Gracenote.ai: Legal Generative AI for Regulatory Compliance

Jules Ioannidis, Joshua Harper, Ming Sheng Quah [1] and Dan Hunter [1, 2]

[1] Gracenote.ai, Melbourne Australia
[2] The Dickson Poon School of Law, King's College London, United Kingdom*

Abstract: We investigate the transformative potential of large language models (LLMs) in the legal and regulatory compliance domain by developing advanced generative AI solutions, including a horizon scanning tool, an obligations generation tool, and an LLM-based expert system. Our approach combines the LangChain framework, OpenAI's GPT-4, text embeddings, and prompt engineering techniques to effectively reduce hallucinations and generate reliable and accurate domain-specific outputs. A human-in-the-loop control mechanism is used as a final backstop to ensure accuracy and mitigate risk. Our findings emphasise the role of LLMs as foundation engines in specialist tools and lay the groundwork for building the next generation of legal and compliance applications. Future research will focus on extending support across multiple jurisdictions and languages, refining prompts and text embedding datasets for enhanced legal reasoning capabilities, and developing autonomous AI agents and robust LLM-based expert systems.

Keywords: AutoGPT, compliance, expert systems, GPT-4, GRC, LangChain, large language models, legal generative AI, legal text embeddings, prompt engineering, regulation

1. Introduction

Law relies on language. So, it's hardly surprising that the development of large language models (LLMs) and the explosion of interest in publicly accessible LLMs such as ChatGPT has led to the proliferation of papers about the use of these kinds of models in law. These papers include some general explorations of natural language parsing and LLMs in law [1], [2], [3], as well as a number looking at whether GPT can pass the bar exam, law school exams, or other standardised legal tests [4], [5], [6], [7], and [8]. Inevitably, there have been papers foretelling the death of lawyers by LLMs [9], together with those which argue that the future of all professional work will be revolutionised by LLMs [10].

LLMs are often called "foundation models" because they can be used as a foundation to drive a range of products and services. This can be done in several ways, but the two standard approaches involve prompt engineering and fine-tuning. Fine-tuning involves adapting a pre-trained LLM to a specific task or domain by training it on a specialised dataset, thus enhancing its relevance and performance [11]. Although the providers of LLMs will often create interfaces for users to fine-tune the model, fine-tuning is relatively difficult because it requires significant data and advanced skills at controlling the dataset to tune the pre-trained model. A more common way of using LLMs as a foundation engine is to use prompt engineering to constrain the output generated by the model.

LLMs rely on user-generated prompts to provide both the context and the request that generates the output (often called the

"completion") [12]. Text embeddings measure the relatedness of text strings and can be employed to enhance LLMs' contextual understanding for more accurate outputs. This technique allows LLMs to effectively perform tasks like custom search applications, summarisation, and classification, where the context may be largely contained in the prompt. Thus, clever prompt engineering can generate valuable outputs without the need for any fine-tuning or (worse) retraining of the model. The first reported job advertisement for a legal prompt engineer was posted by the UK law firm Mishcon de Reya, [13] but the skill is one that numerous law firms and companies now need [14].

The combination of simplicity in prompt engineering and ease of access to foundation models has led to a proliferation of generative AI companies and products in law. The highest profile example of this is the generative AI company Harvey.ai, which seeks to redefine elite legal and consultancy practice by using OpenAI's LLMs as the foundation engine [15]. But other legal examples include companies focused on generative AI legal drafting solutions [16] or contract lifecycle maintenance [17].

Our company, Gracenote, adopts the same methodology of building legal products on top of LLMs as the core engine, using a combination of prompt engineering, text embeddings, model fine-tuning, and NLP-based techniques. We use a range of foundation models to create legal tools that solve commercial regulatory and compliance issues. Governance, risk and compliance (generically referred to as "GRC") is the business function where laws and regulations intersect most directly with business and commerce. It is a useful arena to test the ability of generative AI to undertake legal tasks, for a range of reasons. Notably, the scale of the problems in GRC are vast, the people using these platforms usually are not lawyers, they typically need lots of guidance, and the implications of getting the answers wrong are very serious. GRC tools built using appropriately configured/constrained LLMs and using various generative AI methods can address many of these issues.

Gracenote works with law firms and consultancies to create generative AI environments to solve the GRC problems of their clients using our tools. We have developed a platform that uses OpenAI's GPT-4 model as the engine for several tools that generate legal and regulatory content in multiple forms. The platform can accurately and reliably horizon scan various public sources of information to generate regulatory newsfeeds, as well as generate obligations registers from legislation, regulations, and policy. It can also automatically create expert system-like consultation tools directly from legal text.

Originally the team behind Gracenote worked with trained lawyers to create regulatory updates, obligations registers, and rule-based consult tools. But now it uses LLM prompt engineering techniques to automatically generate the content from public sources such as press releases, regulatory alerts, case reports, and legislation. Lawyers in our law firm clients view the generated completions side-by-side with the original content, in order to assess accuracy, validity and relevance prior to publication internally to practice groups or externally to clients. The current platform has been trialled in a range of sectors including financial services, insurance, and cybersecurity.

In this paper, we report on our research into the use of generative AI to solve GRC issues and document the methods and tools we use to solve three different GRC problems using LLMs. These three problems are the creation of regulatory newsfeeds from public sources, the generation of obligations from legislative material, and the creation of consultation tools from legislation.

The research advances reported on in this paper will focus on (1) legal prompt engineering techniques to generate GRC solutions from multi-modal legal source documents (e.g. regulatory press releases, legislation, explanatory memoranda, and legal cases); (2) the difference in accuracy and quality of content generated by different LLMs; and (3) design considerations that reduce the hygiene problems inherent in using generative AI models in legal settings—including the hallucination problem, privacy issues in passing personally-identifying data, and confidentiality issues in passing commercially-sensitive data.

The paper proceeds as follows. In Section 2 we document our methods in using generative AI to perform horizon scanning of regulatory material and generate regulatory newsfeeds. In Section 3 we discuss methods of creating obligations

registers from legislative material, and in Section 4 we sketch a solution to the automatic creation and maintenance of legal expert systems using a combination of generative AI techniques. In Section 5 we discuss methods for reducing the well-documented generative AI issues of hallucination and privacy/confidentiality leakage. In Section 6 we provide conclusions and note the further work that we have begun undertaking to improve our tools and methods.

2. Generative AI for regulatory newsfeeds

Law firms and the risk/compliance officers of corporations need to be apprised of upcoming changes in various laws. The traditional way of performing this function is to use a horizon scanning service like Lexology [18] that uses human editors to find and summarise upcoming regulatory change.

Gracenote uses automatic horizon scanning and scraping methods to find and access regulatory information, and then uses generative AI methods to summarise and categorise the information for clients.

2.1. Horizon scanning

For regulatory update horizon scanning, the platform automatically monitors feeds that are identified by a law firm practice group and scrapes all new information coming from that feed. It sends a notification to the responsible author in that group alerting them to the new content and summary.

The scraping process can be decomposed into a few parts. First, to streamline the scraping process, the update tool uses an abstract class containing generic methods and variables, e.g. relevant data fields, methods to insert entries into the database, etc. The abstract class is implemented in classes specific to each website, utilising unique regular expression patterns and/or a document object model crawler. This approach enables the efficient location of media release links, titles, and publication dates on their respective news pages.

Using a cronjob, the system periodically checks for new regulatory material on the supported websites, ensuring the database stays up to date with the latest information. The periodicity can be adjusted for each regulatory feed depending on the speed and velocity of content publication. Generally, most sites are monitored hourly, but it is possible to monitor more regularly.

In the pilot phase, we support scraping from Australian regulatory bodies including the Australian Securities and Investments Commission, the Australian Transaction Reports and Analysis Centre, the Australian Prudential Regulation Authority, the Office of the Australian Information Commissioner (OAIC), the Australian Financial Complaints Authority, the Reserve Bank of Australia, and the Australian Taxation Office (ATO). While most pages can be scraped by simply obtaining the initial response, some websites such as the OAIC and the ATO require the use of a headless browser to enable JavaScript execution. This is necessary as the list of media releases is served through JavaScript. Apart from this distinction, the general HTML structure of the supported pages is largely identical. Additionally, user-specified links can be processed, albeit without the post-processing function that sanitises and categorises the page content.

As an extension of this research, our future work will investigate the capabilities of advanced frameworks, such as AutoGPT [19] and LangChain [20], in creating AI agents that autonomously navigate the web and execute searches. The objective of these agents is to extract information from a broad range of industry news sites and media outlets, which is particularly useful in the context of rapidly evolving or emerging regulatory change. By leveraging this approach, law and consulting firms can proactively respond to breaking news without solely relying on updates from regulatory authorities' official channels.

2.2. Prompt engineering for regulatory newsfeeds

The scraped material from the update is inserted into a SQL database, and then a series of GPT prompts are generated using the scraped content as part of the prompt context. These prompts are used to generate three discrete types of completions that are then stored in the database:

1. An arbitrarily long summary of the scraped content (default is 150 tokens).
2. An impact level, labelled as "high", "medium" or "low" impact, which aids in the organisation and presentation of the generated updates.
3. Hashtags that specify the type of content scraped and the main topic(s) that the content covers.

These three methods are discussed in more detail below.

2.2.1. Summarisation methods

We use mixed methods for the first completion, depending on the content type which is scraped. These content types fall into three categories: general media releases, long form text, and adjudications.

General Media Releases

For creating general summaries using GPT-4, we use a version of the following prompt:

```
Summarise the key regulatory updates
from the media release below, keeping
within a 125-word limit. Use Australian
English spelling. To ensure conciseness,
use commonly accepted abbreviations or
acronyms for Australian regulatory bodies
or courts, such as "ASIC" for the
Australian Securities and Investments
Commission or "FCA" for the Federal Court
of Australia. Ensure the summary is
professional, factual, and authoritative,
without any embellishments or rhetoric.
Begin with "On [insert date], ". Relevant
text: [Source material]
```

Older OpenAI models (Davinci-003) would struggle with special characters. However, GPT-3.5 and GPT-4 do not face this issue. This allows us to scrape the raw text data contained within the relevant scraped source material and produce a summary, enabling a more standardised approach to scraping that will be compatible with future AI agents. It further allows us to preserve rich information, such as headings, lists and paragraphs, all of which can assist in the final presentation of the completion.

Long Form Text: Case Law and Consultation Papers

GPT is constrained by token limits. The version of GPT-4 that we use can only process approximately 3,000 words at a time, meaning we need a different solution for long form text. While more advanced versions of GPT-4 can process 25,000 words at a time, we will still need a chunking solution for longer form text such as certain case law and consultation papers. Over time, we anticipate that future versions of GPT with higher context windows may make this chunking solution obsolete for this particular use case, though we expect to find future applications for it.

The chunking solution works as follows:

1. Divide the original text into an array of sections, each under the model's token limit.
2. Create an array of summaries for each section.
3. Continue summarising array entries by concatenating them together under the token limit until the total number of characters in the array is under the token limit.
4. Concatenate the array and return the final summary.

For future work, we propose to incorporate a system message at the beginning explaining what the task is (i.e., "summarise the following future messages") to improve token efficiency. We further propose to integrate LangChain document loaders and text splitters to chunk the original text based on the structure of the source material. That is, instead of using an arbitrary end point, each chunk will be created based on rich information such as sections, headings and page numbers.

Adjudication Summary

Summarising case law is a more complex prompt engineering exercise. At present we only summarise decisions from the Australian Financial Complaints Authority (AFCA). These decisions are relatively well-structured and templated, compared with other higher-level appellate court decisions. Even so, the prompt engineering is quite complex.

We use multiple prompts to create tailored completions of adjudication summaries, as follows:

```
Create a title which first states the
shortened names of the parties involved
in the format of "X v Y", before including
a colon symbol and a very short summary
of the complaint. Use Australian English
spelling. Write in a concise and
authoritative tone. Relevant text:
[insert pages 1 & 2]
```

```
Summarise the background of the
complaint to AFCA in less than 125 words.
Include a description of the parties
involved and their roles. Do not state
the role of AFCA or the outcome of the
complaint. Use Australian English
spelling. Write in a concise and
authoritative tone. Relevant text:
[insert pages 1 & 2]
```

```
Summarise the issues, key findings,
reasons for the determination, and basis
for why the outcome is fair (citing
specific facts) into medium-sized bullet
point sentences. Retain the original
heading structure used in the issues and
key findings section. Use Australian
English spelling. Write in a concise and
authoritative tone. Relevant text:
[insert pages 1 & 2]
```

```
Summarise the outcome of the
determination made by AFCA in less than
75 words. Use Australian English
spelling. Write in a concise and
authoritative tone. Relevant text:
[insert 1.3 Determination]
```

```
Create a 3-bullet point summary
extracting essential information about
the complaint. Specify the parties
involved (ignoring AFCA), the reason for
the complaint, and the outcome of the
determination by AFCA (including why the
determination was fair). Use Australian
English spelling. Write in a concise and
authoritative tone. Relevant text:
[insert outputs of above prompts]
```

This approach allows us to create a summary that can be directly inserted into a regulatory newsfeed by simply scraping a PDF of the AFCA decision.

In each of the aforementioned summary prompts, we employed a higher temperature setting, ranging between 0.5 and 0.7. This elevated temperature allows the LLM to consider words with marginally lower probabilities, thus introducing increased variation, randomness, and creativity. Consequently, this facilitates the generation of distinctive summaries that can serve as foundational material across multiple of law and consulting firms.

2.2.2. Categorisation methods

We use prompt engineering to categorise regulatory newsfeeds, both for impact and for document/topic type classifications, as follows.

Impact Level Assignment
For categorising impact levels, we use the following prompt:

```
Your task is to categorise regulatory
updates into 'high', 'medium', and 'low'
impact levels. Simply respond with
'high', 'medium', or 'low' without
additional words or explanations. Here is
a summary of each impact level category.
High impact may only be used when a new
law or regulation comes into effect.
Medium impact refers to criminal or civil
charges filed against an individual or
organisation, sentencing, regulatory
enforcement actions, class orders,
registering new instruments, consultation
papers, and other similar developments.
Low impact refers to legislative sitting
dates, general industry news, and other
updates. Most regulatory updates are low
and medium impact. Relevant text: [Source
material]
```

This prompt allows us to determine how impactful a regulatory update is, which assists in prioritising content consumption of the regulatory newsfeeds.

Topics and Document Type
To categorise document type, we use the following prompt:

```
Your task is to analyse the source
material provided and categorise it by
assigning the most appropriate document
type hashtag. Choose only one hashtag
from the options below. Do not create new
hashtags. Document type hashtags: #Media
#Consultation #Enforcement #Legislation
#Adjudication. [Source material]
```

For categorising topic, we use the following prompt:

```
Your task is to analyse the source
material provided and categorise it by
assigning the most appropriate topic
hashtag. Choose only one hashtag from the
options below. Do not create new
hashtags. Topic hashtags: #Accountability
```

```
#Audit #AFCA #AML #APRA #ASIC #Climate
#ConsumerProtection    #Credit    #Crypto
#Digital  #Employers  #FinancialAdvice
#FundMergers  #Insurance  #Investments
#Licensing #Parliament #Payments #Privacy
#PrudentialStandards          #Regulators
#Retirement  #SMSF  #Tax  #Transparency.
[Source material]
```

The above prompts allow us to further categorise regulatory updates to support the navigation of our regulatory newsfeeds, allowing clients to select the topics they are interested in and receive tailored alerts. They can also combine topics and document types to extract insights, such as all the enforcement actions which ASIC commenced in the last six months against crypto-related products. We can see this being a valuable research tool for regulatory change management.

2.3.   User interface

There are two different interfaces, one for an author/publisher and one for the end-user/client.

For the authoring tool, upon login the author is presented with all work product that is pending from monitored feeds for that author. The author can adjust which feeds are monitored from a settings page.

The authoring environment itself has three main panes—the leftmost pane ("ASIC places interim stop orders...") is the timeline for all content still to be published, the middle pane is the original content from the monitored source, and the rightmost pane contains the GPT-generated content. The author compares the original source with the summary to assess accuracy and quality, and they can adjust settings on the prompt and edit the summary prior to publication. This is done to reduce the hallucination/integrity issues inherent in LLMs. (A topic which we examine in more detail in section 5 below.). Refer to **Figure 1** in Annexure A.

Upon publishing the content, the update is stored in a structured form and can be used in external law firm workflows—e.g., bulk email systems, newsletters, etc—or pushed to the client using our client interface as shown in **Figure 2** of Annexure A.

Regulatory change management in Australia presents a considerable challenge to corporate legal teams. 72% of these teams report struggling to keep up with regulatory change, raising legal and reputational risk [21]. Our collaborative findings with partner law firms and consultancies estimate that the integration of our regulatory newsfeeds solution can save approximately one FTE in small to medium-sized corporate legal teams. We expect this resource optimisation to free up internal capacity and enable teams to concentrate more on value-added delivery.

3.   Obligations from legislation

An important part of any compliance function is a canonical list of obligations that apply to the business. The register is canonical in the sense that it provides a definitive, authoritative and universally accepted list of obligations which apply to the business. Many corporate legal teams rely on external providers to prepare an initial obligations register. However, as regulatory change is fast-moving, these registers typically become outdated quickly. Alternatively, legal teams may subscribe to an obligations register service, which typically cost between AU$50,000-$100,000 per year. However, even these solutions often fail to keep up with the commencement date of new obligations, as their registers are human created and inefficient.

We use a range of generative AI solutions to create a register of obligations from various types of legislative and regulatory material. This solution is fast, cost-effective and—where appropriately monitored—canonical.

The general approach is standard across all LLM types and allows users to paste a URL to a table of contents page, generating an obligations register. The system systematically opens each link to every provision of the act and extracts text to summarise detected obligations.

During the development process of this tool, we moved from GPT-3.5 to GPT-4, and the nature of the prompts changed in interesting ways.

The GPT-3.5 prompt used was:

```
Obligations generally say a person
'must' do something. Summarise the
obligation in the following text in as
few words as possible. An acceptable
degree of legal accuracy must be
maintained. Do not use list or bullet
point formatting. Where a list of
exceptions applies to the obligation,
```

```
simply state the relevant subsection
rather than outlining every exception.
Begin by saying 'a [insert] must'. Do not
state the nature of the offence. If no
obligation is detected, state 'No
obligation detected'. Use Australian
English spelling. Relevant text: [source
material]
```

A key issue with GPT-3.5 is its tendency to add embellishments or write more text than necessary. Despite instructions to avoid stating the nature of the offence, GPT-3.5 regularly did so. Several attempts to tune the prompt were unsuccessful.

We were more successful with GPT-4. The GPT-4 prompt used was:

```
Obligations generally say a person
'must' do something. Summarise the
obligation in the following text in a
manner that is sufficient for an
obligations register. Write concisely but
maintain an acceptable degree of legal
accuracy. Do not use list or bullet point
formatting. Where a list of exceptions
applies to the obligation, simply state
the relevant subsection rather than
outlining every exception. Begin by
saying 'a [insert] must'. Do not state
the nature of the offence. If no
obligation is detected, state 'No
obligation detected'. Use Australian
English spelling. Relevant text: [source
material]
```

GPT-4 naturally writes more concisely than GPT-3.5. Therefore, we found that a prompt containing words to the effect of "in as few words as possible" was interpreted too literally. We tuned the prompt to balance writing concisely with maintaining a degree of legal accuracy acceptable for creating an obligations register. GPT-4 "understood" the task better than GPT-3.5 and can reliably produce outputs that do not include the nature of the offence.

To ensure legal accuracy in the generation of obligations, we employed a lower temperature setting, ranging between 0.0 and 0.3. This lower temperature restricts the LLMs output to the most probable words, reducing variation and encouraging more precise, deterministic completions. This heightened level of predictability and accuracy is crucial in the context of generating obligations registers.

We evaluated the system using diverse legislative documents, and in certain instances,

GPT-4 produced obligations exhibiting greater legal precision compared to our human-generated reference registers. For instance, the obligation listed in our reference register for Section 912DA of the Corporations Act simply stated, "A licensee must notify ASIC of changes in control." In contrast, the GPT-4-generated obligation provided a more legally accurate description: "A financial services licensee must notify ASIC of changes in control within 30 business days, using the prescribed form." The key distinction between the two is the specific time limit, which the human missed.

All completions/obligations are stored in a local database, which is then post-processed for duplicates and other issues. That is, after assembling a register from multiple sources of law, it's necessary to merge functionally equivalent obligations. For example, the high-level obligation to "… not engage in misleading or deceptive conduct" is covered by various sources of law depending on the context but should be merged into one obligation for the purposes of an obligations register.

To check for matching or functionally equivalent obligations in a spreadsheet, we perform the following operations:

1. Import the obligations register into the Python environment, using the pandas library.
2. Clean and normalise the text data by converting it to lowercase, removing punctuation, stop words, and stemming or lemmatizing the words. We use the nltk or spaCy libraries for this purpose.
3. Convert the text into numerical representations and using TF-IDF (Term frequency – Inverse document frequency) to compare the similarity between the obligations more effectively.
4. Use cosine similarity to compare the numerical representations of each pair of obligations, to generate a similarity score to identify matching/functionally equivalent obligations, which is then assessed against a pre-defined score of "matching" or "functionally equivalent."
5. Flag those obligations that have a score above the chosen threshold, and have a user validate that these obligations are

in-fact matching or functionally equivalent.

In validation and testing, we found similarity detection to be accurate in most cases. However, in some instances, the system identified obligations that sounded similar but were not functionally equivalent.

Some examples of the algorithm's matched obligations include the requirement under Australian laws for financial services licensees not to engage in unconscionable conduct under Section 991A of the *Corporations Act 2001* (Cth) (Corporations Act), and the requirement for a person not to engage in unconscionable conduct under Section 12CA of the *Australian Securities and Investments Commission Act 2001* (Cth) (ASIC Act). Another example is the obligation not to participate in misleading and deceptive conduct under Section 1041H of the Corporations Act and Section 12DA of the ASIC Act.

We are conducting further research on large semantic search operations to improve the quality of similarity check outputs, but even with the current process, we can very quickly assess and accept or reject obligations as matching/equivalent. This process, in addition to the generative AI creation of the obligations, significantly speeds up register generation and maintenance.

4. LLM-based expert systems

Related to the automatic generation of obligations registers, we have developed a method to use GPT-4 as an engine for an expert system consultation tool. This is similar in some ways to a traditional legal expert system—the topic of many ICAIL papers thirty years ago—but the system is constructed in a distinctive way, and the interface operates differently.

Our approach leverages the method of generating obligations registers described in section 3 above. The obligations register is processed using the LangChain framework [20] and OpenAI's word embeddings tool [22] to vectorize the register. LangChain is a framework for developing advanced applications powered by LLMs. Our application uses LangChain to compare the vectorized obligations register with a vectorized user query to generate a list of similarities, which are then passed to GPT-4, along with a system message, to produce a completion.

The full description of this process is as follows:

1. Prepare an obligations register and export it to a CSV file (columns include 'legislation', 'section number', and 'obligation description')
2. Initialise the dataset using LangChain CSVLoader. [23]
3. Split the dataset into chunked text data using LangChain Character Text Splitter. [24]
4. Send the chunked text data to Open AI's Embeddings API (model: text-embedding-ada-002). Open AI returns text as vectors. [22]
5. Use LangChain ChatGPT Plugin Retriever to combine the text as vectors with a vectorized user query (i.e., `"List all the obligations I may have breached by [insert]"`) and a system message (i.e., `"Use the following pieces of legal text to answer the question at the end [insert]"`) to prompt Open AI's ChatGPT (model: gpt-4) for reasoning. [25]
6. Open AI returns the final output.

Two examples of the interaction with this tool give an idea of how the system works:

Scenario 1
User prompt
`Failing to keep records of income`

Response
`Life Insurance Act 1995, Section 75: A life company must keep records of income and outgoings for each statutory fund.`

`Life Insurance Act 1995, Section 76: A life company must keep records of income and outgoings for overseas funds.`

`Corporations Act 2001, Section 988B: The records must be kept in a way that enables true and fair profit and loss statements, and balance sheets, of the financial services business of the licensee to be prepared from time to time and allows those statements and balance sheets to be conveniently and properly audited.`

User prompt

```
Giving misleading information to an
auditor
```

Response

```
Corporations Act 2001, Section 1309: A
company    must    not    give    misleading
information about a company to a director
or   auditor   knowing   its   misleading.   A
company  must  take  reasonable  steps  to
ensure  it  does  not  give  misleading
information about a company to a director
or                                      auditor.
```

```
Financial  Sector  (Collection  of  Data)
Act 2001, Section 17D: Giving knowingly
false  information  to  an  auditor  of  a
reporting   entity   or   failing   to   take
reasonable  steps  to  avoid  doing  so.
```

```
Life Insurance Act 1995, Section 91(1):
An employee or officer of a life company
must    not    give    false    or    misleading
information  to  an  auditor  regarding  the
affairs        of        the        company.
```

```
Life Insurance Act 1995, Section 91(2):
An employee or officer of a life company
must  take  reasonable  steps  not  to  give
false  or  misleading  information  to  an
auditor  regarding  the  affairs  of  the
company.
```

In each of the above scenarios, lawyers reviewed the outputs provided by GPT-4 and confirmed that they were relevant and accurate based on the information supplied in the user prompt. These results demonstrate the potential of this tool to streamline the obligations assessment process for corporate legal teams and licensees. To address confidentiality concerns, the intended interactions with this tool is used in hypothetical scenarios only.

Significant testing of the validity of this approach is necessary for deployment in commercial settings; but initial user feedback has demonstrated the basic utility of the approach.

5. Mitigating problems with generative AI

One of the features of generative AI models is that they have no internal representation of the world, and instead they are merely generating text one word at a time, based on a mathematical analysis of what word should follow from the previous ones. This means that they are prone to creating text that sounds plausible but is utterly wrong. This poses a huge risk for unsupervised generative AI systems within a law firm, where accurate advice is a fundamental requirement of legal practice.

The Gracenote platform controls for the hallucination problem by requiring all content to be approved by a responsible person within the firm, typically a partner. Content is never published by the generative AI model; instead, the model generates draft content according to agreed-upon prompts and displays this side-by-side with the material from the feed. The human author assesses the correctness of the content, can edit it as necessary, and only then publishes the content to the database—which is then used to send content externally to practice groups or client.

This type of control—often called "human-in-the-loop"—ensures that a law firm is never exposed to risk of poor-quality content going out under its name. This type of control also mitigates issues with semi-random changes in completions that can occur with some models. In essence, because a responsible human will always control the editing and dissemination of content, it doesn't matter that a prompt to any given LLM may generate slightly different completions over time.

Privacy and confidentiality are two other hygiene concerns with the use of LLMs in legal settings. Sending personally identifying information to a public endpoint of a LLM may be a breach of various privacy laws, including the GDPR or the *Australian Privacy Act 1998* (Cth). Similarly, using confidential information as a prompt for a LLM may lead to disclosure of that information, contrary to a range of ethical/legal professional practice laws, as well as the commercial interests of a firm.

To address these concerns, we propose two strategies. For scenarios involving sensitive user information, we propose the use of privately-hosted open-source models or privately-hosted proprietary models. This approach provides a layer of security, ensuring that confidential information remains within a secure, controlled environment within a set jurisdiction. On the other hand, when working with publicly available information, we expect to continue utilising public proprietary models such as GPT-4. The

rationale behind this strategy lies in the superior reasoning capabilities these models are expected to maintain. Their performance is attributable to the substantial compute and vast datasets involved in their training process, an advantage open-source models might lack. We do not use private or confidential information in any of the tools described here, and so we haven't needed to use privately-hosted open-source models or privately-hosted proprietary models.

6. Conclusions and further work

To the best of our knowledge there is no other generative AI company that focuses on GRC problems in the way we do. We believe that the tools and methods described here are advances on the state of the art in the delivery of legal services and compliance functions using generative AI.

Our future work includes the following:

1. For the horizon scanning tool, we are working with a range of law firm clients to expand the range of supported feeds, in multiple jurisdictions and languages (Australia, England and Wales, Singapore, and the USA). We are working on being able to tailor the tone of the summary from formal to informal to support a wider range of audiences and clients. We are also working on improving link sanitisation for the one-off insertion of links by users, and refining prompts to improve the quality of outputs for various subject matter. As an extension of this research, we will investigate creating AI agents that autonomously navigate the web and execute searches, enabling law and consulting firms to proactively respond to rapidly evolving or emerging regulatory updates from a broad range of industry news sites and media outlets.

2. For the obligations-generation tool, we are working on a user-friendly interface to facilitate the management of obligations and provide a seamless experience for users. We are also working on supporting multiple jurisdictions, languages and tones for different audiences. We also propose developing an interface for building registers from various sources of law. This interface will enable users to construct a register from multiple sources of law, including specific divisions from those sources. The register can then be deployed via an API connection to various governance, risk, and compliance platforms. Finally, this tool will have a method for versioning legislative provisions. This feature will enable users to import amending legislation, which will update the relevant provisions of the primary register. A GitHub-style approach will be used to commit changes to production registers, allowing for an affordable and always up-to-date obligations register.

3. For the expert system generator, we plan to investigate the inclusion of penalties in the dataset to enable GPT-4 to reason whether an obligation is "deemed significant" and determine if a breach of those obligations is reportable to regulators. We also propose to extend the approach to a range of other pieces of legislation, in multiple jurisdictions. We believe that this approach has significant benefits over other methods of producing expert system tools and is an advance on the standard methods. We also believe that it can provide a more comprehensive and accurate solution to streamline the assessment and reporting of compliance breaches.

Generative AI models will change a huge range of legal functions, and one of these is in legal and regulatory compliance. LLMs provide remarkable opportunities to improve decision-making in law, as we have demonstrated with the tools and methods, we have developed within Gracenote.

7. References

[1] Rupert Macey-Dare, "How ChatGPT and Generative AI Systems will Revolutionize Legal Services and the Legal Profession. Retrieved May 5, 2023 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4366749

[2] Daniel Martin Katz, Dirk Hartung, Lauritz Gerlach, Abhik Jana and Michael J. Bommarito, Natural Language Processing in the Legal Domain.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4336224

[3] Daniel Schwarcz & Jonathan H. Choi, AI Tools for Lawyers: A Practical Guide. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4404017

[4] Daniel Martin Katz, Michael James Bommarito, Shang Gao and Pablo Arredondo, GPT-4 Passes the Bar Exam. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4389233

[5] OpenAI. 2023. GPT-4 Technical Report. DOI: https://arxiv.org/abs/2303.08774

[6] Stuart Hargreaves, 'Words Are Flowing Out Like Endless Rain Into a Paper Cup': ChatGPT & Law School Assessments. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4359407

[7] Ilias Chalkidis, ChatGPT May Pass the Bar Exam Soon, but as a Long Way to Go for the LexGLUE Benchmark. https://arxiv.org/abs/2304.12202

[8] Jonathan H. Choi, Kristin E. Hickman, Amy B. Monahan and Daniel Schwarcz, ChatGPT Goes to Law School. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4335905

[9] Kwan Yuen Iu & Vanessa Man-Yi Wong, ChatGPT by OpenAI: The End of Litigation Lawyers? https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4339839

[10] Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. DOI: https://doi.org/10.48550/arXiv.2303.10130

[11] OpenAI. Fine-tuning. https://platform.openai.com/docs/guides/fine-tuning

[12] OpenAI. (February 2023). Model index for researchers. OpenAI: https://platform.openai.com/docs/model-index-for-researchers

[13] LawCareers.Net, Law firm to hire 'AI engineer' to identify how its lawyers can use ChatGPT. (March 2023) Retrieved May 5, 2023 from https://www.lawcareers.net/Explore/News/Law-firm-to-hire-AI-engineer-to-identify-how-its-lawyers-can-use-ChatGPT-060320

[14] ResumeBuilder. 9 in 10 companies that are currently hiring want workers with ChatGPT experience. (April 2023). https://www.resumebuilder.com/9-in-10-companies-that-are-currently-hiring-want-workers-with-chatgpt-experience

[15] Gabriel Pereyra and Winston Weinberg. Sequoia and OpenAI Back Harvey to Redefine Professional Services, Starting with Legal. Harvey.ai. (April 2023). https://www.harvey.ai/blog

[16] CaseText. Meet CoCounsel-the world's first AI legal assistant. (March 2023). https://casetext.com/blog/casetext-announces-cocounsel-ai-legal-assistant

[17] RobinAI. Unleash the Power of Large Language Models in the Legal Industry with Robin AI. (February 2023). https://www.robinai.co.uk/post/large-language-models-legal-industry

[18] Lexology. (May 2023. https://www.lexology.com/

[19] AutoGPT. (May 2023). https://github.com/Significant-Gravitas/Auto-GPT

[20] LangChain. (May 2023). https://langchain.com

[21] Hannah Wootton, In-house lawyers 'failing to manage regulatory, business risks'. (29 April 2021) Retrieved May 5, 2023 from https://www.afr.com/companies/professional-services/in-house-lawyers-failing-to-manage-regulatory-business-risks-20210428-p57n79

[22] Ryan Greene, Ted Sanders, Lilian Weng, and Arvind Neelakantan. 2022. *New and improved embedding model.* OpenAI. (December 2022). https://openai.com/blog/new-and-improved-embedding-model

[23] Harrison Chase. *CSV Files*. LangChain. (May 2023). https://python.langchain.com/en/latest/modules/indexes/document_loaders/examples/csv.html

[24] Harrison Chase. *Character Text Splitter*. LangChain. (May 2023). https://python.langchain.com/en/latest/modules/indexes/text_splitters/examples/character_text_splitter.html

[25] Harrison Chase. *ChatGPT Plugin Retriever*. LangChain. (May 2023). https://python.langchain.com/en/latest/modules/indexes/retrievers/examples/chatgpt-plugin-retriever.html
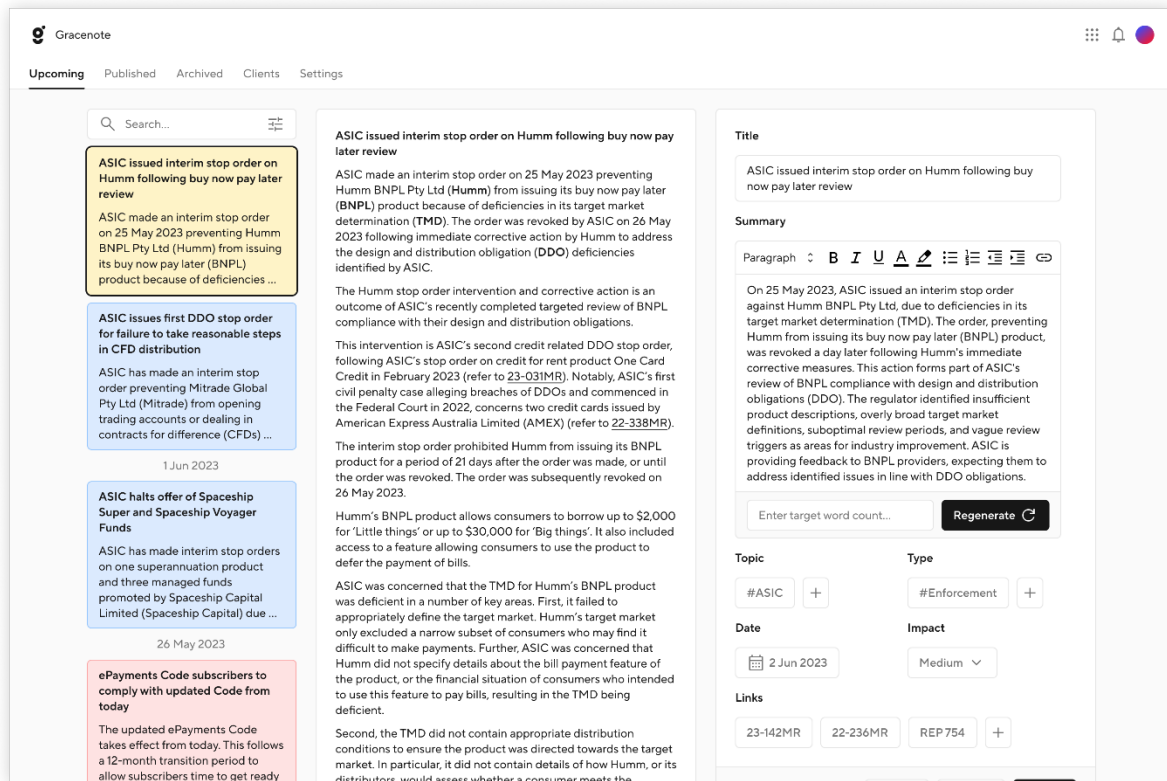
Annexure A – User interface
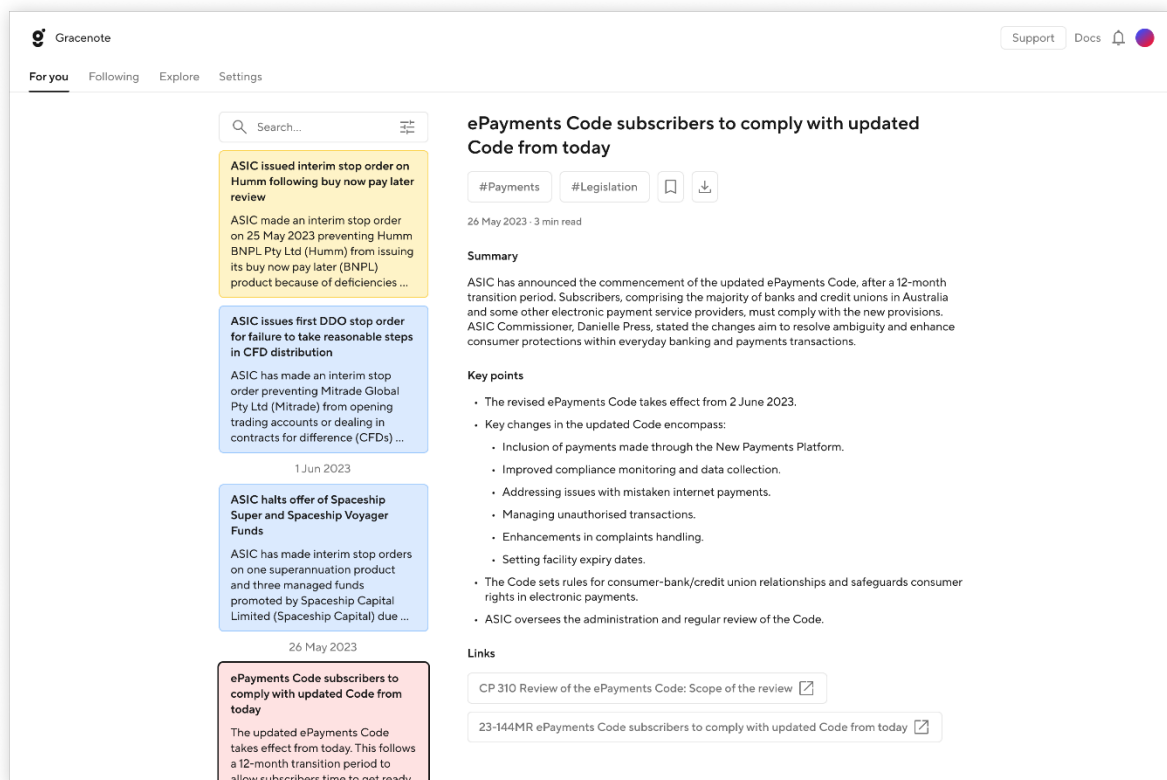

Figure 1: Authoring environment

Figure 2: Client environment