

```
= Considerations
```

```
:order: 2
```

```
:type: lesson
```

```
:slides: true
```

```
[ .slide.discrete]
```

```
== Caution
```

While GenAI and LLMs provide a lot of potential, you should also be cautious.

At their core, LLMs are highly complex predictive text machines. LLMs don't know or understand the information they output; they simply predict the next word in a sequence.

The words are based on the patterns and relationships from other text in the training data.

```
[ .slide]
```

```
== Access to Data
```

The sources for this training data are often the internet, books, and other publicly available text.

The data could be of questionable quality or even incorrect.

Training happens at a point in time, the data is static, and may not reflect the current state of the world or include any private information.

```
[ .slide.discrete]
```

```
== Access to Data
```

When prompted to provide a response, relating to new or data not in the training set, the LLM may provide a response that is not accurate.

```
image::images/llm-missing-data.svg["A diagram of an LLM returning out of data information."]
```

```
[ .slide.col-2]
```

```
== Accuracy
```

```
[ .col]
```

```
=====
```

LLMs are designed to create human-like text and are often fine-tuned to be as helpful as possible, even if that means occasionally generating misleading or baseless content, a phenomenon known as **hallucination**.

For example, when asked to "Describe the moon." an LLM may respond with "The moon is made of cheese.".

While this is a common saying, it is not true.

```
=====
```

```
[ .col]
```

```
image::images/confused-llm.svg["A diagram of a confused LLM with a
```

question mark thinking about the moon and cheese."]

[.transcript-only]

=====

While LLMs can represent the essence of words and phrases, they don't possess a genuine understanding or ethical judgment of the content.

=====

[.slide]

== Temperature

LLMs have a temperature, corresponding to the amount of randomness the underlying model should use when generating the text.

The higher the temperature value, the more random the generated result will become, and the more likely the response will contain false statements.

A higher temperature may be appropriate when configuring an LLM to respond with more diverse and creative outputs, a lower temperature required when responses should be consistent and precise.

How temperature affects the output is dependent on the model due to differences in design and training data.

[TIP]

. Consider the correct temperature

=====

In June 2023,

[link:https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/](https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/) [A US judge sanctioned two US lawyers for submitting an LLM-generated legal brief[^]] that contained six fictitious case citations.

=====

A quick fix may be to reduce the temperature. But more likely, the LLM is hallucinating because it hasn't got the information required.

[.slide.col-2]

== Transparency

[.col]

=====

GenAI Models are often considered "black boxes" due to the difficulty deciphering their decision-making processes.

The LLM is also unable to provide the sources for its output or explain its reasoning.

=====

[.col]

image::images/llm-blackbox.svg["An LLM as a black box, responding to the question 'How did you determine that answer?' with 'I don't know.'"]

```
[ .quiz]
== Check Your Understanding

include::questions/1-hallucination.adoc[leveloffset=+1]
```

```
[ .summary]
== Lesson Summary
```

In this lesson, you learned about the limitations of GenAI models, including hallucination, and access to data, and how these factors can lead to outputs that might be biased, devoid of context, or lack logical coherence.

In the next lesson, you will learn about how to provide context to GenAI models to improve the accuracy of their responses.