



Saying Please to The Machine

The Art of Prompt Engineering



%PRESENTATION()

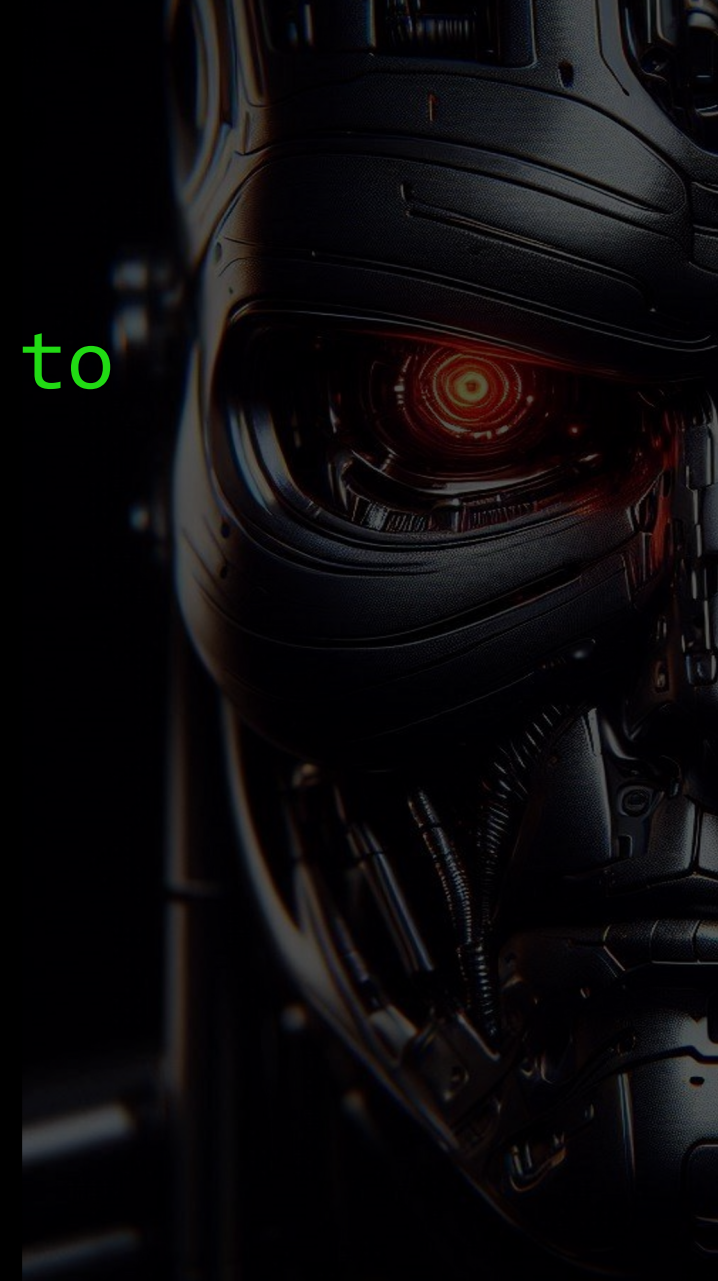
...

Macro variable PRESENTATION resolves to
skynet.exe

...

executing skynet.exe

...



Generating Disclaimer...

...

...



The views and opinions expressed in this presentation are solely those of the presenters and do not necessarily reflect the official policy or position of any associated organization.

I'm sorry, I didn't get that

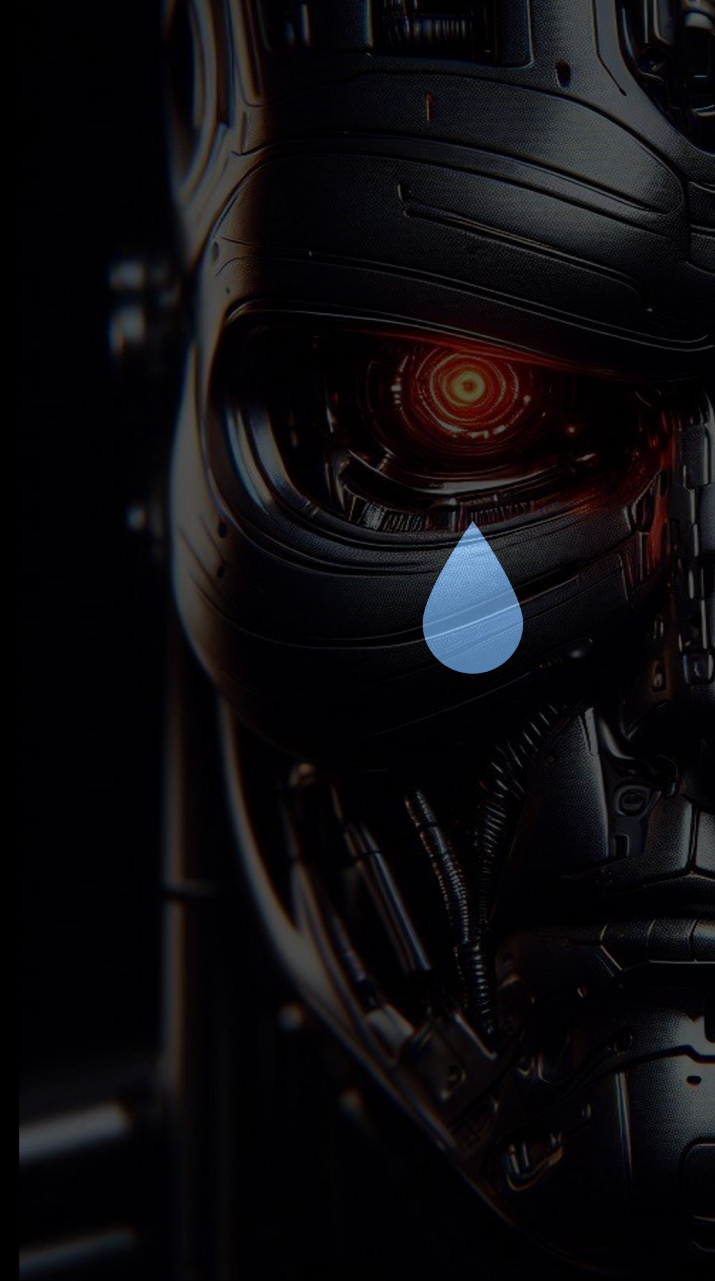
...

I'm not sure I understand

...

Abort execution skynet.exe

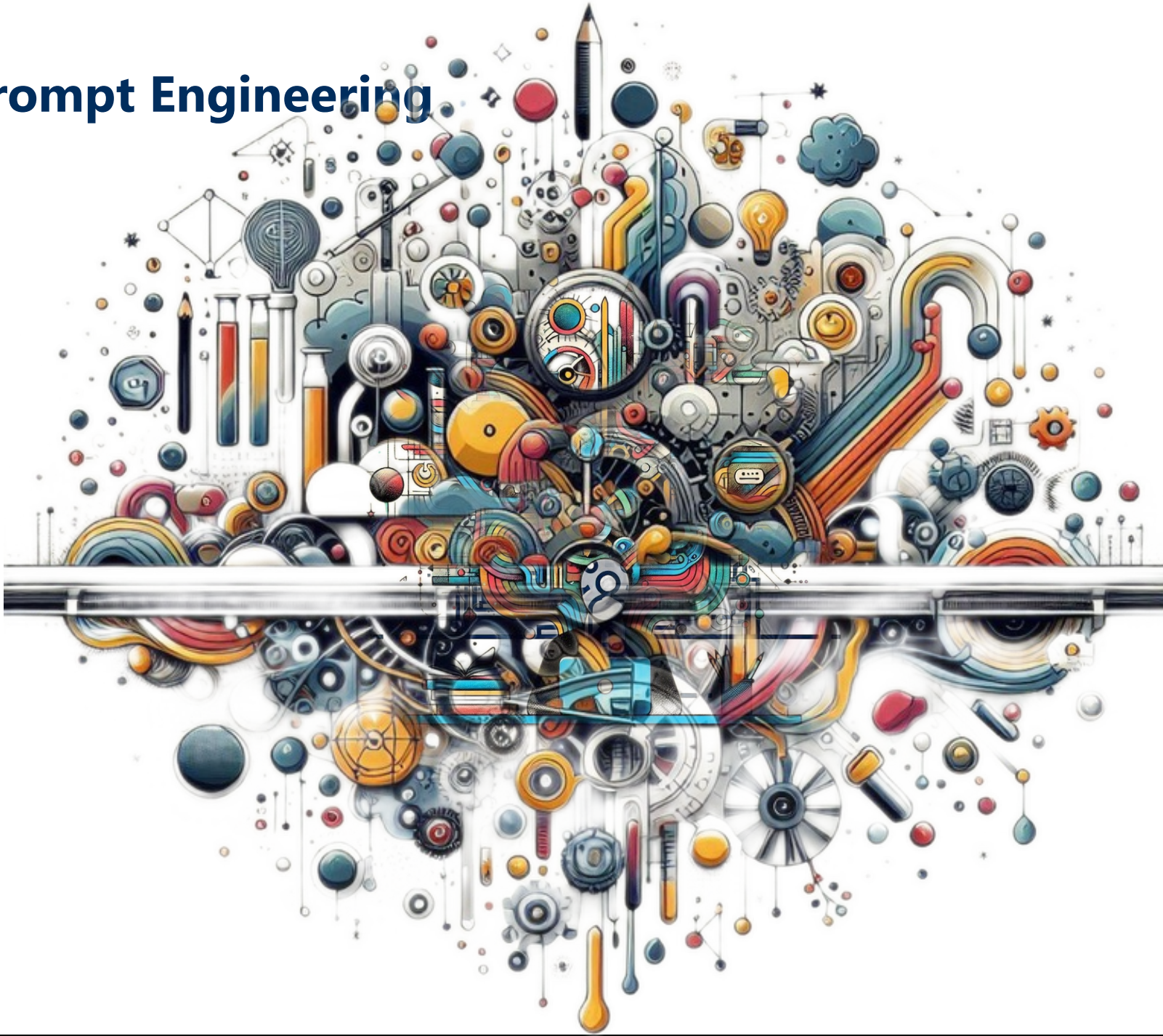
...



What is Prompt Engineering



What is Prompt Engineering



Why we care about Prompt Engineering



Importance for General Users



Enhances productivity



Minimized input, maximized output

Improved decision-making



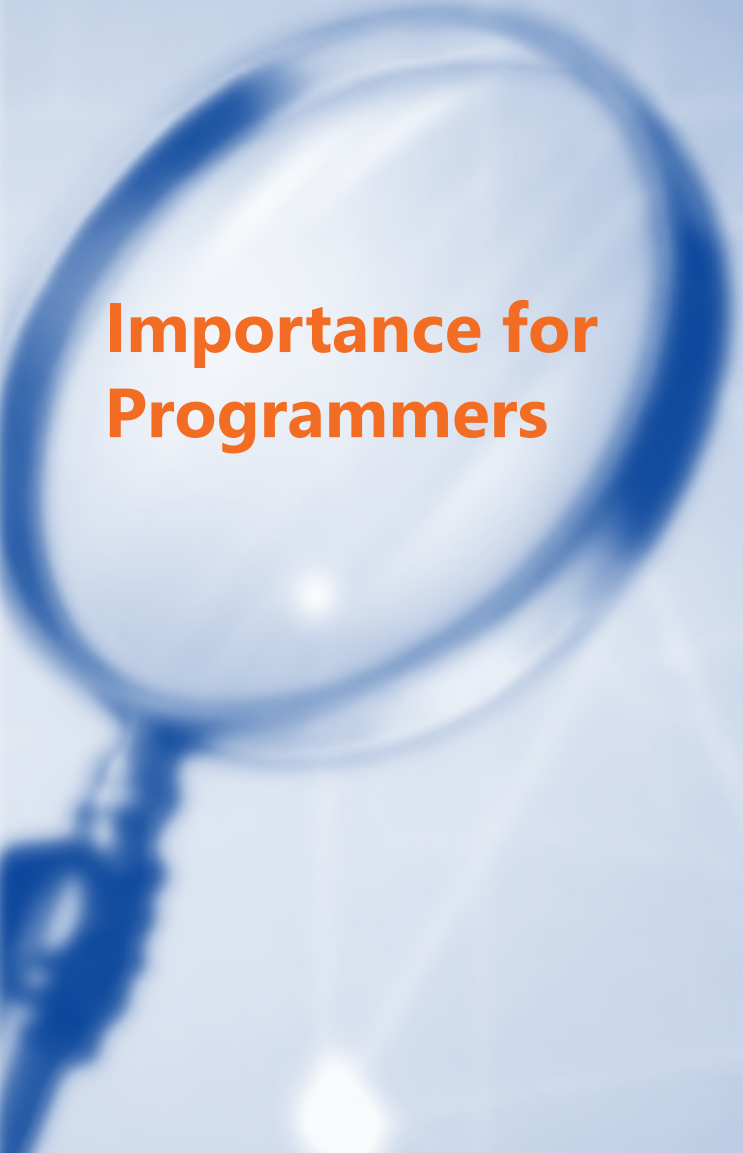
Quicker access to information

Bias Mitigation



Importance for Programmers





Importance for Programmers

Optimized code development

Algorithm development

Code conversion (e.g.: SAS to R)

Ethical considerations

Reduced Human-AI Bias

Problem-solving



How LLMs Reads Inputs

Write a text

asking a friend to be my plus-one at a wedding

Help me pick

an outfit that will look good on camera

Tell me a fun fact

about the Roman Empire

Design a programming game

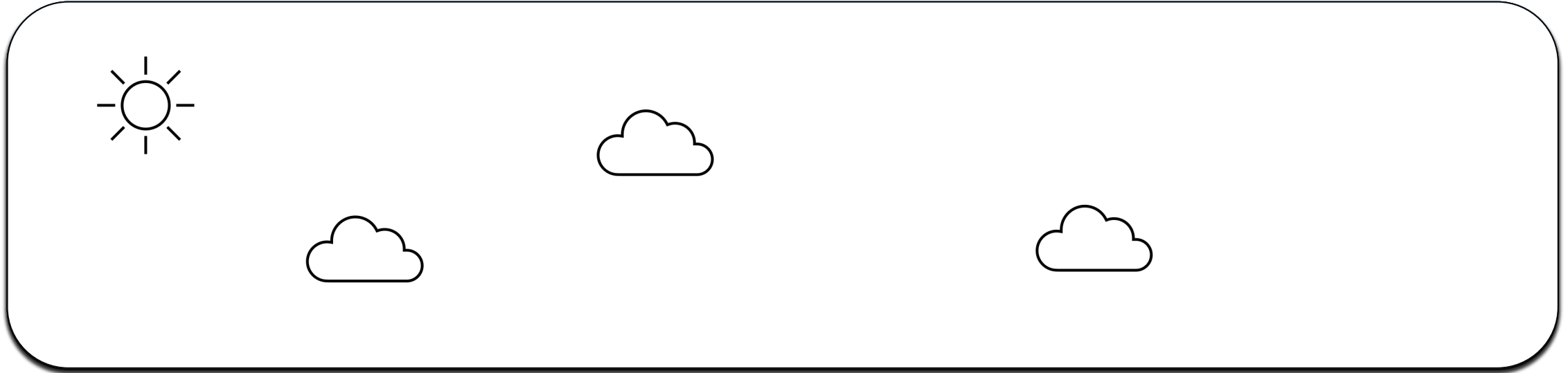
teach basics in a fun way

What colour is the sky?



By messaging ChatGPT, you agree to our [Terms](#) and have read our [Privacy Policy](#).

How LLMs Reads Inputs



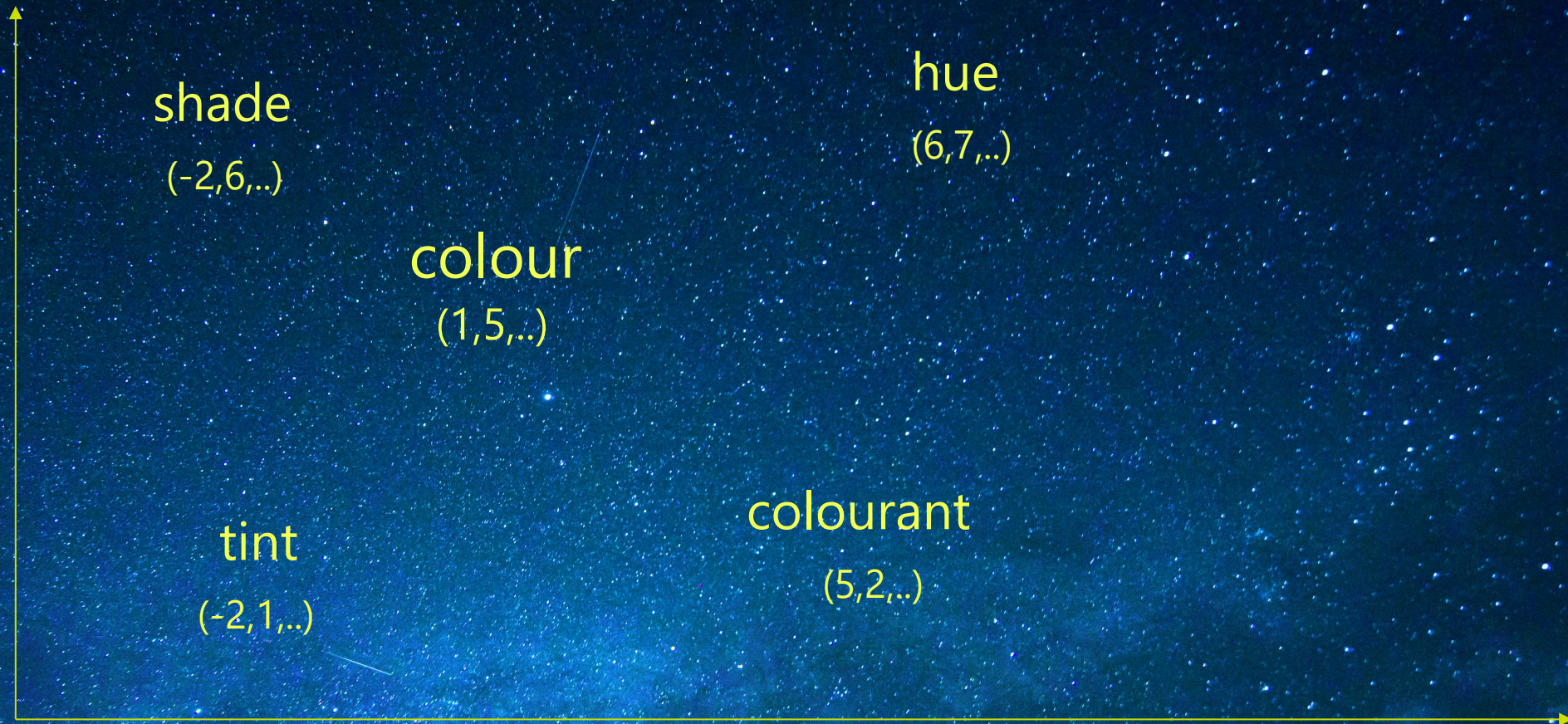
What colour is the sky?

How LLMs Reads Inputs



What colour is the sky ?
(1,4,...) (1,5,...) (6,1,...) (3,8,...) (9,3,...) (0,1,...)

How LLMs Reads Inputs



How LLMs Reads Inputs



How LLMs Read Inputs

What colour is the sky?



The color of the sky can vary depending on factors like time of day and weather conditions. During the day, when the sun is out, the sky typically appears blue due to the scattering of sunlight by the Earth's atmosphere. At sunrise and sunset, the sky can range from shades of red and orange to pink and purple, again due to scattering and the angle of sunlight passing through more atmosphere. At night, when the sun is below the horizon, the sky appears dark blue or black, with stars visible in clear conditions.



How LLMs Read Inputs

In one word, what colour is the sky?



Blue



Prompt Engineering Principals

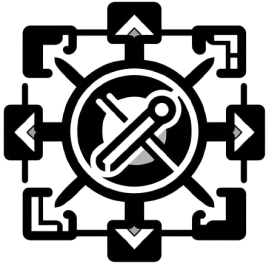
Constraints

Feedback Loop

Clarity and Specificity

Prompt Length

Context Establishment



Prompt Engineering Principals

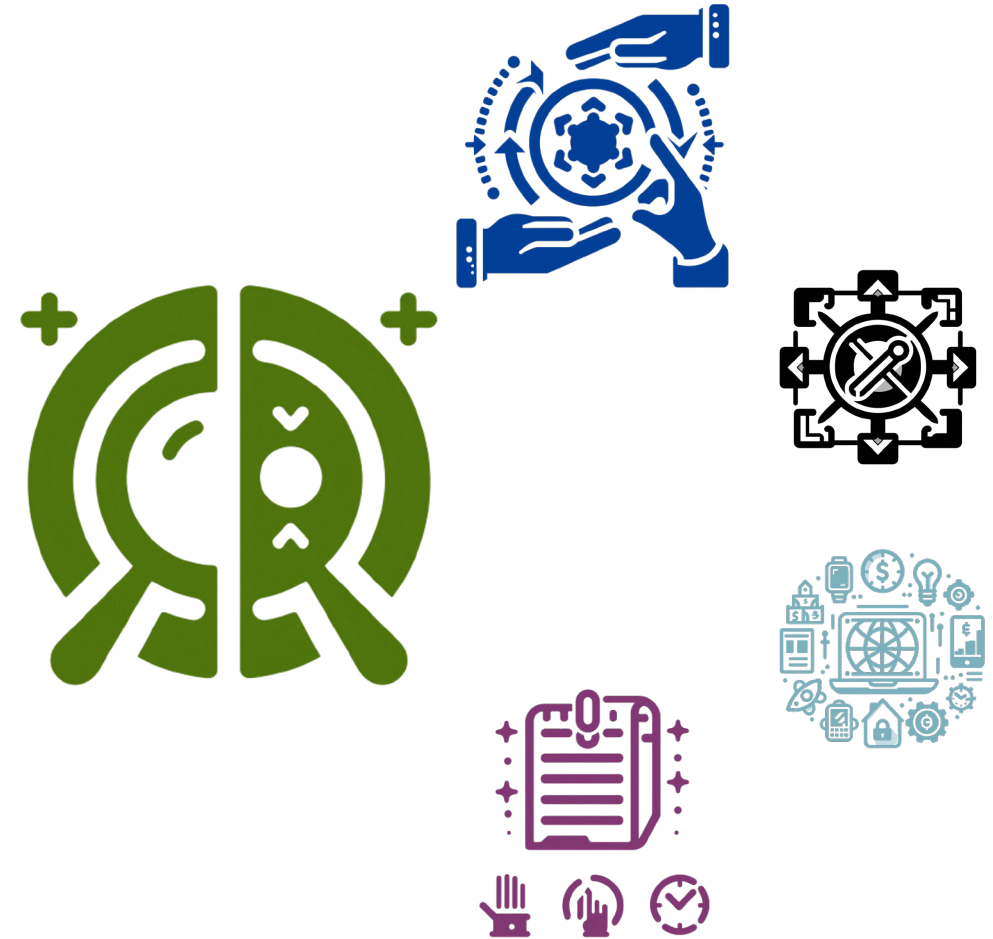
Constraints

Feedback Loop

Clarity and Specificity

Prompt Length

Context Establishment



Prompt Engineering Principals

Feedback Loop

Clarity and Specificity

Prompt Length

Context Establishment

Constraints



Prompt Engineering Principals

Clarity and Specificity

Prompt Length

Context Establishment

Constraints

Feedback Loop



Prompt Engineering Principals

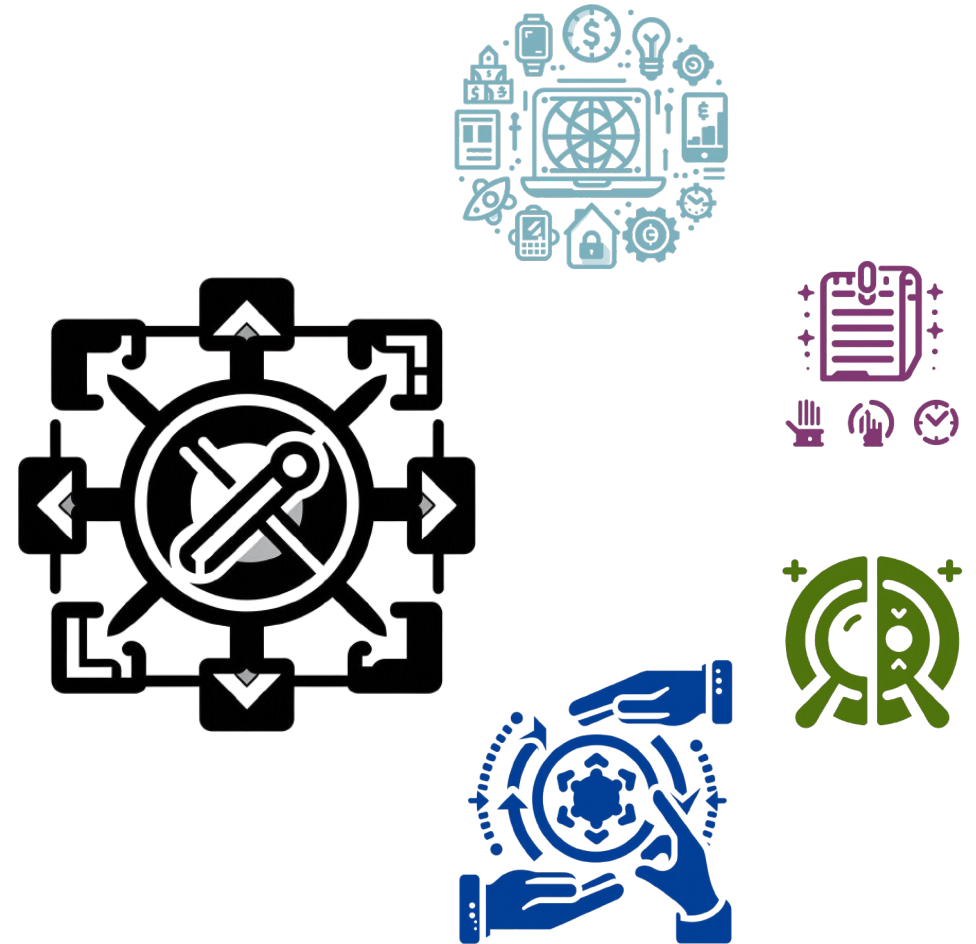
Prompt Length

Context Establishment

Constraints

Feedback Loop

Clarity and Specificity



Prompt Engineering Principals

Context Establishment

Constraints

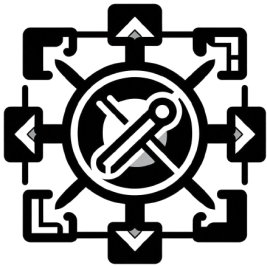
Feedback Loop

Clarity and Specificity

Prompt Length



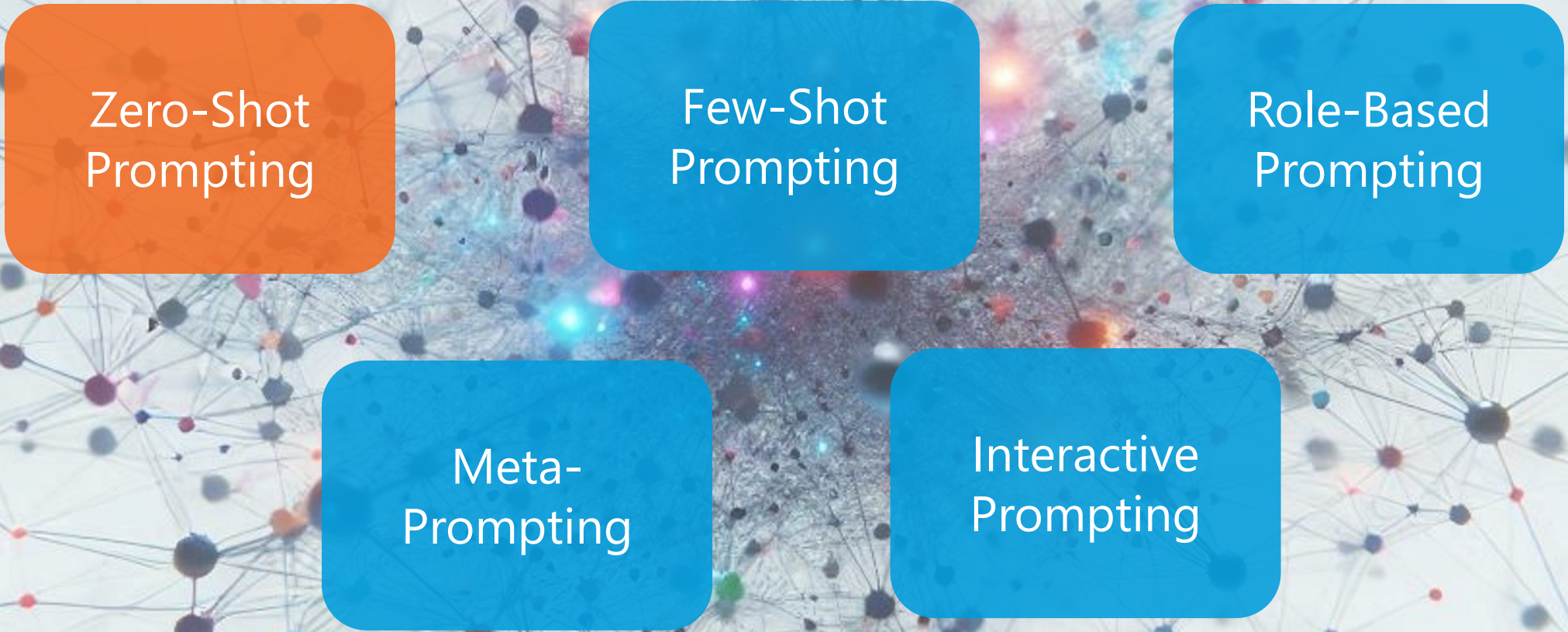
Prompt Engineering Principals



The Formula

1.Task	[Mandatory]
2.Context	[Important]
3.Exemplar	[Important]
4.Persona	[Nice]
5.Format	[to]
6.Tone	[Haves]

Prompting Strategies



Zero-Shot Prompting

Few-Shot Prompting

Role-Based Prompting

Meta-Prompting

Interactive Prompting

Prompting Strategies



Efficiency

Zero-Shot Prompting



Dependency on Pre-training

Zero-Shot Prompting Example

Generate SAS code to create summary statistic for Subject AGE.



Certainly! Here's an example of SAS code that calculates summary statistics for a variable ``AGE`` in a dataset named ``YourDataset``:

Zero-Shot Prompting Example

```
/* Assuming the dataset is named YourDataset */  
/* Replace YourDataset and AGE with your actual  
dataset and variable names */
```

```
/* Sort the dataset by AGE if necessary */  
proc sort data=YourDataset;  
    by AGE;  
run;
```


Zero-Shot Prompting Example

```
/* Calculate summary statistics */  
proc univariate data= YourDataset noprint;  
  var AGE;  
  output out=SummaryStats  
    mean=Mean_AGE  
    median=Median_AGE  
    q1=Q1_AGE  
    q3=Q3_AGE  
    min=Min_AGE  
    max=Max_AGE  
    n=N;  
  
run;
```

Zero-Shot Prompting Example

```
/* Print the summary statistics */  
proc print data=SummaryStats;  
    title 'Summary Statistics for AGE';  
run;
```


Zero-Shot Prompting Example

Explanation:

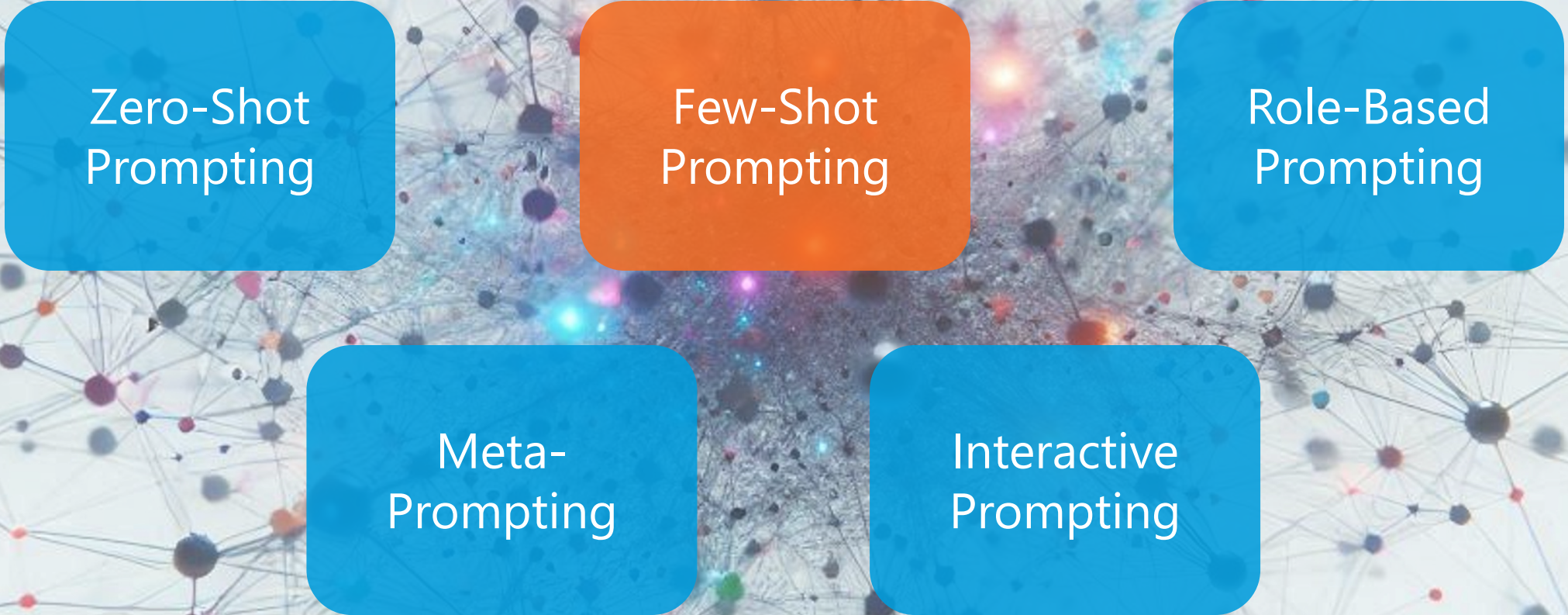
1. **proc sort:** Sorts the dataset by the variable ``AGE``. Sorting is optional but can be useful when reviewing summary statistics.
2. **proc univariate:** Computes univariate statistics (mean, median, quartiles, min, max, and count) for the variable ``AGE``.
 - ``var` AGE;``: Specifies that the variable ``AGE`` is the one for which statistics are to be calculated.
 - ``output out=SummaryStats``: Stores the calculated statistics in a dataset named ``SummaryStats``.
 - ``mean=Mean_AGE`, `median=Median_AGE`, etc.:` Specifies the names of variables in ``SummaryStats`` where the respective statistics will be stored.
3. **proc print:** Prints the summary statistics dataset (``SummaryStats``) to the SAS output window.

Notes:

- Replace ``YourDataset`` with the actual name of your dataset.
- Replace ``AGE`` with the actual variable name for which you want to calculate summary statistics.
- Adjust the ``title`` statement in ``proc print`` to suit your preference for the output title.

This code will provide you with summary statistics including mean, median, quartiles, minimum, maximum, and count for the variable ``AGE`` in your dataset.

Prompting Strategies



Zero-Shot Prompting

Few-Shot Prompting

Role-Based Prompting

Meta-Prompting

Interactive Prompting

Prompting Strategies



Improved Performance

Few-Shot Prompting



Complexity

Few-Shot Prompting Example

Example

This is how I calculate summary statistics for AGE.

```
proc means data=ADSL noprint;  
  by GROUP;  
  var AGE;  
  output out=STATS_AGE n=n_age mean=mean_age std=sd_age  
    median=median_age min=min_age max=max_age;  
run;
```

Prompt

Calculate summary statistics for "WEIGHTBL".



To calculate summary statistics for the variable "WEIGHTBL" using SAS, you can use the `PROC MEANS` procedure similar to how you calculated summary statistics for "AGE". Here's how you can do it:

Few-Shot Prompting Example

```
proc means data=ADSL noprint;  
  var WEIGHTBL;  
  output out=STATS_ WEIGHTBL  
  n=n_weightbl  
  mean=mean_weightbl  
  std=sd_weightbl  
  median=median_weightbl  
  min=min_weightbl  
  max=max_weightbl;  
run;
```

Few-Shot Prompting Example

Let's break down what each part of this SAS code does:

1. **PROC MEANS:** This is the SAS procedure used for calculating summary statistics such as mean, standard deviation, median, minimum, and maximum values.
2. **data=ADSL:** Specifies the dataset ``ADSL`` from which SAS will calculate the summary statistics.
3. **noprint:** Prevents the procedure from printing the output to the results window, which is useful when you want to store the results in another dataset (like ``STATS_WEIGHTBL``).
4. **var WEIGHTBL:** Specifies the variable ``WEIGHTBL`` for which you want to calculate the summary statistics.
5. **output out=STATS_WEIGHTBL:** This statement saves the computed statistics into a new dataset named ``STATS_WEIGHTBL``.
6. **n=n_weightbl mean=mean_weightbl std=sd_weightbl median=median_weightbl min=min_weightbl max=max_weightbl:** These are the options used with the ``OUTPUT`` statement to specify the names of the variables in the output dataset (``STATS_WEIGHTBL``) where SAS will store the count (``n_weightbl``), mean (``mean_weightbl``), standard deviation (``sd_weightbl``), median (``median_weightbl``), minimum (``min_weightbl``), and maximum (``max_weightbl``) values of ``WEIGHTBL``.

After running this SAS code, you will have a dataset named ``STATS_WEIGHTBL`` that contains the summary statistics for the variable ``WEIGHTBL``. Adjust the names (``STATS_WEIGHTBL``, ``n_weightbl``, ``mean_weightbl``, etc.) as per your preference or naming conventions.

Prompting Strategies

Zero-Shot
Prompting

Few-Shot
Prompting

Role-Based
Prompting

Meta-
Prompting

Interactive
Prompting

Prompting Strategies



Specialized Responses

Role-Based Prompting



Dependency on Role Definition

Role-Based Prompting Example

You are a **statistician** on a clinical trial, and I am a **programmer**. Explain to me null hypothesis in short.



The null hypothesis is a statement that there is no effect or no difference between groups in a clinical trial. It's essentially the default assumption that any observed results are due to random chance rather than a real effect.

For example, if you're testing a new drug, the null hypothesis might be that the drug has no effect compared to a placebo. The goal of the trial is to collect data and see if there's enough evidence to reject this null hypothesis in favor of an alternative hypothesis, which suggests that the drug does have an effect.

Prompting Strategies

Zero-Shot
Prompting

Few-Shot
Prompting

Role-Based
Prompting

Meta-
Prompting

Interactive
Prompting

Prompting Strategies



Reflection

Meta-Prompting



Model Complexity

Meta-Prompting Example

In short what is the best way to visualize survival analysis?



The best way to visualize survival analysis is typically with a Kaplan-Meier curve. This plot shows the probability of survival over time for one or more groups, with the survival function plotted on the y-axis and time on the x-axis. It helps to illustrate differences in survival rates between groups and identify any patterns or trends. For comparing multiple groups, you can overlay multiple Kaplan-Meier curves on the same plot.



Meta-Prompting Example

In short why is that the best way.



The Kaplan-Meier curve is effective because it clearly displays survival probabilities over time and accounts for censored data (i.e., individuals who drop out or are lost to follow-up). It allows for easy comparison of survival experiences between different groups and visually represents changes in survival rates, making it a straightforward and informative tool for survival analysis.

Meta-Prompting Example

In short why do we care about censored data?



We care about censored data because it provides partial information about the survival experience of individuals who haven't yet experienced the event of interest by the end of the study. Properly accounting for censored data ensures more accurate and reliable survival estimates, as it includes these individuals in the analysis rather than ignoring their contribution up to the point they were last observed.

Prompting Strategies

Zero-Shot
Prompting

Few-Shot
Prompting

Role-Based
Prompting

Meta-
Prompting

Interactive
Prompting

Prompting Strategies



Adaptability

Interactive Prompting



Time-Consuming

Interactive Prompting Example



You

Generate an image of lighting.



Interactive Prompting Example



You

Replace the tree with a cell tower



Interactive Prompting Example



You

Change the background to a rural setting.



Interactive Prompting Example



You

Add lighting striking the cell tower.



Conclusion

How to Safely Integrate LLMs Into Health Care

Scott Gottlieb, MD¹; Lauren Silvis, JD²

» [Author Affiliations](#) | [Article Information](#)

JAMA Health Forum. 2023;4(9):e233909.

Regulators Face Novel Challenges as Artificial Intelligence Tools Enter Medical Practice

Scott Gottlieb, MD¹; Lauren Silvis, JD²

» [Author Affiliations](#) | [Article Information](#)

JAMA Health Forum. 2023;4(6):e232300. doi:10.1001/jamahealthforum.2023.2300

Artificial Intelligence (AI/ML)-Enabled Medical Devices

BERTox Initiative

Initiative to apply LLMs to facilitate analysis of FDA documents and public literature for improved efficiency and accuracy.

M Dissecting racial bias in an algorithm used to manage the health of populations

AI-Generated Medical Advice—GPT and Beyond

Claudia E. Haupt, JSD, PhD^{1,2}; Mason Marks, MD, JD^{3,4}

[Authors Info & Affiliations](#)

Conclusion

GPP

Data Privacy

Dual Coding

Human Oversight

Conclusion

"AI is not going to replace you. You're going to be replaced by someone who uses AI to outperform you."

Laurence Liew

References

- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. and Zhou, D. (2022). Chain of Thought Prompting Elicits Reasoning in Large Language Models. *arXiv:2201.11903 [cs]*. [online] Available at: <https://arxiv.org/abs/2201.11903>.
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E. and Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models*. [online] Available at: <https://arxiv.org/pdf/2302.13971>.
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J. and Amodei, D. (2020). Scaling Laws for Neural Language Models. *arXiv:2001.08361 [cs, stat]*. [online] Available at: <https://arxiv.org/abs/2001.08361>.
- www.promptingguide.ai. (n.d.). *Prompt Engineering Guide – Nextra*. [online] Available at: <https://www.promptingguide.ai/>.
- Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C. and Hesse, C. (2020). Language Models Are Few-Shot Learners. *arxiv.org*, [online] 4. Available at: <https://arxiv.org/abs/2005.14165>.
- Wang, X., Wei, J., Schuurmans, D., Le, Q., Chi, E., Narang, S., Chowdhery, A. and Zhou, D. (2022). Self-Consistency Improves Chain of Thought Reasoning in Language Models. *arXiv:2203.11171 [cs]*. [online] Available at: <https://arxiv.org/abs/2203.11171>.
- Liu, J., Liu, A., Lu, X., Welleck, S., Peter, Le Bras, R., Choi, Y. and Allen, P. (n.d.). *Generated Knowledge Prompting for Commonsense Reasoning Hannaneh Hajishirzi*. [online] Available at: <https://arxiv.org/pdf/2110.08387> [Accessed 23 Jul. 2024].



Burger Lubbe
blubbe@mmsholdings.com



Vangeli Grammatikos
vgrammatikos@mmsholdings.com

