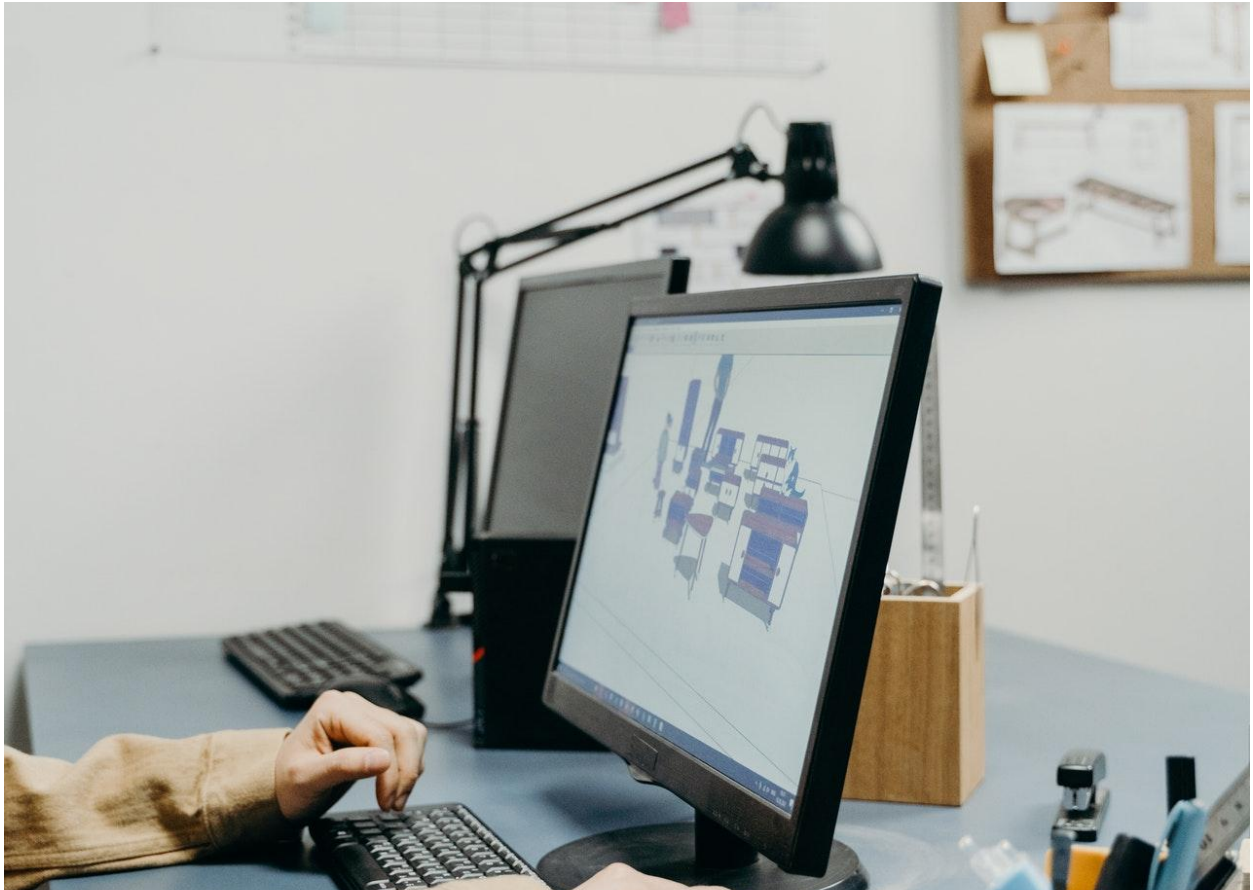# THE REALITY OF AI / MACHINE LEARNING
Title Database Indexing Still Needs Humans

Introduction

The latest headlines in the title industry are all about how machine learning, Natural Language Processing (NLP)feed into artificial intelligence that will change our processes in ways we can't easily measure. This may be true for some aspects of the business. But when it comes to indexing historical and day forward courthouse documents for chain of title research, it isn't entirely applicable. But it DOES increase productivity with a reduced cost to leverage human capacity with proven business processes.

The Time, Cost and Accuracy Challenge

Most title professionals who run a chain of title do so manually at the courthouse, use the courthouse digital records, subscribe to a title plant provider, maintain their own title database OR some combination of all of these. Each presents its own challenges and advantages based on the volume of business done in specific counties.

Why Is the Imersion Implementation Different?

Since 1999, Imersion Global has been a Texas based industry leader in bulk and day forward indexing for chain of title research databases. Imersion has processed tens of millions of pages over the years as part of the document verification and assembly needed to properly inventory and geoindex a county clerk data set. Documents change through the eras of handwritten documents, retyped documents, microfilmed documents, mass scanning vs scanning individual documents, changing of clerk vendors and so on. There are variations in time stamps, volume/page numbering, book types, clerk number assignments, plat recording eras, etc. The list goes on for all of the factors we have come to expect.

Knowing this, Imersion Global created the Key 1 ML Indexing Enhancer platform. The platform applies a baseline rule set built up on the most common factors to extract out legal descriptions and other required fields from documents in bulk.   The platform creates different virtual processing queues from the document processing based on document types, expected outputs, outputs with conflicts or missing data and so on.

The platform processes thousands of documents in hours to create workflow queues for the human indexing teams with the best skill sets to validate, complete and approve each individual record associated with a document. The automated portion of this can be adjusted for additional speed or quantities by adjusting the processors/server loads. The system is designed to be extensible.

By creating queues and exceptions the manual portion of the processing becomes more optimal as instead of having general queues where all experience and skill levels review any document, the process makes it most likely to match documents with the right teams.

As the humans process the document indexing queues, any modifications, changes or notes are tagged for export as feedback to the AI/ML algorithm team. The algorithm team adjusts the models and output criteria for subsequent outputs. By virtue of humans processing queues, a feedback /validation loop is built into the work flows.

## How Can This Help Real Estate Title Organizations?

Producing a chain of Title is a critical part of the title insurance real estate transfer process. Earlier in this document, we referred to the various methods of creating a chain of Title.

- For large title plant providers, this can expedite addition of sovereignty title databases by acquiring fully indexed data sets more quickly to expand coverage for subscribers and underwriters while keeping the cost of doing so more manageable. Keep your subscribers happy by expanding to lesser available counties.
- For independent title agencies, owning your own sovereignty data set may be within reach at a more affordable long-term cost than subscribing to the mega providers, manually doing your own or outsourcing your title research. Why not own your own complete data set for counties where you do more business, have people subscribe to your data and where the premium splits among the number of licensed agencies makes sense?
- Cost consistently increases for employees, taxes, overhead, software, licensing and so on. This hybrid system allows Imersion Global to produce fully indexed complete county data sets at reasonable prices relative to other alternatives with a leadership team specializing in title documents for over 20 years. The cost is reasonable, and timelines stay reasonable as well so you can project your return on the investment.
- Use cutting edge technology without having to manage the feedback loop, the cost of servers, hiring of staff and the unknowns to keep up with the competition.
- You don't have to choose two out of the three: fast, good or low cost. We all get LinkedIn invitations and other random correspondence about how you should outsource your work to the XYZ guy who does "data entry." We know how those usually turn out.
- More affordable digital databases means more competition which keeps overall costs down for the industry and consumers. Lower costs equals happier customers and investors.

For Those Interested in the Underlying Technology

The Key 1 ML Enhanced Processing Platform is a proprietary design with thousands of man hours collectively put into developing, testing and validating the algorithms and rules. The development process was based on data from multiple counties from documents recorded in multiple decades.

The diagram below illustrates the machine process that feeds into the human validation portion.

Service Capability Model Level of Automation in Hybrid Index Capture Model
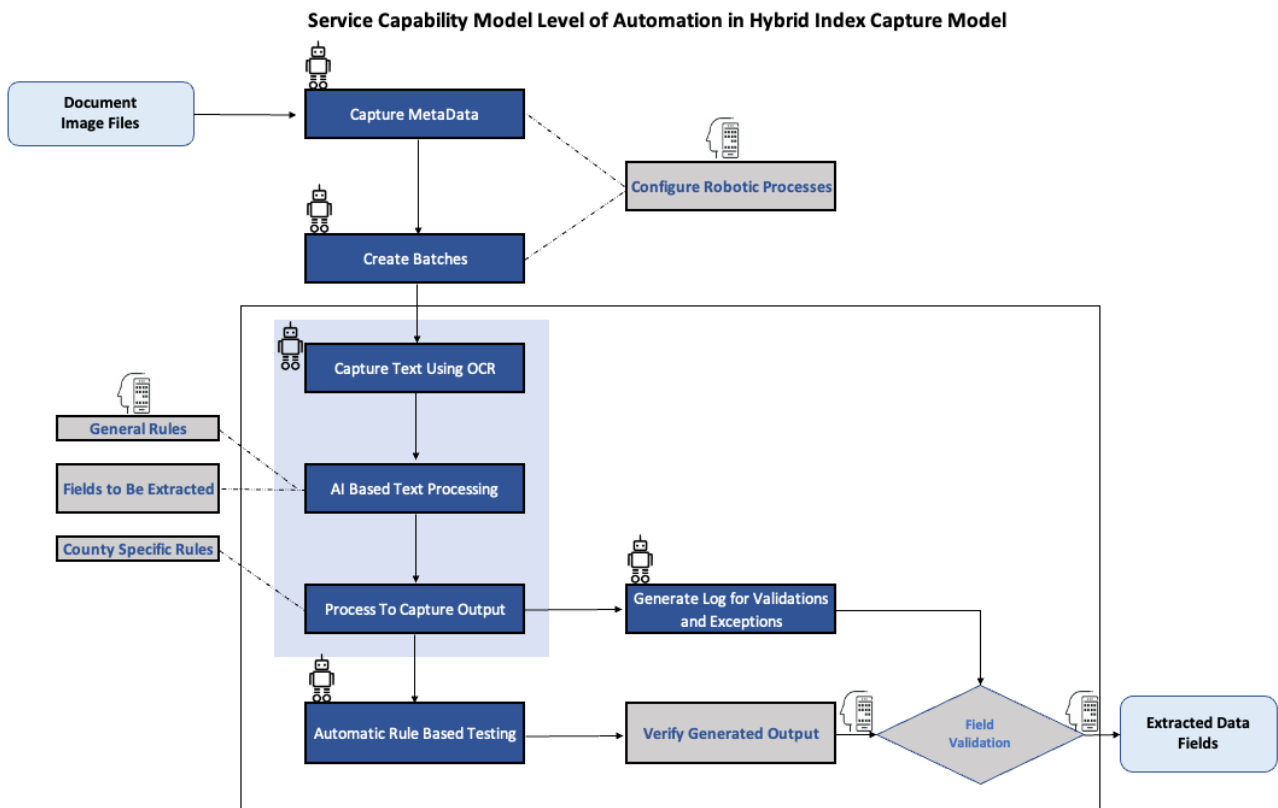
The technology stack is built on advanced AI based text processing algorithms. The processing can be quickly scaled horizontally to handle spikes in demand. Imersion uses the technology to focus on the fields used for creating a chain of titles. The technology can be adjusted to capture complimentary metadata for title related verticals such as mineral research, mortgage tracking, lien tracking and other applications.

Conclusion

Instead of diminishing returns on constantly tweaking AI/ML algorithms for trying to account for all situations, Imersion Global implements a 80/20 hybrid model through the Key 1 ML Indexing Enhancer. Sometimes the cliche of "working smarter, not harder" really does apply.

The lessons learned through this exercise are:

- Training programmers to develop constantly changing rules is not the best use of their time, resources or that of title indexing subject matter experts.
- Imersion Global strikes a balance between AI/ML algorithms and maintenance of these algorithms by using a systemic feedback loop.
- Experienced teams of human indexers cut through the noise, complexity and randomness to efficiently index title documents.
- Technology is not a panacea. It is only as good as the resources who can leverage it to become part of a more productive workflow.

Find Out How

Schedule a time to discuss your current workflows, business goals and future projects. Imersion Global will help you evaluate how to leverage both of our resources to help you reach your goals whether it is reducing subscription costs, owning your own courthouse data, adding new revenue opportunities, expanding your coverage or taking a deeper look at data you already have. Go to www.imersionglobal.com for more information and to submit inquiries.