The Law of Large Documents: Understanding the Structure of Legal Contracts Using Visual Cues

Allison Hegel
Lexion
Seattle, Washington, USA
allison@lexion.ai

Marina Shah
Lexion
Seattle, Washington, USA
marina@lexion.ai

Genevieve Peaslee
Lexion
Seattle, Washington, USA
genevieve@lexion.ai

Brendan Roof
Lexion
Seattle, Washington, USA
brendan@lexion.ai

Emad Elwany
Lexion
Seattle, Washington, USA
emad@lexion.ai

ABSTRACT

Large, pre-trained transformer models like BERT have achieved state-of-the-art results on document understanding tasks, but most implementations can only consider 512 tokens at a time. For many real-world applications, documents can be much longer, and the segmentation strategies typically used on longer documents miss out on document structure and contextual information, hurting their results on downstream tasks. In our work on legal agreements, we find that visual cues such as layout, style, and placement of text in a document are strong features that are crucial to achieving an acceptable level of accuracy on long documents. We measure the impact of incorporating such visual cues, obtained via computer vision methods, on the accuracy of document understanding tasks including document segmentation, entity extraction, and attribute classification. Our method of segmenting documents based on structural metadata out-performs existing methods on four long-document understanding tasks as measured on the Contract Understanding Atticus Dataset.

CCS CONCEPTS

• **Applied computing → Law**; **Optical character recognition**;
• **Computing methodologies → Information extraction**.

KEYWORDS

long document understanding, document structure, OCR, NLP

1 INTRODUCTION

Businesses interact through contracts, and legal departments in many companies are stretched thin managing requests for information about these contracts. Companies need timely reporting on the date contracts are up for renewal, the length of cancellation periods, how much money they should bill, and other details that are embedded in the contract prose. Finding these answers quickly in a large collection of documents, some of which can be hundreds of pages long, can lead to a lot of frustration (and billable hours). At Lexion, we use natural language processing methods to perform this work automatically, extracting key pieces of information from contracts for lawyers to find at a glance [4].

The most powerful recent methods for document understanding are computationally expensive, and they are limited in the length
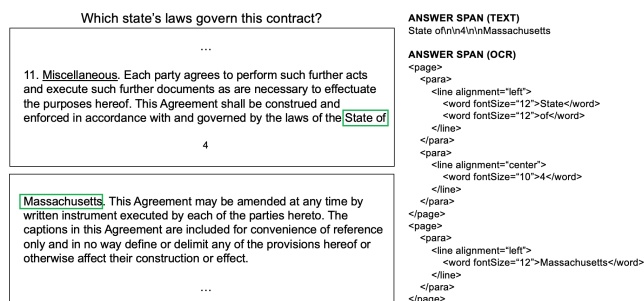

Figure 1: Legal counsel needs to quickly find which state or country's laws will govern an agreement, but this information is often hidden in long documents in sections like "Miscellaneous". In this example contract from CUAD [6], the desired answer, "State of Massachusetts", is broken across a page boundary, with the page number "4" in between. To extract this result correctly, a model would need information about the structure and formatting of the document to know that it should discard the page break, newlines, page number, and any other extraneous information common in legal documents such as headers and footers. OCR metadata provides the rich information necessary to find the correct answer span in difficult cases like this one.

of text they can process at once. To leverage these state-of-the-art model architectures, researchers have developed strategies to segment documents and feed them piecemeal into models. However, these strategies typically discard the document-level context that is essential for complex natural language tasks on long documents. We develop a method that makes use of state-of-the-art models while also maintaining information about the document's structure and formatting. Key to this method is the use of features from Optical Character Recognition (OCR) above and beyond the document's raw text (Figure 1).

We evaluate our method on a publicly-available dataset of commercial legal contracts (Section 3). Our method (Section 4) outperforms prior work on the dataset for four key contract understanding tasks (Section 5). We also quantify the effect of the OCR features and measure the impact of document length on the difficulty of the task (Section 6).

2 RELATED WORK

2.1 Long Document Understanding

When working with long documents, many models are limited in the length of text they can consider at once. To overcome this limitation, the simplest method is to keep only the portion of the document that fits and discard the rest, but many legal documents contain important information that can be located anywhere in the document. Another method is to split the text into equal-sized sections and input them into the model separately, for example using a sliding window over the documents so that each segment fits within the 512-token maximum length of BERT-based models [6]. However, splitting documents into sections without any contextual information makes downstream tasks much more difficult (see Subsection 6.1).

Most work on long documents circumvents the challenge of splitting documents into meaningful sections by using datasets with clearly demarcated sections, such as Wikipedia articles [1, 2], but determining section boundaries is difficult on a corpus of documents with inconsistent structure like legal documents. Research has consistently shown better results on document understanding tasks after splitting documents into sections [2, 5, 12, 13]. As we will show, representing document structure is essential to the section splitting process, and therefore to achieving strong results on downstream tasks.

2.2 OCR

Converting images into text via OCR is a precursor to many NLP tasks, but modern OCR tools provide much more than just textual information. OCR methods produce rich metadata including page partitioning and placement, font style and size, justification, and level of indentation, which become features that allow NLP models to incorporate the document-level context that is essential for understanding long documents.

However, OCR features can be difficult to integrate into NLP pipelines. Existing strategies include encoding them as special characters in models like BERT [3], creating encodings to represent the spatial location of each character [8], or using representations particularly suited to structured data such as graphs [7].

3 DATASET

Legal documents are especially challenging for NLP tasks because of their length. We use the Contract Understanding Atticus Dataset (CUAD), which contains 510 English-language commercial legal contracts from the public domain [6]. Each contract is labeled by trained annotators for 41 attributes. Some of these attributes are phrases that must be extracted from the contract as written (entity extraction), and other attributes are yes-or-no values that can be answered from the contract text (classification). Documents in CUAD are longer than those in most NLP datasets, which are often composed of sentences [10, 14] or short social media posts like tweets limited to 140 characters [11, 15]. Even document-length corpora consist of documents that are much shorter than legal documents: the average length of a Wikipedia article, for example, is 620 words [9], compared to CUAD's average length of 9,594 words. Detailed statistics on CUAD are included in Appendix B.

Since CUAD does not provide section labels, we report section splitting results on our proprietary corpus of contracts, evaluated against section labels created by expert annotators.

4 EXPERIMENTAL SETUP

Our end-to-end system is outlined in Figure 2 below, beginning with a raw PDF contract and ending with a predicted answer for a document understanding task.
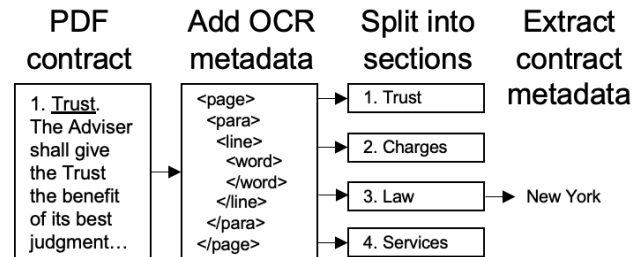

Figure 2: Our end-to-end document understanding system.

In order to provide focused inputs to our information extraction models, we split documents into coherent sections of several types: clauses, sub-clauses, headers, and footers (Figure 3). While clauses and sub-clauses contain the information that lawyers are interested in parsing from contracts, it is necessary to detect other types of sections like headers and footers as well because they contain noise that interrupts the flow of text, making it challenging for models to extract the correct values.
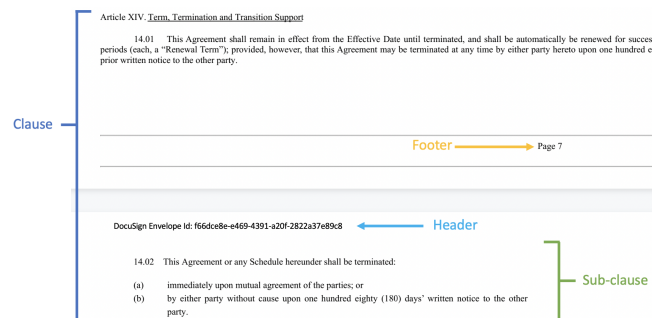

Figure 3: An example document showing each section type.

We first trained a baseline model for section splitting which takes textual and linguistic features as inputs, with no visual or structural information. This model struggled to accurately identify sections that a human could easily distinguish using visual cues not available to our baseline model.

With the hypothesis that the rich visual data available through OCR would improve performance on section splitting, we tested ablations of our baseline model to measure the impact of four types of visual cues:

- **Page layout features**: the proximity of text to various regions of the page (e.g. the beginning, end, center, or margin of a page).

- **Text placement features**: the placement of text on the page (e.g. is it centered or aligned).
- **Visual grouping features**: which words in the text are grouped together into a paragraph, list, or table.
- **Stylistic features**: whether the words are stylized (e.g. bold, italic, or underlined).

We find that these visual features improve performance for all section types, with the largest impact on footer detection. We detail these results in Section 5.

5 RESULTS

5.1 Metrics

We evaluate our models using precision, recall, and F1 score because many contract attributes suffer from class imbalance. For example, only 15 of the 510 contracts in CUAD (3%) forbid price changes (the "Price Restriction" attribute). If a lawyer wanted to find these 15 contracts, a model that returned none of them would achieve 97% accuracy, while the recall of such a model would be a more meaningful 0%.

5.2 Section Splitting

To evaluate the impact of visual cues on the accuracy of our section splitting method, we run multiple experiments in which we activate different groups of visual features, and then measure the accuracy on a held out dataset. We compare each group's performance to a baseline model that relies only on text-based linguistic features and low-level layout information, with no access to visual cues from OCR features.

We observe that visual cues have a strong impact, which varies depending on the section type (Table 1). At a high level, it is clear that footer detection gains the most from OCR features, but clauses and sub-clauses see significant improvements as well. The results for headers are mixed: recall improves, but at a cost to precision. We look more closely at the reasons for these results in Subsection 6.2.

5.3 Document Understanding Tasks

Using the outputs of the section splitting process, we select relevant sections and input them into downstream models trained to extract key metadata from contracts. The four tasks we evaluate reflect four common questions lawyers need to know about their contracts:

- Which contracts are still active? ("Expiration Date", framed as an entity extraction task to find the correct date in the contract text)
- Which state or country's law governs this contract? ("Governing Law", framed as an entity extraction task to find the correct location in the contract text)
- Can I terminate this contract? ("Termination for Convenience", framed as a classification task to answer yes or no correctly based on the contract text)
- Which contracts will survive a merger or acquisition? ("Anti-Assignment", framed as a classification task to answer yes or no correctly based on the contract text)

We evaluate against two competing approaches:

Table 1: Impact of OCR Features on Section Splitting

CLAUSES	Precision	Recall	F1
Baseline	.904	.897	.900
+ Page Layout	.902 (-0.2%)	.897	.899 (-0.1%)
+ Text Placement	.908 (+0.5%)	.897	.902 (+0.3%)
+ Visual Grouping	.912 (+0.9%)	.899 (+0.3%)	.905 (+0.6%)
+ Style	.917 (+1.5%)	**.902 (+0.6%)**	.909 (+1.0%)
+ All Groups	**.919 (+1.7%)**	.901 (+0.5%)	**.910 (+1.1%)**

SUB-CLAUSES	Precision	Recall	F1
Baseline	.901	.913	.907
+ Page Layout	.900 (-0.2%)	.913	.906 (-0.1%)
+ Text Placement	.901	.913	.907
+ Visual Grouping	.904 (+0.3%)	**.914 (+0.1%)**	.909 (+0.2%)
+ Style	.908 (+0.7%)	.913	.910 (+0.4%)
+ All Groups	**.910 (+0.9%)**	.913	**.911 (+0.4%)**

HEADERS	Precision	Recall	F1
Baseline	.900	.956	.927
+ Page Layout	.840 (-6.7%)	**.961 (+0.5%)**	.896 (-3.3%)
+ Text Placement	.845 (-6.1%)	.955 (-0.1%)	.897 (-3.3%)
+ Visual Grouping	**.910 (+1.1%)**	.958 (+0.2%)	**.933 (+0.7%)**
+ Style	.890 (-1.1%)	.956	.922 (-0.6%)
+ All Groups	.858 (-4.7%)	.960 (+0.4%)	.906 (-2.3%)

FOOTERS	Precision	Recall	F1
Baseline	.845	.760	.800
+ Page Layout	.877 (+3.8%)	**.862 (+13.4%)**	.869 (+8.6%)
+ Text Placement	.849 (+0.5%)	.792 (+4.2%)	.820 (+2.4%)
+ Visual Grouping	.855 (+1.2%)	.834 (+9.7%)	.844 (+5.5%)
+ Style	.843 (-0.2%)	.757 (-0.4%)	.798 (-0.3%)
+ All Groups	**.887 (+4.9%)**	.857 (+12.8%)	**.872 (+8.9%)**

- **Expert rules**: Regex-like rules written by trained annotators referencing a corpus of legal documents that does not include the CUAD contracts. For example, an annotator might determine that if the phrase "may terminate at will" appears in a contract, that contract's value for the "Termination for Convenience" attribute is "yes". Because of the differences between the contracts in the annotators' reference corpus and the contracts in CUAD, these rules have high precision but low recall, since they cannot generalize to unseen language in a new corpus.
- **DeBERTa**: The previously best-performing model on CUAD [6], a DeBERTa-xlarge model fine-tuned for the question answering task. We compare to the reported precision at 80% recall for four attributes.

Results for the full end-to-end tasks are detailed in Table 2. Our method, which uses rich visual cues to split contracts into sections for use on downstream tasks, offers the best balance of precision and recall across the four tasks. Additionally, examples of correct and incorrect model predictions on each task are included in Appendix A.

Table 2: Results on End-to-End Long Document Understanding Tasks

| | **ENTITY EXTRACTION TASKS** | | | | | |
| | Expiration Date | | | Governing Law | | |
Model	P	R	F1	P	R	F1
Expert rules	.77	.64	.70	.75	.60	.67
DeBERTa	.86	.80	.83	.97	.80	.88
Our model	**.87**	**.87**	**.87**	**.98**	**.98**	**.98**
	CLASSIFICATION TASKS					
	Term. for Conv.			Anti-Assignment		
Model	P	R	F1	P	R	F1
Expert rules	**.80**	.44	.57	.83	.57	.68
DeBERTa	.37	**.80**	.51	.76	.80	.78
Our model	.77	.75	**.76**	**.89**	**.88**	**.89**

6 ANALYSIS

6.1 Impact of Document Length

To quantify the impact of document length on the difficulty of a document understanding task, we train a CRF model to extract Governing Law from contract segments of increasing length. For each length, we draw a fixed-size window around the correct answer span in every contract. We use a random offset before and after the answer span so that answers are not always located in the same place. As shown in Figure 4, as document length increases, performance sharply falls. Entity extraction on long documents is a difficult task without the benefit of visually-aware, structural features that help the model locate the answer within a document.
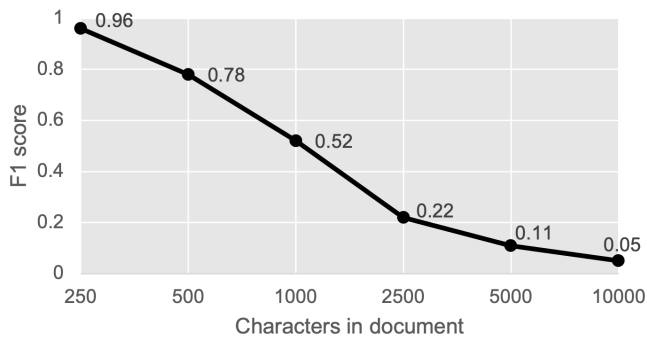

Figure 4: Longer documents are much more difficult for a CRF model extracting Governing Law from contracts of increasing size.

6.2 Impact of OCR Features

Our results show strong performance on end-to-end document understanding tasks, but to what extent can we attribute that performance to the model's understanding of document structure via OCR features, as opposed to the task setup or model architecture? To quantify the impact of visual cues on one of our downstream

tasks, we compare an Anti-Assignment classifier trained using sections that were obtained with and without the help of visual cues. As shown in Table 3, the downstream model performs even better when it has access to document structure via OCR features.

Table 3: Impact of Visual Section Splitting on Anti-Assignment Classification

Model	P	R	F1
Our model	.89	.69	.71
Our model + visual cues	**.93**	**.81**	**.85**

Finally, we provide a more detailed analysis of the performance for each section type from Table 1:

6.2.1 Clauses. Clauses benefit the most from style features. This is intuitive, since clauses very frequently start with a phrase that acts as a heading, set apart in bold or underlined. The visual grouping features also had a positive impact, since clauses often appear together in paragraphs of text.

6.2.2 Sub-clauses. Sub-clauses benefit from the same features as clauses, but it is noteworthy that the impact is lower, particularly for recall. Sub-clause detection is generally a harder problem, even for human annotators, due to deeply nested sub-clauses and a wider variety of possible ways to format them.

6.2.3 Headers. Headers do not see a large benefit from OCR features. In fact, page layout features hurt precision on headers, since they can be misleading when clauses and sub-clauses appear at the beginning of a page like a header. Similarly, text placement features are also less effective because the model tends to overfit to the placement of logos and auto-generated stamps like "DocuSign Envelope Id", which always show up in the same location. To address this weakness, we plan to add more diverse data and regularization in addition to richer features.

6.2.4 Footers. Footers benefit the most from OCR features. This is not surprising, since footers consistently appear at the end of the page, and are much more common than headers. Accurate footer detection is particularly useful since footers often interrupt the flow of clauses, which hurts downstream models that rely on clean section text.

7 CONCLUSION

We discuss a method for extracting valuable metadata from legal documents, which are especially challenging due to their length and their diverse structure and formatting. We show that visual cues are a key component in a successful document understanding pipeline, from section splitting to entity extraction and classification.

This work explores only a subset of the types of visual cues available. Future work might explore additional features such as font information, spacing between lines and blocks of text, superscripting, subscripting, and color. Given the impact of the visual cues we examined, we expect these features to further improve performance on document understanding tasks.

REFERENCES

[1] Ilias Chalkidis, Emmanouil Fergadiotis, Prodromos Malakasiotis, and Ion Androutsopoulos. 2019. Large-Scale Multi-Label Text Classification on EU Legislation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 6314–6322. https://doi.org/10.18653/v1/P19-1636

[2] Eunsol Choi, Daniel Hewlett, Jakob Uszkoreit, Illia Polosukhin, Alexandre Lacoste, and Jonathan Berant. 2017. Coarse-to-Fine Question Answering for Long Documents. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Vancouver, Canada, 209–220. https://doi.org/10.18653/v1/P17-1020

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[4] Emad Elwany, Dave Moore, and Gaurav Oberoi. 2019. BERT Goes to Law School: Quantifying the Competitive Advantage of Access to Large Legal Corpora in Contract Understanding. In *NeurIPS 2019 Workshop on Document Intelligence*. http://arxiv.org/abs/1911.00473

[5] Hongyu Gong, Yelong Shen, Dian Yu, Jianshu Chen, and Dong Yu. 2020. Recurrent Chunking Mechanisms for Long-Text Machine Reading Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 6751–6761. https://doi.org/10.18653/v1/2020.acl-main.603

[6] Dan Hendrycks, Collin Burns, Anya Chen, and Spencer Ball. 2021. CUAD: An Expert-Annotated NLP Dataset for Legal Contract Review. *arXiv preprint arXiv:2103.06268* (2021).

[7] Hailey James, Otkrist Gupta, and Dan Raviv. 2020. OCR Graph Features for Manipulation Detection in Documents. *CoRR* abs/2009.05158 (2020). arXiv:2009.05158 https://arxiv.org/abs/2009.05158

[8] Anoop R Katti, Christian Reisswig, Cordula Guder, Sebastian Brarda, Steffen Bickel, Johannes Höhne, and Jean Baptiste Faddoul. 2018. Chargrid: Towards Understanding 2D Documents. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Brussels, Belgium, 4459–4469. https://doi.org/10.18653/v1/D18-1476

[9] Size of Wikipedia. 2021. Size of Wikipedia — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia [Online; accessed 11-May-2021].

[10] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. 2013. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Seattle, Washington, USA, 1631–1642. https://www.aclweb.org/anthology/D13-1170

[11] Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan. 2015. A Neural Network Approach to Context-Sensitive Generation of Conversational Responses. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Denver, Colorado, 196–205. https://doi.org/10.3115/v1/N15-1020

[12] Sajad Sotudeh Gharebagh, Arman Cohan, and Nazli Goharian. 2020. GUIR @ LongSumm 2020: Learning to Generate Long Summaries from Scientific Documents. In *Proceedings of the First Workshop on Scholarly Document Processing*. Association for Computational Linguistics, Online, 356–361. https://doi.org/10.18653/v1/2020.sdp-1.41

[13] Lulu Wan, George Papageorgiou, Michael Seddon, and Mirko Bernardoni. 2019. Long-length Legal Document Classification. *CoRR* abs/1912.06905 (2019). arXiv:1912.06905 http://arxiv.org/abs/1912.06905

[14] Alex Warstadt, Amanpreet Singh, and Samuel R Bowman. 2018. Neural Network Acceptability Judgments. *arXiv preprint arXiv:1805.12471* (2018).

[15] Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, and Ritesh Kumar. 2019. Predicting the Type and Target of Offensive Posts in Social Media. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 1415–1420. https://doi.org/10.18653/v1/N19-1144

A MODEL PREDICTIONS

We list examples of correct and incorrect model outputs for the four end-to-end tasks: Expiration Date (Table 4), Governing Law (Table 5), Termination for Convenience (Table 6), and Anti-Assignment (Table 7). "Relevant Section Text" refers to the output of our section splitting model, which identifies clauses and sub-clauses and then labels which sections are relevant to a given downstream task.

B DATASET

B.1 Details on CUAD

Table 8 lists detailed statistics on the contracts in CUAD.

CUAD provides both PDF and text versions of contracts. For our experiments, we use the PDF versions so that we can extract OCR metadata from the formatting and structure of the contracts.

B.2 Attributes Chosen from CUAD

While CUAD includes labels for the two most useful contract metadata attributes, Document Name and Parties, we choose not to evaluate using these labels because both of these attributes occur multiple times throughout each document, and the label offsets in CUAD do not consistently come from the same location in the document. For example, some Document Name labels are selected from an Appendix rather than the beginning of the document, making evaluation on these labels misleading.

B.3 Test Set

To facilitate future work on CUAD, we list the filenames of the 51 documents in our test set in Table 9.

Table 4: Model Predictions for "Expiration Date"

Relevant Section Text	Correct Answer	Model Prediction
2.1 **This Agreement shall be effective from the date first above written and shall continue indefinitely until terminated by either Party in accordance with the provisions of this Agreement.**	bold text	bold text ✓
The term of this Agreement shall be five (5) years.	full text	full text ✓
(b)The channel **is expected to be uploaded on August 1, 2010.**	bold text	full text ✗
(i) the date on which the aggregate Required Capital Contributions paid by the Investor, and received by the Borrowers, in accordance with the terms of this Agreement equal $5,000,000; (ii) the Release Date (as defined in Section 19) or (iii) payment in full, in cash, of all Obligations and the termination of the Financing Agreement; provided, however, that this Agreement shall continue to be effective, or be reinstated, as the case may be, if at any time the aggregate Required Capital Contributions paid by the Investor, and received by the Borrowers, in accordance with the terms of this Agreement is less than $5,000,000 and any payment, or any part thereof, on account of any of the Obligations is invalidated, declared to be fraudulent or preferential, set aside, rescinded or must otherwise be restored or returned by the Agent or the Lenders upon the insolvency, bankruptcy, liquidation, dissolution or reorganization of any Borrower or upon or as a result of the appointment of a receiver, intervenor or conservator of, or trustee or similar officer for any Borrower, or any substantial part of this property, or otherwise, all as though such payment had not been made.	full text	bold text ✗
17. TERM AND TERMINATION (a) **This Agreement and the license granted under this Agreement shall remain in effect perpetually as long as fees are paid by Sparkling in accordance with the Fee Schedule and the Agreement is not otherwise terminated in accordance with this Section.**	bold text	bold text ✓

Table 5: Model Predictions for "Governing Law"

Relevant Section Text	Correct Answer	Model Prediction
14. Governing Law. **This Agreement shall be construed and interpreted in accordance with the laws of the State of Ohio, without recourse to any principles of law governing conflicts of law, which might otherwise be applicable.**	bold text	bold text ✓
This Agreement and performance under this Agreement shall be governed by the laws of the United States of America and of the Commonwealth of Pennsylvania as applied to agreements entered into and to be performed entirely within Pennsylvania between Pennsylvania residents, excluding its conflicts of law provisions.	full text	full text ✓
12.2 **This Agreement and all matters pertaining hereto shall be governed by and construed under the laws of the State of Louisiana, except to the extent** that the conflict of law rules of said state would require that the laws of another state would govern its validity, construction, or interpretation.	full text	bold text ✗
This Agreement shall be governed by the laws of the Province of Ontario and the federal laws of Canada applicable therein.	full text	full text ✓
Without reference to choice or conflict of law principles, this Agreement shall be governed by and construed in accordance with the laws of the State of California, USA.	full text	full text ✓
(a) **This Agreement and all Actions (whether in contract or tort) that may be based upon, arise out of or relate to this Agreement or the negotiation, execution, or performance hereof or thereof shall be governed by and construed in accordance with the Law of the State of Delaware, without regard to any Laws or principles thereof that would result in the application of the Laws of any other jurisdiction.**	bold text	bold text ✓

Table 6: Model Predictions for "Termination for Convenience"

Relevant Section Text	Correct Answer	Model Prediction
8.2 Termination. Either Party may terminate this Agreement for the material breach or default of any of the terms or conditions of this Agreement by the other Party upon thirty (30) days' written notice and the opportunity to cure during such notice period; and such termination shall be in addition to any other remedies that it may have at law or in equity. Additionally, LBIO may terminate this Agreement if MD Anderson is declared insolvent or enters into liquidation or has a receiver or an administrator appointed over all or any part of its assets or ceases or threatens to cease to carry on business, or a resolution is passed or a petition presented to any court for the winding up of the Party or for the granting of an administration order in respect of MD Anderson, or any proceedings are commenced relating to the insolvency or possible insolvency of MD Anderson.	No	Yes ✕
7.2. This Agreement may be terminated by either party with cause upon thirty (30) days written notice. Upon Marketing Affiliate's default in payment or other breach of this Agreement, Equidata may terminate this Agreement without notice to Marketing Affiliate. Upon termination for any reason, Equidata reserves the right to deactivate Marketing Affiliate's access to the services including the Equidata Web Site. Termination does not release Marketing Affiliate from paying all amounts owed to Equidata.	No	Yes ✕
8 1 TERM OF AGREEMENT. This Agreement shall continue in force for a term of twelve (12) months from the Effective Date, unless terminated earlier under the provisions of this Article 8 (the "Term"); PROVIDED that TouchStar shall have the right to terminate this Agreement at any time after the Effective Date upon not less than fifteen (15) days' prior written notice to Reseller. Prior to the end of the Term, each of TouchStar and Reseller may notify the other if it desires to negotiate a further agreement by written request received at least ninety (90) days in advance of the termination of this Agreement. If both parties desire to negotiate a further agreement, they may consider the terms of this Agreement in coming to an understanding. Nothing in this Agreement shall be construed to obligate either party to renew or extend the term of this Agreement. Renewals for additional terms, if any, shall not cause this Agreement to be construed as an agreement of indefinite duration.	Yes	No ✕
18.1 The Company may terminate the Executive's employment under this Agreement with immediate effect without notice and with no liability to make any further payment to the Executive (other than in respect of amounts accrued at the Termination Date) if in the reasonable opinion of the Company the Executive:	Yes	No ✕
8.2. Termination for Cause. Either party may terminate this Agreement immediately upon written notice to the other party in the event any material breach of a material term of this Agreement by such other party that remains uncured 30 days in the case of a breach of a payment obligation, or 45 days for all other breaches, after notice of such breach was received by such other party; provided, however that if such breach is not reasonably capable of cure within the applicable cure period, the breaching party shall have an additional 180 days to cure such breach so long as the cure is commenced within the applicable cure period and thereafter is diligently prosecuted to completion as soon as possible.	No	Yes ✕
4.2 The term of this Agreement is for a period of five (5) years (the "Term") commencing on the Effective Date and, unless terminated earlier in accordance with the termination provisions of this Agreement, ending on January 31, 2025.	Yes	No ✕
Section 2 — Term. This Agreement shall commence as of the Effective Date and shall continue in full force and effect for an initial term of three (3) years from the Promotion Commencement Date, divided into three one-year periods. Unless terminated in accordance with the provisions of Section 18, this Agreement shall automatically renew for each subsequent one-year term.	No	No ✓
TERM AND TERMINATION. A. By either party as a result of default by the other party under this Agreement and failure to cure said default within thirty (30) days after notice of said default is given.	No	Yes ✕
Terms and Termination: The term of this agreement will begin on April 1, 2018 and continue until April 30, 2018 at 11:59pm.	No	No ✓

Table 7: Model Predictions for "Anti-Assignment"

Relevant Section Text	Correct Answer	Model Prediction
VII. ASSIGNMENT. Neither this Agreement nor any rights or obligations or licenses hereunder may be assigned, pledged, transferred or encumbered by either party without the express prior written approval of the other party, except that either HEMISPHERX or SCIEN may assign this Agreement to any successor by merger or sale of substantially all of its business or assets to which this Agreement pertains, without any such consent. Any assignment in violation hereof is void.	Yes	Yes ✓
ASSIGNMENT: NOW, THEREFORE, for good and valuable consideration, the receipt and adequacy of which are hereby acknowledged, Seller does hereby transfer, sell, assign, convey and deliver to Backhaul all right, title and interest in, to and under the Assigned Intellectual Property, including, without limitation, the Trademarks and Patents set forth on Schedules A and B hereof, respectively, and all goodwill of the Purchased Business associated therewith. Seller hereby covenants and agrees, that from time to time forthwith upon the reasonable written request of Backhaul or Buyer, that Seller will, at Backhaul's cost and expense, do, execute, acknowledge and deliver or cause to be done, executed, acknowledged and delivered, each and all of such further acts, deeds, assignments, transfers, conveyances and assurances as may reasonably be required by Backhaul or Buyer in order to transfer, assign, convey and deliver unto and vest in Backhaul title to all right, title and interest of Seller in, to and under the Assigned Intellectual Property.	No	Yes ✗
(no relevant section found)	Yes	No ✗
4.4 Assignments and Transfers by Seller Trusts. The provisions of this OMA shall be binding upon and inure to the benefit of the Seller Trusts and their respective successors and assigns. A Seller Trust may transfer or assign, in whole or from time to time in part, to one or more liquidating trusts its rights hereunder in connection with the transfer or resale of Stock held by such Seller Trust, provided that such Seller Trust complies with all laws applicable thereto and provides written notice of assignment to GWG promptly after such assignment is effected, and provided further that such liquidating trust and each beneficiary thereof executes a joinder to this OMA effective as of the date of such assignment or transfer.	No	Yes ✗
133 Assignment. Neither Party shall assign this Development Agreement or the obligations contained herein without the express written consent of the other Party.	Yes	Yes ✓
15.7 Successors and Assigns. This Agreement shall inure to the benefit of and be binding upon the Parties and their respective successors and assigns, including, but not limited to, any chapter 11 or chapter 7 trustee; provided, however, that this Agreement may not be assigned by any of the Parties without the prior written consent of the other, provided further that notwithstanding the foregoing, GA and Tiger may each collaterally assign this Agreement and their rights thereunder to their respective lenders.	Yes	Yes ✓
9.2 Assignment 51	Yes	Yes ✓

Table 8: Statistics on the Contract Understanding Atticus Dataset (CUAD)

Total documents	510
Documents in train set (80%)	408
Documents in dev/test sets (10%/10%)	51
Average characters per contract	52,563
Characters in shortest contract	645
Characters in longest contract	338,211
Average words per contract	9,594
Words in shortest contract	109
Words in longest contract	103,923

Table 9: Filenames of documents from CUAD in our test set

Monsanto Company - SECOND A&R EXCLUSIVE AGENCY AND MARKETING AGREEMENT .PDF
IdeanomicsInc_20151124_8-K_EX-10.2_9354744_EX-10.2_Content License Agreement.pdf
REGANHOLDINGCORP_03_31_2008-EX-10-LICENSE AND HOSTING AGREEMENT.PDF
GridironBionutrientsInc_20171206_8-K_EX-10.1_10972555_EX-10.1_Endorsement Agreement.pdf
BLACKBOXSTOCKSINC_08_05_2014-EX-10.1-DISTRIBUTOR AGREEMENT.PDF
OLDAPIWIND-DOWNLTD_01_08_2016-EX-1.3-AGENCY AGREEMENT2.pdf
ClickstreamCorp_20200330_1-A_EX1A-6_MAT CTRCT_12089935_EX1A-6_MAT CTRCT_Development Agreement.pdf
NYLIACVARIABLEANNUITYSEPARATEACCOUNTIII_04_10_2020-EX-99.8.KK-SERVICE AGREEMENT.PDF
Columbia Laboratories (Bermuda)Ltd. - AMEND NO. 2 TO MANUFACTURING AND SUPPLY AGREEMENT.PDF
MERITLIFEINSURANCECO_06_19_2020-EX-10.(XIV)-MASTER SERVICES AGREEMENT.PDF
NATIONALPROCESSINGINC_07_18_1996-EX-10.4-SPONSORSHIP AGREEMENT.PDF
MJBIOTECHINC_12_06_2018-EX-99.01-JOINT VENTURE AGREEMENT.PDF
KUBIENTINC_07_02_2020-EX-10.14-MASTER SERVICES AGREEMENT_Part2.pdf
ArcGroupInc_20171211_8-K_EX-10.1_10976103_EX-10.1_Sponsorship Agreement.pdf
WORLDWIDESTRATEGIESINC_11_02_2005-EX-10-RESELLER AGREEMENT.PDF
SPOKHOLDINGSINC_06_19_2020-EX-10.1-COOPERATION AGREEMENT.PDF
LegacyEducationAllianceInc_20200330_10-K_EX-10.18_12090678_EX-10.18_Development Agreement.pdf
HEALTHGATEDATACORP_11_24_1999-EX-10.1-HOSTING AND MANAGEMENT AGREEMENT - Escrow Agreement.pdf
LECLANCHE S.A. - JOINT DEVELOPMENT AND MARKETING AGREEMENT.PDF
THERAVANCEBIOPHARMA,INC_05_08_2020-EX-10.2-SERVICE AGREEMENT.PDF
HUBEIMINKANGPHARMACEUTICALLTD_09_19_2006-EX-10.1-OUTSOURCING AGREEMENT.PDF
GSITECHNOLOGYINC_11_16_2009-EX-10.2-INTELLECTUAL PROPERTY AGREEMENT between SONY ELECTRONICS INC.
    and GSI TECHNOLOGY, INC..PDF
SECURIANFUNDSTRUST_05_01_2012-EX-99.28.H.9-NET INVESTMENT INCOME MAINTENANCE AGREEMENT.PDF
HarpoonTherapeuticsInc_20200312_10-K_EX-10.18_12051356_EX-10.18_Development Agreement.PDF
InnerscopeHearingTechnologiesInc_20181109_8-K_EX-10.6_11419704_EX-10.6_Distributor Agreement.pdf
SoupmanInc_20150814_8-K_EX-10.1_9230148_EX-10.1_Franchise Agreement1.pdf
EdietsComInc_20001030_10QSB_EX-10.4_2606646_EX-10.4_Co-Branding Agreement.pdf
FerroglobePlc_20150624_F-4A_EX-10.20_9154746_EX-10.20_Outsourcing Agreement.pdf
BIOCEPTINC_08_19_2013-EX-10-COLLABORATION AGREEMENT.PDF
HertzGroupRealtyTrustInc_20190920_S-11A_EX-10.8_11816941_EX-10.8_Trademark License Agreement.pdf
MEDALISTDIVERSIFIEDREIT,INC_05_18_2020-EX-10.1-CONSULTING AGREEMENT.PDF
BIOFRONTERAAG_04_29_2019-EX-4.17-SUPPLYAGREEMENT.PDF
MSCIINC_02_28_2008-EX-10.10-.PDF
FEDERATEDGOVERNMENTINCOMESECURITIESINC_04_28_2020-EX-99.SERV AGREE-SERVICES AGREEMENT_POWEROF.pdf
CUROGROUPHOLDINGSCORP_05_04_2020-EX-10.3-SERVICING AGREEMENT.PDF
NOVOINTEGRATEDSCIENCES,INC_12_23_2019-EX-10.1-JOINT VENTURE AGREEMENT.PDF
TALCOTTRESOLUTIONLIFEINSURANCECO-SEPARATEACCOUNTTWELVE_04_30_2020-EX-99.8(L)-SERVICE AGREEMENT.PDF
CORALGOLDRESOURCES,LTD_05_28_2020-EX-4.1-CONSULTING AGREEMENT.PDF
ChinaRealEstateInformationCorp_20090929_F-1_EX-10.32_4771615_EX-10.32_Content License Agreement.pdf
GLOBALTECHNOLOGIESLTD_06_08_2020-EX-10.16-CONSULTING AGREEMENT.PDF
SalesforcecomInc_20171122_10-Q_EX-10.1_10961535_EX-10.1_Reseller Agreement.pdf
MANUFACTURERSSERVICESLTD_06_05_2000-EX-10.14-OUTSOURCING AGREEMENT.PDF
ONEMAINHOLDINGS,INC_02_20_2020-EX-99.D-JOINT FILING AGREEMENT.PDF
BerkshireHillsBancorpInc_20120809_10-Q_EX-10.16_7708169_EX-10.16_Endorsement Agreement.pdf
VAXCYTE,INC_05_22_2020-EX-10.19-SUPPLY AGREEMENT.PDF
SEASPINEHOLDINGSCORP_10_10_2018-EX-10.1-SUPPLY AGREEMENT.PDF
WaterNowInc_20191120_10-Q_EX-10.12_11900227_EX-10.12_Distributor Agreement.pdf
ElPolloLocoHoldingsInc_20200306_10-K_EX-10.16_12041700_EX-10.16_Development Agreement.pdf
TubeMediaCorp_20060310_8-K_EX-10.1_513921_EX-10.1_Affiliate Agreement.pdf
IOVANCEBIOTHERAPEUTICS,INC_08_03_2017-EX-10.1-STRATEGIC ALLIANCE AGREEMENT.PDF
VertexEnergyInc_20200113_8-K_EX-10.1_11943624_EX-10.1_Marketing Agreement.pdf