

# Machine Learning for Tax or Customs Administrations Case Study: Answering Questions using Semantic Search

## 1 Introduction

The techniques that are presented in the **Online Workshop: Practical Induction to Machine Learning and Data Sciences for Tax Administrations** are, by definition, simplified exercises that demonstrate specific techniques and algorithms in machine learning. It is sometimes difficult to see how these can be used in production. In this document we present a series of case studies of how the Tax or Customs Administration can use the techniques on real-life applications, showing all the steps from start to finish.

As is mentioned multiple times in the workshop, machine learning applications are only possible if the organization has reliable data that we can trust. Data is then used by skilled human resources to create tools that augment a tax administration to make it more efficient, more intelligent, and productive.



Machine Learning is used to a) correct deficiencies in the revenue administrations, such as unreliability, slowness, and inaccuracies, b) Detect irregularities and errors, c) predict fraudulent behavior and d) proactively assist taxpayers,

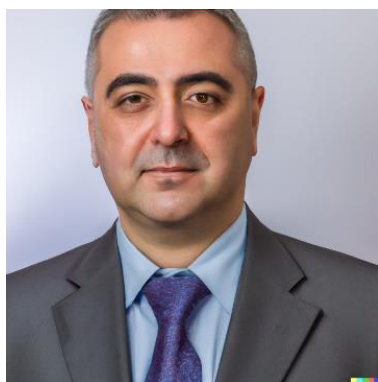
Machine learning is most effective when it is used across organization to solve many small problems, rather than have one single application such as identifying fraudulent tax or import declarations. Think of your cellular phone and you will see that it uses artificial intelligence in many small areas that collectively improve the experience of using your phone. It recognizes your fingerprint with one application, it recognizes your face with a different application, it predicts the word you will type next with a separate application, it recognizes a voice command with a separate application, etc.

Similarly, the Tax or Customs administration can use a wide variety of small machine learning applications to solve many inefficiencies or enable many advanced applications. Small applications add up and enable levels of performance that are hard to reach with traditional systems.

## 2 Case Study

In this case study, we will focus on using machine learning to help tax officials find information they need quickly and efficiently, and hence increasing the effectiveness of their work, and taxpayers find information using the website, avoiding calling the tax administration's helpline and hence reducing operational costs. The source code for the entire exercise is Open Source and is available in <https://github.com/TaxAIExamples/semantic-search-using-chatgpt>. There is a video that demonstrates the resulting system, available in [https://youtu.be/QwzII\\_KK1KE](https://youtu.be/QwzII_KK1KE).

The example that is shown here took two weeks to complete, from data collection to final production-level prediction, giving us a reasonable idea of the level of complexity of building machine learning algorithms using state of the art public domain software. Examine the code carefully and you will be amazed how simple it is to build a system that, before large language models existed, would have taken an army of programmers and a very long time to produce.



<sup>1</sup> Abdullah bin Ahmed, is the manager in the Tax Administration of Laurania<sup>2</sup> (TAL) who is responsible for maintaining and updating the agency's internal knowledge base and oversees the development of the public-facing website. He is in need of a more effective search solution and has decided to explore the potential of ChatGPT and semantic search to improve the user experience.

The tax agency's internal knowledge base and public-facing website are filled with complex information, making it difficult for users to find specific answers or guidance quickly. Currently, the Tax Administration relies on traditional text searches for their knowledge base, but tax officials struggle to find the information they need quickly and efficiently, and taxpayers find it difficult to

access tax information using the website, resulting in frequent calls to the tax administration's helpline, leading to increased operational costs.

Traditional text searches in a knowledge base suffer from several limitations, which can lead to disappointing performance, and Abdullah hopes that semantic searches using AI can overcome many of these problems and offer more relevant, accurate results.

Currently, the Tax Administration has a knowledge base that uses traditional text searches that rely on keyword matching. The current approach has the following limitations:

1. Searches look for exact matches of the words entered in the search query. This approach often fails to capture the essence of the query and misses out on relevant information. For example, if a taxpayer

---

<sup>1</sup> Note that Machine Learning was used to generate this fictitious identity. This person does not exist. Photo from DALL-E <https://labs.openai.com/> using a Generative Pre-trained Transformer (GPT) model that was initially developed by OpenAI in 2018 using a Transformer architecture.

<sup>2</sup> Laurania is a fictional European state that is used for examples in the Tax Diamond tool.

searches for "retirement benefits," a traditional search might miss a document discussing the "deductions available for retirement" because the exact keywords don't match.

2. Searches have difficulty dealing with ambiguous queries, as they can't discern the intended meaning based on context. For example, if a taxpayer searches for "Tax Deduction," the existing search engine will return results related to both personal income tax deductions and business tax deductions, making it hard to find the information needed.
3. Searches struggle with complex queries, especially those that involve multiple concepts or require an understanding of relationships between ideas. For example, if a taxpayer searches "effects of tax policies on small businesses and economic growth," the existing search engine might return results discussing either tax policies, small businesses, or economic growth, but not necessarily the intersection of all three topics.
4. Searches currently use simple algorithms like term frequency or keyword density to rank results. This often leads to suboptimal rankings, with less relevant or outdated information appearing higher in the search results.

From what Abdullah was able to discern, Semantic searches using AI, on the other hand, can overcome many of these problems and offer more relevant, accurate results.

If semantic searches use AI to understand the meaning and context of a search query, the search engine should be capable of identifying synonyms and related concepts, leading to more comprehensive and relevant results. A semantic search would understand that "tax deduction" and "tax credit" are closely related and return documents discussing the advantages of tax credits when the search is about tax deductions.

Abdullah also hopes that semantic searches can leverage AI to analyze the context of the query and determine the most relevant meaning. If a tax official searches for "Apple revenue," Abdullah hopes that the AI would understand that she is interested in the technology company's financial information, not the fruit. If semantic search can analyze complex queries and identify the relationships between concepts. It can understand that someone is looking for information on how tax policies impact both small businesses and economic growth and return relevant results accordingly.

To illustrate the difference between traditional text searches and semantic searches, consider the analogy of searching for a needle in a haystack. Traditional text search is like sifting through the haystack by hand, looking for a needle that matches a specific shape or size. Semantic search, on the other hand, is like using a metal detector that can find any needle, regardless of its shape or size, and even alert to the presence of other related metal objects.

Abdullah's vision is that the Tax Administration of Laurania (TAL) can implement an intelligent search engine that can be trained to serve information on two domains, as can be seen in the figure on the right.

The Website Search tool would be used by taxpayers and the internal search engine would be used by tax officials. The semantic search engine would process requests from both domains and answer questions that are specific to each. By processing requests from these two different domains, the engine can provide accurate and relevant information specific to each user group. This dual-use system not only would streamline the tax information search process but also help bridge the knowledge gap between taxpayers and tax officials, leading to better tax compliance and more efficient tax administration.



Abdullah then contacted the IT department, that allocated a Data Science specialist to work on this issue.



<sup>3</sup> Kavya Gupta, the Data Science specialist assigned to the case, began by analyzing the Tax Administration's internal knowledge base and the public-facing website. She realized that the knowledge base contained a vast amount of unstructured text data, while the public-facing website and internal search interface needed to present information in a user-friendly and conversational manner.

Because of the recommendation of Abdullah bin Ahmed, the ideal solution should be able to understand, interpret, and generate human-like responses for taxpayers and tax officials. She decided to use text embeddings and a large language model to address these challenges.

Text embeddings are a technique used to convert text into numerical representations, making it easier for machine learning models to process and understand the text. By creating embeddings for the Tax Administration's internal knowledge base, Kavya could efficiently analyze and extract relevant information. However, text embeddings alone would not be sufficient to generate user-friendly responses on the public-facing website. This is where a large language model comes into play. Large language models are pre-trained on massive amounts of text data and can generate contextually relevant and coherent responses. By combining text embeddings with a large language model, Kavya could not only analyze the knowledge base efficiently but also generate user-friendly responses.

---

<sup>3</sup> Identity fictitious. This person does not exist. Photo from Realistic Vision V2.0 <https://civitai.com/models/4201/realistic-vision-v20> image generated by a fine-tuned Stable Diffusion Transformer Network.

## 2.1 About Text Embeddings

Embeddings are mathematical representations of words or phrases in a multi-dimensional vector space. They capture the semantic meaning and relationships between words, allowing AI models to understand and process natural language queries more effectively. Embeddings are generated using advanced algorithms that analyze vast amounts of text data, identifying patterns and relationships between words to create dense vector representations.

To better understand how embeddings work, imagine that every word or phrase in a language is a point on a map. Embeddings essentially assign a set of coordinates to each word, placing them at specific locations on the map. Words with similar meanings or related concepts are placed closer together, while words with different meanings or unrelated concepts are placed further apart.

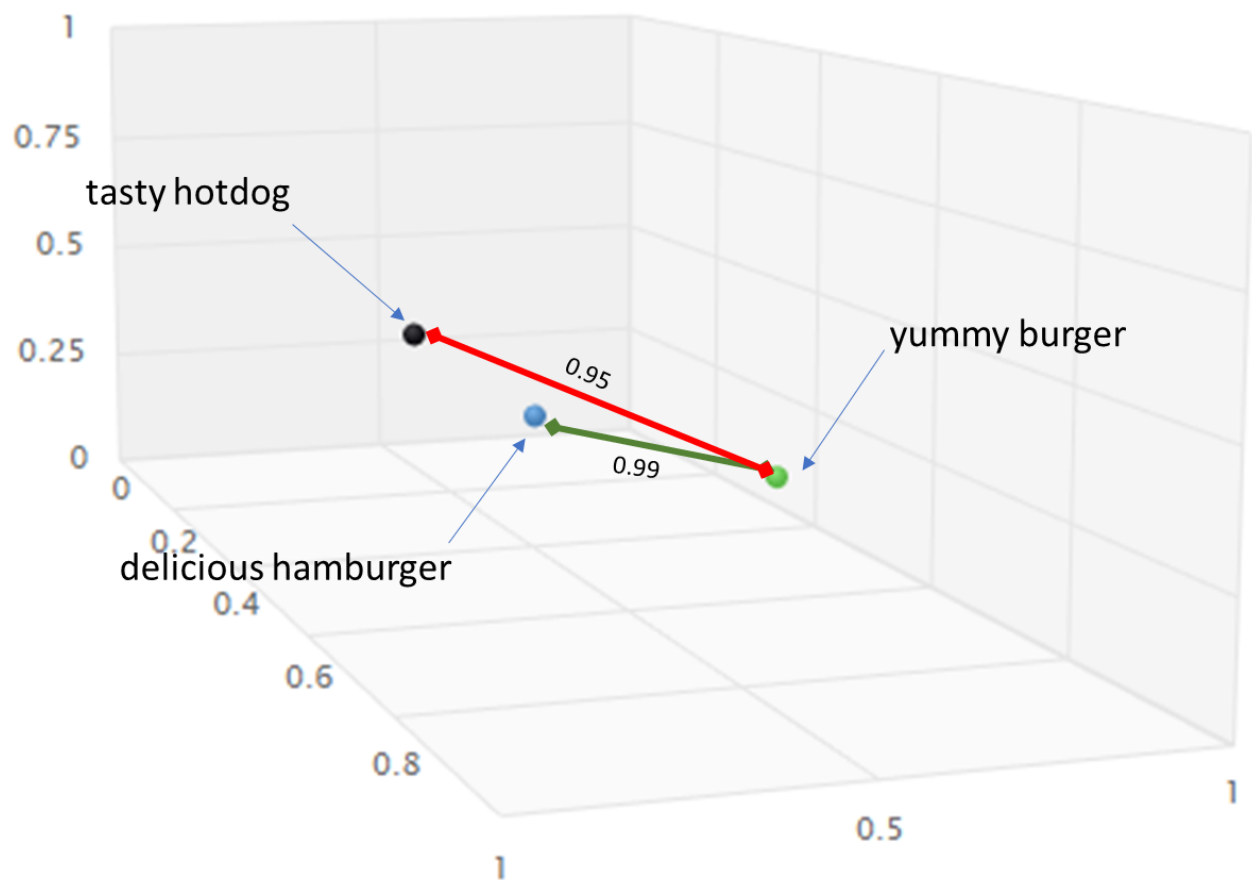
For instance, on this hypothetical map, the words "car," "automobile," and "vehicle" would be positioned close together, as they share similar meanings. Conversely, the words "car" and "apple" would be placed further apart, as they represent unrelated concepts.

When searching for similar concepts, the AI model calculates the distance between these vectors to determine their semantic similarity. The shorter the distance between two vectors, the more closely related their meanings are. For example, let's consider the words "bank," "financial institution," and "riverbank." While the word "bank" can refer to both a financial institution and the side of a river, its vector representation will be closer to "financial institution" due to their shared context and meaning. Consequently, when searching for information related to "bank," the AI model will prioritize results related to financial institutions over those related to riverbanks.

For example, let's say we have two phrases, "delicious hamburger" and "tasty hotdog". We can represent each of these phrases as a vector, such as:

"delicious hamburger" = [0.8, 0.5, 0.2]

"tasty hotdog" = [0.7, 0.6, 0.1]



Here, to illustrate the concept, each vector has three dimensions, but in reality, embeddings can have even hundreds or thousands of dimensions.

To find the most related phrase to a given query, we can calculate the distance between the query's vector and the vectors of other phrases. One common way to measure distance is the cosine similarity, which gives a value between -1 and 1, where values closer to 1 indicate more similarity.

For example, let's say we have a query "yummy burger". We can represent this as a vector, such as:

"yummy burger" = [0.85, 0.4, 0.15]

We can then calculate the cosine similarity between "yummy burger" and the vectors for "delicious hamburger" and "tasty hotdog":

$\text{cosine similarity}(\text{"yummy burger"}, \text{"delicious hamburger"}) = 0.99$

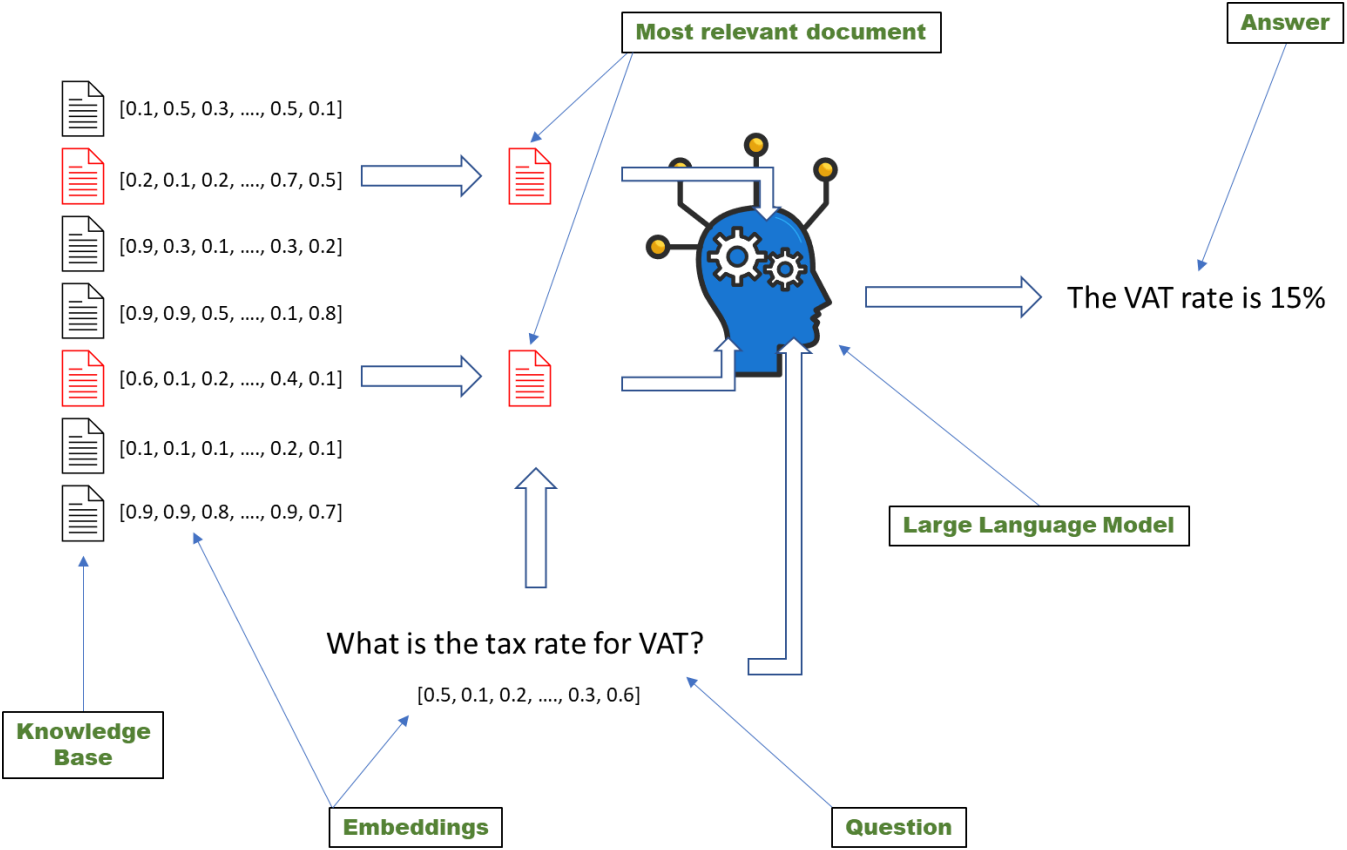
$\text{cosine similarity}(\text{"yummy burger"}, \text{"tasty hotdog"}) = 0.95$

Here, we can see that "yummy burger" is more similar to "delicious hamburger" than to "tasty hotdog", based on the cosine similarity scores. Therefore, "delicious hamburger" would be the most related phrase to the query "yummy burger" in this example.

Embeddings are generated using advanced machine learning algorithms that analyze vast amounts of text data. These algorithms identify patterns and relationships between words, learning their contextual usage and semantic associations. By processing this information, the algorithm can create dense vector representations for each word or phrase, capturing their semantic meanings in a multi-dimensional vector space. The result is a language model that is capable of understanding and processing natural language queries more effectively.

**Use of Large Language Models in Semantic Search**

Semantic search is a multi-step process that provides users with relevant and precise information by understanding the intent and context behind a query. The first step in this process is creating embeddings to efficiently retrieve documents from a knowledge base. The second step involves using a large language model (LLM) to understand the query, extract the required information from the retrieved documents, and present the answer in an easily understandable form, as can be seen in the figure below:



If the LTA aims to improve the search functionality of its online taxpayer knowledge base, it needs two primary components: document embeddings and a large language model (LLM).

As we discussed before, the LTA gathers all tax-related information into a comprehensive database, referred to as the knowledge base. Document embeddings are then created for each piece of content within the knowledge base. These embeddings capture the semantic meaning of the content, allowing for efficient indexing and retrieval of documents based on similarity.

For example, if a taxpayer searches for "tax deductions for homeowners," the system will retrieve documents that are semantically related, even if they do not contain the exact phrase "tax deductions for homeowners."

Once the relevant documents have been retrieved, the LLM comes into play. It is responsible for understanding the user's query, extracting information from the retrieved documents, and presenting the answer in a human-readable format.

For example, if a taxpayer issues the query "What are the tax deductions available for homeowners?"

The system should:

1. Use embeddings to interpret the taxpayer's query and identifies the key components, such as "tax deductions" and "homeowners.". We do this simply by computing the embeddings of the question.
2. Scans the documents retrieved through embeddings, analyzing the content for the necessary information to answer the query. We find the best documents simply by computing the "distance" between the question vector and the vector for each of the documents in the database.
3. Feed the LLM the original question and the relevant documents to generate a response. The LLM constructs a natural language response that conveys the extracted information in a manner that the taxpayer can understand, such as: "Homeowners may be eligible for the following tax deductions: 1. Mortgage interest, 2. Property taxes, 3. Home office expenses (if applicable), and 4. Energy-efficient home improvements."

## 2.2 Selection of the best technologies

In recent years, the advancements in natural language processing (NLP) have led to the development of various text embedding and large language model (LLM) techniques that can significantly improve the performance of semantic search applications. The AI community Hugging Face has emerged as a popular platform offering a wide range of pre-trained models and tools for NLP tasks, including text embeddings and LLMs such as BERT, GPT-2, RoBERTa, and DistilBERT<sup>4</sup>. These models have demonstrated their efficacy in various applications, including semantic searches on knowledge bases.

Apart from Hugging Face, there are other potential sources for text embedding and LLM models. TensorFlow Hub<sup>5</sup>, developed by Google, provides pre-trained models and embeddings like Universal Sentence Encoder, which can be used for semantic search tasks. The Allen Institute for AI's AllenNLP library<sup>6</sup> is another resource that offers pre-trained models, including ELMo and OpenAI's GPT<sup>7</sup>, suitable for text-based applications. SpaCy<sup>8</sup>, a popular Python NLP library, also supports word embeddings like Word2Vec, GloVe, and fastText, which can be used to enhance semantic search applications.

While these sources offer a variety of models and embeddings, Kavya decided to use ChatGPT developed by OpenAI<sup>9</sup>. ChatGPT, based on the GPT-4 architecture, stands out as the best choice for implementing semantic searches on a knowledge base for a Tax Administration. There are many reasons behind this choice.

The current incarnation of ChatGPT demonstrates an exceptional understanding of natural language, owing to its training on vast amounts of text data. This enables it to generate highly accurate and contextually relevant search results.

---

<sup>4</sup> See <https://huggingface.co/docs/transformers/index>

<sup>5</sup> See <https://tfhub.dev/s?module-type=text-embedding,text-language-model>

<sup>6</sup> See <https://github.com/allenai/allennlp>

<sup>7</sup> See <https://github.com/allenai/allennlp-models>

<sup>8</sup> See <https://spacy.io/universe>

<sup>9</sup> See <https://platform.openai.com/docs/models>



As a state-of-the-art LLM, ChatGPT is capable of handling complex queries and multi-turn conversations<sup>10</sup>, which is a vital requirement for an efficient semantic search system. This feature would allow Tax Administration users to ask follow-up questions or provide additional context for their queries, resulting in more accurate search results.

Finally, ChatGPT has proven its versatility by showcasing its capabilities in various domains, including Tax Administrations. Its ability to understand and process domain-specific terminology, rules, and regulations makes it an ideal choice for enhancing the search experience of Tax Administration users.

In conclusion, although there are multiple sources for text embeddings and LLM models, ChatGPT's advanced language understanding, contextual relevance, and domain-specific expertise make it the optimal choice for implementing a semantic search system for a Tax Administration knowledge base.

## 2.3 Implementation Details

### 2.3.1 Technologies

Within the ChatGPT constellation of solutions, the OpenAI embeddings API is detailed in <https://platform.openai.com/docs/api-reference/embeddings> and the details of the technologies are documented in <https://platform.openai.com/docs/guides/embeddings/what-are-embeddings>. For the project Kavya selected V2 of the text-embedding-ada-002 model, that is capable of indexing a piece of text up to 8191 tokens.

For the LLM component, the API is called Completions and documented in <https://platform.openai.com/docs/api-reference/completions> and the list of possible models we can use is documented in <https://platform.openai.com/docs/models>. As of this writing, GPT-3.5 is a good candidate since it performs well with the least cost associated.

We will need to store our knowledgebase in a database that stores the knowledge articles and the corresponding text embeddings. The Website for the LTA is built on .NET Core, so we will build a .NET Core solution to manage our knowledgebase and to store it in a local SQL Database. Our .NET Core solution will use Entity Framework to manage the database, simplifying the process of creating the database considerably.

For the initial proof of concept, we will create two Windows Forms applications. The first will allow us to manage the knowledgebase documents, and the second application will allow us to query the semantic search system, simulating a call from the Website. Once the system has been tested, migrating the search query component to the Website is a trivial task.

### 2.3.2 Constructing the Knowledgebase

The number of words in 8,191 tokens can vary depending on the text, as tokens do not necessarily correspond to whole words. In English text, a token can represent a word, a part of a word, or a punctuation mark. For instance, the word "don't" is tokenized into two tokens: "don" and "'t". Similarly, contractions, compound words, or words with special characters may also be split into multiple tokens. On average, however, in English

---

<sup>10</sup> A multi-turn conversation is an interaction between two or more participants, typically a user and an AI system, where the conversation consists of multiple exchanges or "turns" between the parties. In contrast to single-turn conversations, where the user asks a question and receives a one-time response, multi-turn conversations involve back-and-forth exchanges that allow for clarifications, follow-up questions, or additional context to be provided.

text, 8,191 tokens would generally correspond to about 8,000 words, considering variations due to contractions, compound words, and punctuation.

This means that every document in the Knowledge Base must have at most 8,000 words. The number of pages that 8,000 words will take up depends on several factors, such as the font size, line spacing, and margins used. However, a general estimate is that a single-spaced page with 12-point font and one-inch margins will contain approximately 500 words. Based on this estimate, 8,000 words would take up approximately 16 single-spaced pages or 32 double-spaced pages with standard formatting.

Even though ChatGPT Embeddings can encode approximately 16 single-spaced pages, it is crucial to ensure that the documents in a knowledge base for a semantic search system are significantly smaller. The primary reason for this is that very large documents often cover multiple topics, which may reduce the effectiveness of the semantic search system. To build a truly effective semantic search system using Text Embeddings and ChatGPT as a Language Model (LLM), it is best to have each article in the knowledge base focus on a single topic.

To build a semantic search knowledge base, Kavya's team will start from zero and incrementally add new smaller articles that conform to the following good practices:

- **Focus on one topic per article:** Limit each article to a single topic to improve the precision and relevance of search results. This helps the LLM understand the context and provide more accurate responses.
- **Keep it concise:** Write clear and concise content, avoiding unnecessary jargon or filler. This will make it easier for the model to comprehend the text and deliver relevant results.
- **Use clear headings and subheadings:** Organize the content with clear headings and subheadings to improve readability and make it easier for the model to identify key points within the article.
- **Incorporate relevant keywords:** Use relevant keywords and phrases in the content to enhance its discoverability. However, avoid keyword stuffing, as this can negatively impact the model's performance.
- **Maintain a consistent structure:** Follow a consistent structure and format across all articles in the knowledge base. This will help the model better understand the context and improve its ability to provide accurate responses.
- **Regularly update content:** Keep the knowledge base up-to-date by revising and updating content as needed. This ensures that the semantic search system provides the most current and accurate information.
- **Curate high-quality content:** Prioritize the use of reliable and authoritative sources to ensure the accuracy and credibility of the knowledge base. This will help the model deliver reliable search results and boost user trust.

One of the greatest advantages of building a knowledge base following these best practices is that it can be done incrementally, allowing the Tax Administration of Laurania (TAL) to compute the embedding vectors for each new document as it is added to the system. This means that, initially, the semantic search will have knowledge about some basic topics, but as time goes on and more articles are added, the knowledge base will expand and become more comprehensive.

This incremental approach offers several benefits. The TAL doesn't have to wait for a massive dataset before its semantic search system becomes useful. Even with a small initial knowledge base, the system can provide valuable search results and insights. As new documents are added and existing ones are updated, the

performance and relevance of the semantic search system will improve over time, offering TAL's users better and more accurate results.

Additionally, an incremental knowledge base allows the TAL to quickly add new information, as it becomes available. Building and maintaining a knowledge base is a significant task, but by developing it incrementally, TAL can distribute the workload over time, making it more manageable and sustainable.

Kavya's team identified that the following sources of information can be used to seed the initial knowledge base:

**Regulations:**

- Laurania Income Tax Act (LITA)
- Laurania Corporate Tax Regulation (LCTR)
- Laurania Value Added Tax (LVAT) Law
- Laurania Capital Gains Tax (LCGT) Regulation
- Laurania Inheritance and Gift Tax (LIGT) Code
- Laurania Taxation of Non-Residents Act (LTNR)
- Laurania Tax Credits and Deductions Regulation (LTCD)
- Laurania Payroll Tax and Social Security Contributions Law (LPTSS)
- Laurania Tax Dispute Resolution and Appeals Act (LTDRA)
- Laurania Tax Compliance and Reporting Requirements Regulation (LTCRR)

**Internal Documents:**

- LITA Section 5 Interpretation and Guidance
- LVAT Exempt Goods and Services List
- LCTR Corporate Tax Rate Schedules
- LTNR Tax Treaty and Double Taxation Agreements
- LIGT Inheritance and Gift Tax Thresholds
- LTCRR Record Keeping Requirements for Businesses
- LPTSS Employee and Employer Contribution Rates
- LTDRA Tax Dispute Resolution Process
- LTCD Eligible Tax Deductions for Individuals
- LCGT Capital Gains Tax Calculation Methodology

**FAQs:**

- How do I register for a Laurania Tax Identification Number (LTIN)?
- What are the income tax rates for individuals in Laurania?
- How do I claim a tax credit for my children in Laurania?
- When is the annual tax return due in Laurania?
- What is the Laurania Value Added Tax (LVAT) rate?
- How do I calculate my capital gains tax in Laurania?
- What are the reporting requirements for non-resident taxpayers in Laurania?
- How do I submit my tax return electronically in Laurania?
- What are the penalties for late tax return submission in Laurania?
- Can I claim my home office expenses as a tax deduction in Laurania?

### Other Relevant Content:

- Laurania Tax Administration's Guide to Filing Taxes for Individuals
- Laurania Tax Administration's Guide to Filing Taxes for Businesses
- Laurania Tax Administration's Guide to Tax Exemptions and Rebates
- Laurania Tax Administration's Guide to Reporting Foreign Income
- Laurania Tax Administration's Guide to Taxation of Rental Income
- Laurania Tax Administration's Guide to Taxation of Freelancers and Self-Employed Individuals
- Laurania Tax Administration's Guide to Taxation of Pension Income
- Laurania Tax Administration's Guide to Taxation of Dividends and Interest Income
- Laurania Tax Administration's Guide to Taxation of Cryptocurrency Gains
- Laurania Tax Administration's Guide to Tax Audit and Assessment Procedures

### Additional FAQs:

- What tax relief options are available for senior citizens in Laurania?
- How do I report my cryptocurrency transactions for tax purposes in Laurania?
- Are education expenses tax-deductible in Laurania?
- What is the process for amending a previously submitted tax return in Laurania?
- How long should I keep my tax records in Laurania?
- Are non-profit organizations exempt from taxes in Laurania?
- How can I request an extension for filing my tax return in Laurania?
- What are the tax implications of receiving an inheritance in Laurania?
- How do I determine my residency status for tax purposes in Laurania?
- Are there any tax benefits for environmentally friendly practices in Laurania?

Appendix 1 of this case study has a sample of the content that was initially selected from these documents, conforming with the stated best practices and used as the initial corpus for the technical evaluation of the effectiveness of the semantic search system.

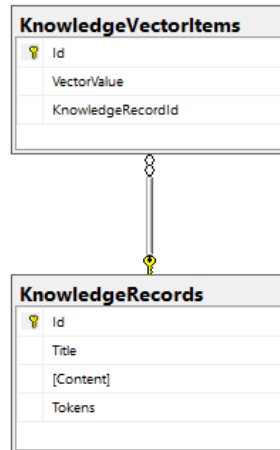
The knowledgebase articles are stored in a SQL Server database using Entity Framework. The following classes are used to define the record structure:

```
public class KnowledgeRecord
{
    public int Id { get; set; }
    public string Title { get; set; }
    public string Content { get; set; }
    public int Tokens { get; set; }
    public List<KnowledgeVectorItem> KnowledgeVector { get; set; }
}

public class KnowledgeVectorItem
{
    public int Id { get; set; }
    public double VectorValue { get; set; }
```

```
}
```

When the database is instantiated, the corresponding SQL database is very simple, as shown in the following diagram. This contributes substantially to the reduction of complexity of the end solution.



### 2.3.3 Implementing the Embedding System

The implementation is centered around the ChatGPT class. This class has only two public methods. One method calculates the embeddings for a particular knowledgebase article, and one method obtains an answer from the ChatGPT system for any given question we may ask.

```
public interface IChatGPT
{
    (string answer, List<KnowledgeRecordBasicContent> contextList)
        GetChatGPTAnswerForQuestion(string question, bool briefDetails);
    EmbeddingResponse GetEmbedding(string embedding);
}
```

The function called `GetEmbedding` calculates the embeddings for a particular knowledgebase article<sup>11</sup>.

```
public EmbeddingResponse GetEmbedding(string embedding)
{
    string apiKey = Environment.GetEnvironmentVariable("OPENAI_API_KEY");
    string apiUrl = _configuration.GetSection("Embeddings").GetValue<string>("URL");

    // Create a new HttpClient instance
    HttpClient client = new HttpClient();
```

<sup>11</sup> We eliminated all error checking for clarity. Check the OpenSource repository to see the full code.

```

// Set the API key in the Authorization header
client.DefaultRequestHeaders.Authorization =
    new System.Net.Http.Headers.AuthenticationHeaderValue("Bearer", apiKey);

// Set the content type to JSON
client.DefaultRequestHeaders.Accept.Add(
    new System.Net.Http.Headers.MediaTypeWithQualityHeaderValue(
        "application/json"));

// Create the JSON payload
var payload = new
{
    model = _configuration.GetSection("Embeddings").GetValue<string>("Model"),
    input = embedding
};

var jsonPayload = Newtonsoft.Json.JsonConvert.SerializeObject(payload);

// Create the request content
var content = new StringContent(jsonPayload, Encoding.UTF8, "application/json");

// Send the request
var response = client.PostAsync(apiUrl, content).GetAwaiter().GetResult();

// Read the response content as a string
string responseContent = response.Content.ReadAsStringAsync()
    .GetAwaiter().GetResult();

// Deserialize the JSON response into a C# class
EmbeddingResponse embeddingResponse = Newtonsoft.Json.JsonConvert
    .DeserializeObject<EmbeddingResponse>(responseContent);
return embeddingResponse;
}

```

All we are really doing is creating an HTTP Client that has the text of the article in the message payload, and the definition of the model that we want to use within the class of embeddings in OpenAI. This model is defined in the application configuration file:

```

"Embeddings": {
  "URL": "https://api.openai.com/v1/embeddings",
  "Model": "text-embedding-ada-002"
}

```

We are using the `"text-embedding-ada-002"` model. This is an efficient and cheap model that minimizes our operation cost. The API call will return a Jason object that we can deserialize into a numeric vector containing the 1,536 dimensions.

#### 2.3.4 Implementing the ChatGPT Quesiton Answering System

The function called `GetChatGPTAnswerForQuestion` allows us to ask ChatGPT a question and retrieves a reasonable answer, given a context. This code is shown below:

```
public (string answer, List<KnowledgeRecordBasicContent> contextList)
    GetChatGPTAnswerForQuestion(string question, bool briefDetails)
{
    string apiKey = Environment.GetEnvironmentVariable("OPENAI_API_KEY");

    string apiUrl = _configuration.GetSection(
        "Completions").GetValue<string>("URL");

    // Create a new HttpClient instance
    HttpClient client = new HttpClient();

    // Set the API key in the Authorization header
    client.DefaultRequestHeaders.Authorization = new
        System.Net.Http.Headers.AuthenticationHeaderValue("Bearer", apiKey);

    // Set the content type to JSON
    client.DefaultRequestHeaders.Accept.Add(new
        System.Net.Http.Headers.MediaTypeWithQualityHeaderValue(
            "application/json"));

    // Create the JSON payload
    (string prompt, List<KnowledgeRecordBasicContent> contextList) =
        GetPromptForQuestion(question, briefDetails);

    if (String.IsNullOrEmpty(prompt))
    {
        return ("I don't know", new List<KnowledgeRecordBasicContent>());
    }

    var payload = new
    {
        model = _configuration.GetSection(
            "Completions").GetValue<string>("Model"),
        prompt = prompt,
        max_tokens = _configuration.GetSection(
            "Completions").GetValue<int>("MaxTokens"),
        temperature = _configuration.GetSection(
            "Completions").GetValue<int>("Temperature")
    }
```

```

};

var jsonPayload = Newtonsoft.Json.JsonConvert.SerializeObject(payload);

// Create the request content
var content = new StringContent(jsonPayload,
    Encoding.UTF8, "application/json");

// Send the request
var response = client.PostAsync(apiUrl, content).GetAwaiter().GetResult();

// Read the response content as a string
string responseContent = response.Content.ReadAsStringAsync()
    .GetAwaiter().GetResult();

// Deserialize the JSON response into a C# class
CompletionResponse completionResponse = Newtonsoft.Json.JsonConvert
    .DeserializeObject<CompletionResponse>(responseContent);
return (completionResponse.Choices[0].Text, contextList);
}

```

Again, all we are really doing is creating an HTTP Client that has the prompt we will pass to ChatGPT, and the definition of the model that we want to use within the completion engine in OpenAI. Like in the embeddings function, this model is defined in the application configuration file.

The trick here is constructing the prompt. The prompt needs to give precise instructions and these need to include the context and the question. Let's look at the function in line 246 called get prompt for question.

Constructing this prompt is what makes ChatGPT work. Let's look at an example of what a prompt looks like. The text in blue is the context, extracted from our knowledgebase. The text in red is the question that we want an answer to. The text in black are the instructions provided to ChatGPT. If we submit this to ChatGPT, it will answer 60 days:

Answer the question as honestly as possible using the context text provided and if the answer is not included in the text below, just say "I don't know". Avoid repeating the question and provide as much detail as possible.

The Context ends with the string ">>>>>"

Context: <<<<<

In the Laurania Tax Dispute Resolution and Appeals Act (LTDRA), the tax dispute resolution process begins when a taxpayer submits a written request to the Lauranian Revenue Service (LRS) within 60 days of receiving a tax assessment or decision. The LRS then reviews the request and supporting documentation, and issues a written determination within 90 days. If the taxpayer disagrees with the LRS's determination, they may submit a written appeal to the Tax Dispute Resolution Board (TDRB) within 30 days of receiving the determination.

>>>>>

Question: How long do I have to submit a dispute resolution?

Answer:



The code that constructs this prompt is shown below.

```
private (string prompt, List<KnowledgeRecordBasicContent> contextList)
    GetPromptForQuestion(string question, bool briefDetails)
{
    (string context, List<KnowledgeRecordBasicContent> contextList) =
        GetContextForQuestion(question);

    string header = "Answer the question as honestly as possible using the"+
        "context text provided " +
        "and if the answer is not included in the text below, "+
        " just say \"I don't know\".  " +
        "Avoid repeating the question";

    if (briefDetails)
    {
        header = header + " and be as brief as possible.";
    }
    else
    {
        header = header + " and provide as much detail as possible.";
    }

    header = header + "  The Context ends with the string" +
        "\">>>>\"\\r\\n\\r\\nContext: <<<< \\r\\n";

    return (header + context + "\\r\\n>>>>\\r\\n\\r\\n Question: " +
        question + "\\r\\n Answer:", contextList);
}
```

We need to get, from our database, the context by finding those knowledge articles that are similar, conceptually, to our question. We do this in the function called `GetContextForQuestion`.

```
private (string context, List<KnowledgeRecordBasicContent> contextList)
    GetContextForQuestion(string question)
{
    List<Similarity> resultList = GetSimilarityScoreForQuestion(question);

    List<KnowledgeRecordBasicContent> theContextList = new
        List<KnowledgeRecordBasicContent>();

    double similarityScoreThreshold = _configuration.GetSection("Completions")
        .GetValue<double>("MinSimilarityThreshold");

    // Add the number of tokens to each result
    foreach (var item in resultList)
    {
        KnowledgeRecord theRecord = _context.KnowledgeRecords
            .Where(p => p.Id == item.KnowledgeRecordId).AsNoTracking().FirstOrDefault();
        item.Tokens = theRecord.Tokens;
        item.Text = theRecord.Content;

        if(item.SimilarityScore >= similarityScoreThreshold)
        {
            theContextList.Add(new KnowledgeRecordBasicContent()
            {
                ID = item.KnowledgeRecordId,
                Title = theRecord.Title,
                Content = theRecord.Content,
                SimilarityScore = item.SimilarityScore
            });
        }
    }

    int maxSectionLen = _configuration.GetSection("Completions")
        .GetValue<int>("MaxSectionLen");
    string separator = "\n* ";

    List<int> tokens = GPT3Tokenizer.Encode(separator);
    int separatorTokens = tokens.Count;

    int totalLenght = 0;

    StringBuilder context = new StringBuilder();
    foreach (var item in resultList)
```

```

{
    if (context.Length == 0)
    {
        context.Append(item.Text);
    }
    else
    {
        context.Append(separator);
        context.Append(item.Text);
    }

    totalLenght += item.Tokens;
    if (totalLenght > maxSectionLen)
    {
        break;
    }
}

return (context.ToString(), theContextList);
}

```

All this function does is obtain from the database the knowledgebase articles that have the greatest similarity score to our question, and then constructs a string with the contents of each article. The `GetSimilarityScoreForQuestion` function is at the heart of this magic.

```

private List<Similarity> GetSimilarityScoreForQuestion(string question)
{
    double[] questionEmbeddings = GetEmbeddingScores(question);
    string stringVector = string.Join(",", questionEmbeddings);

    // Set up connection string
    string connectionString = _configuration.GetConnectionString("DefaultConnection");

    // Set up SQL command
    string commandText = _configuration.GetSection("Similarity")
        .GetValue<string>("SimilarityScoreSPName");
    SqlCommand command = new SqlCommand(commandText);
    command.CommandType = System.Data.CommandType.StoredProcedure;

    // Add parameters
    SqlParameter csvListParam = new SqlParameter("@csvList",
        System.Data.SqlDbType.VarChar);
    csvListParam.Value = stringVector;
    command.Parameters.Add(csvListParam);
}

```

```

SqlParameter maxResultsParam = new SqlParameter("@maxResults",
    System.Data.SqlDbType.Int);
maxResultsParam.Value = _configuration.GetSection("Similarity")
    .GetValue<int>("MaxResults");
command.Parameters.Add(maxResultsParam);

// Set up connection
SqlConnection connection = new SqlConnection(connectionString);
command.Connection = connection;

// Open connection and execute command
connection.Open();
SqlDataReader reader = command.ExecuteReader();

// Read data into a list of objects
List<Similarity> resultList = new List<Similarity>();
while (reader.Read())
{
    resultList.Add(new Similarity
    {
        KnowledgeRecordId = Convert.ToInt32(reader["id"]),
        SimilarityScore = Convert.ToDouble(reader["similarity"])
    });
}

// Close connection and reader
reader.Close();
connection.Close();

return resultList;
}

```

In this function two things are happening. The first, is that we obtain the embedding vector for our question, convert this vector of 1,536 numbers into a coma separated string. The second is that we pass this vector to a stored procedure that will find knowledge base articles that are similar to the question.

The stored procedure that handles searching for conceptually similar articles is shown below:

```

CREATE PROCEDURE [dbo].[sp_CalculateLargestCosineSimilarities]
    @csvList varchar(max),
    @maxResults int
AS
BEGIN
    -- SET NOCOUNT ON added to prevent extra result sets from

```

```

-- interfering with SELECT statements.
SET NOCOUNT ON;

declare @rec as table(id int)
declare @result as table(id int, similarity float)

insert into @rec
select distinct [Id]
from [dbo].[KnowledgeRecords]

declare @id as int

create table #tempa(row_num int, val float)
create table #tempb(row_num int, val float)
CREATE CLUSTERED INDEX temp_indexa ON #tempa (row_num)
CREATE CLUSTERED INDEX temp_indexb ON #tempb (row_num)

insert into #tempb
SELECT ROW_NUMBER() OVER (ORDER BY (SELECT NULL)) AS row_num,
       CAST(value AS float) AS val
FROM STRING_SPLIT(@csvList, ',')

DECLARE allRecordsCursor CURSOR FOR
SELECT id FROM @rec
OPEN allRecordsCursor

FETCH NEXT FROM allRecordsCursor INTO @id

WHILE @@FETCH_STATUS = 0
BEGIN

    truncate TABLE #tempa

    insert into #tempa
    SELECT ROW_NUMBER() OVER (ORDER BY [Id]) as row_num, [VectorValue] as val
    FROM [dbo].[KnowledgeVectorItems]
    where KnowledgeRecordId = @id

    DECLARE @dotProduct float, @aMagnitude float, @bMagnitude float,
           @similarity float

    -- ChatGPT Embeddings are normalized, so the similarity is just
    -- the dot product of the two vectors. If you are using a
    -- non-normalized embedding, then uncomment these lines
    -- to use cosine similarity.

```

```

--SELECT @dotProduct = SUM(a.val * b.val),
--      @aMagnitude = SQRT(SUM(a.val * a.val)),
--      @bMagnitude = SQRT(SUM(b.val * b.val)),
--      @similarity = @dotProduct / (@aMagnitude * @bMagnitude)
--FROM #tempa a
--JOIN #tempb b ON a.row_num = b.row_num

SELECT @dotProduct = SUM(a.val * b.val),
       @similarity = @dotProduct
FROM #tempa a
JOIN #tempb b ON a.row_num = b.row_num

insert into @result (id, similarity) values(@id, @similarity)

FETCH NEXT FROM allRecordsCursor INTO @id
END

CLOSE allRecordsCursor
DEALLOCATE allRecordsCursor

drop table #tempa
drop table #tempb

select top(@maxResults) id, similarity from @result order by similarity desc

END

```

All that this procedure does is get a list of all articles in the database, and inside a cursor, for each article, obtain a similarity score. We typically use a mathematical function called Cosine Similarity and you can see how we would determine this score in the commented code. Since ChatGPT embeddings normalize the result, the similarity score is given only by the dot product of the question vector and the knowledge article vector. This is a fast operation that is easy to compute.

The stored procedure returns the top n results to the search function, that then displays the answer that ChatGPT gives us.

## 2.4 Concluding Remarks

This is pretty much the extent of the magic behind semantic search using ChatGPT. There is some syntactic sugar that is added to complete the cycle, and we encourage you to download the source code and try the system. Examine the code carefully and you will be amazed how simple it is to build a system that, before large language models existed, would have taken an army of programmers and a very long time to produce.

## Appendix 1 - Knowledge Base Initial Content

Article Title	Article Content
Are education expenses tax-deductible in Laurania?"	<p>In Laurania, certain education expenses may be eligible for tax deductions, depending on the nature of the expenses and the taxpayer's eligibility. Eligible education expenses include tuition fees, books, supplies, and certain fees related to enrollment and attendance at an accredited educational institution. This encompasses primary, secondary, and higher education institutions, as well as vocational training programs. It is important to note that the deductions apply only to the taxpayer, their spouse, or their dependents.</p> <p>To claim education expense deductions in Laurania, taxpayers must meet specific requirements. The student must be enrolled at least part-time in a qualifying educational program. Additionally, the expenses must be directly related to the student's course of study and necessary for the completion of the program. Personal expenses, such as room and board, transportation, or insurance, are not considered eligible for tax deductions. Taxpayers should carefully review the Laurania Tax Administration's guidelines on education expense deductions to ensure they meet all the necessary criteria and to understand any limitations or maximum deduction amounts.</p>
What tax relief options are available for senior citizens in Laurania?	<p>In Laurania, several tax relief options are available for senior citizens to help reduce their tax burden and provide financial support during retirement. These options include an increased personal allowance, deductions for medical expenses, and a reduced property tax rate for qualifying seniors. To be eligible for senior citizen tax relief benefits, individuals must be aged 65 or older and meet certain income and residency requirements set forth by the Laurania Tax Administration.</p> <p>The increased personal allowance for senior citizens provides a higher tax-free threshold, allowing seniors to earn more income before being subject to taxation. Additionally, qualifying seniors can claim deductions for medical expenses exceeding a certain percentage of their adjusted gross income, including prescription medications, hospitalization costs, and long-term care expenses. Furthermore, seniors who own their primary residence may be eligible for a reduced property tax rate, provided they meet specific income and residency criteria. It is essential for senior citizens in Laurania to consult the Laurania Tax Administration's guidelines on tax relief options to ensure they take full advantage of the available benefits and comply with all requirements.</p>
How do I report my cryptocurrency transactions for tax purposes in Laurania?	<p>In Laurania, cryptocurrency transactions are subject to taxation, and taxpayers must report them accordingly on their annual tax returns. Cryptocurrency transactions that need to be reported include trading, mining, staking, airdrops, and any other income-generating activities involving cryptocurrencies. Additionally, capital gains or losses resulting from the sale or exchange of cryptocurrencies must also be reported.</p>

	<p>To report cryptocurrency transactions in Laurania, taxpayers must maintain accurate records of their transactions throughout the year, including the dates, amounts, and fair market values of the transactions in Lauranian currency at the time of occurrence. This information is used to calculate the capital gains or losses and any taxable income derived from the transactions. Taxpayers should consult the Laurania Tax Administration's guidelines on reporting cryptocurrency transactions to ensure they accurately report their transactions and comply with all tax regulations. The guidelines also provide information on any specific forms or schedules that may need to be completed and submitted along with the annual tax return.</p>
Are education expenses tax-deductible in Laurania?	<p>In Laurania, certain education expenses may qualify for tax deductions, depending on the nature of the expenses and the taxpayer's eligibility. Eligible education expenses include tuition fees, books, supplies, and specific fees related to enrollment and attendance at an accredited educational institution. These institutions encompass primary, secondary, and higher education institutions, as well as vocational training programs. It is important to note that the deductions apply only to the taxpayer, their spouse, or their dependents.</p> <p>To claim education expense deductions in Laurania, taxpayers must meet specific requirements. The student must be enrolled at least part-time in a qualifying educational program. Additionally, the expenses must be directly related to the student's course of study and necessary for the completion of the program. Personal expenses, such as room and board, transportation, or insurance, are not considered eligible for tax deductions. Taxpayers should carefully review the Laurania Tax Administration's guidelines on education expense deductions to ensure they meet all the necessary criteria and to understand any limitations or maximum deduction amounts.</p> <p>It is worth mentioning that Laurania also offers tax credits for education expenses, which can directly reduce the amount of tax owed. The eligibility criteria for these credits may differ from those for deductions, and taxpayers may need to choose between claiming a deduction or a credit for the same education expenses. To make an informed decision and maximize their tax savings, taxpayers should familiarize themselves with the Laurania Tax Administration's guidelines on both education expense deductions and tax credits.</p>
What is the process for amending a previously submitted tax return in Laurania?	<p>In Laurania, taxpayers may need to amend a previously submitted tax return if they discover errors, omissions, or changes in their tax situation after the original submission. The process for amending a tax return in Laurania involves submitting a correction request to the Laurania Tax Administration, providing all the necessary documentation to support the changes, and revising the figures on the tax return accordingly.</p> <p>To initiate the amendment process, taxpayers must first obtain the appropriate forms and instructions from the Laurania Tax Administration's website or local tax office. They will need to complete the amendment form, providing details about the changes they wish to make and the</p>



	<p>reasons for the amendments. Taxpayers should attach any supporting documentation, such as revised income statements, receipts, or other relevant records, to substantiate the changes. It is important to submit the amendment request within the specified time frame, typically within three years from the original filing deadline or two years from the date the tax was paid, whichever is later.</p> <p>Once the amendment request is submitted, the Laurania Tax Administration will review the changes and may request additional information or clarification if needed. Taxpayers should be prepared to respond to any inquiries promptly to avoid delays in the processing of their amended return. If the amendment results in a refund, the taxpayer will receive the refund after the amended return is processed. Conversely, if the amendment results in additional tax owed, taxpayers must pay the outstanding amount by the specified due date to avoid penalties and interest. It is crucial for taxpayers to familiarize themselves with the Laurania Tax Administration's guidelines on amending tax returns.</p>
How long should I keep my tax records in Laurania?	<p>In Laurania, it is essential to keep tax records for a minimum of seven years from the date of filing your tax return or the tax return's due date, whichever is later. This requirement ensures that taxpayers maintain accurate and comprehensive records, enabling the Lauranian Tax Administration (LTA) to assess and verify income, deductions, and credits effectively. By retaining your tax records for the specified period, you can provide necessary documentation to the LTA in case of an audit or any discrepancies that may arise in your tax filing history.</p> <p>Tax records you should retain include, but are not limited to, income statements (such as W-2s or 1099 forms), expense receipts, invoices, bank and investment statements, canceled checks, mileage logs, and any other documentation that supports your tax return's income, deductions, and credits. Additionally, keep copies of your filed tax returns, as they serve as a summary of your financial activities during the tax year and can be useful in answering any questions from the LTA or for personal reference.</p> <p>It is important to note that certain circumstances may require you to keep your tax records for longer than the standard seven-year period. For example, if you have claimed a loss from worthless securities or a bad debt deduction, you must keep the relevant records for a period of ten years. If you are unsure whether your situation requires extended record-keeping, consult with a qualified tax professional.</p>
Are non-profit organizations exempt from taxes in Laurania?	<p>In Laurania, non-profit organizations that meet specific criteria are generally exempt from certain taxes. To qualify for tax-exempt status, a non-profit organization must primarily serve a charitable, educational, religious, scientific, or other public interest purpose. Upon meeting these requirements, the organization can apply for tax-exempt status with the Lauranian Tax Administration (LTA). Once granted, the non-profit organization will be exempt from paying income tax on income related to its tax-exempt purpose.</p>

	<p>While non-profit organizations may be exempt from income tax, they are not necessarily exempt from all taxes in Laurania. For instance, these organizations may still be subject to sales tax, property tax, or payroll taxes for their employees. Non-profits must also be cautious when generating income from unrelated business activities, as this income could be subject to taxation. If a non-profit organization conducts substantial unrelated business activities, it may jeopardize its tax-exempt status.</p> <p>It is essential for non-profit organizations in Laurania to maintain accurate financial records and regularly file the necessary forms and reports with the LTA. This includes an annual information return, which details the organization's revenue, expenses, and other financial activities. By remaining compliant with reporting requirements and adhering to the guidelines for tax-exempt organizations, non-profits can continue to benefit from their tax-exempt status and focus on their mission to serve the public interest.</p>
How can I request an extension for filing my tax return in Laurania?	<p>If you find yourself unable to file your tax return in Laurania by the established deadline, you can request an extension to give you additional time to complete and submit your paperwork. The Lauranian Tax Administration (LTA) offers an extension period of up to 90 days, during which you will not be subject to late filing penalties. However, it's important to note that an extension does not grant additional time to pay any taxes due; interest will still accrue on any unpaid tax balance during the extension period.</p> <p>To request an extension, you will need to complete the Form LTA-EXT, "Application for Automatic Extension of Time to File Lauranian Tax Return." This form can be found on the official LTA website or at any local tax office. Make sure to accurately fill out all required fields, including your taxpayer identification number, contact information, and the estimated amount of tax due. Submit the completed form either electronically through the LTA's online portal or by mailing it to the appropriate address listed on the form instructions.</p> <p>Once your Form LTA-EXT is received and processed by the LTA, you will be notified of the approval status of your extension request. If approved, you will have an additional 90 days to file your tax return without incurring late filing penalties. It is crucial to meet this extended deadline, as failing to do so may result in penalties and additional interest charges. Remember, if you owe taxes, it is recommended to pay them as soon as possible to minimize interest accumulation, even if you have been granted an extension for filing your return.</p> <p>If you find that you need more time beyond the initial 90-day extension, you may be eligible for an additional extension under certain circumstances. To request this, you will need to file Form LTA-EXT2, "Application for Additional Extension of Time to File Lauranian Tax Return." Be aware that additional extensions are not automatically granted and</p>

	<p>require a reasonable cause for the request, such as serious illness or natural disaster. The LTA will review your application and notify you of the decision.</p> <p>During the extension period, it's essential to gather all the necessary documentation and information to complete your tax return accurately. If you need assistance, the LTA website offers a range of resources and tools to help taxpayers navigate the process. Additionally, you may consider seeking professional tax advice or utilizing tax preparation software to ensure that your return is filed correctly and on time.</p> <p>In summary, if you need more time to file your Lauranian tax return, you can request an extension using Form LTA-EXT. Keep in mind that extensions only provide additional time to file, not to pay taxes owed. Be diligent in submitting your return within the extension period to avoid penalties, and don't hesitate to seek assistance if needed.</p>
What are the tax implications of receiving an inheritance in Laurania?	<p>In Laurania, when you receive an inheritance, you may be subject to various taxes, depending on the type and value of the assets you inherit. The two main taxes you may face are inheritance tax and capital gains tax.</p> <p>Inheritance tax is a tax on the total value of the assets you inherit, above a certain tax-free threshold, known as the "nil-rate band." As of 2023, the nil-rate band in Laurania is set at LAR 325,000, meaning that any inheritance above this amount is subject to tax at a rate of 40%. However, certain exemptions and reliefs may apply, such as spousal transfers and charitable donations, which can reduce your overall tax liability.</p> <p>Capital gains tax may also apply if you sell the inherited asset at a later date. In Laurania, the tax is calculated based on the difference between the asset's market value at the time of inheritance and its selling price, less any associated costs. The capital gains tax rate is currently set at 20%, but the tax liability may be reduced if you have held the asset for a long period, due to the "step-up in basis" rule. This rule allows the inherited asset's value to be adjusted to its fair market value at the time of the benefactor's death, rather than its original purchase price.</p> <p>In addition to these taxes, you may also face property tax if you inherit real estate property in Laurania. Property tax is assessed annually, based on the property's assessed value and the prevailing tax rate in the jurisdiction where the property is located. The tax rate varies depending on the property's location, but it typically ranges from 1% to 2% of the property's assessed value. As the new owner, it's crucial to ensure timely payment of property taxes to avoid penalties and legal complications.</p>
How do I determine my residency status for tax purposes in Laurania?	<p>In Laurania, an individual's residency status for tax purposes is determined by two primary factors: the number of days spent within the country during the tax year and the individual's personal connections to Laurania. To be considered a tax resident, you must meet one of the following criteria: (1) reside in Laurania for at least 183 days in the tax year, or (2) have substantial personal, economic, and social ties to the country.</p>

	<p>Substantial personal, economic, and social ties to Laurania may include owning a permanent home, having immediate family members residing in the country, holding a long-term or permanent job, and participating in social or cultural activities. If you meet one or more of these criteria, even if you have not spent 183 days in Laurania during the tax year, you may still be considered a tax resident. If you are unsure about whether you meet these criteria, it is advisable to consult with a tax professional or the Lauranian Tax Administration directly for guidance.</p> <p>If you are not considered a tax resident of Laurania but have earned income within the country, you may be classified as a non-resident taxpayer. Non-resident taxpayers are subject to taxation on their Lauranian-sourced income only, while tax residents are taxed on their worldwide income. It is important to accurately determine your residency status to ensure proper tax compliance and avoid potential penalties. If your residency status changes during the tax year, you may be required to file as a part-year resident, which would involve reporting and paying taxes on your Lauranian-sourced income for the portion of the year you were a non-resident, and on your worldwide income for the portion of the year you were a tax resident. Part-year residents should also take note of any applicable tax credits or deductions that may be available to them, as these may differ depending on the specific circumstances.</p> <p>It is essential to maintain accurate records of your time spent in Laurania and the nature of your ties to the country, as this information will be necessary when filing your tax return. You should also be aware of any tax treaties or agreements that may exist between Laurania and your home country, as these can potentially impact your tax obligations and the overall process of determining your residency status.</p>
Are there any tax benefits for environmentally friendly practices in Laurania?	<p>In Laurania, the government is committed to promoting environmentally friendly practices and sustainable development. To encourage individuals and businesses to adopt green initiatives, the tax administration offers a range of tax benefits, credits, and deductions related to environmentally responsible behavior. These incentives aim to support the adoption of clean energy, the reduction of carbon emissions, and overall environmental conservation efforts.</p> <p>For individuals, there are tax credits available for the purchase and installation of energy-efficient equipment or systems in their homes, such as solar panels, geothermal heat pumps, and energy-efficient windows. Additionally, tax benefits may be available for those who drive electric or hybrid vehicles, with credits offered for the purchase of qualifying vehicles or the installation of electric vehicle charging stations. Lauranian taxpayers may also be eligible for deductions if they contribute to approved environmental conservation organizations or participate in tree-planting initiatives.</p> <p>Businesses in Laurania can also benefit from a variety of environmentally-focused tax incentives. These may include credits or deductions for</p>

	<p>investments in energy-efficient equipment, waste reduction, or pollution control measures. In some cases, businesses that engage in research and development of environmentally friendly technologies may qualify for special tax relief programs. Furthermore, the Lauranian government offers preferential tax rates to companies operating in certain green industries, such as renewable energy production, sustainable agriculture, or eco-tourism. By taking advantage of these tax benefits, both individuals and businesses can contribute to a greener, more sustainable future for Laurania while enjoying financial savings.</p>
The Laurania Income Tax Act (LITA)	<p>The Laurania Income Tax Act (LITA) is a comprehensive legislative framework outlining the income tax policies and procedures applicable to both individuals and businesses operating within the jurisdiction of Laurania. The Act establishes the criteria for tax residency, defines the various types of taxable income, and sets forth the progressive tax rates applicable to different income levels. LITA also provides detailed guidance on deductions, exemptions, and tax credits that taxpayers can utilize to reduce their tax liability.</p> <p>LITA mandates that all residents of Laurania, as well as non-residents earning income in the country, are required to file an annual tax return with the Lauranian Tax Administration (LTA). The Act sets forth the filing deadlines and prescribes penalties for late filing, non-compliance, or tax evasion. It also outlines the process for filing amended returns and disputing tax assessments, including the appeals process and the taxpayer's rights during an audit.</p> <p>The Act provides for several categories of income, such as employment income, business income, investment income, and capital gains, each with its own set of rules and tax treatment. LITA includes provisions aimed at preventing tax avoidance, such as transfer pricing rules and anti-avoidance measures for controlled foreign corporations. Furthermore, the Act recognizes the importance of international cooperation in tax matters and sets out provisions for information exchange and collaboration with other tax authorities to ensure a fair and efficient tax system for all Lauranian taxpayers.</p>
Treatment of employment income in the Laurania Income Tax Act (LITA)	<p>The "employment income" section of the Laurania Income Tax Act (LITA) defines the various types of income that are considered as originating from employment or the provision of personal services. This includes, but is not limited to, salaries, wages, bonuses, commissions, tips, and fringe benefits. The Act provides guidance on how to determine the taxable portion of these income sources and the appropriate withholding requirements for employers.</p> <p>The LITA stipulates that certain allowances and reimbursements, such as those covering travel expenses, relocation costs, and work-related tools or equipment, may be excluded from taxable employment income, provided they meet specific criteria set forth in the Act. The section also outlines the tax treatment of non-cash benefits, such as employer-provided housing or</p>

	<p>vehicles, and clarifies the valuation method for determining their taxable value.</p> <p>In order to promote fairness and discourage tax avoidance, the employment income section of the LITA contains provisions aimed at ensuring that compensation arrangements between employers and employees adhere to the arm's length principle. This means that any transactions or compensation agreements between related parties, such as family members or affiliated companies, must be consistent with the terms and conditions that would be expected in a similar arrangement between independent parties. The Act grants the Lauranian Tax Administration the authority to adjust employment income in cases where the arm's length principle has not been observed, potentially resulting in increased tax liabilities for the parties involved. The LITA also provides guidance on the tax treatment of stock options, restricted stock units, and other equity-based compensation plans, with specific rules regarding the timing of income recognition and valuation of the equity awards.</p> <p>In addition to these provisions, the employment income section of the LITA offers information on the deductibility of certain employment-related expenses. These deductions may include, but are not limited to, union or professional dues, expenses for maintaining a home office, and costs associated with required work uniforms. However, the Act specifies that these deductions are only permissible if they are directly related to the taxpayer's employment and have not been reimbursed by the employer.</p>
Treatment of deductibility of certain employment-related expenses in the Laurania Income Tax Act (LITA)	<p>The Laurania Income Tax Act (LITA) outlines specific provisions related to the deductibility of employment-related expenses, which may help reduce a taxpayer's taxable income. To qualify for these deductions, the expenses must be directly related to the taxpayer's employment, incurred for the purpose of earning employment income, and not reimbursed by the employer. Taxpayers must also retain documentation, such as receipts or invoices, as proof of the expenses incurred, in case of an audit or inquiry by the Lauranian Tax Administration.</p> <p>One deductible expense under LITA is the cost of union or professional dues, which taxpayers are required to pay as a condition of their employment. Additionally, expenses associated with maintaining a home office may be deductible, provided that the office space is either the primary place of work or is exclusively used for work purposes and for meeting clients or customers on a regular basis. Eligible home office expenses may include a portion of rent, utilities, property taxes, and home insurance, proportional to the area of the home dedicated to the office space.</p> <p>Furthermore, the LITA permits deductions for the cost of required work uniforms or protective clothing, as long as these items are not suitable for everyday use and are not reimbursed by the employer. Expenses related to work-related travel, such as transportation, accommodations, and meal costs, may also be deductible, provided they are incurred while temporarily</p>

	<p>away from the taxpayer's regular place of work. It is important to note that these deductions do not apply to commuting expenses between an employee's home and their regular workplace. By carefully reviewing the LITA's provisions on employment-related expenses and maintaining accurate records, taxpayers can ensure they are maximizing their deductions and minimizing their tax liabilities.</p>
<p>Tax treatment of employer-provided housing or vehicles in the Laurania Income Tax Act (LITA)</p>	<p>In the Laurania Income Tax Act (LITA), employer-provided housing and vehicles are considered taxable benefits, subject to specific rules and regulations. For employer-provided housing, the taxable value of the benefit is calculated as 10% of the employee's annual salary, with a cap of 40,000 Lauranian Dollars (LD). If the employee contributes to the housing costs, the amount of their contribution is deducted from the calculated taxable value. For example, if an employee with an annual salary of 100,000 LD receives employer-provided housing and contributes 5,000 LD towards the housing costs, the taxable value of the housing benefit would be <math>(10\% \times 100,000) - 5,000 = 5,000</math> LD.</p> <p>Employer-provided vehicles are also considered taxable benefits, with the taxable value determined by the distance driven for personal use and the vehicle's fair market value. Personal use is any travel that is not directly related to business activities. For each kilometer driven for personal use, the employee must pay a rate of 0.20 LD, up to a maximum of 20,000 LD per year. The fair market value of the vehicle is multiplied by a 2% annual rate to determine the taxable value, which is then added to the amount calculated based on the distance driven for personal use. For instance, if an employee uses a company vehicle with a fair market value of 50,000 LD for 10,000 kilometers of personal use, the taxable value would be <math>(10,000 \times 0.20) + (50,000 \times 2\%) = 2,000 + 1,000 = 3,000</math> LD. In this example, the employee would be taxed on a vehicle benefit of 3,000 LD for the year.</p> <p>It is important for employees receiving these benefits to keep accurate records of their personal use of employer-provided vehicles, as well as any contributions made towards employer-provided housing. These records will be necessary for accurately reporting the taxable value of these benefits on their annual Lauranian income tax return. Employers are also required to report the value of these benefits on the employees' T4L tax slips, which are used to report income and deductions for tax purposes. By understanding and adhering to the guidelines set forth in the Laurania Income Tax Act (LITA), both employees and employers can ensure compliance with tax laws and avoid potential penalties.</p>
<p>The Laurania Payroll Tax and Social Security Contributions Law (LPTSS)</p>	<p>The Laurania Payroll Tax and Social Security Contributions Law (LPTSS) governs the collection of taxes and contributions from both employers and employees to fund the country's social security system. Employers are required to withhold payroll taxes from their employees' wages and remit them to the tax authorities on a monthly basis. The LPTSS sets the standard payroll tax rate at 10% of an employee's gross salary, with a maximum taxable wage base of L\$100,000 per year. Employers are also required to make additional social security contributions equal to 8% of the employee's</p>

	<p>gross salary, which covers programs like pension, healthcare, and unemployment benefits.</p> <p>Employees in Laurania are also required to contribute to the social security system through the LPTSS. Employee contributions are withheld from their wages and amount to 5% of their gross salary. These contributions, along with the employer's 8%, are directly allocated to the country's social security programs. An employee's maximum annual contribution is capped at L\$5,000, which corresponds to the 5% contribution rate applied to the L\$100,000 maximum taxable wage base.</p> <p>To ensure compliance with the LPTSS, employers are required to file monthly reports with the Lauranian Tax Administration (LTA), detailing the payroll taxes and social security contributions withheld from their employees' wages. In addition, employers must submit an annual reconciliation report to the LTA, which summarizes the total payroll taxes and social security contributions remitted during the year. Employers who fail to comply with these reporting requirements or remit the appropriate taxes and contributions may face penalties, which can include fines, interest charges, and potential legal actions.</p>
The Laurania Corporate Tax Regulation (LCTR)	<p>The Laurania Corporate Tax Regulation (LCTR) establishes the rules and procedures for the taxation of corporate entities within the country. Under the LCTR, corporations are subject to a flat corporate income tax rate of 20% on their net income, which is calculated as revenues minus allowable deductions and exemptions. The LCTR applies to both domestic corporations, which are taxed on their worldwide income, and foreign corporations with a permanent establishment in Laurania, which are taxed on their Lauranian-source income. Companies are required to file annual tax returns with the Lauranian Tax Administration (LTA), providing a detailed breakdown of their income, deductions, and tax liabilities.</p> <p>The LCTR provides for a range of deductions and exemptions that corporations can use to reduce their taxable income. These include, but are not limited to, ordinary and necessary business expenses, depreciation of tangible assets, research and development costs, and employee wages and benefits. Additionally, the LCTR offers tax incentives to encourage investment in certain industries, such as renewable energy and advanced manufacturing. Companies that qualify for these incentives may benefit from reduced tax rates, tax credits, or accelerated depreciation allowances.</p> <p>To promote transparency and compliance with the LCTR, corporations must maintain accurate financial records and documentation to support their tax filings. The LTA has the authority to audit corporate taxpayers and request additional information or documentation as needed. Corporations found to be in violation of the LCTR, whether through underreporting of income or improper use of deductions and exemptions, may face penalties such as fines, interest charges, and in some cases, criminal prosecution. To avoid these consequences, companies are encouraged to consult with qualified</p>



	tax professionals to ensure they are in full compliance with the Laurania Corporate Tax Regulation.
Deductions and exemptions that corporations can use to reduce their taxable income under the Laurania Corporate Tax Regulation (LCTR)	<p>Under the Laurania Corporate Tax Regulation (LCTR), corporations can take advantage of a variety of deductions and exemptions to minimize their taxable income. One such deduction is for ordinary and necessary business expenses, which include costs directly related to the operation of the business, such as rent, utilities, office supplies, and advertising expenses. To qualify as an ordinary and necessary business expense, the cost must be both common and accepted in the taxpayer's industry and directly contribute to the generation of income. It's important for companies to maintain detailed records and documentation to support these deductions in the event of an audit.</p> <p>Another key deduction under the LCTR is the depreciation of tangible assets, which allows corporations to recover the cost of capital investments in property, plant, and equipment over their useful lives. Laurania uses the Modified Accelerated Cost Recovery System (MACRS) to determine the depreciation schedule for various asset classes, which generally range from 3 to 30 years. For example, a company that purchases a piece of machinery worth L\$50,000 with a 7-year MACRS class life would be able to claim a depreciation deduction of L\$7,143 per year for seven years. This helps to spread the cost of the investment over its useful life, thereby reducing the company's annual taxable income.</p> <p>The LCTR also offers deductions for research and development (R&amp;D) costs, which aim to incentivize innovation and technological advancement. Companies engaged in qualified R&amp;D activities, such as the development of new products, processes, or software, may deduct these costs in the year they are incurred. Additionally, the LCTR provides a tax credit for R&amp;D expenditures, which allows companies to claim a credit of 20% of their eligible R&amp;D costs, reducing their tax liability directly. Lastly, corporations may deduct employee wages, salaries, and benefits as an expense, provided they are reasonable and necessary for the operation of the business. This includes not only base wages and salaries but also bonuses, commissions, and other forms of compensation tied to the employee's work.</p>
Standards that corporations must maintain accurate financial records and documentation to support their tax filings under the Laurania Corporate Tax Regulation (LCTR)	Under the Laurania Corporate Tax Regulation (LCTR), corporations are required to maintain accurate financial records and documentation that support their tax filings. This obligation is critical for ensuring transparency and compliance with Laurania's tax laws. Accurate record-keeping is essential for substantiating income, deductions, exemptions, and credits claimed on tax returns. To meet this requirement, companies should develop and maintain a comprehensive accounting system that captures all financial transactions and enables the preparation of accurate financial statements, including balance sheets, income statements, and cash flow statements.

	<p>In addition to maintaining a robust accounting system, corporations must retain documentation that supports the figures reported on their tax returns. This includes invoices, receipts, expense reports, contracts, and any other records that substantiate the company's reported income and expenses. For example, if a corporation claims a L\$10,000 deduction for office supplies, it must retain receipts or other documentation that verifies the purchase and use of the supplies for business purposes. Similarly, if a corporation is claiming a tax credit for research and development costs, it should maintain detailed records of the R&amp;D activities, the expenses incurred, and the staff involved in the project.</p> <p>The LCTR mandates that corporations retain their financial records and supporting documentation for a minimum of six years from the filing date of the tax return or the due date of the return, whichever is later. This retention period allows the Laurantian Tax Administration (LTA) to audit corporate taxpayers and verify the accuracy of their reported income, deductions, and credits. Failure to maintain proper records and documentation can result in the disallowance of claimed deductions and credits, leading to additional tax liability, interest charges, and potential penalties. In more severe cases of noncompliance, corporations may face criminal prosecution. To avoid these consequences, companies are encouraged to work with qualified tax professionals and implement a robust record-keeping system that adheres to the LCTR's requirements.</p>
<p>Tax Administration's Guide to Taxation of Rental Income</p>	<p>The Laurantian Tax Administration (LTA) requires taxpayers to report rental income on their annual tax returns, as outlined in the "Tax Administration's Guide to Taxation of Rental Income." Rental income is defined as the total amount received from leasing or subleasing property, including payments made in the form of goods or services. Taxpayers must include all rental income received during the tax year, regardless of whether the property is residential or commercial. Taxpayers are also required to report any security deposits that are not returned to the tenant, but can exclude rental income that is received as a refundable security deposit.</p> <p>Taxpayers are allowed to deduct certain expenses related to their rental properties, such as mortgage interest, property taxes, insurance, maintenance, repairs, and depreciation. To calculate depreciation, taxpayers must determine the property's cost basis, which is typically the purchase price plus any improvements made to the property. Depreciation is calculated using the Modified Accelerated Cost Recovery System (MACRS) over a 27.5-year period for residential properties and a 39-year period for commercial properties. Taxpayers must allocate their expenses between personal use and rental use if they are using the property for both purposes. The LTA allows a deduction of up to 20% of the net rental income for eligible taxpayers as a Qualified Business Income (QBI) deduction.</p> <p>In Laurantia, taxpayers may fall under one of three tax brackets for rental income: 10%, 20%, and 30%. Taxpayers with a taxable income of up to L\$50,000 are subject to a 10% tax rate on rental income. Those with a taxable income between L\$50,001 and L\$150,000 are subject to a 20% tax</p>

	<p>rate. Finally, taxpayers with a taxable income exceeding L\$150,000 must pay a 30% tax rate on their rental income. To ensure compliance with the tax laws, landlords are required to provide the LTA with a statement of rental income earned and taxes withheld for each tenant. Failure to report rental income or underreporting income can result in penalties, including fines and imprisonment.</p>
About Laurania	<p>Laurania is a fictional country created for the purpose of examples and case studies in various fields, such as tax administration, economics, legal systems, and public administration.</p> <p>Laurania serves as a neutral and imaginary territory, allowing professionals and educators to develop and discuss hypothetical scenarios without referring to any specific real-world country or jurisdiction. In the context of tax administration, Laurania may have its own tax laws, regulations, and policies that can be used to illustrate concepts or explore issues related to taxation, compliance, enforcement, and international tax treaties.</p> <p>By using Laurania as an example, professionals can avoid potential biases and focus on the principles and best practices that can be applied universally.</p> <p>Please note that any details about Laurania's geography, political structure, or economic system are entirely fictitious and are created for educational purposes only.</p>
Details of the country of Laurania	<p>Laurania is a fictional country created for the purpose of examples and case studies in various fields, such as tax administration, economics, legal systems, and public administration.</p> <p><b>Geography:</b></p> <p>Laurania is located on a large island with a diverse landscape that includes coastal areas, mountains, plains, and forests. Its climate ranges from Mediterranean along the coast to temperate inland, with some high-altitude regions experiencing alpine conditions.</p> <p><b>Currency:</b></p> <p>The currency of Laurania is the Lauranian Dollar (LAD), which is used for all transactions within the country. The Central Bank of Laurania manages the nation's monetary policy and regulates its financial institutions.</p> <p><b>Population:</b></p> <p>Laurania has a population of approximately 20 million people. The population is diverse, with a mix of ethnic groups and cultures that contribute to its rich heritage. The official language of Laurania is Lauranian, but many citizens are bilingual or multilingual, speaking various other languages.</p> <p><b>Economy:</b></p> <p>Laurania has a mixed economy, with key industries including agriculture, manufacturing, technology, and tourism. The country has a well-developed</p>

	<p>infrastructure, including a network of roads, railways, and airports, which facilitates domestic and international trade. The government of Laurania has a progressive taxation system that supports public services such as healthcare, education, and social welfare programs.</p> <p>Political Structure:</p> <p>Laurania is a democratic republic with a parliamentary system. The country is divided into several administrative regions, each governed by a local council. The president serves as the head of state and is elected every five years, while the prime minister, who is the head of government, is chosen by the parliament.</p> <p>Education:</p> <p>Laurania has a strong emphasis on education, with a high literacy rate and a well-funded public education system. The country is home to several prestigious universities, which attract students from around the world.</p> <p>Culture:</p> <p>The culture of Laurania is a blend of various influences, reflecting the diverse population of the country. Traditional arts, music, and dance are celebrated and preserved, while contemporary culture is also thriving, with a vibrant arts scene and a growing film industry.</p>
Tax Dispute Resolution Process	<p>In the Laurania Tax Dispute Resolution and Appeals Act (LTdra), the tax dispute resolution process begins when a taxpayer submits a written request to the Lauranian Revenue Service (LRS) within 60 days of receiving a tax assessment or decision. The LRS then reviews the request and supporting documentation, and issues a written determination within 90 days. If the taxpayer disagrees with the LRS's determination, they may submit a written appeal to the Tax Dispute Resolution Board (TDRB) within 30 days of receiving the determination.</p>
The Taxpayer Advocate	<p>The Taxpayer Advocate is an independent office within the Lauranian Revenue Service (LRS) that assists taxpayers in resolving issues with the LRS, identifies systemic problems affecting taxpayer rights, and proposes solutions to improve the tax system. Reporting directly to the President of Laurania and appointed for a five-year term, the Taxpayer Advocate can be contacted by taxpayers for assistance in resolving tax disputes, especially when they have exhausted other administrative remedies or face significant financial hardship.</p>
The Tax Dispute Resolution Board	<p>The Tax Dispute Resolution Board (TDRB) is an independent entity responsible for reviewing tax appeals and issuing final decisions on tax disputes. Comprising five members appointed by the President of Laurania for five-year terms and confirmed by the Senate, the TDRB members possess expertise in tax law, accounting, and/or economics, and are required to maintain no conflicts of interest with taxpayers or the tax administration.</p>