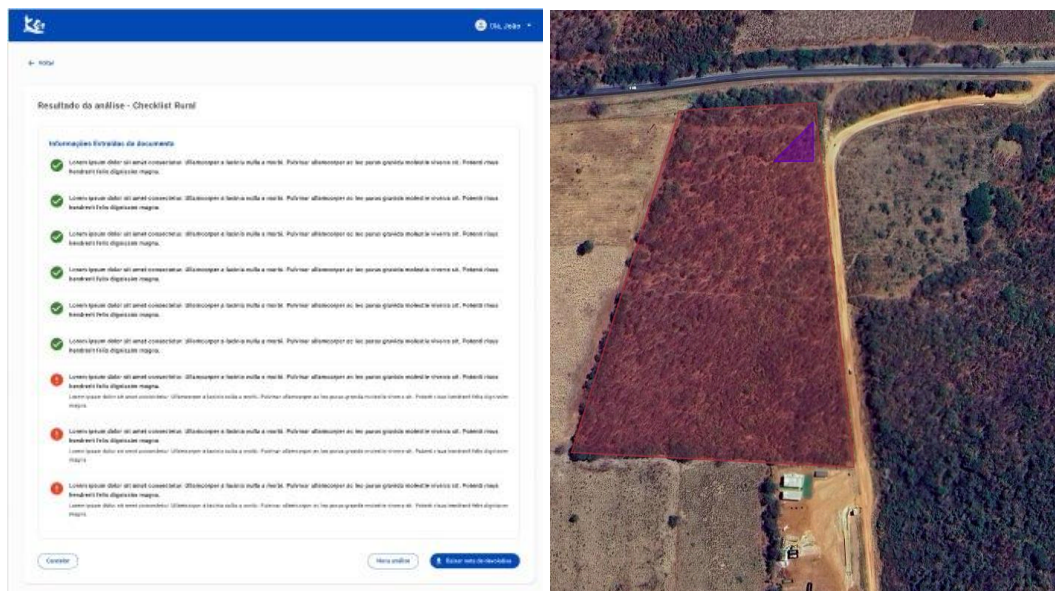


A) Digitalização dos Dados da Matrícula

B) Chat Inteligente sobre Matrículas

Figura 1. Uso de Técnicas de OCR para Extração e IA para Busca

Um dos diferenciais centrais da solução é a aplicação de *checklists* jurídicos configuráveis (Figura 2A), que permitem a identificação de inconformidades nas matrículas analisadas e a geração automática de notas devolutivas em formato editável, apoiando diretamente a atividade cartorial. Além disso, o sistema oferece visualização geográfica das coordenadas dos imóveis, possibilitando a identificação de sobreposição de áreas (Figura 2B).



A) Checklist de Qualificação de Matrículas

B) Visualização de Sobreposição de Áreas

**Figura 2. Funcionalidades de Análise de Inconformidades e de Geolocalização**

A aplicação da IARA nos cartórios parceiros evidenciou ganhos relevantes em termos de eficiência operacional. A comparação entre o tempo médio necessário para a análise manual de matrículas (cerca de 2 análises por dia) e o tempo requerido com apoio do sistema (1 análise em 3 minutos) demonstrou uma redução significativa no esforço humano, especialmente nas etapas de extração de dados e verificação de inconformidades. Além da redução de tempo, observou-se melhoria na padronização das análises e maior rastreabilidade das informações utilizadas no processo decisório [Tafti et al. 2016], contribuindo para a qualidade do serviço prestado.

Do ponto de vista organizacional, o IARA pode possibilitar maior escalabilidade das operações cartoriais, sem aumento proporcional de custos operacionais. No entanto, em sua versão atual, o IARA inclui o escopo restrito a dois cartórios baianos, sem testes em escala nacional. Como trabalhos futuros, recomenda-se expansão com treinamento em *datasets* nacionais (INCRA). Para o mercado de software, o trabalho demonstra o potencial de soluções baseadas em Inteligência Artificial aplicadas a domínios regulados e tradicionalmente conservadores, como o registro de imóveis, evidenciando oportunidades para inovação responsável, em conformidade com a LGPD.

Do ponto de vista de SI, a solução pode ser compreendida como um sistema sociotécnico, alinhado às dimensões de pessoas, processos e tecnologia propostas pelo referencial GrandSI-BR [SBC 2025], ao integrar práticas cartoriais, fluxos de análise registral e técnicas de IA em um ambiente unificado de apoio à decisão.

## Referências

- Brasil (2018) “Lei Nº 13.709, de 14 de Agosto de 2018”, [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/113709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm), Janeiro.
- Brasil (1973) “Lei Nº 6.015, de 31 de Dezembro de 1973”, [https://www.planalto.gov.br/ccivil\\_03/leis/l6015compilada.htm](https://www.planalto.gov.br/ccivil_03/leis/l6015compilada.htm), Janeiro.
- CNJ - Conselho Nacional de Justiça (2023) “Provimento N. 149, de 30 de Agosto de 2023”, <https://atos.cnj.jus.br/files/compilado1806222023111665565a1e0fc83.pdf>, Janeiro.
- Costa, F. e Silva, F. (2024). Registro de Imóveis no Brasil: Origem, Evolução e Princípios. In *Aracê*, v. 6, n. 3, páginas 7124–7136.
- Davenport, T., e Ronanki, R. (2018). Artificial Intelligence for the Real World. In *Harvard Business Review*, v. 96, n. 1, páginas 108–116.
- Filho, E., Falavigna, J., e Lima, A. (2019) “Hemeroteca Digital”, In: Anais Estendidos do XV Simpósio Brasileiro de Sistemas de Informação, páginas 132–133.
- Hegghammer, T. (2022). OCR with Tesseract, Amazon Textract, and Google Document AI: A Benchmarking Experiment. In *Journal of Computational Social Science*, v. 5, páginas 861–882.
- SBC (2025) “II GranDSI-Br Grandes Desafios de Sistemas de Informação no Brasil 2026-2036”, In: Publicações SBC, Organização de Renata Araujo, Sean Siqueira, Tadeu Classe, Rita Maciel e Clodis Boscaroli, SBC, Porto Alegre.
- Souza, M. (2005). Imóvel Rural, Função Social e Produtividade. In *Revista da Faculdade de Direito UFPR*, v. 43, páginas 1–19.
- Tafti, A. P., et al. (2016). OCR as a Service: An Experimental Evaluation of Google Docs OCR, Tesseract, ABBYY FineReader, and Transym. In: Advances in Visual Computing. ISVC 2016. Lecture Notes in Computer Science, v. 10072. Springer.
- Vijayarani, S. e Sakila, S. (2015). Performance Comparison of OCR Tools. In *International Journal of UbiComp*, v. 6, n. 3, páginas 19–30.