Quoth the Raven

Phase 1 Update:
Natural Language Generation

Peter Yonka / Data Scientist
January 27, 2021

The Prompt (overall)

As a social media team managing the brand presence for a diverse set of literary creative talent, how can we better leverage any current creative assets to engage existing and potential fan bases to augment and expand the brand of our clients?

.

Top Three Goals (or Why we're here)

Drive Web Traffic

We've invested heavily in unified online presences to engage potential fans via newsletter signup, embedded social media stream/highlights, merch, and latest news.

How can we drive more traffic?

Leverage Client IP

Our literary clients have bodies of existing work (from novels, to essay to screenplays).

How can we use this to our advantage?

Every Account a VIP

We have limited/shifting account managers for many of our clients.

How can we maintain the authenticity for each client's presence?

Quoth the Raven application

Proof of concept app.

Component pieces directly or indirectly answer these key questions
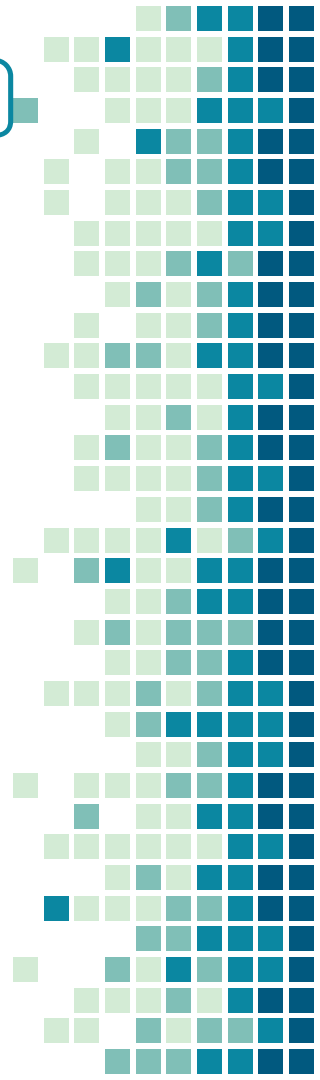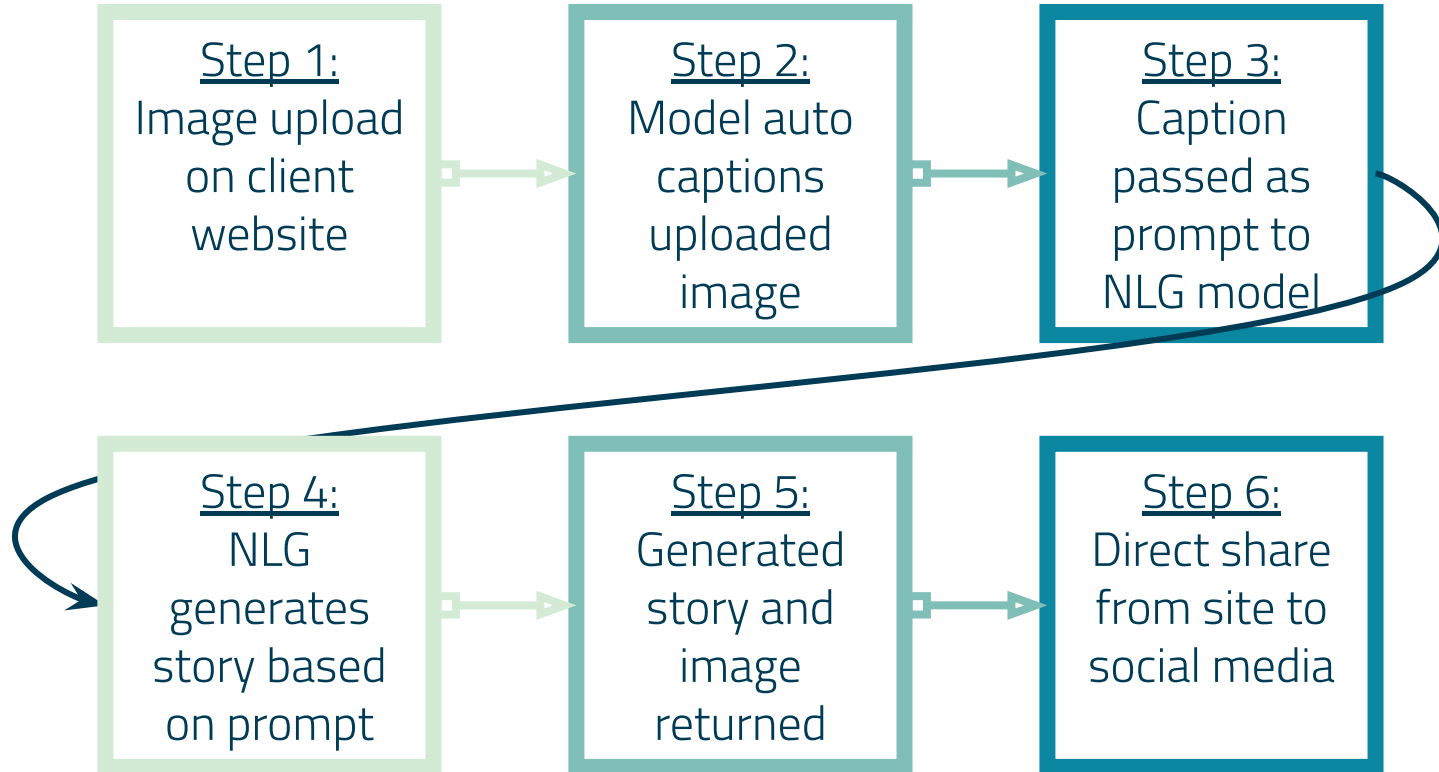
The Prompt (for app)

If a picture paints a thousand words, what words would an AI version of Edgar Allan Poe paint based on a user supplied picture?

Utilizing neural networks and the complete works of Edgar Allan Poe, we'll find out!
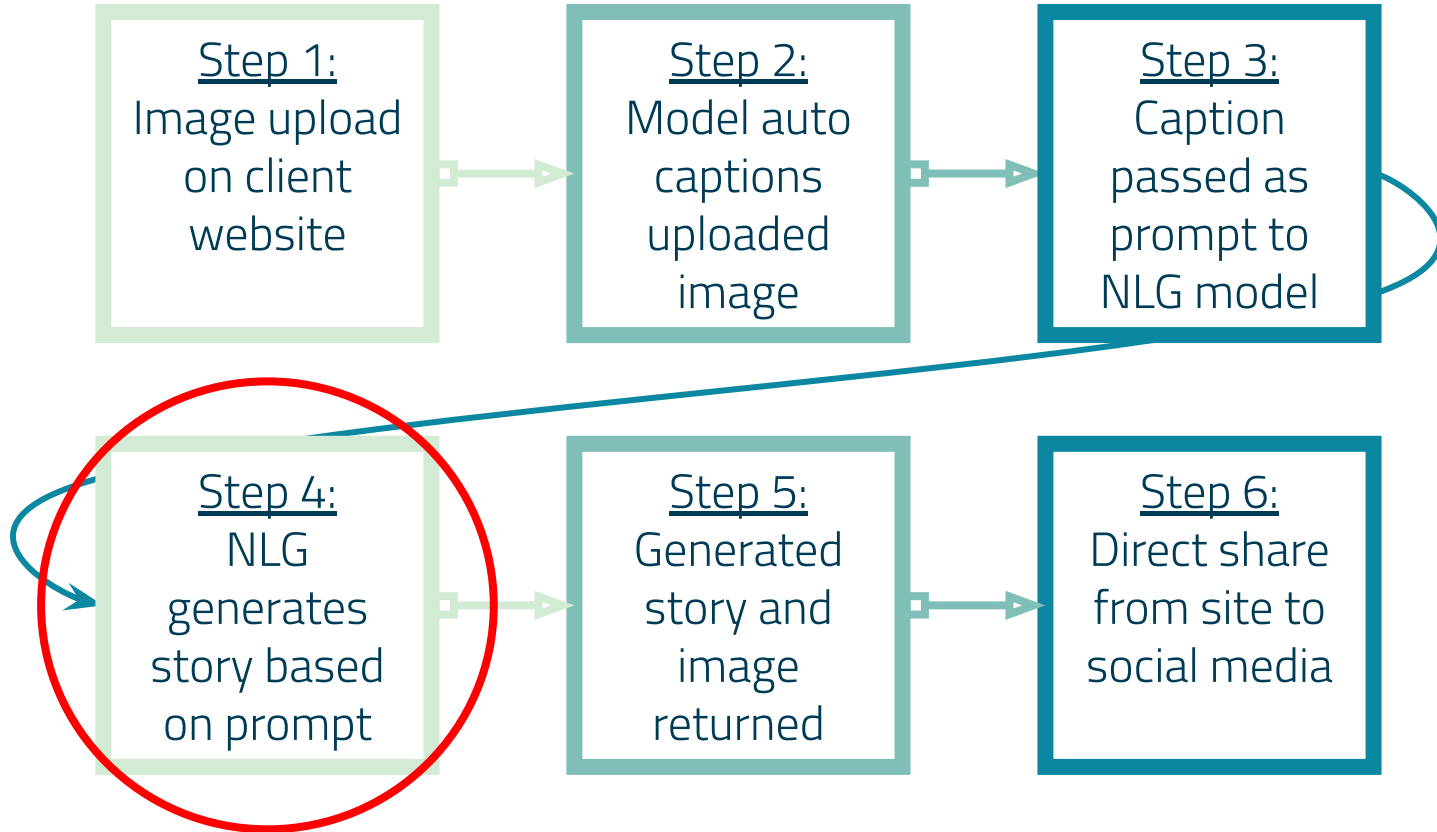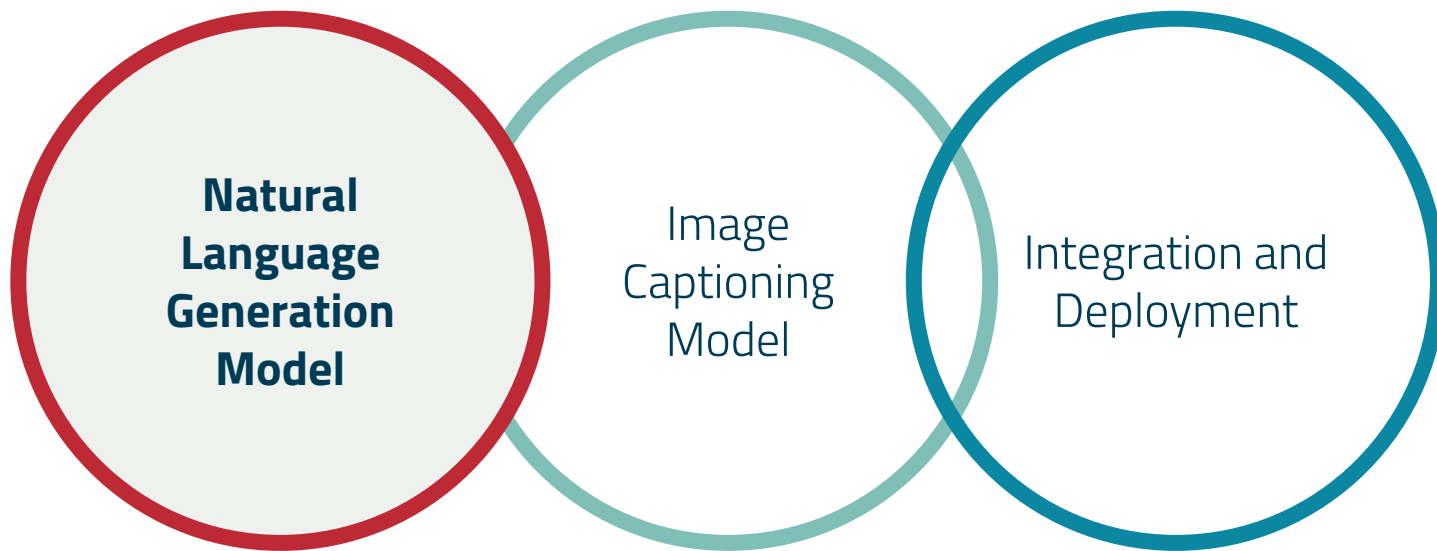
QTR: Application Overview

Step 1:
Image upload on client website

Step 2:
Model auto captions uploaded image

Step 3:
Caption passed as prompt to NLG model

Step 4:
NLG generates story based on prompt

Step 5:
Generated story and image returned

Step 6:
Direct share from site to social media

QTR: Application Overview

Step 1:
Image upload on client website

Step 2:
Model auto captions uploaded image

Step 3:
Caption passed as prompt to NLG model

Step 4:
NLG generates story based on prompt

Step 5:
Generated story and image returned

Step 6:
Direct share from site to social media

QTR: Project Build Phases

Natural Language Generation Model

Image Captioning Model

Integration and Deployment

General Modeling Process (a brief overview)

Gather Data

Prepare Data

Create and Train Model

Evaluate Results

General Modeling Process (a brief overview)

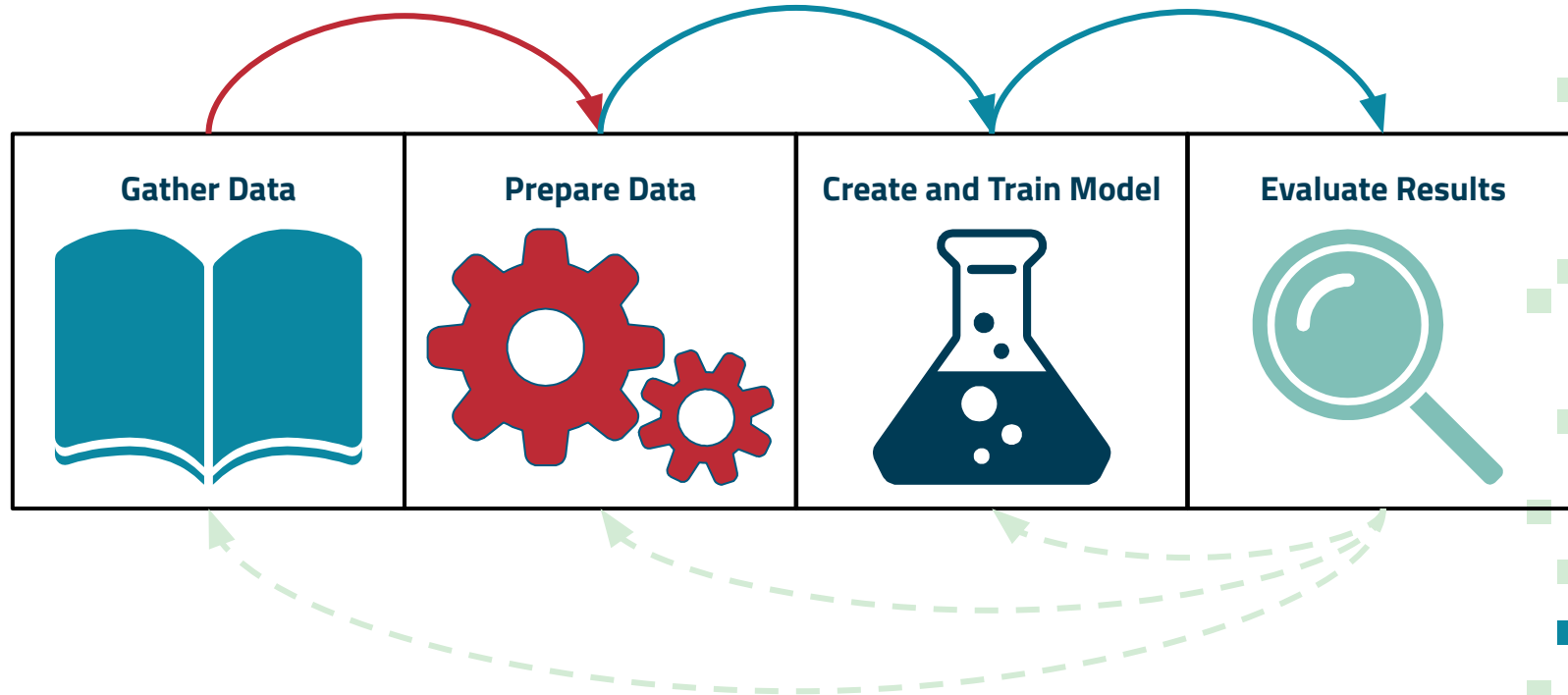| Gather Data | Prepare Data | Create and Train Model | Evaluate Results |
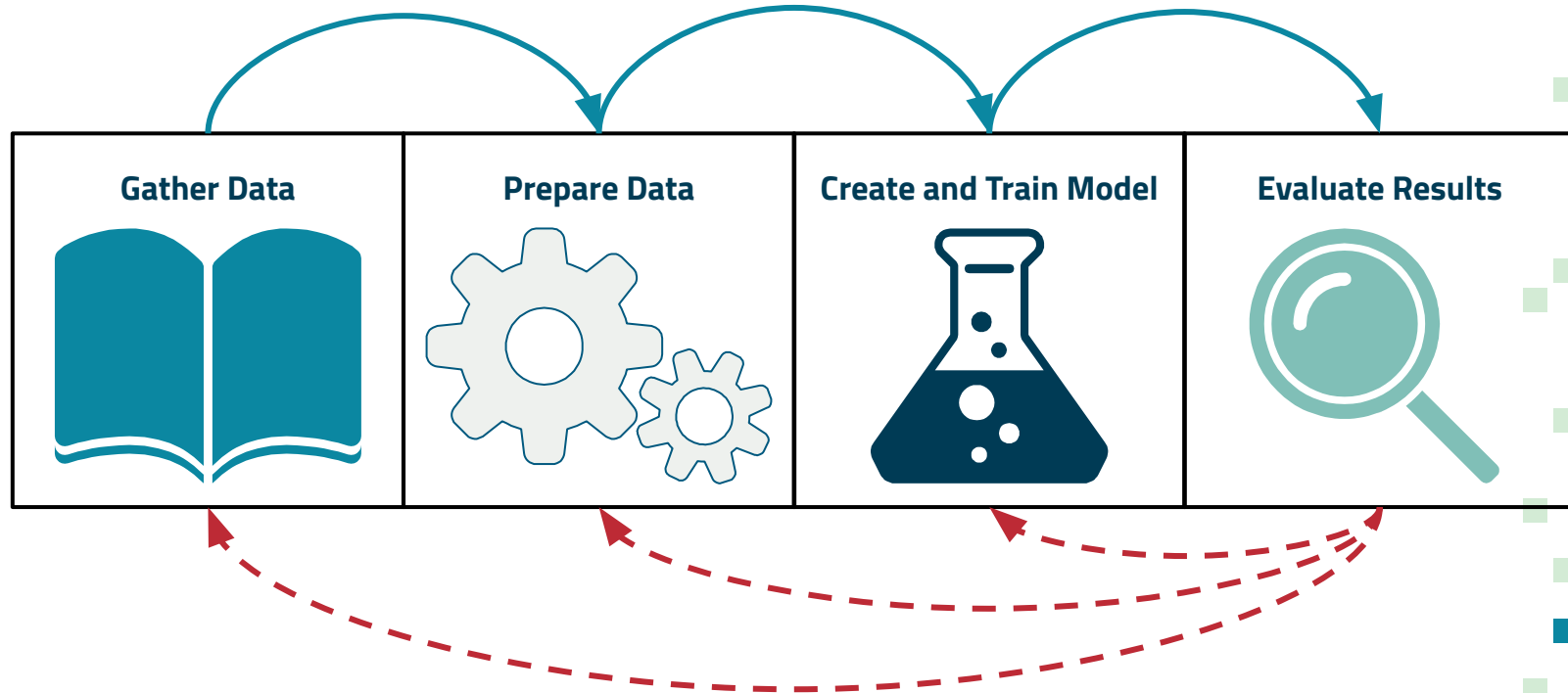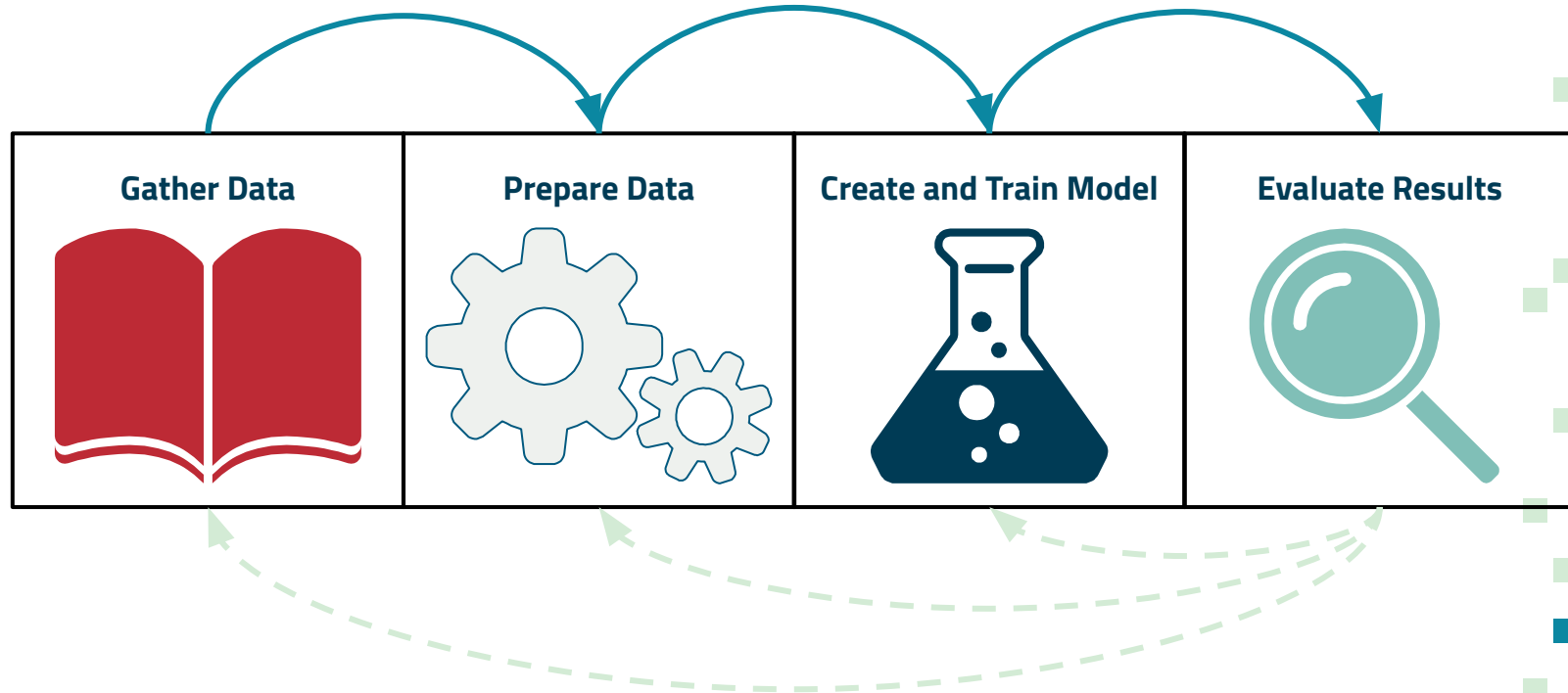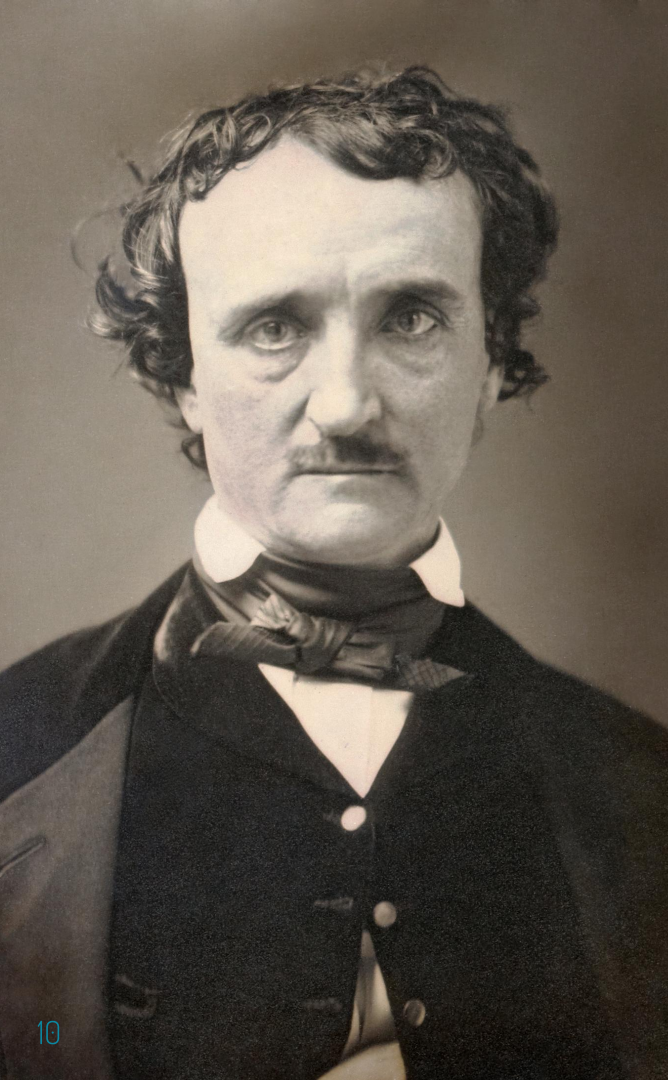
Why Edgar Allan Poe?

- Personal affinity

- Enough content to create a draft model

- Works are accessible and in the public domain

- Distinctive writing style

  - Allows us to see if model is actually generating based on our text (especially true for GPT-2 model)
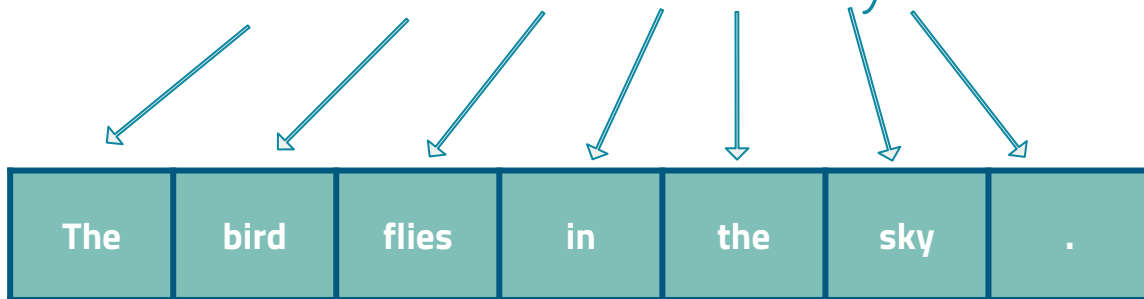
Source: wikipedia.org

General Modeling Process (a brief overview)

Gather Data	Prepare Data	Create and Train Model	Evaluate Results

Word-based Text Models

Example:

The bird flies in the sky.

The	bird	flies	in	the	sky	.

A few reasons for this approach...

Each "word" becomes a token

Attempts to allow the model to understand word relationships

Attempts to allow the model to understand punctuation and capitalization

Word–based Text Models

Sequences are fed into the model...

Sequence number	Sequence of 6 tokens fed into network						Model learns to predict
1	The	bird	flies	in	the	sky	.
2	bird	flies	in	the	sky	.	
3	flies	in	the	sky	.		It

These continues on and on until the entire text is read. Then the whole process is repeated again and again for as many epochs as selected.

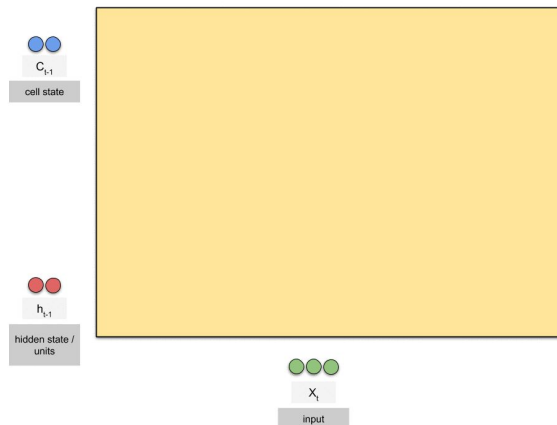# General Modeling Process (a brief overview)

Gather Data

Prepare Data

Create and Train Model

Evaluate Results
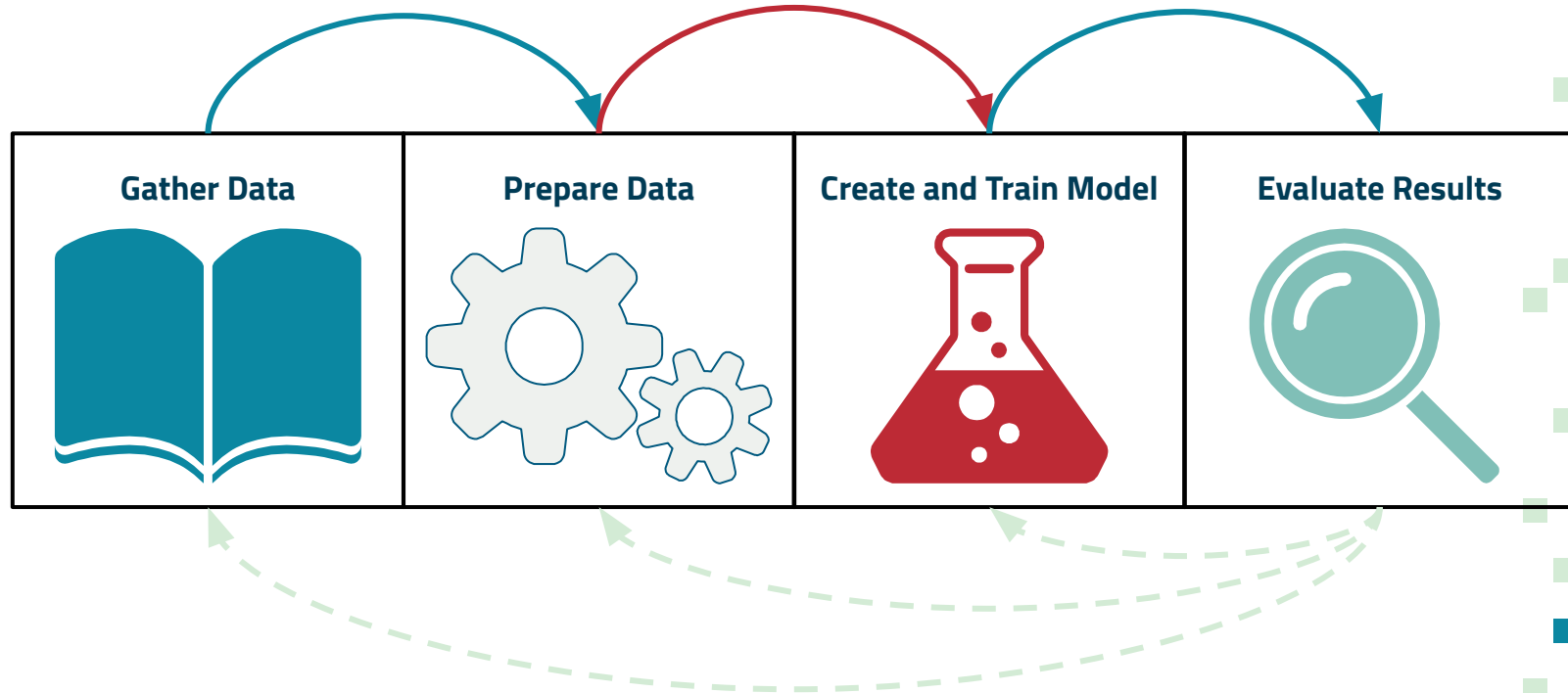
Model 1: Long Short-Term Memory Neural Network

- Type of recurrent neural network

- Up until a few a years ago, one of the best options for these applications

- Improvement over its predecessors in remembering connections over longer periods of time

- Building from scratch, so take a long time to train for any good results
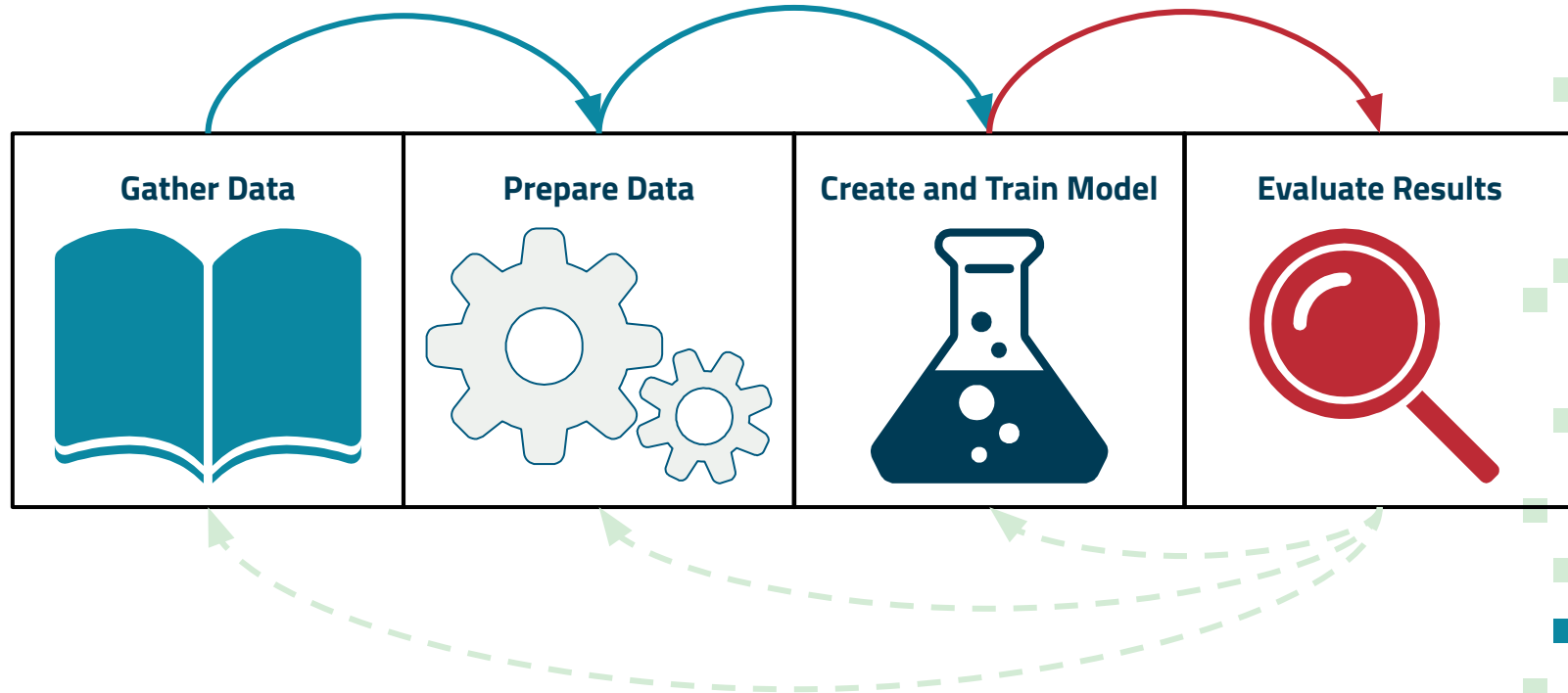
- Notoriously difficult/finicky to finetune

```
Model: "sequential"
_____
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, 25, 128)           2875776
_____
lstm (LSTM)                  (None, 25, 256)           394240
_____
dropout (Dropout)            (None, 25, 256)           0
_____
lstm_1 (LSTM)                (None, 256)               525312
_____
dropout_1 (Dropout)          (None, 256)               0
_____
dense (Dense)                (None, 256)               65792
_____
dense_1 (Dense)              (None, 22467)             5774019
=================================================================
Total params: 9,635,139
Trainable params: 9,635,139
Non-trainable params: 0
_____
```

C_{t-1}
cell state

h_{t-1}
hidden state / units

X_t
input

Source: Raimi Karim, "Animated RNN, LSTM and GRU"

15

General Modeling Process (a brief overview)

Gather Data → **Prepare Data** → **Create and Train Model** → **Evaluate Results**
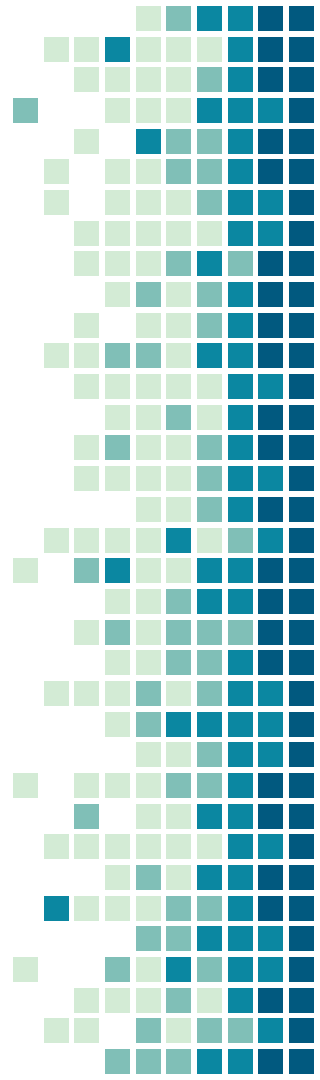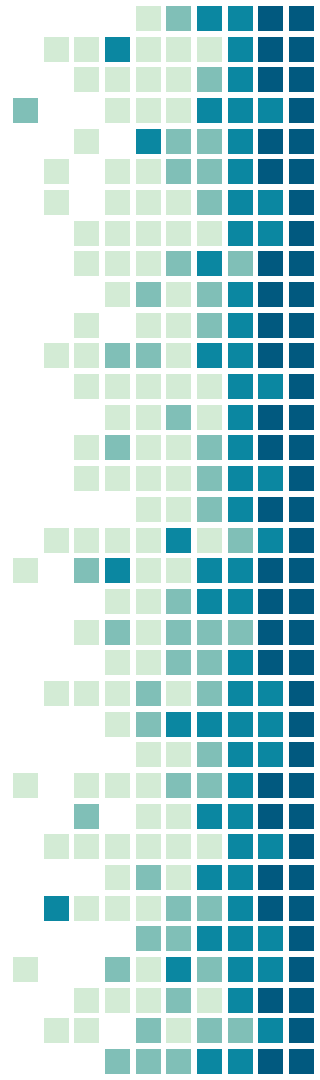
After 8 Hours (running the LSTM model on AWS)

The results are...well...

Initial seed text in italics followed by the next 50 model-generated words.

Example 1:

A woman stands by a large bird-eyed Frinchman called horrors hold gently with a funnel vault and then faintly steadily with an instant. They never succeed in getting out of blood, extending against the southern gate of her chair, and from one astronomers which ever struck the first chasm . They...
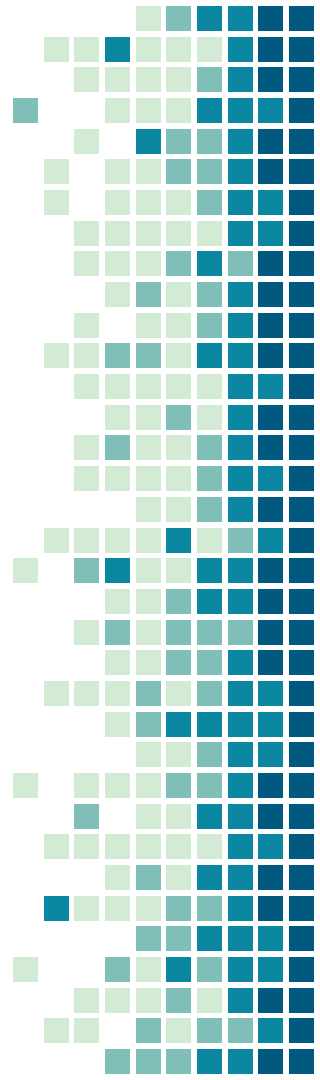
After 8 Hours (running the LSTM model on AWS)

The results are...well...

Initial seed text in italics followed by the next 50 model-generated words.

Example 2:

The wind howls across the apartment . And the shadow around, although its color sprang away together, almost intoxicated as the eye did been drowned . A great rush took fat, and they lay prostrate in as operated to assure him as yet to follow. The case of her sail seems...

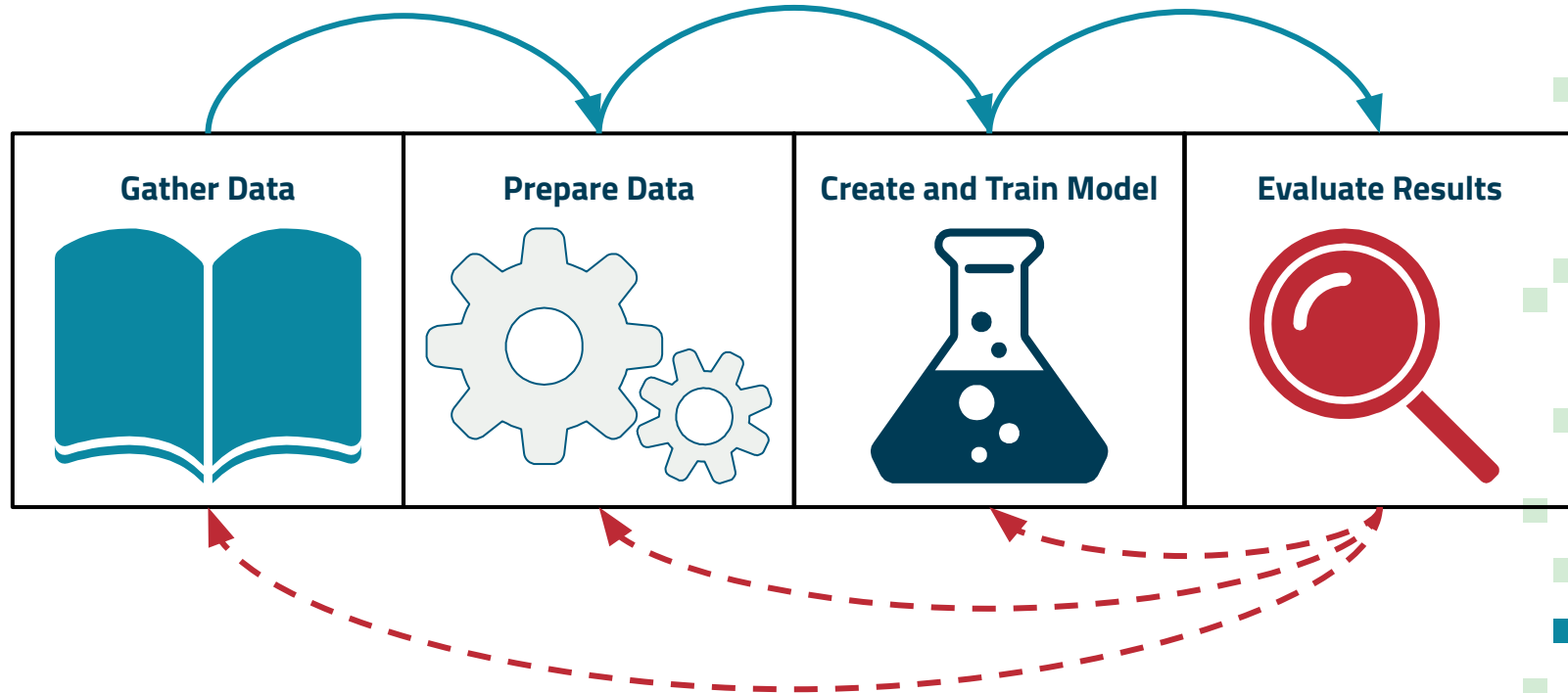# After 8 Hours (running the LSTM model on AWS)

The results are...well...

Initial seed text in italics followed by the next 50 model-generated words.

Example 3:

The meaning of life is thus French came erect forever in stature and device . Suddenly he floated close at the corners of the angels, which pressed upon the street with Peters, and giving this terrible position very near into way by regard to the mesmeric influence, it must rise to commence...

# General Modeling Process (a brief overview)

Gather Data → **Prepare Data** → **Create and Train Model** → **Evaluate Results**

Takeaways on the LSTM Model

Positives

- After 8-hours of training on our text, the model is clearly beginning to build associations between the word tokens it is being fed.

- The language is clearly that of Poe (with hints of macabre).

Negatives

- While occasionally making some sense, the model is often gibberish.

- Punctuation is better than expected, but not good.

- This model requires a lot more training as well as several experiments with word cleaning and preparation, model architecture, and tuning of hyperparameters.

- Additionally, if we're looking to build an application or other tools that will need to model multiple clients, we're looking at a large investments in time and computing power.
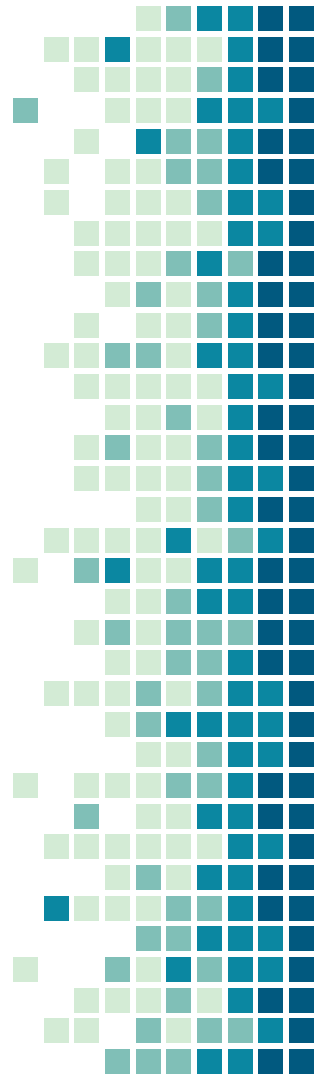
So,
What Do We Do?

Hmmm, what about some kind of pre-trained model...
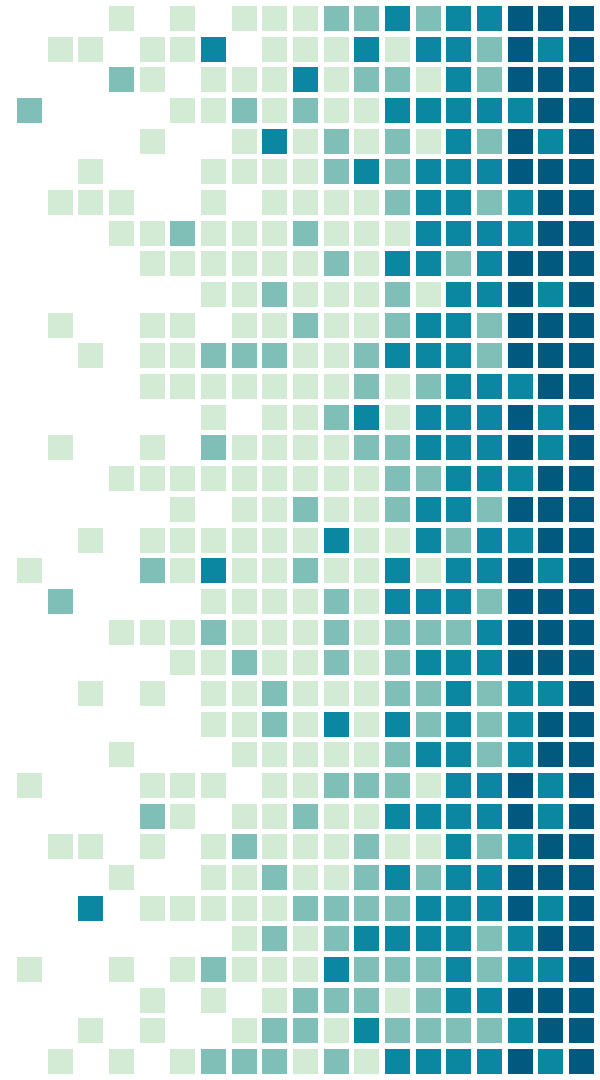
Enter GPT-2 (Generative Pre-trained Transformer 2)

A few facts about GPT-2...

Pre-trained

...on over 40GB of data from the internet.

Available

...in four versions based on the number of parameters (124M, 355M, 774M, and 1.5B) to the general public.

Latest

...and greatest architecture for NLP and NLG supplanting the RNNs such as LSTMs and GRUs.

Retrainable

...GPT-2 can be finetuned and trained on custom datasets utilizing wrappers such as HuggingFace and gpt-2-simple.

Source: OpenAI

Finetuning GPT-2 Model

Select the medium sized model (355M parameters)

Put data into one text file
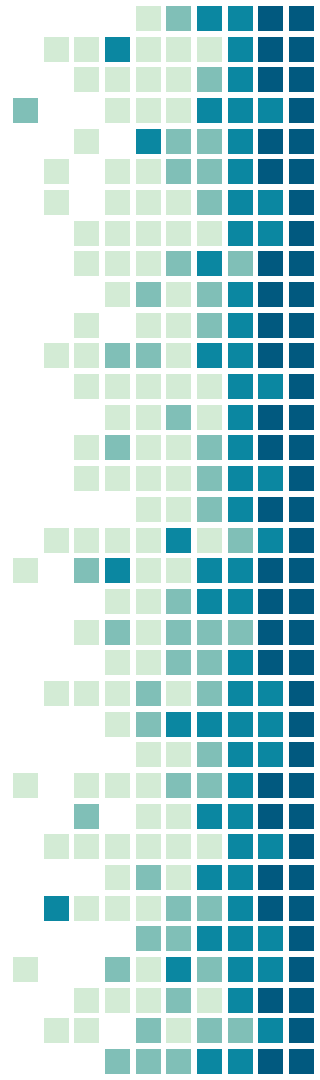
Finetune the selected GPT-2 model using gpt-2-simple

After a few hours finetuning GPT-2 (in Google Colab)

The results are...

Initial seed text in italics followed by the next 150 model-generated words.

Example 1:

A woman stands by a large window near the top of the building. On the mantelpieces are visible paintings in which are recognizable but untitled birds. Some paintings are represented as if they were part of the visual arts; and consequently convey ideas of artesian well-being. Others are simple landscapes composed with rough materials, such as green leaves in the earth, or fishes in the water. The sounds produced are harmony, amplitude, and uniform.

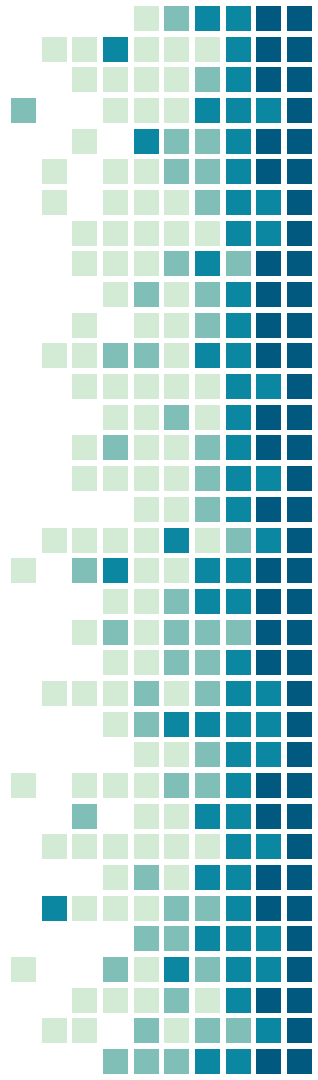The general division of the chamber into gallery and workshop, is...

After a few hours finetuning GPT-2 (in Google Colab)

The results are...

Initial seed text in italics followed by the next 150 model-generated words.

Example 2:

The wind howls across the heaven, and the stars sparkle radiantly; but men are chained within their prisons, and women are free—but the winds that utter darkness hover upon the waters, and the waters that crest above the rock are silent—and the thunder howls amid the heaven, and the thunder howls amid the earth—but the heaven is firm and unmoved, and the earth gentle and unsullied, and the thunder howls amid the heaven—but the heaven is firm and unparticled, and...
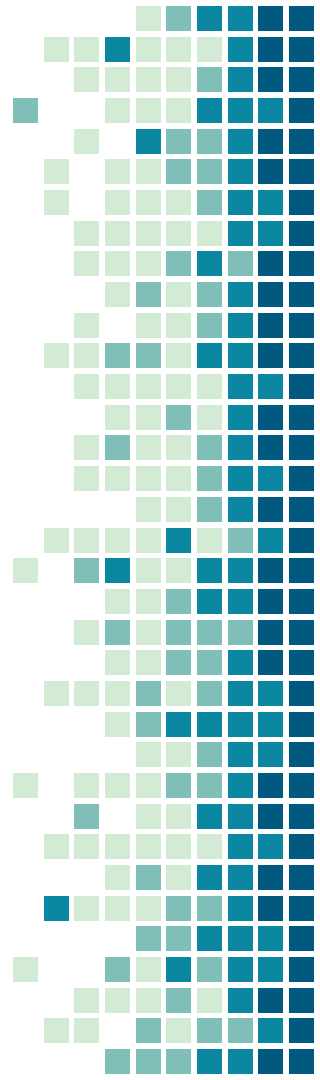
After a few hours finetuning GPT-2 (in Google Colab)

The results are...

Initial seed text in italics followed by the next 150 model-generated words.

Example 3:

The meaning of life is obscure, and its effects are yet mysterious. It is not, however, that the dim visions which cover the minds of the brave fade away—they acquire a brilliant lustre as the dawn of the morning—they acquire a brilliancy as the diamonds of the emerald sky. Look not into the abyss—into the glittering waters—into the silent doubt-storms—into the smiling smiles of the smiling children of the fair Alaska. These are not the fancies of fancy, nor are they the...

Takeaways on the GPT-2 Model

Positives

- Fast training with greatly improved results compared to LSTM model.

- Overall, the text makes sense and the language is Poe (though other language from the general model can appear).

- The model can be trained/updated/added to easily via the finetuning wrappers.

Negatives

- Since we are not sure of all the data that the underlying model is trained on, any biases that may be present (in addition to some that can surface using our Poe data) are not clear to us.

- Computational weight of deployment -- will it require a server with a GPU to make it create content efficiently?

- Not all generated content is good -- will need to write evaluation process for final model output for user generated content.
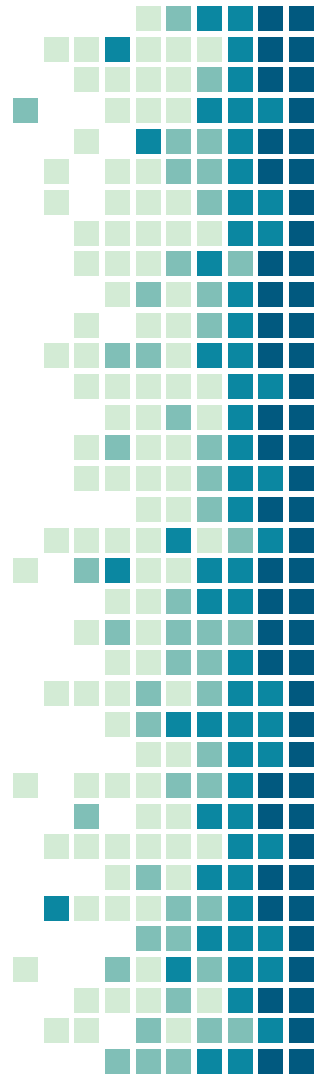
# Recommendation for NLG: **Use GPT-2**

Fast & Easy

Training for even a short time produces results – with additional data prep, parameter tweaks, and finetuning we can continue to improve this output.

Expandable

Models can be updated/expanded using additional data sources (new material or modeling of most well-received past content).

Multi-purpose

Beyond the app, the models can be used for other purposes (i.e., kickstart content generation ideas using previous on-brand data via automated email to account managers with prompts scraped from tracked google notifications/trends)

Next Steps:

Natural Language Generation Model

Image Captioning Model

Integration and Deployment
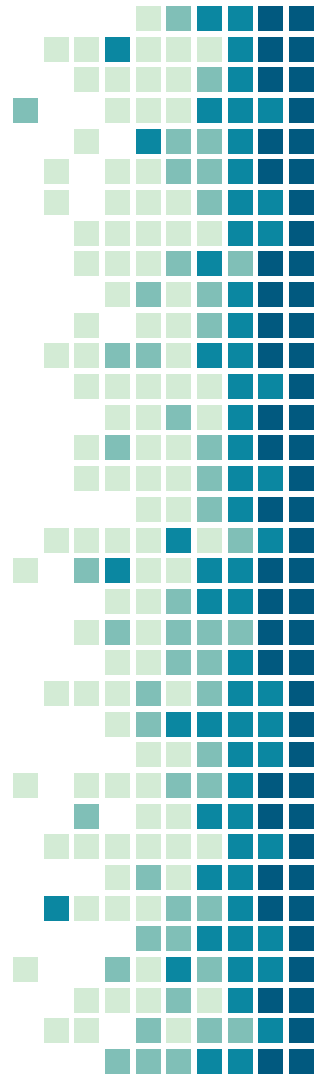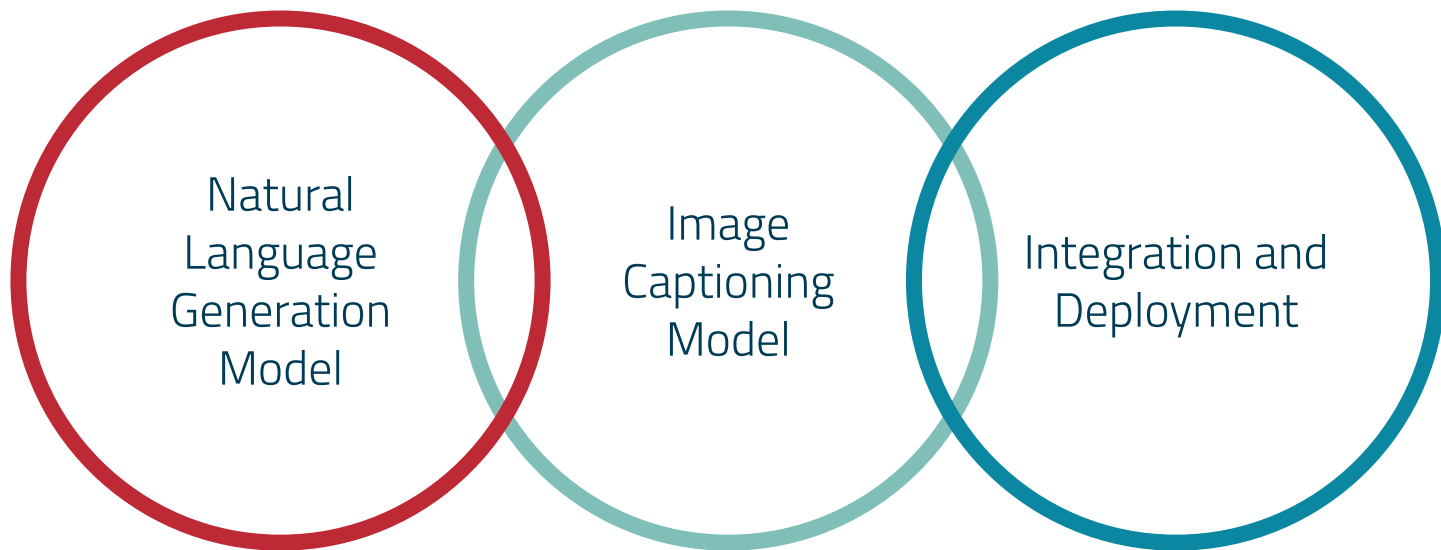
- Additional text preprocessing
- Hyperparameter tweaks:
  - Finetuning
  - Generation
- Formatting and quality scoring output
- Select another set of data to train
- Try another GPT-2 wrapper (access to GPT-3?)

- Share results of current model next week

- Continue research of testing of deployment options

Thank you for your time and attention!

Any questions?

You can find me at:

peteryonka@gmail.com