MetricStream

WHITE PAPER:
AI and Contract Management
Summer 2020

Collaboration By:
MetricStream and Golden Gate University

By Saloni Jain | Brandon Pak

Foreword

The age of artificial intelligence has made tools more accessible than ever before. Organizations are automating manual processes to achieve quicker time to insight and eliminating the tediousness of repetitive tasks. In the realm of Contract Lifecycle Management, organizations are utilizing natural language processing to extract relevant contract information to ensure information veracity and to capture anomalies that manual review might miss.

About Us (Golden Gate University ~ MetricStream Collaboration)

MetricStream is an independent market leader in enterprise cloud applications for governance, risk, compliance (GRC), and quality management. Through the enterprise platform and cloud software, MetricStream enables enterprises across industries to drive exceptional business performance based on the foundation of good governance, trust and integrity.

Founded in 1901 in San Francisco, **Golden Gate University** offers undergraduate and graduate programs in business and management, accounting, taxation, and law.

Contract Lifecycle Management

Contracts are a legally binding agreement. Two different parties, such as companies, enter contracts usually to exchange of goods and services for currency. The starting point of a contract's life is usually a contract template for contracts entered into many times with different parties. The terms of the contract are sometimes negotiated between two parties and agreed upon. Once agreement occurs, the final contract version is signed and both parties are responsible for delivering what they agreed to in the contract. The legal department at a company usually drafts the contract before it is negotiated but contracts entered can impact many departments of a company. For example, sales, finance, and accounting, operations, supply chain, manufacturing and more. The life of a contract continues through to amending a contract, renewal, or termination of a contract. Contract Lifecycle Management is the management of the contracts through the process of initiation through execution, performance and renewal/expiry.

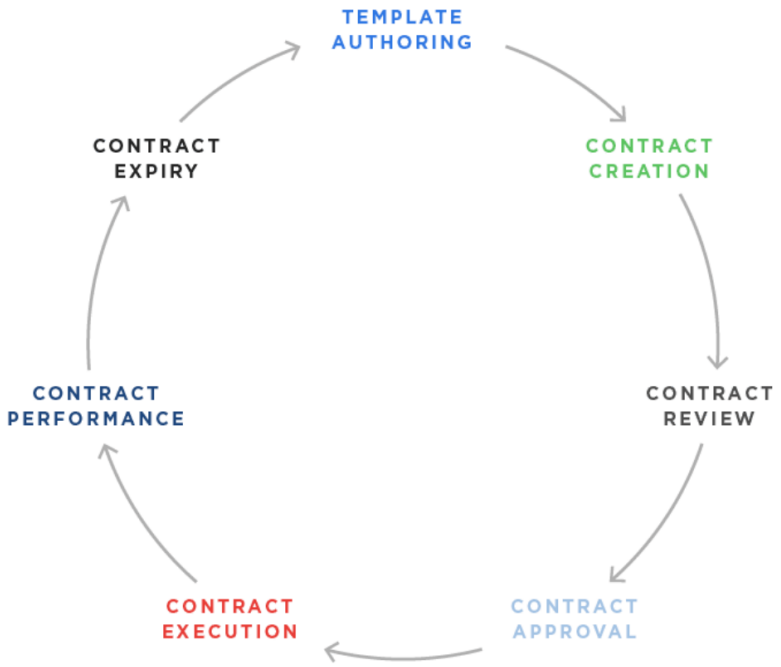


Figure 1: Steps in the Contract Lifecycle Management

The correct CLM solution connects the contracts process to order fulfillment and billing, eliminating errors, reducing risk and ensuring that revenue can be forecasted more accurately.

Introduction to Machine Learning

Machine Learning is an application of Artificial Intelligence (AI) that helps the system to automatically learn, perform and improve from experience. Machine Learning focuses on the development of the computer programs that can access the data and use it for themselves. **Primary aim is to allow the computers to learn automatically without human intervention and act accordingly.** Machine

Learning enables analysis of massive quantities of data, while generating faster and more accurate results in order to identify profitable opportunities or dangerous risks pertaining to the organization. Combining ML with AI and cognitive technologies can make it even more effective in terms of processing large volumes of information.

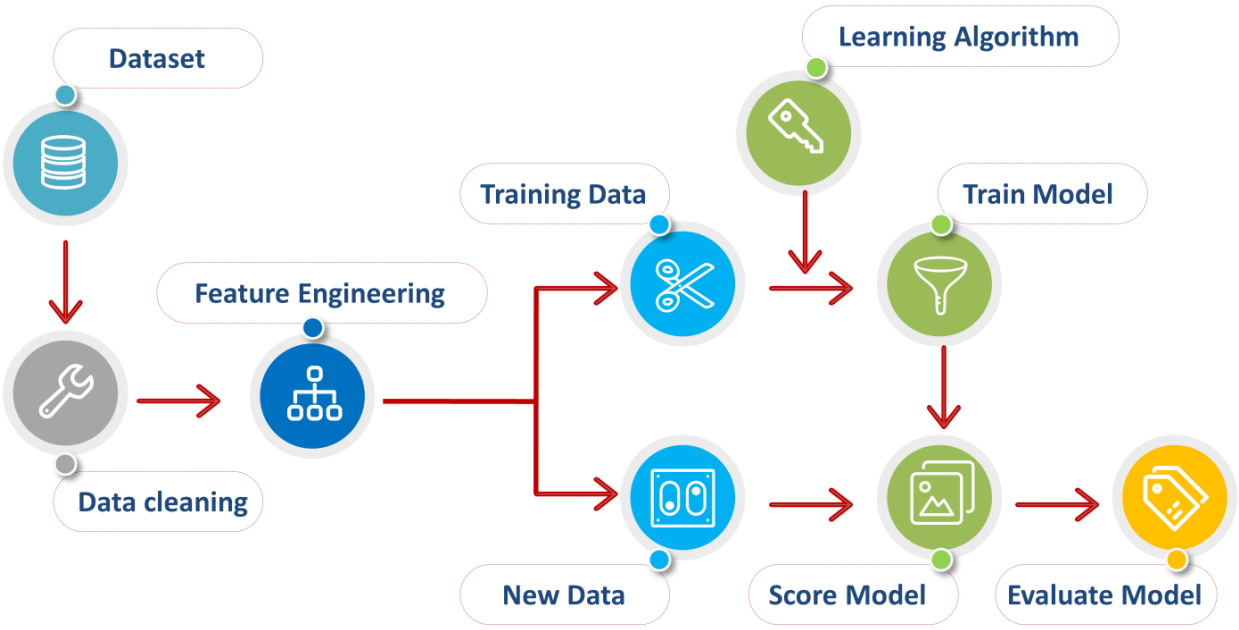


Figure 10: Machine Learning Process overview

It is very important for companies who practice CLM to understand the importance of applying Machine Learning combined with AI and produce customized outputs and perform the analysis.

Artificial Intelligence and Contract Management

Key terms and conditions in a contract typically include the contract name (i.e. the type of contract), the parties involved, effective date, any fees and / or rates, contract milestones, renewal or termination dates, and so on. Companies must understand the key terms and conditions after signing a contract to ensure that they are meeting their responsibility, especially as the company grows and enters more and more contracts. If they do not, they may incur legal fees, financial obligations, unsatisfied customers, early termination of contracts, and much more.

In order to track these key terms and conditions, the company may manually input all that information somewhere to track it, such as into a system or spreadsheet. As the variety and number of contracts entered increases, the more time-consuming and tedious this task becomes. Importantly, customers may negotiate changes to the standard template which must be tracked and met on an on-going basis. This is how contract lifecycle management (CLM) becomes challenging. If this process is not well managed, a customer may not be billed or underbilled, a contract renewal may be missed, the services rendered may not meet a service-level agreement noted in the contract, or worse.

One solution to this problem is to automate the extraction of key words and phrases in contracts so that they are all shown in one file or output. These key words and phrases will be categorized by the relevant actor who is responsible for this area. For example, if a clause is mainly related to finance or a service-level

agreement then a team member on one of those teams or both should know about it and take appropriate action. Using this type of output, employees may be assigned tasks and be more aware of deliverable requirements before, during, and after the products and/or services are rendered, contracts that may be expiring soon, and more.

CLM is one use case in extracting key data from sometimes lengthy documents**,** including Microsoft Word documents or PDFs, and turning them into actionable insight potentially saving the company time and money in the process. This type of automated data extraction can be applied to other documents, such as lease agreements, line of credit agreements, invoices, purchase orders, financial statements, and so on.

Usefulness of Machine Learning Model to MetricStream

Aside from the benefits of automating manual processes, machine learning is especially useful in domain specific oddities. Contracts often use highly specific terminology to define an action or state. Contracts are structured with odd breaks in lines and use a variety of punctuation and characters. Capturing all these structural issues in a set of rules can be difficult, but through machine learning, these oddities can be dealt with via training a contract specific model.

The out-of-box natural language processing components of the “spacy” library are utilized in several sections throughout our tool pipeline. Firstly, “sentencizer” is a natural language processing tool that breaks a document into sentence-like chunks rather than on specific demarcations. Named entity recognition uses specific data labeling that can be used to extract relevant metadata terminology. Dependency parsing looks at the relationship between words and aids in grabbing relevant terminology based on its relationship to other relevant terminology. Synonym usage and lemmatization allows us to broaden search and capture the relevant metadata terminology as it appears in contracts.

These out-of-box tools can be further trained and specialized towards our contracts as we generate labeled training sets for our model. This will further strengthen our tool the longer it is used.

“Sentencizer”	Breaks up document into sentence-like chunks
Named Entity Recognition	Label words with entity tags
Dependency Parser	Displays relationship between words
Synonym and Lemmatization	Finds synonym and similar words for search

Progress to Date (Use cases of AI and Contract Management)

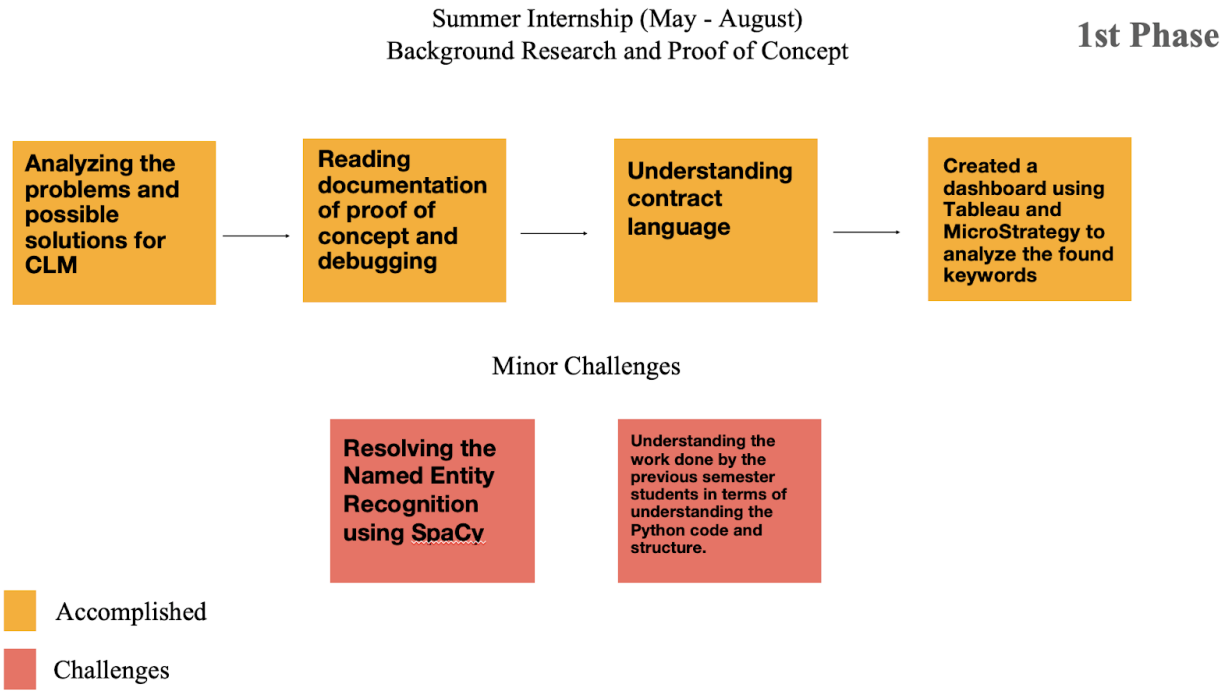
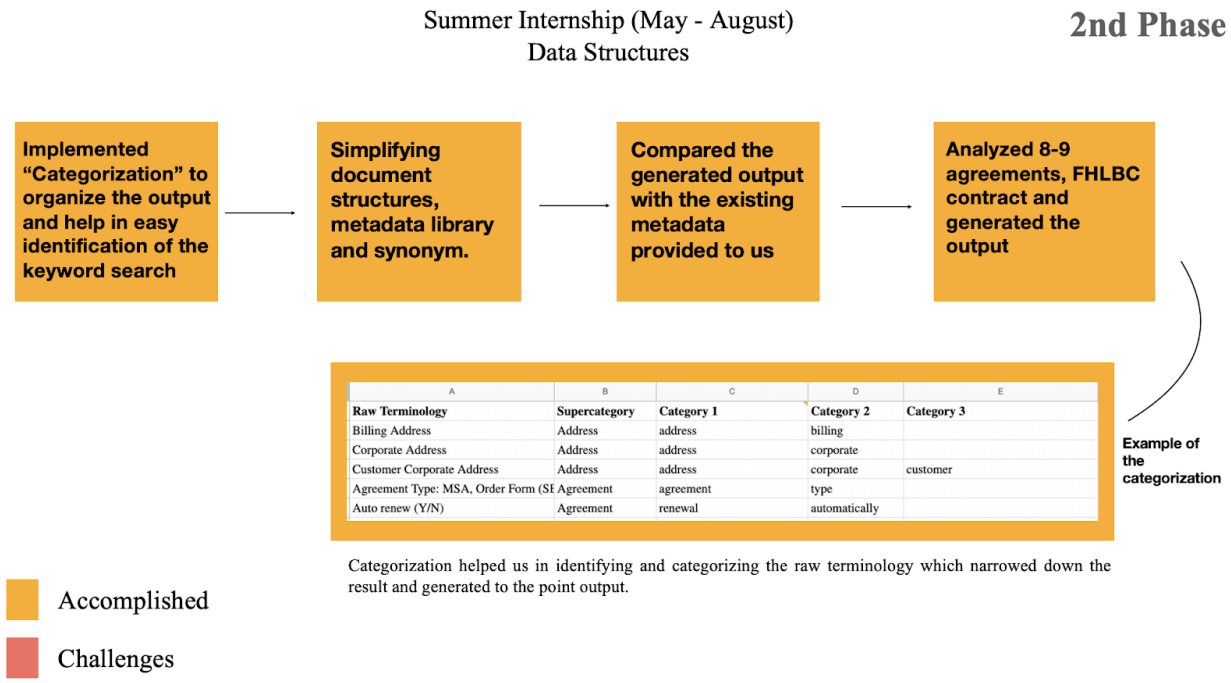


Figure 2: Starting Point

This phase highlights the initial stage of our internship wherein we were trying to understand the work done by the previous semester students as well as analyzing the possible problems and solutions for the contract lifecycle management. We read the documentation of the proof of concept and debugged the errors from the Python code/structure to suit the requirement. By gradually understanding the contract language, we created a drill down dashboard of the demo contracts using Tableau and MicroStrategy that helped us analyze the found keywords. Some minor challenges included resolving the named entity recognition using SpaCy library and deeper understanding of the work done by previous semester students in terms of the Python code and structure.

Phase 2:



	A	B	C	D	E
	Raw Terminology	**Supercategory**	**Category 1**	**Category 2**	**Category 3**
	Billing Address	Address	address	billing	
	Corporate Address	Address	address	corporate	
	Customer Corporate Address	Address	address	corporate	customer
	Agreement Type: MSA, Order Form (SE	Agreement	agreement	type	
	Auto renew (Y/N)	Agreement	renewal	automatically	

Figure 3: 2nd Phase

In this second phase, we concentrated categorizing various terminologies that will help us in the easy identification of the keywords search. Then we focused on simplifying the document structures and the metadata library. Categorizing the raw terminology narrowed down the result and generated the desired output.

Overview of the Categorization List

Initial Metadata Categorization

A	B	C	D
Combined Metadata Categorization	**Category**	**Sub-Category**	**Category (Detailed)**
Customer Corporate Address	Address	Customer	
Corporate Address	Address	Corporate	Organization
Billing Address	Address	Billing	String
Agreement Type: MSA, Order Form (SBD), SOW, CR, Work Order, amendments	Agreement	Amendments	Contract
Type of Support	Agreement		
Churn ACV	Agreement	Churn/AVC	
Churn Type (What is being churned)- text	Agreement	churn	type
breach	Agreement	Breach	Contract
backup	Agreement	backup	Contract
breach	Agreement	breach	Contract
Renewal Price Increase	Financial	Value	Renewal
Escrow Y/N	Financial	escrow	
uptime	Agreement/Time	Day/Date	
License Fee	Fee	Value	License
Cloud Fee	Fee	Value	
Platform Fee	Fee	Value	
Support Fee	Fee	Value	
Third party product fee	Fee	Value	
Auto fee increase (Y/N)	Fee	Value	

Figure 6: Categorization List (Initial Metadata Categorization)

Full Metadata Library

Category	Sub-Category	Metadata_Original	Metadata	Term1	Term2	Term3
Value		99	99	99		
Value		99.8	99.8	99		
Value		99.9	99.9	99		
Time	Hours	99.99	99.99	99		
Time	Days	12 hours	12 hours	hours	12	
Time	Hours/days	15 days	15 days	days	15	
Time	Hours/days	24 hours	24 hours	hours	24	
Time	Hours/Days	30 days	30 days	days	30	
Time	Hours/Days	48 hours	48 hours	hours	48	
Person	Name	Account Owner	Account Owner	owner	account	
Agreement		Agreement Type	Master Service Agreement	Master Service Agreement		
Agreement		Agreement Type	Order Form	Order Form		
Agreement		Agreement Type	Statement of Work	Statement of Work		
Agreement		Agreement Type	Cost Reimbursable	Cost Reimbursable		
Agreement		Agreement Type	Amendment	Amendment		
Agreement		Agreement Type	Work Order	Work Order		
Financial	Value	Annual Renewal Value	Annual Renewal Value	value	renewal	annual
Fee	Value	Auto fee increase (Y/N)	Auto fee increase	fee	automatically	
Time	String	Auto renew (Y/N)	Auto renew	renewal	automatically	
Agreement	backup	backup	backup	backup		
Address	Billing	Billing Address	Billing Address	address	billing	
Agreement	Breach	breach	breach	breach		
Agreement	Churn/AVC	Churn ACV	Churn ACV	churn	acv	
Time	Date	Churn Effective Date	Churn Effective Date	date	churn	effective
Time	Date	Churn Notification Date	Churn Notification Date	date	churn	notification
Agreement	churn	Churn Type (What is being churned)- text	Churn Type	churn	type	
Person	Name/ Yes or No	Client Consent required to assign - Y/N	Client Consent	client	consent	
Person	Name	Client Name	Client Name	client	name	
Product	SKU	Cloud Application SKU	Cloud Application SKU	sku	cloud	application
Fee	Value	Cloud Fee	Cloud Fee	fee	cloud	
Legal		Contract Term- 1 year, multiple years, optional years	Contract Term	contract	term	
Legal	contract	Contractual Relationship	End User License Agreement	End User License Agreement		
Legal	contract	Contractual Relationship	Cancellation Letter	Cancellation Letter		

Figure 7: Full Metadata library (including key terms and categorization)

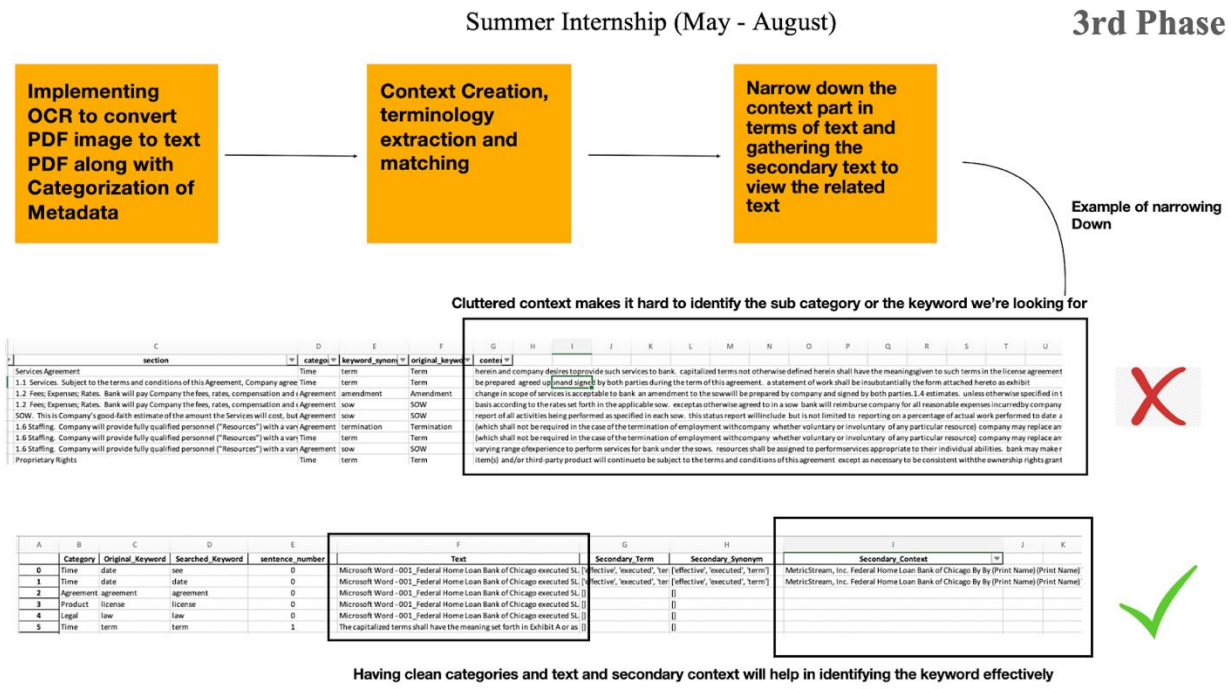


Figure 4: 3rd Phase

The third phase focused on narrowing down the context part of the output generated. The cluttered context made it hard to identify the categories as well as the keywords we were looking for. After running the code, the output generated had the text, secondary term, secondary synonyms and the secondary context. This was done in order to gather the surrounding text which the original text/context missed out. This gathering of the secondary text will help in better understanding of the output. `

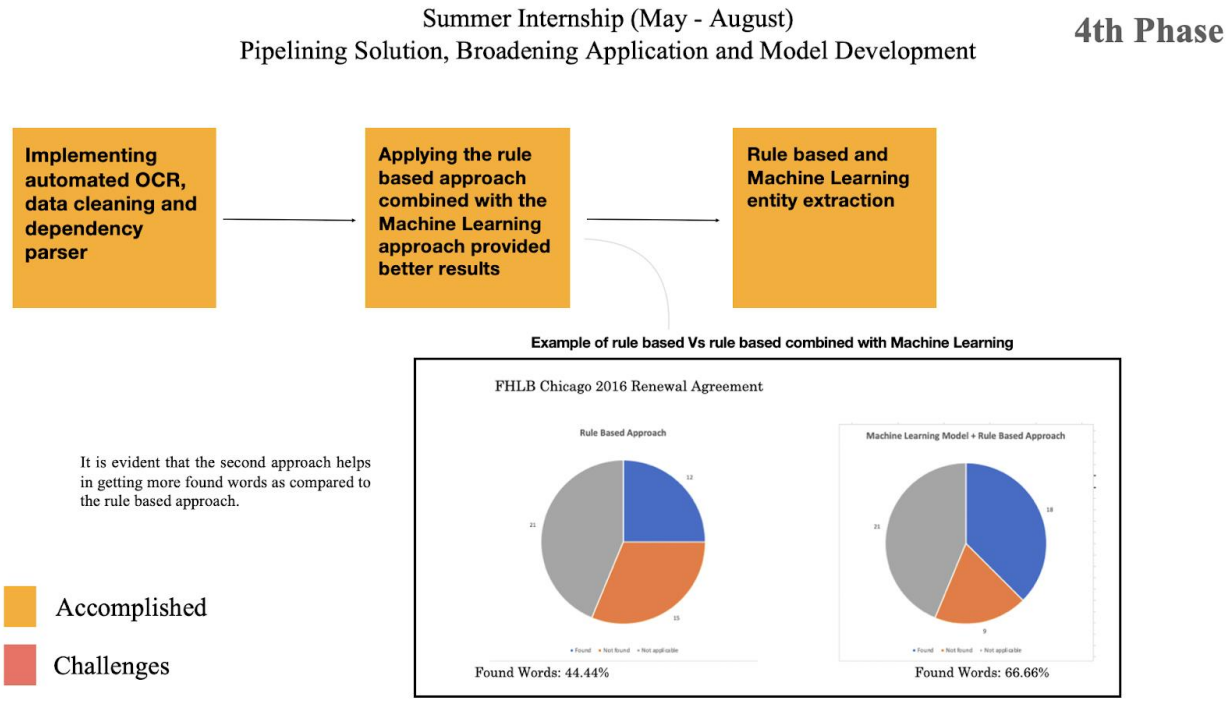


Figure 5: 4th Phase

This phase was more focused on implementing an automated OCR system. After implementing that, there were two approaches to provide the output: a) Rule based; b) Rule based combined with Machine Learning entity extraction. After applying both the approaches, the results were better in the second approach (as mentioned in the figure).

Overview of the Categorization List

Initial Stage

A	B	C	D
Combined Metadata Categorization	**Category**	**Sub-Category**	**Category (Detailed)**
Customer Corporate Address	Address	Customer	
Corporate Address	Address	Corporate	Organization
Billing Address	Address	Billing	String
Agreement Type: MSA, Order Form (SBD), SOW, CR, Work Order, amendments	Agreement	Amendments	Contract
Type of Support	Agreement		
Churn ACV	Agreement	Churn/AVC	
Churn Type (What is being churned)- text	Agreement	churn	type
breach	Agreement	Breach	Contract
backup	Agreement	backup	Contract
breach	Agreement	breach	Contract
Renewal Price Increase	Financial	Value	Renewal
Escrow Y/N	Financial	escrow	
uptime	Agreement/Time	Day/Date	
License Fee	Fee	Value	License
Cloud Fee	Fee	Value	
Platform Fee	Fee	Value	
Support Fee	Fee	Value	
Third party product fee	Fee	Value	
Auto fee increase (Y/N)	Fee	Value	

Figure 6: Categorization List (Initial Stage)

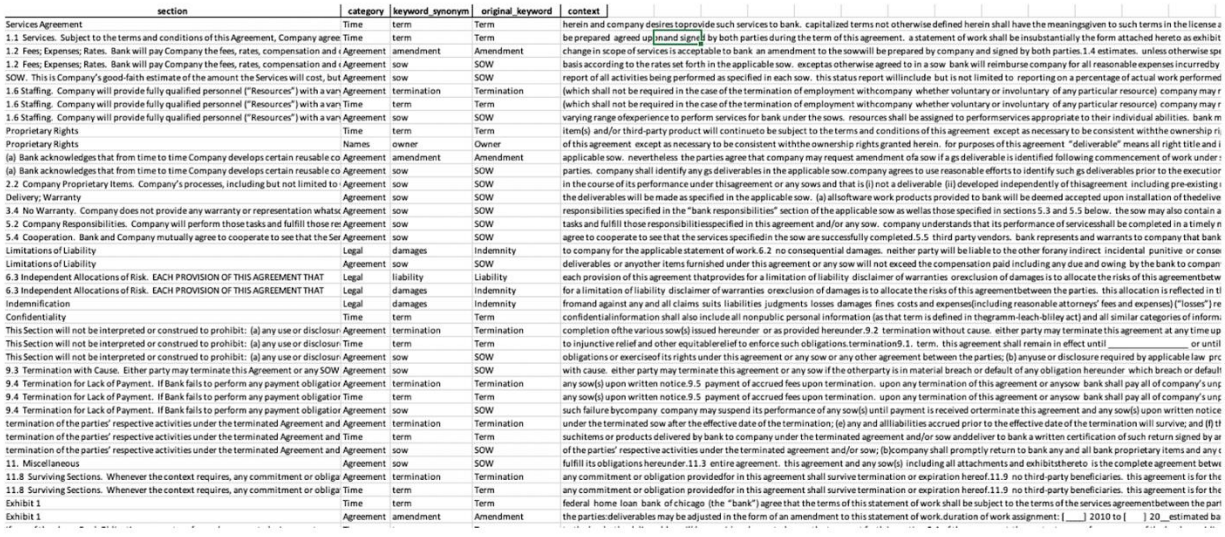
After improvising

Category	Sub-Category	Metadata_Original	Metadata	Term1	Term2	Term3
Value		99	99	99		
Value		99.8	99.8	99		
Value		99.9	99.9	99		
Time	Hours	99.99	99.99	99		
Time	Days	12 hours	12 hours	hours	12	
Time	Hours/days	15 days	15 days	days	15	
Time	Hours/days	24 hours	24 hours	hours	24	
Time	Hours/Days	30 days	30 days	days	30	
Time	Hours/Days	48 hours	48 hours	hours	48	
Person	Name	Account Owner	Account Owner	owner	account	
Agreement		Agreement Type	Master Service Agreement	Master Service Agreement		
Agreement		Agreement Type	Order Form	Order Form		
Agreement		Agreement Type	Statement of Work	Statement of Work		
Agreement		Agreement Type	Cost Reimbursable	Cost Reimbursable		
Agreement		Agreement Type	Amendment	Amendment		
Agreement		Agreement Type	Work Order	Work Order		
Financial	Value	Annual Renewal Value	Annual Renewal Value	value	renewal	annual
Fee	Value	Auto fee increase (Y/N)	Auto fee increase	fee	automatically	
Time	String	Auto renew (Y/N)	Auto renew	renewal	automatically	
Agreement	backup	backup	backup	backup		
Address	Billing	Billing Address	Billing Address	address	billing	
Agreement	Breach	breach	breach	breach		
Agreement	Churn/AVC	Churn ACV	Churn ACV	churn	acv	
Time	Date	Churn Effective Date	Churn Effective Date	date	churn	effective
Time	Date	Churn Notification Date	Churn Notification Date	date	churn	notification
Agreement	churn	Churn Type (What is being churned)- text	Churn Type	churn	type	
Person	Name/ Yes or No	Client Consent required to assign - Y/N	Client Consent	client	consent	
Person	Name	Client Name	Client Name	client	name	
Product	SKU	Cloud Application SKU	Cloud Application SKU	sku	cloud	application
Fee	Value	Cloud Fee	Cloud Fee	fee	cloud	
Legal		Contract Term- 1 year, multiple years, optional years	Contract Term	contract	term	
Legal	contract	Contractual Relationship	End User License Agreement	End User License Agreement		
Legal	contract	Contractual Relationship	Cancellation Letter	Cancellation Letter		

Figure 7: Categorization List (After Improving)

Sample work of the output

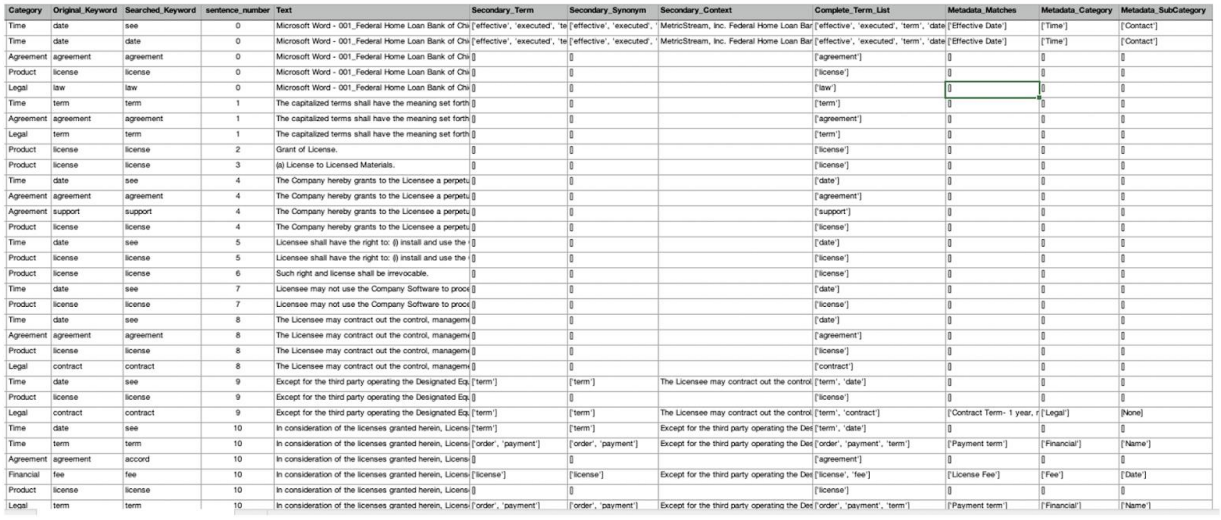
Initial Stage



section	category	keyword_synonym	original_keyword	context
Services Agreement	Time	term	Term	herein and company desires toprovide such services to bank. capitalized terms not otherwise defined herein shall have the meaningsgiven to such terms in the license a
1.1 Services. Subject to the terms and conditions of this Agreement, Company agree	Time	term	Term	be prepared agreed upon and signed by both parties during the term of this agreement. a statement of work shall be insubstantially the form attached hereto as exhibit
1.2 Fees; Expenses; Rates. Bank will pay Company the fees, rates, compensation and	Agreement	amendment	Amendment	change in scope of services is acceptable to bank an amendment to the sowwill be prepared by company and signed by both parties.1.4 estimates. unless otherwise spe
1.2 Fees; Expenses; Rates. Bank will pay Company the fees, rates, compensation and	Agreement	sow	SOW	basis according to the rates set forth in the applicable sow. exceptas otherwise agreed to in a sow bank will reimburse company for all reasonable expenses incurredby
SOW. This is Company's good-faith estimate of the amount the Services will cost, but	Agreement	sow	SOW	report of all activities being performed as specified in each sow. this status report willinclude but is not limited to reporting on a percentage of actual work performed
1.6 Staffing. Company will provide fully qualified personnel ("Resources") with a van	Agreement	termination	Termination	(which shall not be required in the case of the termination of employment withcompany whether voluntary or involuntary of any particular resource) company may r
1.6 Staffing. Company will provide fully qualified personnel ("Resources") with a van	Time	term	Term	(which shall not be required in the case of the termination of employment withcompany whether voluntary or involuntary of any particular resource) company may r
1.6 Staffing. Company will provide fully qualified personnel ("Resources") with a van	Agreement	sow	SOW	varying range ofexperience to perform services for bank under the sows. resources shall be assigned to performservices appropriate to their individual abilities. bank m
Proprietary Rights	Time	term	Term	item(s) and/or third-party product will continueto be subject to the terms and conditions of this agreement except as necessary to be consistent withthe ownership ri
Proprietary Rights	Names	owner	Owner	of this agreement except as necessary to be consistent withthe ownership rights granted herein. for purposes of this agreement "deliverable" means all right title and i
(a) Bank acknowledges that from time to time Company develops certain reusable co	Agreement	amendment	Amendment	applicable sow. nevertheless the parties agree that company may request amendment ofa sow if a gs deliverable is identified following commencement of work under
(a) Bank acknowledges that from time to time Company develops certain reusable co	Agreement	sow	SOW	parties. company shall identify any gs deliverables in the applicable sow.company agrees to use reasonable efforts to identify such gs deliverables prior to the executior
2.2 Company Proprietary Items. Company's processes, including but not limited to	Agreement	sow	SOW	in the course of its performance under thisagreement or any sows and that is (i) not a deliverable (ii) developed independently of thisagreement including pre-existing
Delivery; Warranty	Agreement	sow	SOW	the deliverables will be made as specified in the applicable sow. (a) allsoftware work products provided to bank will be deemed accepted upon installation of thedelive
3.4 No Warranty. Company does not provide any warranty or representation whatso	Agreement	sow	SOW	responsibilities specified in the "bank responsibilities" section of the applicable sow as wellas those specified in sections 5.3 and 5.5 below. the sow may also contain a
5.2 Company Responsibilities. Company will perform those tasks and fulfill those re	Agreement	sow	SOW	tasks and fulfill those responsibilitiesspecified in this agreement and/or any sow. company understands that its performance of servicesshall be completed in a timely n
5.4 Cooperation. Bank and Company mutually agree to cooperate to see that the Ser	Agreement	sow	SOW	agree to cooperate to see that the services specifiedin the sow are successfully completed.5.5 third party vendors. bank represents and warrants to company that bank
Limitations of Liability	Legal	damages	Indemnity	to company for the applicable statement of work.6.2 no consequential damages. neither party will be liable to the other forany indirect incidental punitive or conse
Limitations of Liability	Agreement	sow	SOW	deliverables or anyother items furnished under this agreement or any sow will not exceed the compensation paid including any due and owing by the bank to compan
6.3 Independent Allocations of Risk. EACH PROVISION OF THIS AGREEMENT THAT	Legal	liability	Liability	each provision of this agreement thatprovides for a limitation of liability disclaimer of warranties orexclusion of damages is to allocate the risks of this agreementbetw
6.3 Independent Allocations of Risk. EACH PROVISION OF THIS AGREEMENT THAT	Legal	damages	Indemnity	for a limitation of liability disclaimer of warranties orexclusion of damages is to allocate the risks of this agreementbetween the parties. this allocation is reflected in t
Indemnification	Legal	damages	Indemnity	fromand against any and all claims suits liabilities judgments losses damages fines costs and expenses(including reasonable attorneys' fees and expenses) ("losses") re
Confidentiality	Time	term	Term	confidentialinformation shall also include all nonpublic personal information (as that term is defined in thegramm-leach-bliley act) and all similar categories of inform
This Section will not be interpreted or construed to prohibit: (a) any use or disclosur	Agreement	termination	Termination	completion ofthe various sow(s) issued hereunder or as provided hereunder.9.2 termination without cause. either party may terminate this agreement at any time up
This Section will not be interpreted or construed to prohibit: (a) any use or disclosur	Time	term	Term	to injunctive relief and other equitablerelief to enforce such obligations.termination9.1. term. this agreement shall remain in effect until _____ or until
This Section will not be interpreted or construed to prohibit: (a) any use or disclosur	Agreement	sow	SOW	obligations or exerciseof its rights under this agreement or any sow or any other agreement between the parties; (b) anyuse or disclosure required by applicable law prc
9.3 Termination with Cause. Either party may terminate this Agreement or any SOW	Agreement	sow	SOW	with cause. either party may terminate this agreement or any sow if the otherparty is in material breach or default of any obligation hereunder which breach or default
9.4 Termination for Lack of Payment. If Bank fails to perform any payment obligation	Agreement	termination	Termination	any sow(s) upon written notice.9.5 payment of accrued fees upon termination. upon any termination of this agreement or anysow bank shall pay all of company's unp
9.4 Termination for Lack of Payment. If Bank fails to perform any payment obligation	Time	term	Term	any sow(s) upon written notice.9.5 payment of accrued fees upon termination. upon any termination of this agreement or anysow bank shall pay all of company's unp
9.4 Termination for Lack of Payment. If Bank fails to perform any payment obligation	Agreement	sow	SOW	such failure bycompany company may suspend its performance of any sow(s) until payment is received orterminate this agreement and any sow(s) upon written notice
termination of the parties' respective activities under the terminated Agreement and	Agreement	termination	Termination	under the terminated sow after the effective date of the termination; (e) any and allliabilities accrued prior to the effective date of the termination will survive; and (f) t
termination of the parties' respective activities under the terminated Agreement and	Time	term	Term	suchitems or products delivered by bank to company under the terminated agreement and/or sow anddeliver to bank a written certification of such return signed by a
termination of the parties' respective activities under the terminated Agreement and	Agreement	sow	SOW	of the parties' respective activities under the terminated agreement and/or sow; (b)company shall promptly return to bank any and all bank proprietary items and any
11. Miscellaneous	Agreement	sow	SOW	fulfill its obligations hereunder.11.3 entire agreement. this agreement and any sow(s) including all attachments and exhibitsthereto is the complete agreement betwe
11.8 Surviving Sections. Whenever the context requires, any commitment or obliga	Agreement	termination	Termination	any commitment or obligation providedfor in this agreement shall survive termination or expiration hereof.11.9 no third-party beneficiaries. this agreement is for the
11.8 Surviving Sections. Whenever the context requires, any commitment or obliga	Time	term	Term	any commitment or obligation providedfor in this agreement shall survive termination or expiration hereof.11.9 no third-party beneficiaries. this agreement is for the
Exhibit 1	Time	term	Term	federal home loan bank of chicago (the "bank") agree that the terms of this statement of work shall be subject to the terms of the services agreementbetween the part
Exhibit 1	Agreement	amendment	Amendment	the parties:deliverables may be adjusted in the form of an amendment to this statement of work.duration of work assignment: [____] 2010 to [ ] 20__estimated ba

Figure 8: Sample output (Initial stage)

After improving



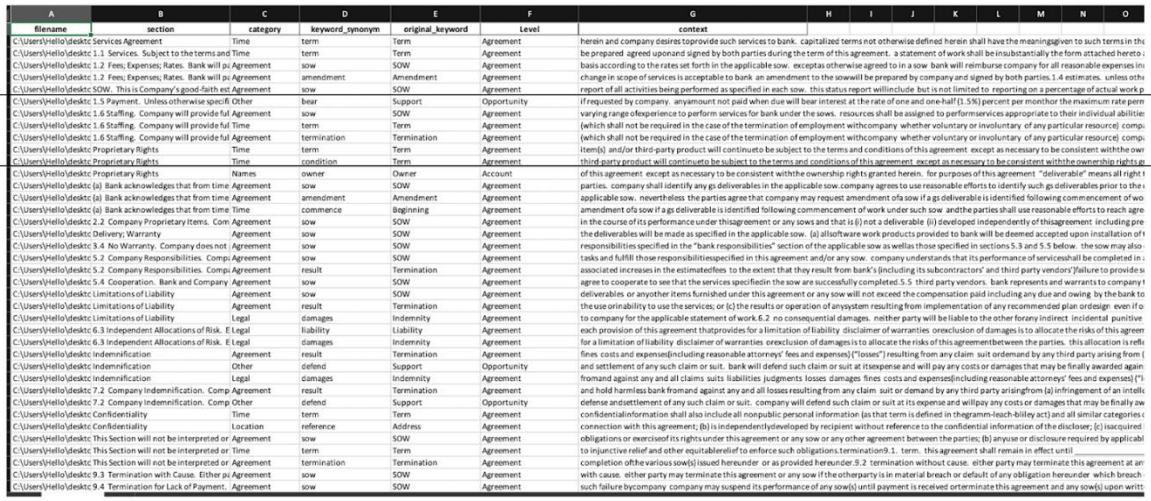
Category	Original_Keyword	Searched_Keyword	sentence_number	Text	Secondary_Term	Secondary_Synonym	Secondary_Context	Complete_Term_List	Metadata_Matches	Metadata_Category	Metadata_SubCategory
Time	date	see	0	Microsoft Word - 001_Federal Home Loan Bank of Chi	['effective', 'executed', 'te	['effective', 'executed', '	MetricStream, Inc. Federal Home Loan Bar	['effective', 'executed', 'term', 'date	['Effective Date']	['Time']	['Contact']
Time	date	date	0	Microsoft Word - 001_Federal Home Loan Bank of Chi	['effective', 'executed', 'te	['effective', 'executed', '	MetricStream, Inc. Federal Home Loan Bar	['effective', 'executed', 'term', 'date	['Effective Date']	['Time']	['Contact']
Agreement	agreement	agreement	0	Microsoft Word - 001_Federal Home Loan Bank of Chi	[]	[]		['agreement']	[]	[]	[]
Product	license	license	0	Microsoft Word - 001_Federal Home Loan Bank of Chi	[]	[]		['license']	[]	[]	[]
Legal	law	law	0	Microsoft Word - 001_Federal Home Loan Bank of Chi	[]	[]		['law']	[]	[]	[]
Time	term	term	1	The capitalized terms shall have the meaning set forth	[]	[]		['term']	[]	[]	[]
Agreement	agreement	agreement	1	The capitalized terms shall have the meaning set forth	[]	[]		['agreement']	[]	[]	[]
Legal	term	term	1	The capitalized terms shall have the meaning set forth	[]	[]		['term']	[]	[]	[]
Product	license	license	2	Grant of License.	[]	[]		['license']	[]	[]	[]
Product	license	license	3	(a) License to Licensed Materials.	[]	[]		['license']	[]	[]	[]
Time	date	see	4	The Company hereby grants to the Licensee a perpetu	[]	[]		['date']	[]	[]	[]
Agreement	agreement	agreement	4	The Company hereby grants to the Licensee a perpetu	[]	[]		['agreement']	[]	[]	[]
Agreement	support	support	4	The Company hereby grants to the Licensee a perpetu	[]	[]		['support']	[]	[]	[]
Product	license	license	4	The Company hereby grants to the Licensee a perpetu	[]	[]		['license']	[]	[]	[]
Time	date	see	5	Licensee shall have the right to: (i) install and use the	[]	[]		['date']	[]	[]	[]
Product	license	license	5	Licensee shall have the right to: (i) install and use the	[]	[]		['license']	[]	[]	[]
Product	license	license	6	Such right and license shall be irrevocable.	[]	[]		['license']	[]	[]	[]
Time	date	see	7	Licensee may not use the Company Software to proce	[]	[]		['date']	[]	[]	[]
Product	license	license	7	Licensee may not use the Company Software to proce	[]	[]		['license']	[]	[]	[]
Time	date	see	8	The Licensee may contract out the control, managem	[]	[]		['date']	[]	[]	[]
Agreement	agreement	agreement	8	The Licensee may contract out the control, managem	[]	[]		['agreement']	[]	[]	[]
Product	license	license	8	The Licensee may contract out the control, managem	[]	[]		['license']	[]	[]	[]
Legal	contract	contract	8	The Licensee may contract out the control, managem	[]	[]		['contract']	[]	[]	[]
Time	date	see	9	Except for the third party operating the Designated Eq	['term']	['term']	The Licensee may contract out the control	['term', 'date']	[]	[]	[]
Product	license	license	9	Except for the third party operating the Designated Eq	[]	[]		['license']	[]	[]	[]
Legal	contract	contract	9	Except for the third party operating the Designated Eq	['term']	['term']	The Licensee may contract out the control	['term', 'contract']	['Contract Term- 1 year, r	['Legal']	[None]
Time	date	see	10	In consideration of the licenses granted herein, Licens	['term']	['term']	Except for the third party operating the Des	['term', 'date']	[]	[]	[]
Time	term	term	10	In consideration of the licenses granted herein, Licens	['order', 'payment']	['order', 'payment']	Except for the third party operating the Des	['order', 'payment', 'term']	['Payment term']	['Financial']	['Name']
Agreement	agreement	accord	10	In consideration of the licenses granted herein, Licens	[]	[]		['agreement']	[]	[]	[]
Financial	fee	fee	10	In consideration of the licenses granted herein, Licens	['license']	['license']	Except for the third party operating the Des	['license', 'fee']	['License Fee']	['Fee']	['Date']
Product	license	license	10	In consideration of the licenses granted herein, Licens	[]	[]		['license']	[]	[]	[]
Legal	term	term	10	In consideration of the licenses granted herein, Licens	['order', 'payment']	['order', 'payment']	Except for the third party operating the Des	['order', 'payment', 'term']	['Payment term']	['Financial']	['Name']

Figure 9: Sample output (After improving)

After improving, we have categorized the metadata field into several columns such as category, original key word, searched keyword, text, secondary term, secondary synonym, secondary context etc which helped in gathering the surrounding text

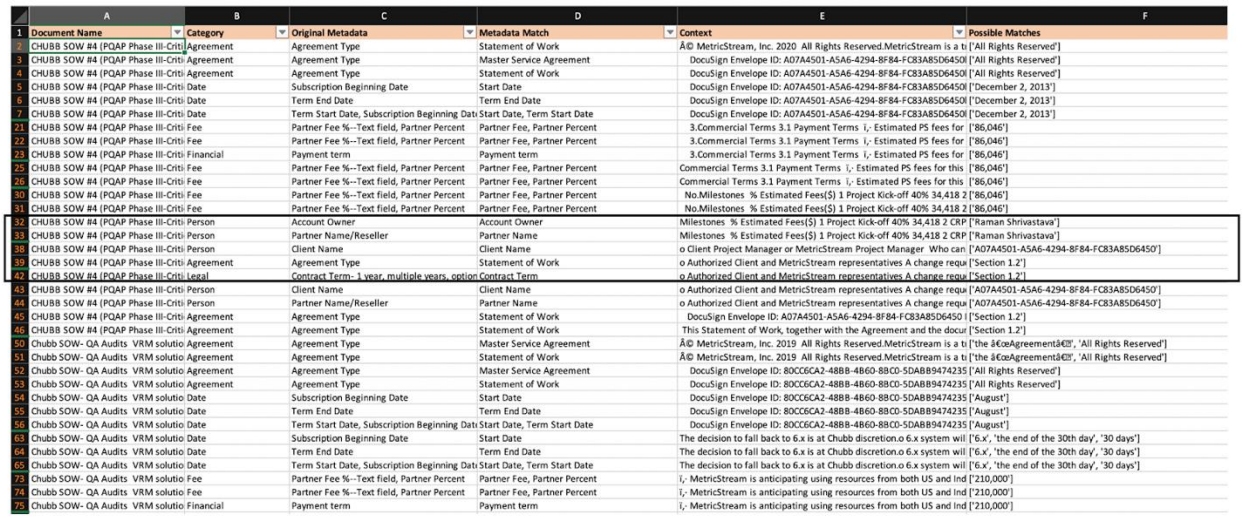
which the original context/synonym did not pick up. Apart from this, when this output was analyzed in Tableau, it not only helped in better understanding of the output but also gave it a cleaner look.

Key Findings:



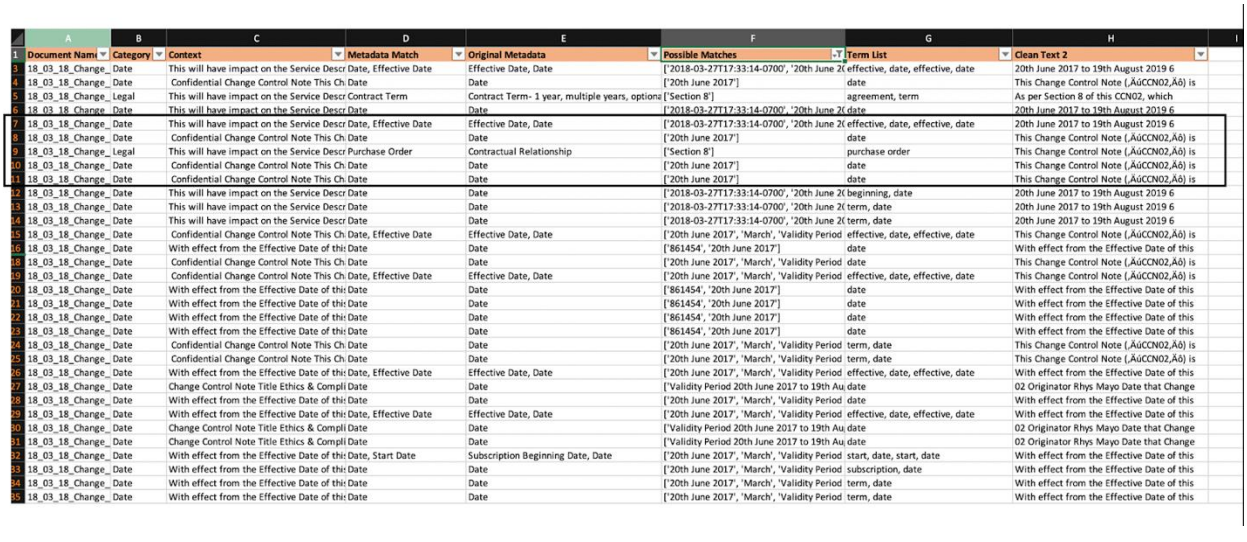
filename	section	category	keyword_synonym	original_keyword	Level	context
C:\Users\Hello\desktc	Services Agreement	Time	term	Term	Agreement	herein and company desires toprovide such services to bank. capitalized terms not otherwise defined herein shall have the meaningsgiven to such terms in the
C:\Users\Hello\desktc	1.1 Services. Subject to the terms and	Time	term	Term	Agreement	be prepared agreed uponand signed by both parties during the term of this agreement. a statement of work shall be insubstantially the form attached hereto
C:\Users\Hello\desktc	1.2 Fees; Expenses; Rates. Bank will p	Agreement	sow	SOW	Agreement	basis according to the rates set forth in the applicable sow. exceptas otherwise agreed to in a sow bank will reimburse company for all reasonable expenses in
C:\Users\Hello\desktc	1.2 Fees; Expenses; Rates. Bank will p	Agreement	amendment	Amendment	Agreement	change in scope of services is acceptable to bank an amendment to the sowwill be prepared by company and signed by both parties.1.4 estimates. unless oth
C:\Users\Hello\desktc	SOW. This is Company's good-faith est	Agreement	sow	SOW	Agreement	report of all activities being performed as specified in each sow. this status report willinclude but is not limited to reporting on a percentage of actual work p
C:\Users\Hello\desktc	1.5 Payment. Unless otherwise specifi	Other	bear	Support	Opportunity	if requested by company. anyamount not paid when due will bear interest at the rate of one and one-half (1.5%) percent per monthor the maximum rate perm
C:\Users\Hello\desktc	1.6 Staffing. Company will provide ful	Agreement	sow	SOW	Agreement	varying range ofexperience to perform services for bank under the sows. resources shall be assigned to performservices appropriate to their individual abilities
C:\Users\Hello\desktc	1.6 Staffing. Company will provide ful	Time	term	Term	Agreement	(which shall not be required in the case of the termination of employment withcompany whether voluntary or involuntary of any particular resource) compa
C:\Users\Hello\desktc	1.6 Staffing. Company will provide ful	Agreement	termination	Termination	Agreement	(which shall not be required in the case of the termination of employment withcompany whether voluntary or involuntary of any particular resource) compa
C:\Users\Hello\desktc	Proprietary Rights	Time	term	Term	Agreement	item(s) and/or third-party product will continueto be subject to the terms and conditions of this agreement except as necessary to be consistent withthe own
C:\Users\Hello\desktc	Proprietary Rights	Time	condition	Term	Agreement	third party product will continueto be subject to the terms and conditions of this agreement except as necessary to be consistent withthe ownership rights g
C:\Users\Hello\desktc	Proprietary Rights	Names	owner	Owner	Account	of this agreement except as necessary to be consistent withthe ownership rights granted herein. for purposes of this agreement "deliverable" means all right t
C:\Users\Hello\desktc	(a) Bank acknowledges that from time	Agreement	sow	SOW	Agreement	parties. company shall identify any gs deliverables in the applicable sow.company agrees to use reasonable efforts to identify such gs deliverables prior to the
C:\Users\Hello\desktc	(a) Bank acknowledges that from time	Agreement	amendment	Amendment	Agreement	applicable sow. nevertheless the parties agree that company may request amendment ofa sow if a gs deliverable is identified following commencement of wo
C:\Users\Hello\desktc	(a) Bank acknowledges that from time	Time	commence	Beginning	Agreement	amendment ofa sow if a gs deliverable is identified following commencement of work under such sow andthe parties shall use reasonable efforts to reach agre
C:\Users\Hello\desktc	2.2 Company Proprietary Items. Com	Agreement	sow	SOW	Agreement	in the course of its performance under thisagreement or any sows and that is (i) not a deliverable (ii) developed independently of thisagreement including pre
C:\Users\Hello\desktc	Delivery; Warranty	Agreement	sow	SOW	Agreement	the deliverables will be made as specified in the applicable sow. (a) allsoftware work products provided to bank will be deemed accepted upon installation of t
C:\Users\Hello\desktc	3.4 No Warranty. Company does not	Agreement	sow	SOW	Agreement	responsibilities specified in the "bank responsibilities" section of the applicable sow as wellas those specified in sections 5.3 and 5.5 below. the sow may also
C:\Users\Hello\desktc	5.2 Company Responsibilities. Compa	Agreement	sow	SOW	Agreement	tasks and fulfill those responsibilitiesspecified in this agreement and/or any sow. company understands that its performance of servicesshall be completed in
C:\Users\Hello\desktc	5.2 Company Responsibilities. Compa	Agreement	result	Termination	Agreement	associated increases in the estimatedfees to the extent that they result from bank's (including its subcontractors' and third party vendors')failure to provide s
C:\Users\Hello\desktc	5.4 Cooperation. Bank and Company	Agreement	sow	SOW	Agreement	agree to cooperate to see that the services specifiedin the sow are successfully completed.5.5 third party vendors. bank represents and warrants to company t
C:\Users\Hello\desktc	Limitations of Liability	Agreement	sow	SOW	Agreement	deliverables or anyother items furnished under this agreement or any sow will not exceed the compensation paid including any due and owing by the bank to
C:\Users\Hello\desktc	Limitations of Liability	Agreement	result	Termination	Agreement	the use orinability to use the services; or (c) the results or operation of anysystem resulting from implementation of any recommended plan ordesign even if o
C:\Users\Hello\desktc	Limitations of Liability	Legal	damages	Indemnity	Agreement	to company for the applicable statement of work.6.2 no consequential damages. neither party will be liable to the other forany indirect incidental punitive
C:\Users\Hello\desktc	6.3 Independent Allocations of Risk. E	Legal	liability	Liability	Agreement	each provision of this agreement thatprovides for a limitation of liability disclaimer of warranties orexclusion of damages is to allocate the risks of this agreem
C:\Users\Hello\desktc	6.3 Independent Allocations of Risk. E	Legal	damages	Indemnity	Agreement	for a limitation of liability disclaimer of warranties orexclusion of damages is to allocate the risks of this agreement between the parties. this allocation is refl
C:\Users\Hello\desktc	Indemnification	Agreement	result	Termination	Agreement	fines costs and expenses(including reasonable attorneys' fees and expenses) ("losses") resulting from any claim suit ordemand by any third party arising from (
C:\Users\Hello\desktc	Indemnification	Other	defend	Support	Opportunity	and settlement of any such claim or suit. bank will defend such claim or suit at itsexpense and will pay any costs or damages that may be finally awarded again
C:\Users\Hello\desktc	Indemnification	Legal	damages	Indemnity	Agreement	fromand against any and all claims suits liabilities judgments losses damages fines costs and expenses(including reasonable attorneys' fees and expenses) ("l
C:\Users\Hello\desktc	7.2 Company Indemnification. Comp	Agreement	result	Termination	Agreement	and hold harmless bank fromand against any and all losses resulting from any claim suit or demand by any third party arisingfrom (a) infringement of an intell
C:\Users\Hello\desktc	7.2 Company Indemnification. Comp	Other	defend	Support	Opportunity	defense andsettlement of any such claim or suit. company will defend such claim or suit at its expense and willpay any costs or damages that may be finally aw
C:\Users\Hello\desktc	Confidentiality	Time	term	Term	Agreement	confidentialinformation shall also include all nonpublic personal information (as that term is defined in thegramm-leach-bliley act) and all similar categories
C:\Users\Hello\desktc	Confidentiality	Location	reference	Address	Agreement	connection with this agreement; (b) is independentlydeveloped by recipient without reference to the confidential information of the discloser; (c) isacquired
C:\Users\Hello\desktc	This Section will not be interpreted or	Agreement	sow	SOW	Agreement	obligations or exercisesof its rights under this agreement or any sow or any other agreement between the parties; (b) anyuse or disclosure required by applicabl
C:\Users\Hello\desktc	This Section will not be interpreted or	Time	term	Term	Agreement	to injunctive relief and other equitablerelief to enforce such obligations.termination9.1. term. this agreement shall remain in effect until _____
C:\Users\Hello\desktc	This Section will not be interpreted or	Agreement	termination	Termination	Agreement	completion ofthe various sow(s) issued hereunder or as provided hereunder.9.2 termination without cause. either party may terminate this agreement at an
C:\Users\Hello\desktc	9.3 Termination with Cause. Either pa	Agreement	sow	SOW	Agreement	with cause. either party may terminate this agreement or any sow if the otherparty is in material breach or default of any obligation hereunder which breach
C:\Users\Hello\desktc	9.4 Termination for Lack of Payment.	Agreement	sow	SOW	Agreement	such failure bycompany company may suspend its performance of any sow(s) until payment is received orterminate this agreement and any sow(s) upon writt

Initial Stage: In the initial stage, we were able to extract the section of the document along with its category, keyword synonym, original keyword and the context. The context section however needed to be changed and to be on point.



	A	B	C	D	E	F
1	Document Name	Category	Original Metadata	Metadata Match	Context	Possible Matches
2	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Statement of Work	Â© MetricStream, Inc. 2020 All Rights Reserved.MetricStream is a ti	['All Rights Reserved']
3	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Master Service Agreement	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450(	['All Rights Reserved']
4	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Statement of Work	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450(	['All Rights Reserved']
5	CHUBB SOW #4 (PQAP Phase III-Criti	Date	Subscription Beginning Date	Start Date	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450(	['December 2, 2013']
6	CHUBB SOW #4 (PQAP Phase III-Criti	Date	Term End Date	Term End Date	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450(	['December 2, 2013']
7	CHUBB SOW #4 (PQAP Phase III-Criti	Date	Term Start Date, Subscription Beginning Dat	Start Date, Term Start Date	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450(	['December 2, 2013']
21	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	3.Commercial Terms 3.1 Payment Terms Ï¿· Estimated PS fees for	['86,046']
22	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	3.Commercial Terms 3.1 Payment Terms Ï¿· Estimated PS fees for	['86,046']
23	CHUBB SOW #4 (PQAP Phase III-Criti	Financial	Payment term	Payment term	3.Commercial Terms 3.1 Payment Terms Ï¿· Estimated PS fees for	['86,046']
25	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	Commercial Terms 3.1 Payment Terms Ï¿· Estimated PS fees for this	['86,046']
26	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	Commercial Terms 3.1 Payment Terms Ï¿· Estimated PS fees for this	['86,046']
30	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	No.Milestones % Estimated Fees($) 1 Project Kick-off 40% 34,418 2	['86,046']
31	CHUBB SOW #4 (PQAP Phase III-Criti	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	No.Milestones % Estimated Fees($) 1 Project Kick-off 40% 34,418 2	['86,046']
32	CHUBB SOW #4 (PQAP Phase III-Criti	Person	Account Owner	Account Owner	Milestones % Estimated Fees($) 1 Project Kick-off 40% 34,418 2 CRP	['Raman Shrivastava']
33	CHUBB SOW #4 (PQAP Phase III-Criti	Person	Partner Name/Reseller	Partner Name	Milestones % Estimated Fees($) 1 Project Kick-off 40% 34,418 2 CRP	['Raman Shrivastava']
38	CHUBB SOW #4 (PQAP Phase III-Criti	Person	Client Name	Client Name	o Client Project Manager or MetricStream Project Manager Who can	['A07A4501-A5A6-4294-8F84-FC83A85D6450']
39	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Statement of Work	o Authorized Client and MetricStream representatives A change requ	['Section 1.2']
42	CHUBB SOW #4 (PQAP Phase III-Criti	Legal	Contract Term- 1 year, multiple years, option	Contract Term	o Authorized Client and MetricStream representatives A change requ	['Section 1.2']
43	CHUBB SOW #4 (PQAP Phase III-Criti	Person	Client Name	Client Name	o Authorized Client and MetricStream representatives A change requ	['A07A4501-A5A6-4294-8F84-FC83A85D6450']
44	CHUBB SOW #4 (PQAP Phase III-Criti	Person	Partner Name/Reseller	Partner Name	o Authorized Client and MetricStream representatives A change requ	['A07A4501-A5A6-4294-8F84-FC83A85D6450']
45	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Statement of Work	DocuSign Envelope ID: A07A4501-A5A6-4294-8F84-FC83A85D6450 I	['Section 1.2']
46	CHUBB SOW #4 (PQAP Phase III-Criti	Agreement	Agreement Type	Statement of Work	This Statement of Work, together with the Agreement and the docur	['Section 1.2']
50	Chubb SOW- QA Audits VRM solutio	Agreement	Agreement Type	Master Service Agreement	Â© MetricStream, Inc. 2019 All Rights Reserved.MetricStream is a ti	['the â€œAgreementâ€', 'All Rights Reserved']
51	Chubb SOW- QA Audits VRM solutio	Agreement	Agreement Type	Statement of Work	Â© MetricStream, Inc. 2019 All Rights Reserved.MetricStream is a ti	['the â€œAgreementâ€', 'All Rights Reserved']
52	Chubb SOW- QA Audits VRM solutio	Agreement	Agreement Type	Master Service Agreement	DocuSign Envelope ID: 80CC6CA2-48BB-4B60-88C0-5DABB9474235	['All Rights Reserved']
53	Chubb SOW- QA Audits VRM solutio	Agreement	Agreement Type	Statement of Work	DocuSign Envelope ID: 80CC6CA2-48BB-4B60-88C0-5DABB9474235	['All Rights Reserved']
54	Chubb SOW- QA Audits VRM solutio	Date	Subscription Beginning Date	Start Date	DocuSign Envelope ID: 80CC6CA2-48BB-4B60-88C0-5DABB9474235	['August']
55	Chubb SOW- QA Audits VRM solutio	Date	Term End Date	Term End Date	DocuSign Envelope ID: 80CC6CA2-48BB-4B60-88C0-5DABB9474235	['August']
56	Chubb SOW- QA Audits VRM solutio	Date	Term Start Date, Subscription Beginning Dat	Start Date, Term Start Date	DocuSign Envelope ID: 80CC6CA2-48BB-4B60-88C0-5DABB9474235	['August']
63	Chubb SOW- QA Audits VRM solutio	Date	Subscription Beginning Date	Start Date	The decision to fall back to 6.x is at Chubb discretion.o 6.x system wil	['6.x', 'the end of the 30th day', '30 days']
64	Chubb SOW- QA Audits VRM solutio	Date	Term End Date	Term End Date	The decision to fall back to 6.x is at Chubb discretion.o 6.x system wil	['6.x', 'the end of the 30th day', '30 days']
65	Chubb SOW- QA Audits VRM solutio	Date	Term Start Date, Subscription Beginning Dat	Start Date, Term Start Date	The decision to fall back to 6.x is at Chubb discretion.o 6.x system wil	['6.x', 'the end of the 30th day', '30 days']
73	Chubb SOW- QA Audits VRM solutio	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	Ï¿· MetricStream is anticipating using resources from both US and Ind	['210,000']
74	Chubb SOW- QA Audits VRM solutio	Fee	Partner Fee %--Text field, Partner Percent	Partner Fee, Partner Percent	Ï¿· MetricStream is anticipating using resources from both US and Ind	['210,000']
75	Chubb SOW- QA Audits VRM solutio	Financial	Payment term	Payment term	Ï¿· MetricStream is anticipating using resources from both US and Ind	['210,000']

Mid-Point: In the mid-point, we focused on getting the possible matches which will help in identifying the keyword we're looking for. For example, we want to know the Account Owner of which is mentioned in the contract, rather than reading the entire context, the possible matches helped us in identifying the account owner. One can surely read the context for more information.



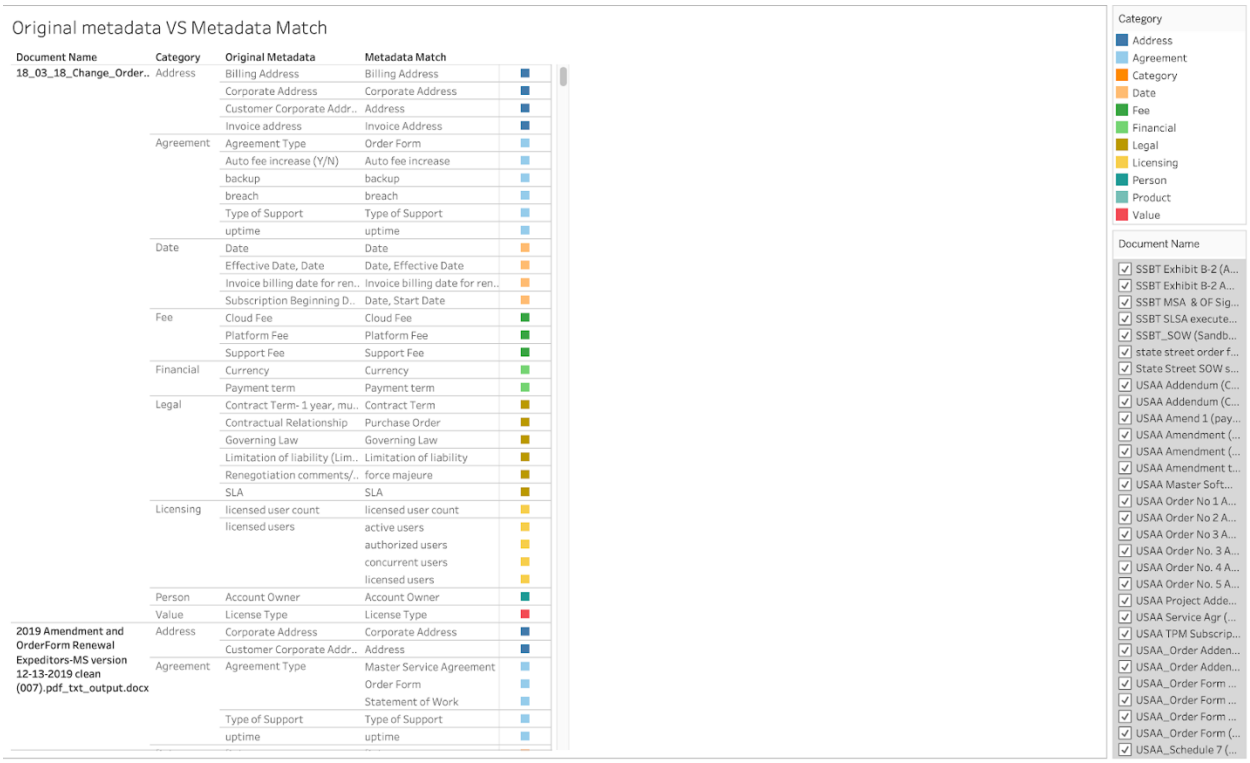
	A	B	C	D	E	F	G	H	I
1	Document Nam	Category	Context	Metadata Match	Original Metadata	Possible Matches	Term List	Clean Text 2	
3	18_03_18_Change_	Date	This will have impact on the Service Descr	Date, Effective Date	Effective Date, Date	['2018-03-27T17:33:14-0700', '20th June 2(	effective, date, effective, date	20th June 2017 to 19th August 2019 6	
4	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017']	date	This Change Control Note (,ÃúCCN02,Âô) is	
5	18_03_18_Change_	Legal	This will have impact on the Service Descr	Contract Term	Contract Term- 1 year, multiple years, option	['Section 8']	agreement, term	As per Section 8 of this CCN02, which	
6	18_03_18_Change_	Date	This will have impact on the Service Descr	Date	Date	['2018-03-27T17:33:14-0700', '20th June 2(	date	20th June 2017 to 19th August 2019 6	
7	18_03_18_Change_	Date	This will have impact on the Service Descr	Date, Effective Date	Effective Date, Date	['2018-03-27T17:33:14-0700', '20th June 2(	effective, date, effective, date	20th June 2017 to 19th August 2019 6	
8	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017']	date	This Change Control Note (,ÃúCCN02,Âô) is	
9	18_03_18_Change_	Legal	This will have impact on the Service Descr	Purchase Order	Contractual Relationship	['Section 8']	purchase order	This Change Control Note (,ÃúCCN02,Âô) is	
10	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017']	date	This Change Control Note (,ÃúCCN02,Âô) is	
11	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017']	date	This Change Control Note (,ÃúCCN02,Âô) is	
12	18_03_18_Change_	Date	This will have impact on the Service Descr	Date	Date	['2018-03-27T17:33:14-0700', '20th June 2(	beginning, date	20th June 2017 to 19th August 2019 6	
13	18_03_18_Change_	Date	This will have impact on the Service Descr	Date	Date	['2018-03-27T17:33:14-0700', '20th June 2(	term, date	20th June 2017 to 19th August 2019 6	
14	18_03_18_Change_	Date	This will have impact on the Service Descr	Date	Date	['2018-03-27T17:33:14-0700', '20th June 2(	term, date	20th June 2017 to 19th August 2019 6	
15	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date, Effective Date	Effective Date, Date	['20th June 2017', 'March', 'Validity Period	effective, date, effective, date	This Change Control Note (,ÃúCCN02,Âô) is	
16	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['861454', '20th June 2017']	date	With effect from the Effective Date of this	
18	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017', 'March', 'Validity Period	date	This Change Control Note (,ÃúCCN02,Âô) is	
19	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date, Effective Date	Effective Date, Date	['20th June 2017', 'March', 'Validity Period	effective, date, effective, date	This Change Control Note (,ÃúCCN02,Âô) is	
20	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['861454', '20th June 2017']	date	With effect from the Effective Date of this	
21	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['861454', '20th June 2017']	date	With effect from the Effective Date of this	
22	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['861454', '20th June 2017']	date	With effect from the Effective Date of this	
23	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['861454', '20th June 2017']	date	With effect from the Effective Date of this	
24	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017', 'March', 'Validity Period	term, date	This Change Control Note (,ÃúCCN02,Âô) is	
25	18_03_18_Change_	Date	Confidential Change Control Note This Ch	Date	Date	['20th June 2017', 'March', 'Validity Period	term, date	This Change Control Note (,ÃúCCN02,Âô) is	
26	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date, Effective Date	Effective Date, Date	['20th June 2017', 'March', 'Validity Period	effective, date, effective, date	With effect from the Effective Date of this	
27	18_03_18_Change_	Date	Change Control Note Title Ethics & Compli	Date	Date	['Validity Period 20th June 2017 to 19th Au	date	02 Originator Rhys Mayo Date that Change	
28	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['20th June 2017', 'March', 'Validity Period	date	With effect from the Effective Date of this	
29	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date, Effective Date	Effective Date, Date	['20th June 2017', 'March', 'Validity Period	effective, date, effective, date	With effect from the Effective Date of this	
30	18_03_18_Change_	Date	Change Control Note Title Ethics & Compli	Date	Date	['Validity Period 20th June 2017 to 19th Au	date	02 Originator Rhys Mayo Date that Change	
31	18_03_18_Change_	Date	Change Control Note Title Ethics & Compli	Date	Date	['Validity Period 20th June 2017 to 19th Au	date	02 Originator Rhys Mayo Date that Change	
32	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date, Start Date	Subscription Beginning Date, Date	['20th June 2017', 'March', 'Validity Period	start, date, start, date	With effect from the Effective Date of this	
33	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['20th June 2017', 'March', 'Validity Period	subscription, date	With effect from the Effective Date of this	
34	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['20th June 2017', 'March', 'Validity Period	term, date	With effect from the Effective Date of this	
35	18_03_18_Change_	Date	With effect from the Effective Date of thi	Date	Date	['20th June 2017', 'March', 'Validity Period	term, date	With effect from the Effective Date of this	

End Point: In the final stage, we focused on getting the outcome as precise as possible. For example, if we are looking for "effective date", the outcome will help

in identifying the possible matches (as date) and the term list. Along with the context, it also shows the clean text which helps in understanding the outcome in a better way.

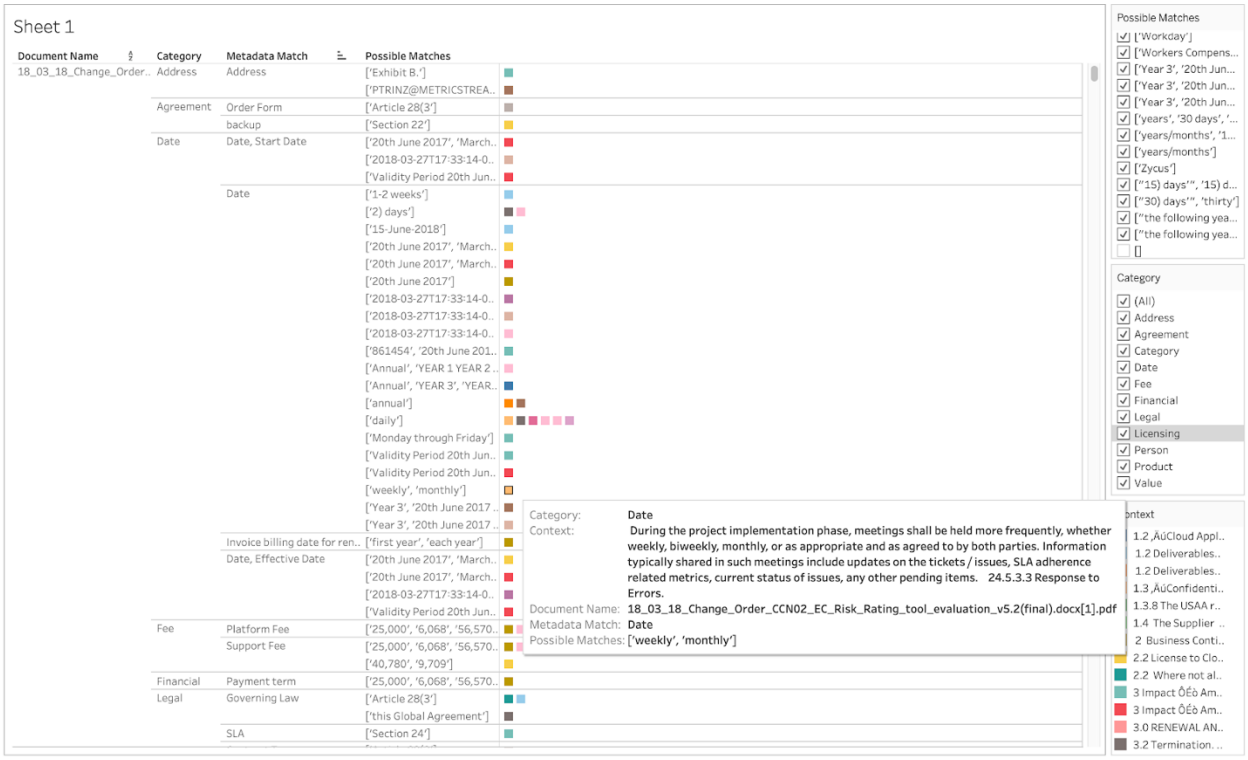
As compared to the initial stage, we're in a better position to identify the outcome. Instead of reading the entire context, we have the possible matches along with the context and the clean text.

Tableau Dashboard



This dashboard focused on the Original Metadata Vs the Metadata Match. We categorized the terminology into various categories and assigned them metadata's. This metadata was then matched by the original metadata and the outcome was as

follows which had the context, along with the metadata matches and possible matches.



The aim of the dashboard is to be able to identify the categories we're looking for according to the document/contract name and find the keyword/synonym. The dashboard helps us to select the category through the filter and find the context in the colored box for more reference.

Conclusion

Applying AI to CLM has already been occurring. In fact, AI can be used to manage a contract's life from beginning to end. From drafting a contract, obtaining approval, alerting and reminding teams of action items on an on-going basis, and more. This can be done by integrating an automatic extraction

algorithm with other systems such as a CRM system. Beyond applying automatic extraction to other types of documents (I.e. invoices, purchase orders, financial statements, etc.), this process can be pushed to the edge of computing, so that the extraction is performed locally.

Future Work and Implementations

We suggest performing these next steps in the provided order.

1.) OCR Solution

The current implementation of Tesseract OCR has several faults. Depending on document quality, the output text may be not be represented in text properly. If we continue to utilize open source technology, future work must be done to improve the text output. Improving the actual optical character recognition technology itself is not a feasible goal and therefore we are limited to what we are able to accomplish with the current output.

One feature of the current tool searches for characters that are clearly incorrect and either removes or replaces them. Future work can look to implement a form of spell checking and word recognition to improve upon current text output.

2.) Customization of metadata library and synonym finder

Current metadata library consists of the initial set of metadata terms provided to us from the contract desk team. An exhaustive list is yet to be

compiled. Adding new metadata fields is as easy as adding a new row to the metadata library excel file.

The terminology list and synonym finder however can be developed as well. For example, synonyms for words such as ‘date’ or ‘uptime’ may be relevant to our metadata field. However, synonyms for ‘cloud’ in our circumstances would not be relevant and therefore updating the synonym library to be exclusive to contract language may be an avenue worth exploring. The nltk wordnet synonym library contains a large database of synonyms and the current tool only makes use of up to 3 synonyms. In certain cases this is too many and in other cases, too few. A more sophisticated method of synonym usage may help in finding more diverse sets of text for a given metadata field.

3.) Named Entity Recognition, Extraction, and the Dependency Parser

Named entity recognition applies labels to words. For example, it will look for dates and apply the tag DATE. In simple cases, this feature of the spacy library works well. However, in cases that are more specific to contracts such as ‘limitation of liability’ the entity recognizer has no tag capable of finding relevant terms. The contract itself must be manually labeled as training data and fed into the spacy machine learning model updater so that these terms can be automatically labeled in the future.

The ability to automatically label very specific pieces of text allow the dependency parser to answer questions directly. While the current implementation of the spacy model allows one to label monetary values and extract them, if the nearby text is unclear about what the monetary value relates to, we must manually perform that. For example, a piece of text may contain the word ‘cloud fee’, ‘license fee’, ‘$10,000’, ’$20,000’. The tool can find these two words and two dollar amounts but determining which word relates to which dollar amount is a manual process. Training the dependency parser using user labeled data will allow the tool to automatically relate the correct term to the correct dollar amount.