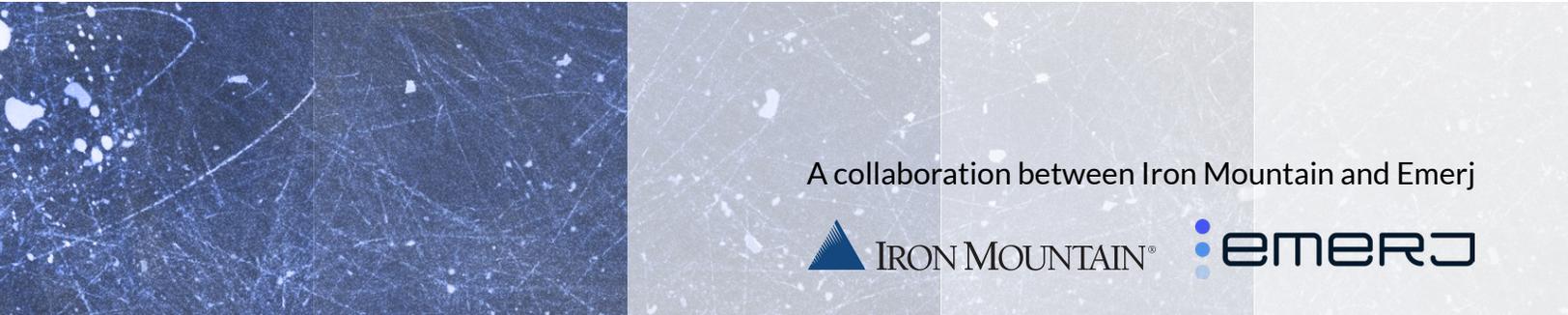


AI-Enabled Search and Discovery for Mortgage Processing and Underwriting



A collaboration between Iron Mountain and Emerj



IRON MOUNTAIN®



emerj

AI-ENABLED SEARCH AND DISCOVERY FOR MORTGAGE PROCESSING AND UNDERWRITING

An Introduction

The financial sector was among the first to adopt artificial intelligence in business by automating fraud prevention with anomaly detection technology. Now financial institutions, including lenders, stand to benefit from automating back-end processes by digitizing documents and eliminating manual data entry.

In mortgage lending, NLP-enabled search and discovery applications, also called enterprise search or information extraction applications, help expedite the processing and underwriting stages. But most lenders today are not using AI for these key functions. According to a [2018 survey by Fannie Mae](#), only 27% of mortgage businesses had experience with AI solutions. This reveals an opportunity for lenders that want to use AI to gain an advantage over their competitors.

One possible first step toward taking advantage of this opportunity lies in understanding what Emerj, the AI research company, calls the [AI Opportunity Landscape](#): a map of what is possible and what is working with AI in a given industry. This can be a jumping off point for mortgage lenders looking to select their first AI project.

Although integrating enterprise search solutions can be challenging, lenders could decrease costs in the long run by using AI to create efficiencies in mortgage processing, underwriting, and compliance workflows. But AI requires digital data, and before taking advantage of AI for search and discovery, lenders need to ensure that the data they want to search is stored in a digital format.

THE PAPER PROBLEM IN THE MORTGAGE INDUSTRY

While most lenders have digital workflows, the mortgage industry as a whole still struggles with a vast amount of physical documents. Anke Conzelmann, Director of Product Management at Iron Mountain, describes the current state of mortgage processing and underwriting:

Sometimes...at the end of that process, you still get a stack of papers....the back-end processes often times still involve literally printing that stuff out and having physical paper sort of moving around.

This process...takes a long time. It takes longer than it should, and part of the reason for that is that there is a lot of manual intervention. Really, what are you trying to do? You're trying to figure out whether you have all of the information that's required to make a decision about providing someone with a loan.

In order to rectify this state of affairs, lenders will need to digitize documents in order to adopt an enterprise

search application. However, not all digitization solutions are created equal.

Traditional optical character recognition (OCR) software often struggles with accurately digitizing paper documents. Some OCR solutions have trouble recognizing text below certain font sizes and text with formatting such as italics and underlines. Others produce documents of pure text instead of filling in digital forms, forcing lenders to manually copy and paste information into digital forms themselves.

AI, however, can help lenders overcome these challenges. Machine vision software, a type of artificial intelligence that can “understand” images and video, can be trained to better recognize:

- Different font sizes
- Text formatting differences
- Handwritten signatures
- The location of information within a form

As a result, the use of machine vision software could drastically reduce and in some cases eliminate the need for man-

ual copy and paste data entry. However, in order to reap the benefits of integrating AI into their processes, lenders will need to ensure they have a document digitization workflow that allows them to transform the paper documents into a digital format.

Digitizing Documents to Expedite Audit Requests

Lenders may also consider digitizing their loan documents to expedite audit requests. In order to meet the audit request timeframe, many lenders will need to pull employees away from their regular production work to focus on locating the requested loan files. This effort can be very time consuming as the requested files may be in different locations. Some may be stored in their office location while others may be stored off-site. Digitizing these documents can provide lenders easier access to them, thus reducing the time it takes lenders to produce documents for the audit and allowing their employees to continue their regular work.

AI-ENABLED SEARCH AND DISCOVERY FOR MORTGAGE PROCESSING

Once a lender digitizes the paper documents it receives, its loan processors and underwriters may still struggle to find relevant information within those documents without manually searching through them.

They may be able to read them faster and use basic search functionality, but AI would allow them to search for specific concepts within many different documents at once. An AI-en-

abled search and discovery software allows loan processors to find information such as:

- A borrower’s social security number, date of birth, and address
- Proof of income and employer information within tax documents and paystubs

- Proof of homeowners insurance within relevant insurance documents
- Proof of available funds within a bank account
- Details about assets and liabilities, including account numbers, within supporting documents

In addition, loan processors are tasked with ensuring all required documents are in the file and complete, as well as ensuring that like data is consistent across all documents within the loan file. As such, the ability to identify missing or inconsistent information within a loan file may reduce the time needed to process a loan file. By introducing AI into their workflows, lenders can add an additional layer of quality control, which can reduce the overall time to close.

For example, a loan processor could use an enterprise search application to extract an applicant's name from numerous documents they've submitted to the lender during the application process. A person named James may write their name as "Jim" or "Jimmy." The AI software may be able to flag the file for an additional review since the applicant's name appears differently across multiple documents.

An executive at a global bank told Emerj the bank is using AI for a similar use-case in which its customers spell their names differently on multiple documents. As a result, the bank needs to determine if the documents are from the same customer:

Matching of names—that looks simple. As a human I can see if a name [is] spelled wrong...but a machine might [not know because it's spelled wrong]. That is the current state of affairs. [We] wrote an algorithm that agnostically looks at two [documents] and says 'Okay, they might be related. The words of those [documents] may be in different orders and there may be spelling errors, but...those [documents] are related.'

Without AI, a loan processor would need to manually search through all of these documents and identify where the applicant's name does not match their legal name or where their name is spelled incorrectly. Instead, AI could do this for the loan processor, saving them time and allowing them to focus on higher value tasks.

Enterprise search software could also reduce the number of errors loan processors make when manually collecting data from documents. A loan processor may accidentally skip over a form that was missing a signature. As a result, the file may be returned to them by underwriting due to the missing signature thus delaying the approval process.

AI software can be trained to automatically flag missing information; verification would no longer rely on the human eye, which is much more error-prone. By reducing manual search effort and human error, AI could reduce the time it takes a loan processor to perform their tasks as well and increase the accuracy of their work.

1 Searching Mortgage Documents to Comply with CCPA

Lenders can find it difficult to stay in compliance with data privacy regulations if they do not have easy access to all of their documents. In January 2020, the [California Consumer Privacy Act](#) (CCPA) went into effect, and if lenders fail to comply within 30 days of receiving notice of a violation, they can face fines of up to \$2,500 for each record involved in the violation. If the violation is intentional, that fine increases to \$7,500 for each record. This amount adds up quickly in the event of a data breach. AI-enabled search and discovery applications help lenders mitigate this risk.

In the banking industry, executives and solution providers are focusing their attention on AI for automating compliance processes. Emerj's research shows that, as of 2019, approximately 17% of the venture capital raised across the AI in banking landscape was for companies that provide compliance solutions. This makes compliance the second-most funded area of banking when it comes to AI. An executive at one of America's top five banks told Emerj:

The industry in general now [understands] there are high prices to pay for a lack of adherence. Rewind maybe ten, fifteen years ago—a lot of the focus would be on massive revenue generation opportunities. But now the downside of behavior that is not compliant with regulators is meaningful to the bottom line...This really made staying compliant a major focus and where a lot of the applications of AI, and specifically, NLP, make sense.

According to Emerj, banking executives are focused on automating compliance workflows above most other processes, for the following reasons:

- Compliance is a required sunk cost; lenders aren't making more money or finding new revenue opportunities when they invest in compliance, but they are required to do it.
- New regulations are making it harder to remain in compliance using traditional, manual methods. Californians can now request that a lender produce all of the data it has on them and/or remove it from its system. The lender needs to be able to prove that it has done this, which can be difficult when so much of its historical customer data is stored in paper documents off-site and in disparate digital storage systems.

AI-enabled search and discovery software allows lenders to review historical documents so they can comply with CCPA and other data privacy regulations. For example, lenders may need to search for mortgage documents within certain parameters to provide them to customers upon request, including:

- Mortgage documents belonging to a specific customer
- Mortgage applications filed within a certain date range
- Mortgage applications filed within a certain state

2 Using AI to Amend Mortgages Before LIBOR is Discontinued

Dan Courtright, Senior Director of Product Management at Iron Mountain, told Emerj that AI can be used to find references to LIBOR within mortgage documents. He stresses the risks that come with failing to update mortgages that use LIBOR to calculate their interest rates before its sunset in 2021:

If a contract is pegged on LIBOR and is going to extend past 2021, that [lender] will have exposure. It will need to do some research to figure out...what [index] it wants to use. The risk is two parts: One is from a borrower perspective—damages associated with that change in the index. Let's say it moves from LIBOR to some other index and that costs the borrower a lot more money. Most likely the borrower is going to file some sort of litigation against the company if they're unhappy with what that is.

From a regulatory perspective, in Europe... they potentially might start fining some of these institutions if they aren't gradually moving away from LIBOR...Getting a grasp on the exposure is where [AI and machine learning] can be used.

Lenders that want to amend mortgages that reference LIBOR could use AI-enabled search and discovery software to extract entities from mortgage documents, including the fallback language present in contracts that reference LIBOR. This may be imperative because lenders have less than two years before the LIBOR is discontinued. John Williams, President of the Federal Reserve Bank of New York, said of LIBOR, ["We need a mindset shift where firms realize that every new U.S. dollar LIBOR contract written digs a deeper hole that will be harder to climb out of."](#)

USE-CASES Enterprise Search for Compliance in Mortgage Lending

AI-enabled search and discovery applications could also help speed up the mortgage underwriting process. Underwriters need to review a mortgage applicant's loan history, credit history, and employment history, among other broad categories of information. This requires them to search and

read through scores of documents both digital and physical before they can make a decision on whether or not to approve someone for a mortgage. Underwriters can use AI-enabled search and discovery software for two broad categories of search: external and internal.

3 Searching External Databases to Underwrite Those With Thin or No Credit

Some AI-enabled search and discovery applications allow for external search, or the ability to scour the web for information on particular topics. Traditional lending is largely based on credit scores, but they don't always tell a complete story about a mortgage applicant's risk of default. Although uncommon today, lenders might consider new data sources during the underwriting process. AI-enabled search and discovery software can help with this, allowing underwriters to search for:

- **Public records on the property the applicant is looking to buy.** Underwriters can then feed this data into a predictive analytics algorithm that could factor in data about the property: its location, its style, how its water is supplied, and so on. For example, a predictive analytics algorithm might find that mortgage applicants in a certain part of the country are more likely to default if they live in ranch-style homes and get their water from a well.

- **News about an applicant.** For example, an applicant may have been featured in a local newspaper for charity work they conducted in their community. This may speak to the applicant's character, which an underwriter could use to determine whether or not to approve them.
- **An applicant's social media presence.** Social posts may reference times of financial prosperity or hardship in the applicant's life. An underwriter could use this to determine if the applicant is likely to default or not.

While use of this type of data is uncommon in the mortgage industry today, there are those that consider this data to be another form of alternate documentation that could be used to underwrite loans for applicants with thin or no credit history.

4 Searching for Internal Documents to Approve More Loan Applicants

Traditional underwriting requires that the underwriter review documents related to an applicant's income, employment, assets, and liabilities.

For nuanced cases, they could use an AI-enabled search and discovery software to retrieve approved mortgages from applicants who are similar to the current applicant. For example, an underwriter might determine that a loan applicant falls just short of a minimum income requirement. Instead of outright rejecting the applicant, the underwriter could more easily search for similar borrowers within the lender's system to get an idea of how successful they were at making their monthly mortgage payments. If these

borrowers never missed a payment, the mortgage underwriter may feel more comfortable approving the applicant at hand.

This means more business for the lender, and although large lenders don't engage in this practice, it's the primary value proposition for many AI-enabled underwriting solution providers and fintech companies. Many of these companies offer predictive analytics applications, but an AI-enabled search and discovery applications could be an alternative to those predictive analytics applications can suffer from a classic machine learning problem: It's in large part a black box.

5 Searching for Internal Documents to Approve More Loan Applicants

Lenders need to ensure they're staying in compliance with fair lending laws, and this can be difficult when they can't discern how a predictive analytics-based underwriting software made the decision to approve or reject a mortgage applicant.

Underwriters that use enterprise search software to speed up their work can still provide their reasons for rejecting an applicant, as required by law. This is not always possible with predictive analytics applications that output an "approve" or "reject" without providing much rationale for the decision. Some prominent companies have claimed to circumvent this problem by providing greater trans-

parency into their algorithms, but this ability is still rather nascent.

In addition, regulators are starting to notice the trend toward using algorithms for underwriting. Presidential candidate Elizabeth Warren and Alabama Senator Doug Jones wrote a [letter](#) to the Consumer Financial Protection Bureau criticizing fintech companies and some traditional lenders for "using algorithms to supplement or replace their underwriting process."

Lenders that use AI for information extraction may attract less criticism because their underwriting processes would involve a human in the decision process.

WHERE LENDERS SHOULD FOCUS THEIR AI EFFORTS

Emerj breaks down the potential ROI of AI into three broad types: revenue improvement, cost reduction, and risk reduction.

Lenders that want to leverage AI to their advantage should focus on cost reduction and risk reduction. The former involves efficiency gains in mortgage processing and audit requests, reducing the time it takes lenders to extract information from mortgage documents and produce documents for audits without pulling employees away from their higher value work; the latter involves developing a greater ability to respond to audits and customers that exercise their rights under data privacy

regulations, reducing the likelihood that the lender could incur large fines.

Cost and risk reduction are where large lenders, particularly banks, are focusing their AI efforts. In particular, adopting AI to automate compliance may be the most effective way a lender can differentiate itself from its competitors. This is because compliance is necessary, and it must be addressed before a lender can focus on securing an advantage that will allow them to win market share. Otherwise, the lender will continue to be distracted by regulatory issues. Emerj asked an AI-focused executive at one of the world's top 5 banks why financial institutions are so focused on leveraging AI for risk reduction:

Since the financial crisis, [financial institutions] are being hammered with fines—tens of billions of dollars in fines. So there is an unbelievably large financial incentive to make sure those fines stop and don't keep coming.

There's another layer on top of that, which is regulatory pressure. Because [financial institutions] are being fined so much, it's not just that the regulators are saying, "I'm going to give you a fine." They're saying, "I'm going to give you a fine and make sure you don't do this again because if you do I'm going to 'ten X' the fine. I might even revoke your banking license for certain areas of your business."

Once a lender frees up the time and money it spends on compliance, it can then start to focus its AI efforts on increasing revenue. For lenders, this type of ROI will largely come from AI-enabled underwriting solutions, such as search and discovery solutions, that could allow underwriters to make use of new data sources.

In the near future, there will likely be much more regulatory guidance on how to use AI algorithms for underwriting while remaining in compliance with fair lending laws. Until then, lenders can prepare themselves for experimenting with these applications in the new decade by first digitizing their documents and then adopting AI for cost and risk reduction. This effort will allow them to develop the skills necessary to successfully adopt AI in the future.

About Emerj

At Emerj, the AI Research and Advisory Company, we help banks and financial services firms find the biggest opportunities for AI at their companies, pick their first AI projects, decide which AI vendors to work with, and build winning AI strategies. Our flagship service, the [AI Opportunity Landscape](#), provides clients with the data, advisory, and decision support tools they need to win market share, save money, and reduce risk with artificial intelligence.

Contact

services@emerj.com

1-617-945-8567

About Iron Mountain

Iron Mountain Incorporated (NYSE: IRM) is the global leader for storage and information management services. Trusted by more than 225,000 organizations around the world, the company helps customers lower cost and risk, comply with regulations, recover from disaster, and enable a more digital way of working.

Iron Mountain InSight is a content services platform that provides actionable business insights and predictive analytics through Machine Learning (ML)-based classification of a company's physical and digital information, which adds structure, context, and meta-data to information to make it more usable. The resulting enriched content can then enable enhanced automated governance and workflows at a lender.