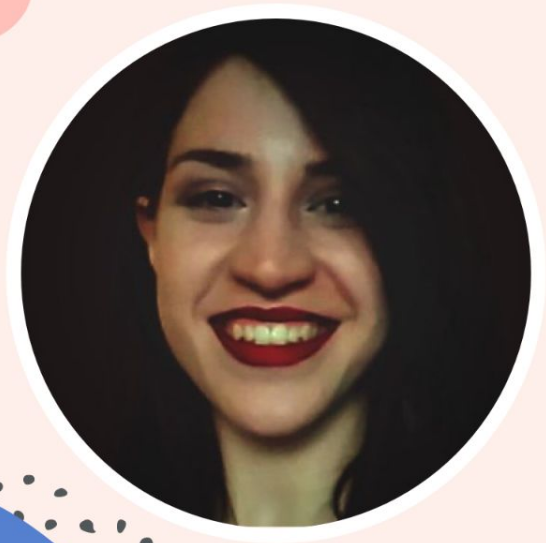


://WOMEN IN  
TECH SEO/

INNOVATE



## A Beginner's Guide to Machine Learning for SEOs

**Lazarina Stoy**

Technical SEO Analyst  
Skale



first,  
thank  
you all.

You inspire me every, single day.



The background is a light cream color with several abstract, organic shapes in muted orange and sage green. In the top left, there is a thin, dark green wavy line. In the top right, a thick, dark green wavy line is present. In the bottom left, a sage green shape is filled with small, dark green dots. In the bottom right, there are four vertical dark green lines of varying heights.

**Before we  
begin...**

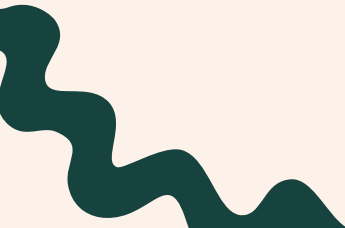
The background is a light beige color with several abstract, organic shapes in muted orange and teal. In the top left, there is a thin, dark teal wavy line. In the top right, there is a thick, dark teal wavy line. In the bottom left, there is a teal shape filled with small, dark teal dots. In the bottom right, there is a thick, dark teal wavy line. The text is centered in the middle of the image.

**...just want to set the  
tone**



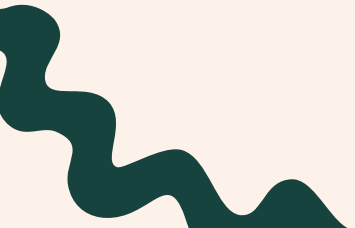



# what I am:

- ✓ Technical SEO Manager
  - ✓ Content Creator
  - ✓ Growth & efficiency obsessed
- 



## what I am not:

- ✗ a Python developer
  - ✗ a machine learning expert
  - ✗ an 'expert'
  - ✗ a 'guru'
- 

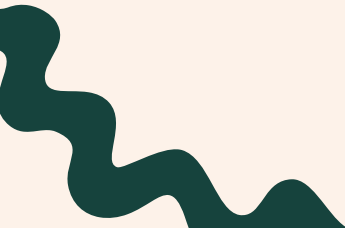
The background features a mix of organic, flowing shapes in shades of light orange, cream, and light teal. In the top left, there is a pattern of small dark teal dots. In the top right, a thin, dark teal line swirls. In the bottom right, a thick, dark teal wavy line is visible. The text is centered on the left side of the image.

**“I have  
no idea  
what I’m  
doing”**

- Ruth Everett
- Britney Muller
- me

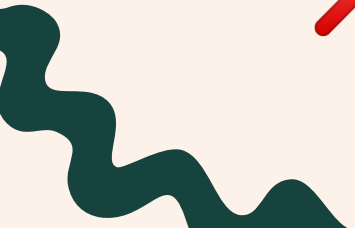


# what I will talk about:

- ✓ Challenges of ML in SEO
  - ✓ Processes to overcome them
  - ✓ Tools & resources to kickstart your journey
- 



# what I will not talk about:

- ✗ textbook definitions - we can all google things
  - ✗ 'Every SEO needs to learn Python'
  - ✗ 'ML is the future of SEO'
- 

The background features a light beige color with several abstract elements: a dark teal wavy line in the top left, a dark teal wavy line in the top right, a light teal organic shape in the top right, a light teal organic shape with dark teal dots in the bottom left, a light teal organic shape in the bottom right, and a dark teal wavy line in the bottom right.

Enough about me...

The background is a light cream color with several abstract, organic shapes in muted teal and terracotta. In the top left, there is a thin, dark teal wavy line. In the top right, a thick, dark teal wavy line is present. A large, solid teal shape is on the right side. In the bottom left, a teal shape contains a pattern of small, dark teal dots. In the bottom right, there are four vertical, dark teal lines of varying heights.


**Let's talk about  
you.**



# Waking up on a Monday...



@lazarinastoy | #WTSFest







**You look  
at your  
schedule.**



@lazarinastoy | #WTSFest



Another one  
of those  
weeks.



@lazarinastoy | #WTSFest

## TOP CATEGORIES



Documenting

**16%** 6 hr 17 min

Last week:

7%

Change:

↑ 9%



Data

**9%** 3 hr 37 min

Last week:

0%

Change:

↑ 9%



Miscellaneous

**7%** 2 hr 44 min

Last week:

6%

Change:

↑ <1%



# 36.3%

of SEOs spend  
most of their  
time on keyword  
research\*.



33.4%

of SEOs spend most  
of their time on  
on-page  
optimisations\*.





**Despite all  
the tools we  
have at our  
disposal...**

**...most  
processes are  
fairly manual.**





**And even  
when a new  
tool comes  
along...**





We  
struggle  
to find  
time.



@lazarinastoy | #WTSFest

**But I'm here to  
argue there are  
other factors at  
play here.**





01.

# **Common challenges in getting started with ML in SEO**

# 1. Limiting beliefs might be holding you back.



**Recognise any  
of these?**



The background features a mix of organic, hand-drawn shapes in light orange, teal, and cream. In the top left, there's a cluster of small dark dots. On the right, a thin, dark, swirling line is visible. The overall aesthetic is modern and artistic.

**You have to be**

**a**

Python

Coding


Maths

Data

Science

**”**

**expert  
to start**

The background features a mix of organic, flowing shapes in light orange, cream, and light teal. In the top left, there is a cluster of small black dots. In the top right, a thin black line forms a series of loops. In the bottom right, a thick black wavy line is visible. The text is arranged in two main blocks, separated by a large quote mark.

**You have to  
know what  
each ML  
algorithm  
does**

**”**

**to start**

The background features a mix of organic, flowing shapes in light orange, cream, and light green. In the top left, there is a cluster of small black dots. In the top right, a thin black line forms a series of loops. In the bottom right, a thick black wavy line is visible. The text is arranged in two columns, with a large closing quotation mark at the top right.

**You need to  
have plenty of  
free time to  
learn from  
scratch**

**”**

**to start**



The background features a mix of organic, flowing shapes in shades of light orange, cream, and light green. In the top left, there is a cluster of small dark green dots. In the top right, a thin, dark green wavy line is visible. In the bottom right, a larger, light green shape contains a dark green wavy line. The text is centered in the middle of the composition.

**It is**

difficult  
unattainable  
scary  
unnecessary

**to start**

”

# What to do

## **Waiting To Get Started?**

Search machine learning in 10 minutes. Follow along.  
Start small but do something today.

## **Awaiting Perfect Conditions?**

Build a habit and track your progress.  
Start small and remain consistent.

## **Struggling or Tried and Failed?**

cut scope or change direction.  
Start small to get back into it.



@lazarinastoy | #WTSFest

# What to do

## Waiting To Get Started?

Search machine learning in 10 minutes. Follow along.

**Start small** but something today.

## Awaiting Perfect Conditions?

Build a habit and track your progress.

**Start small** and remain consistent.

## Struggling or Tried and Failed?

cut scope or change direction.  
**Start small** to get back into it.



@lazarinastoy | #WTSFest

**2. Lack of context might be obstructing your progress.**

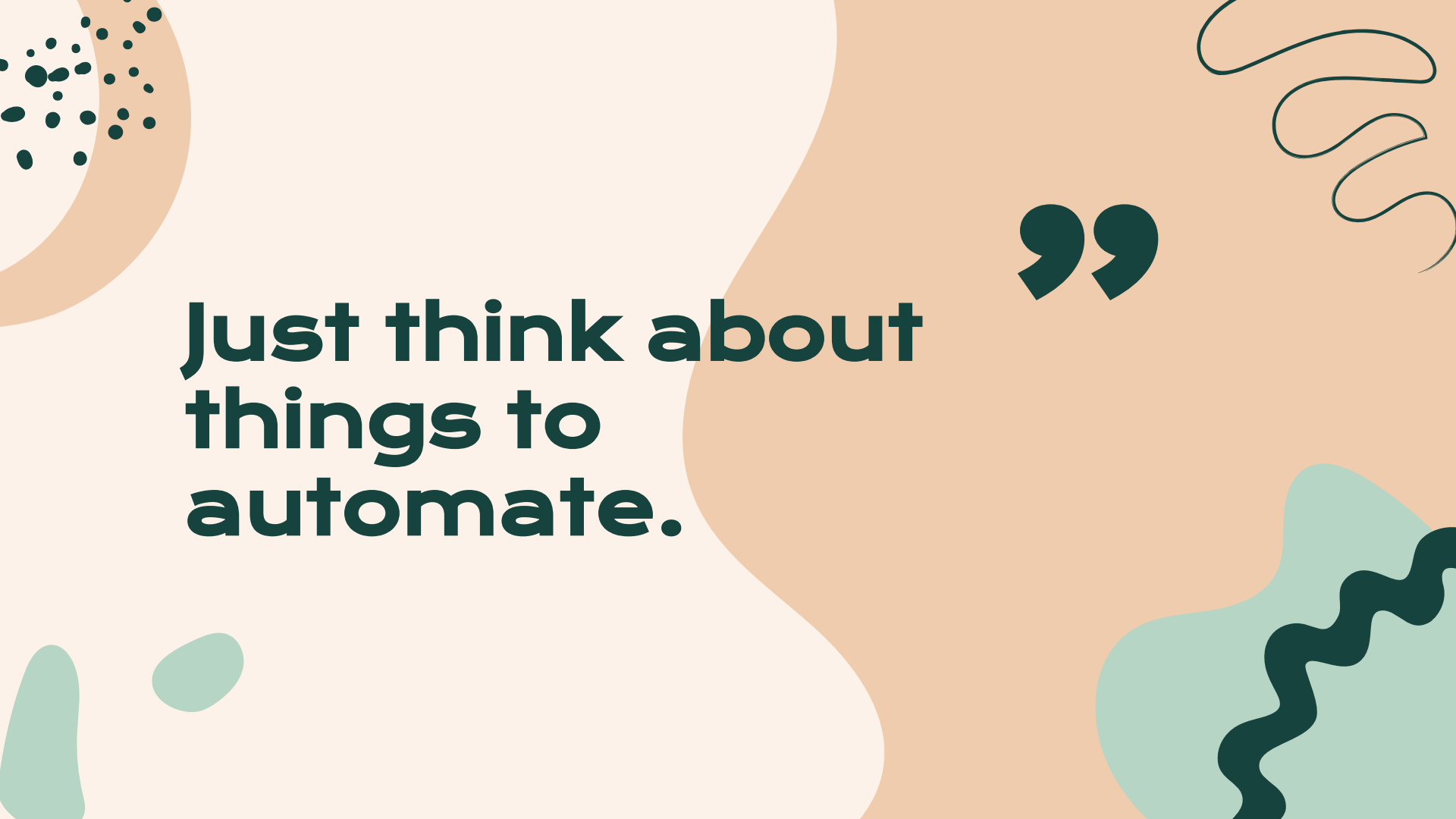




**Often you might  
hear things like...**

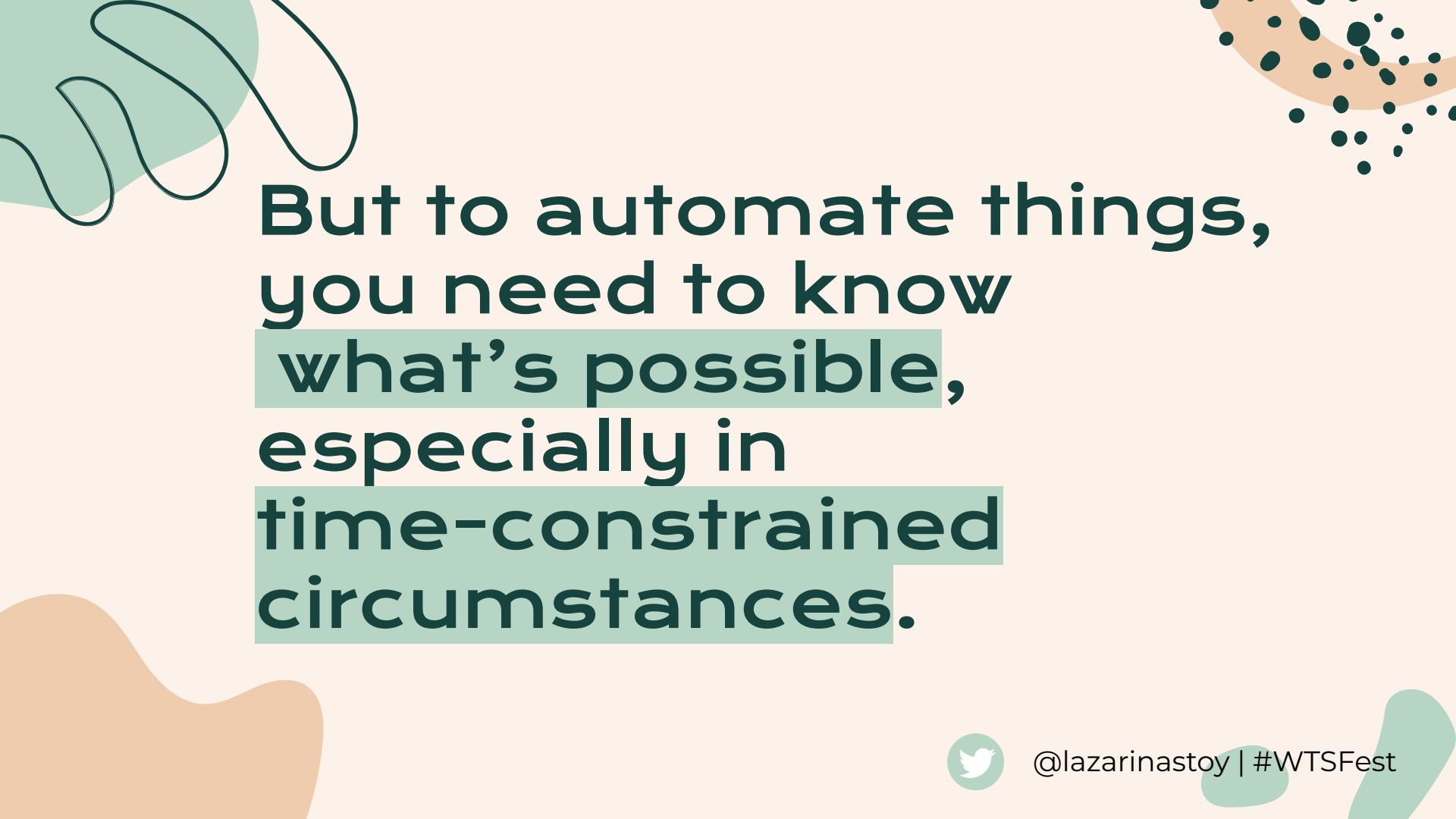


@lazarinastoy | #WTSFest

The background features a mix of organic, flowing shapes in light orange, cream, and light green. In the top left, there's a pattern of small dark dots. In the top right, a thin, dark, swirling line is visible. A large, dark teal quotation mark is positioned to the right of the main text. In the bottom right, a dark teal wavy line is present. The overall aesthetic is modern and minimalist.

**Just think about  
things to  
automate.**

”



But to automate things,  
you need to know  
what's possible,  
especially in  
time-constrained  
circumstances.



@lazarinastoy | #WTSFest

**otherwise it turns  
into this...**



@lazarinastoy | #WTSFest





**Zhuowei Zhang**

@zhuowei

...

Never spend 6 minutes doing something by hand when you can spend 6 hours failing to automate it

9:58 AM - Apr 26, 2020 · Twitter Web App



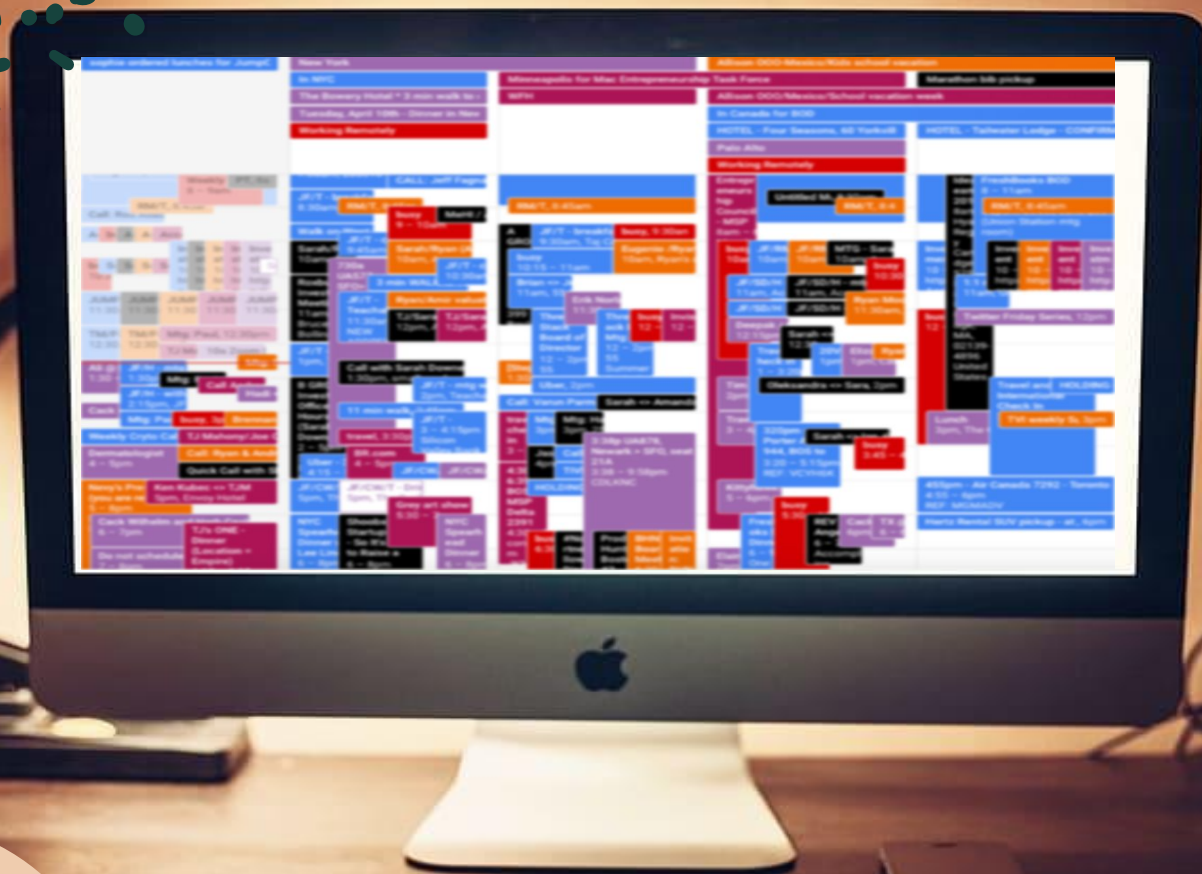
@lazarinastoy | #WTSFest




The beauty of this  
quote is lost when  
your day-to-day  
looks like this...



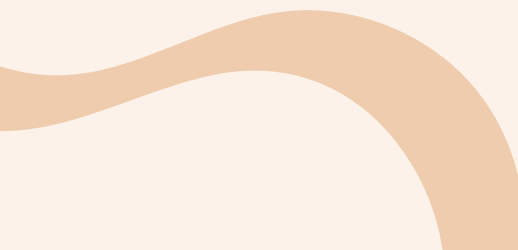
@lazarinastoy | #WTSFest



@lazarinastoy | #WTSFest



**Often, there is no  
need to reinvent the  
wheel, especially as  
a beginner.**

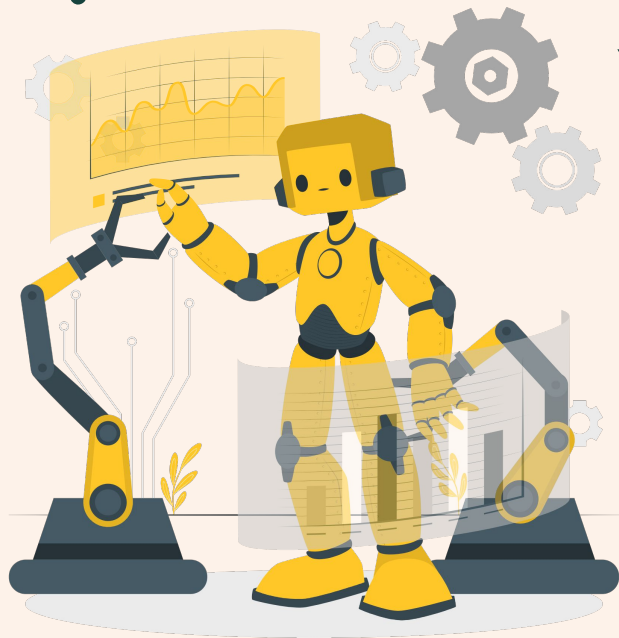


@lazarinastoy | #WTSFest

**You don't need to  
know a ton of  
theory.**



@lazarinastoy | #WTSFest



**You don't need  
end-to-end,  
custom-built,  
automated  
solutions.**



@lazarinastoy | #WTSFest

**You don't need  
to spend a ton  
of time.**



@lazarinastoy | #WTSFest

The background features a warm, earthy color palette of light orange, cream, and muted teal. Large, flowing organic shapes in these colors create a layered effect. In the top right, there are thin, dark teal wavy lines. In the bottom left, a teal shape is filled with a pattern of small dark teal dots. In the bottom right, there are five vertical dark teal bars of equal height.

**Aim to drive value  
and have fun with  
the process.**





# Knowing

- what model to use
- how to find and implement it quickly
- how to drive value via ML

**is the perfect way to start.**



@lazarinastoy | #WTSFest



02.

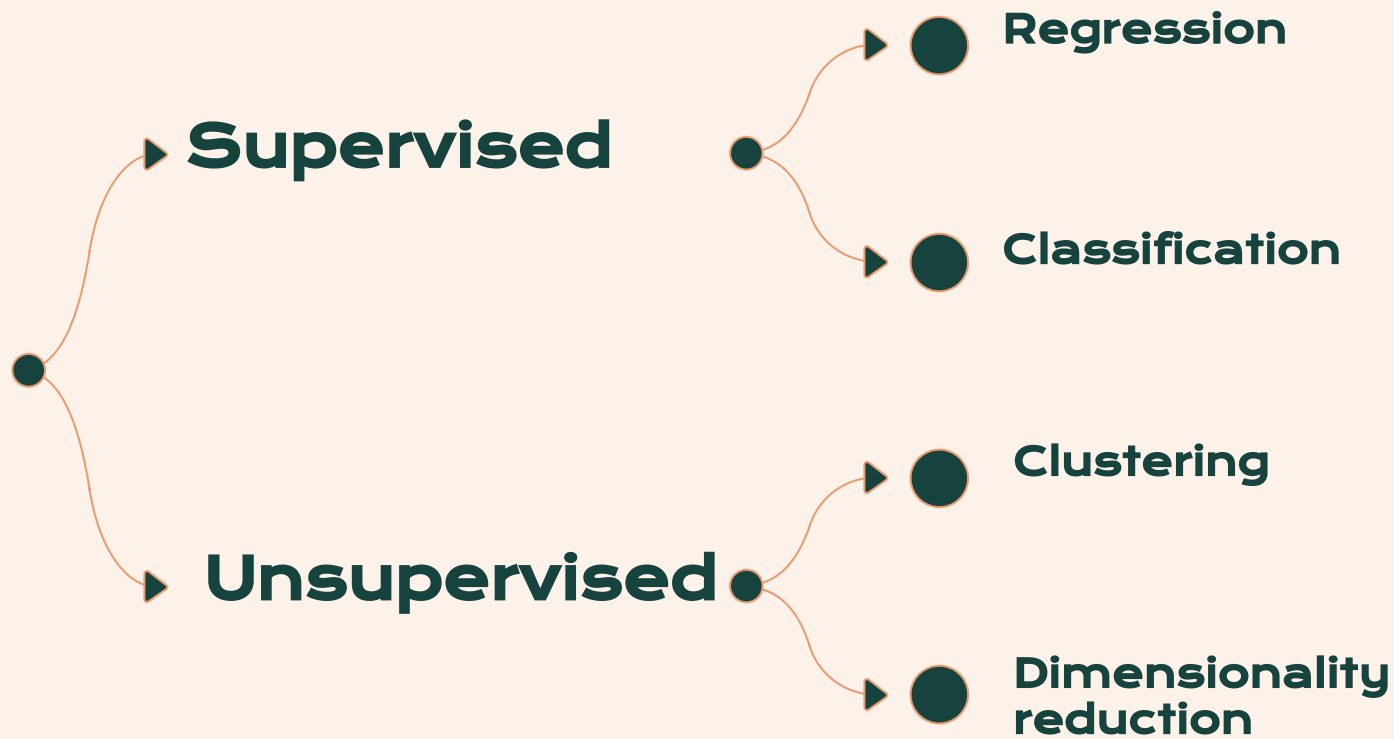
# **When to search for ML in SEO**

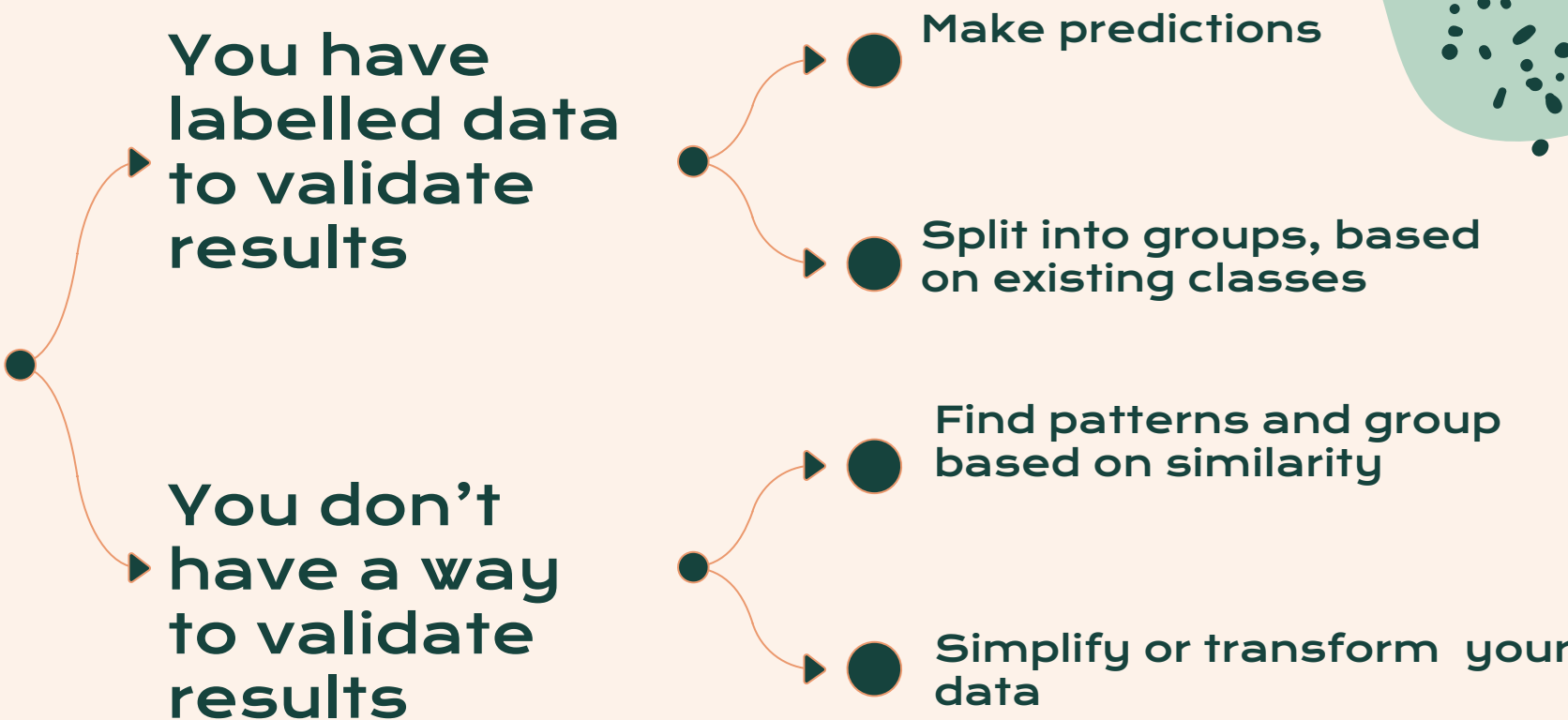


**Let's start  
with the  
basics.**



# **Task characteristics**





**You have  
labelled data  
to validate  
results**

**Make predictions**

**Split into groups, based  
on existing classes**

**You don't  
have a way  
to validate  
results**

**Find patterns and group  
based on similarity**

**Simplify or transform your  
data**



@lazarinastoy | #WTSFest

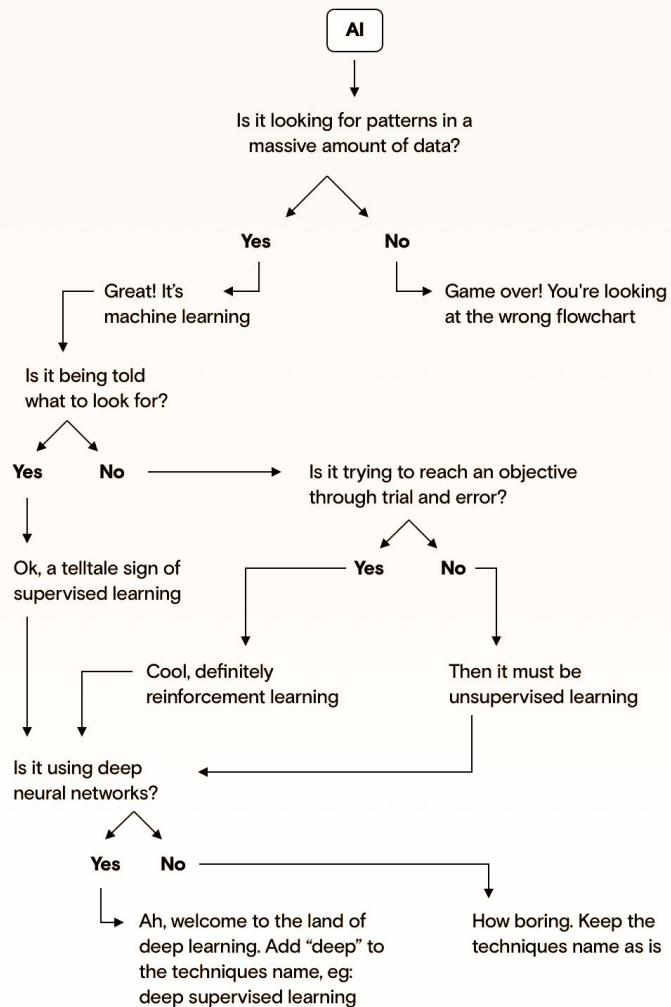


# Beyond the basics

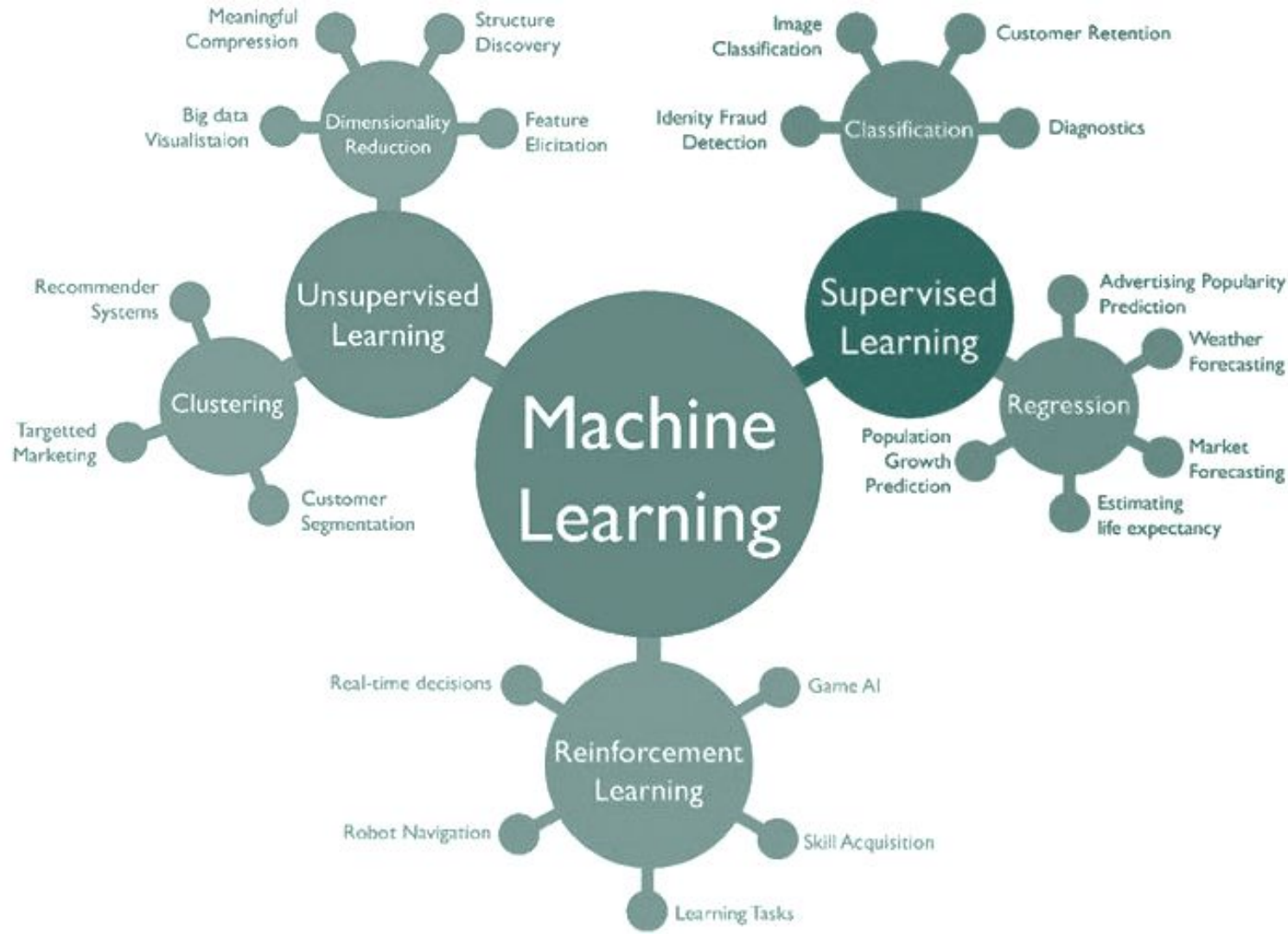
You might find  
something like  
this...



@lazarinastoy | #WTSFest







The background is a light beige color with several abstract decorative elements. In the top left, there is a thick, wavy orange line. In the top right, there is a dark teal wavy line. In the bottom left, there is a light teal shape filled with small dark teal dots. In the bottom right, there is a large orange shape with a light teal arc passing over it. The text "Data characteristics" is centered in a bold, dark teal font.

# **Data characteristics**

# Is your input data:



**Textual**

**Numeric**



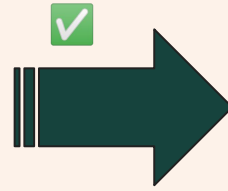
**Image-based**



# **Solution characteristics**

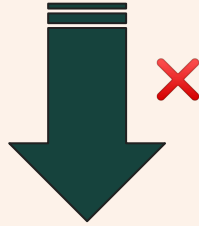


**Is this task  
mission critical?**



Don't rely  
on AI.

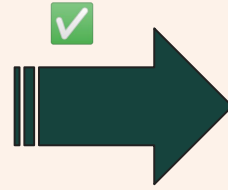
(seriously)



@lazarinastoy | #WTSFest

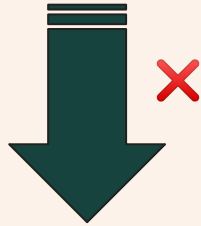


**Must the results  
remain  
consistent  
every time?**



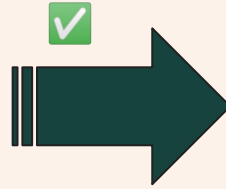
Don't rely  
on AI.

(yeah,  
really)

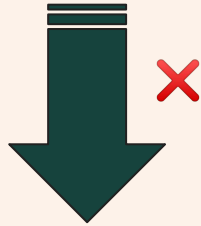


@lazarinastoy | #WTSFest

**Must the results  
remain easy to  
understand and  
relay to the  
stakeholder?**




Don't go  
into the  
deep.



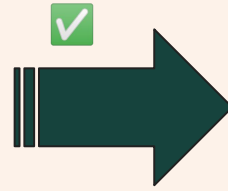
(skip deep  
learning)



@lazarinastoy | #WTSFest



**Is it okay that  
simply on  
average the  
output  
outperforms  
existing  
methods?**



Okay,  
then.

Take a  
look at ML  
options.



@lazarinastoy | #WTSFest





# Assess using multiple factors



insights



complexity



accuracy



bottom  
line



scalability



assets  
(data)



resources



@lazarinastoy | #WTSFest

**Let's put  
this into  
context...**





03.

**How to find suitable ML  
tools?**

The background is a solid light orange color. It features several abstract elements: a large, irregular white shape in the upper left; a teal shape in the lower left containing a cluster of small dark teal dots; a white wavy line in the lower right; a series of five vertical dark teal bars in the bottom right corner; and a dark teal line drawing of a hand in the top right corner.

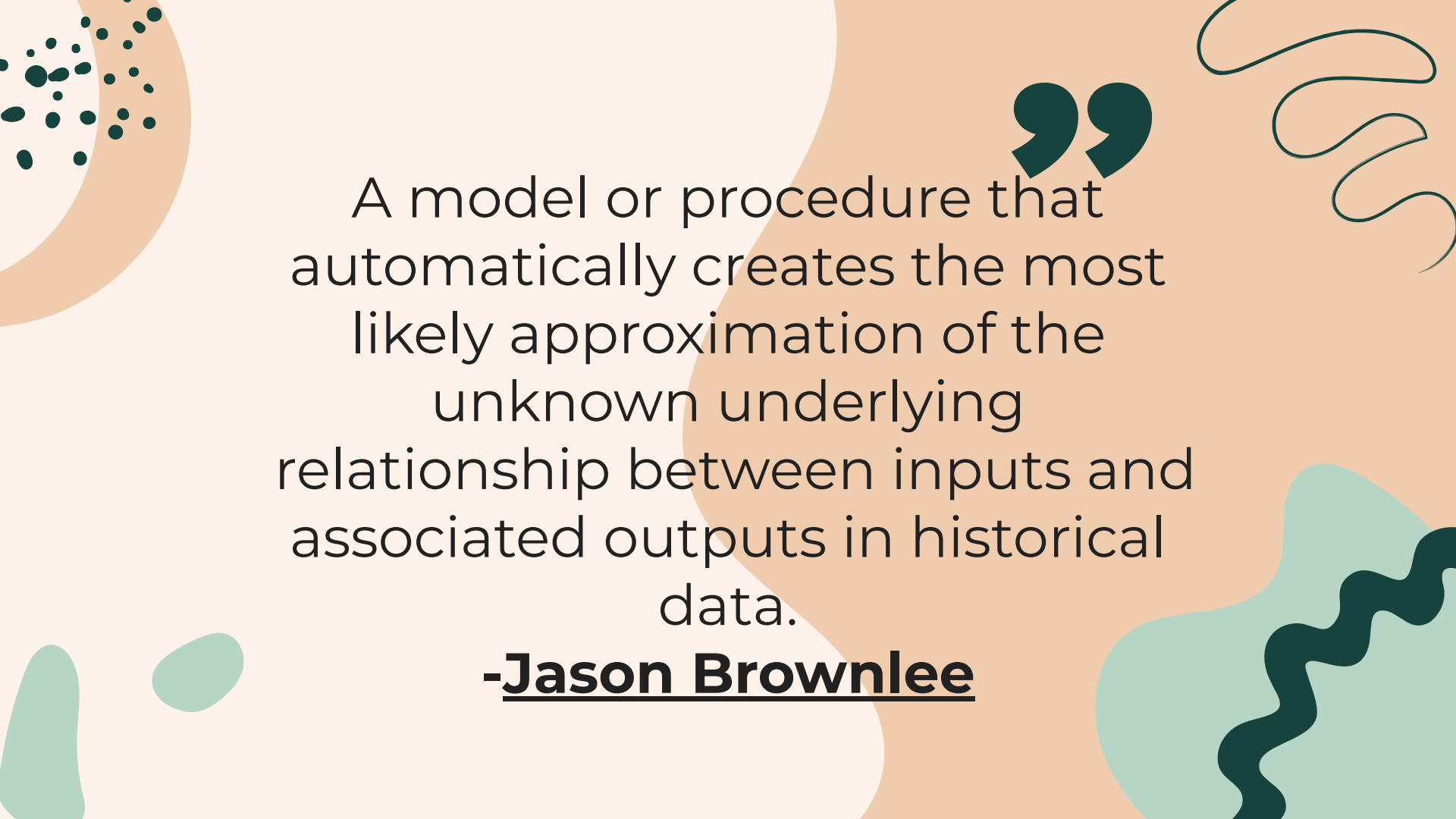
**Keep queries  
specific to data,  
task, and solution.**



”

Find a model or procedure that makes best use of historical data comprised of inputs and outputs in order to skillfully predict outputs given new and unseen inputs in the future.

**-Jason Brownlee**

The background features a light cream color with several organic, hand-drawn style shapes. In the top left, there's a cluster of dark teal dots. In the top right, a thin, dark teal wavy line. In the bottom left, a small teal oval. In the bottom right, a larger teal shape with a dark teal wavy line. A large, dark teal quotation mark is positioned in the upper right area of the text.

A model or procedure that automatically creates the most likely approximation of the unknown underlying relationship between inputs and associated outputs in historical data.

**-Jason Brownlee**

A typical  
on-page  
optimisation  
project might  
include  
mini-projects  
like...



@lazarinastoy | #WTSFest

# 1. Writing Meta Descriptions

- Input data? **Textual** (page content).
- Supervised or unsupervised? **Unsupervised**
  - It is **transformational** (Page content to Page Summary in less than 160 characters)
  - It can also be **generative** (write them from scratch)
- Is it mission critical? **No.**
- Different results okay? **Yes.**
- Explanation of process needed? **Not really.**
- Outperforms average methods? Yes, much **faster.**





**Python script meta descriptions**



Google Search

I'm Feeling Lucky



**Python script unsupervised machine learning text summarisation**



Google Search

I'm Feeling Lucky

## 2. Title / H1 Optimisations

- Input data? **Textual** (page content).
- Supervised or unsupervised? **Unsupervised**
  - It is **transformational** (Page content to Page Summary in less than 60 characters)
  - It can also be **generative** (write them from scratch)
- Is it mission critical? **Hm, debatable (ask HMRC 🙄)**
- Different results okay? **Again, debatable.**
- Explanation of process needed? **Kind of.**
- Outperforms average methods? Yes, much **faster. Not better.**



**Python script python script title optimisation**



Google Search

I'm Feeling Lucky

# 3. Image Alt tag Generation

- Input data? **Image**
- Supervised or unsupervised? **Supervised**
- **Generative**
- Is it mission critical? **No**
- Different results okay? **Yes.**
- Explanation of process needed? **Not really.**
- Outperforms average methods? Yes, much **faster.**



**Python script image alt text caption generation machine learning**



Google Search

I'm Feeling Lucky

# But might also involve...

- **Predicting traffic / revenue** based on presence of keyword in the title/ h1 to get buy in on proposed changes
- Updating **internal links**
- **Researching keywords** for new content updates
- **FAQ** schema implementation

The background features a mix of organic, hand-drawn shapes in muted colors. There are large, soft-edged shapes in light orange and cream. In the top left, a cluster of small dark green dots is arranged in a circular pattern. In the top right, a thin, dark green line forms a series of loops. In the bottom right, a larger light green shape contains a dark green wavy line. The overall style is modern and artistic.

**“Surely a  
script can’t do  
all that?!”**





**No.**



**But a few can.**



# Predict SEO Organic CTR Based on Position Using Machine Learning

Know if your metatags and CTR% are good or bad.



Michael Van Den Reym

Following



Dec 27, 2021 · 4 min read



## Automate Keyword Research with Google



Michael Van Den Reym

Jun 5, 2021 · 3 min read

SEO keyword research is time consuming, so I'm automating this process!

Run 50 iterations

Iterations: 0

Train with 25 topics

[0] government american federal peace war america country economic made security

[1] government america country war federal national american work peace united

[2] government federal american country america states made security peace national

[3] government war america nation states security made peace work national

Topic Documents

Topic Correlations

Time Series

Vocabulary

Downloads

Documents are sorted by their proportion of the currently selected topic, biased to prefer longer documents.

[1953-53/5.5%] Because the building of a completely impenetrable defense against attack is still not possible, total defensive strength must include civil defense preparedness. Because we have incontrovertible evidence that Soviet Russia possesses atomic weapons, this kind of protection becomes sheer necessity...

[1946-44/5.3%] Contract settlement and surplus property disposal.--The winding up of war procurement is the second most important liquidation job. By the end of November a total of 301,000 prime contracts involving commitments of 64 billion dollars had been terminated. Of

Use a different collection:

Documents Choose file No file chosen

Stoplist Choose file No file chosen

Upload

.P

## String Matching with BERT, TF-IDF, and more!



Charly Wagnier · Jul 25, 2021 · 1 min read

## Generate FAQs for your pages automatically with What The FAQ!

Updated: Jul 26, 2021



**Bottom  
line?**

The background features a warm, orange-toned palette with large, flowing organic shapes in white and light beige. In the bottom left, there is a teal-colored area with a pattern of small, dark dots. In the top right, there are thin, dark, wavy lines. In the bottom right, there are four vertical, dark teal bars of varying heights.

**Adding value doesn't  
necessarily mean a  
fully-automated,  
autonomous solution.**

**Incremental  
improvements  
can lead to a  
compounding  
effect.**



@lazarinastoy | #WTSFest

# Leaving time for revenue- & career-advancing projects:

- Leadership
- Growth Strategy
- User Personas
- Personalisation
- Scalable content systems
- ...



@lazarinastoy | #WTSFest

**Fired up?**

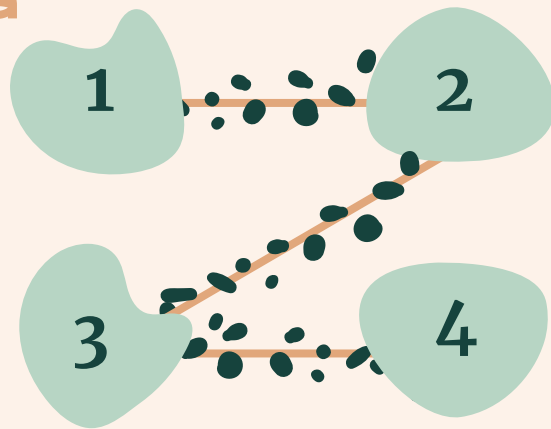




# get ready to practice

Install Anaconda  
and pip. Set up  
Colab.

Get your foot  
through the  
door with some  
daily practice

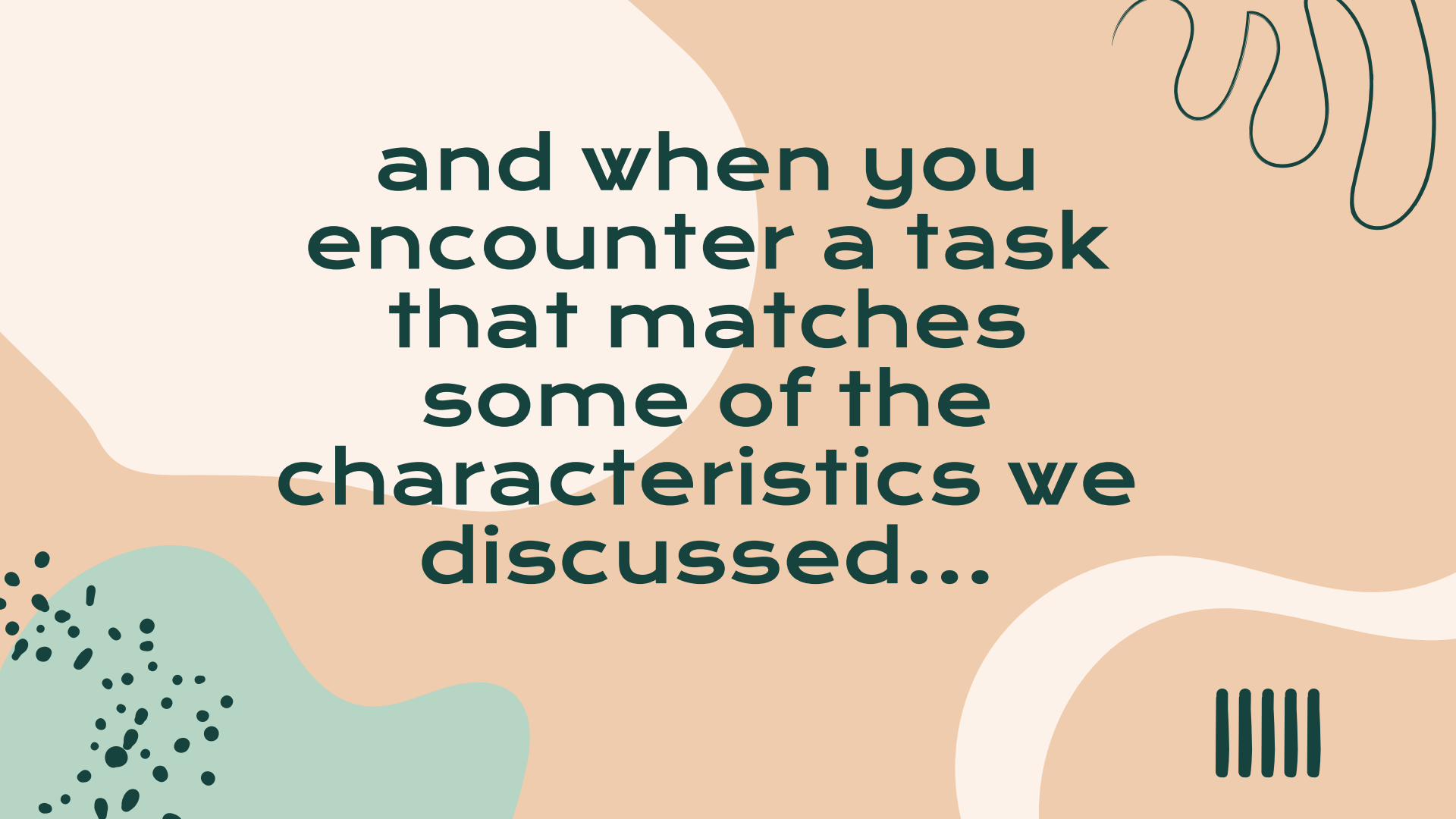


Install the  
main ML  
libraries

Then get on  
with your role  
as per usual



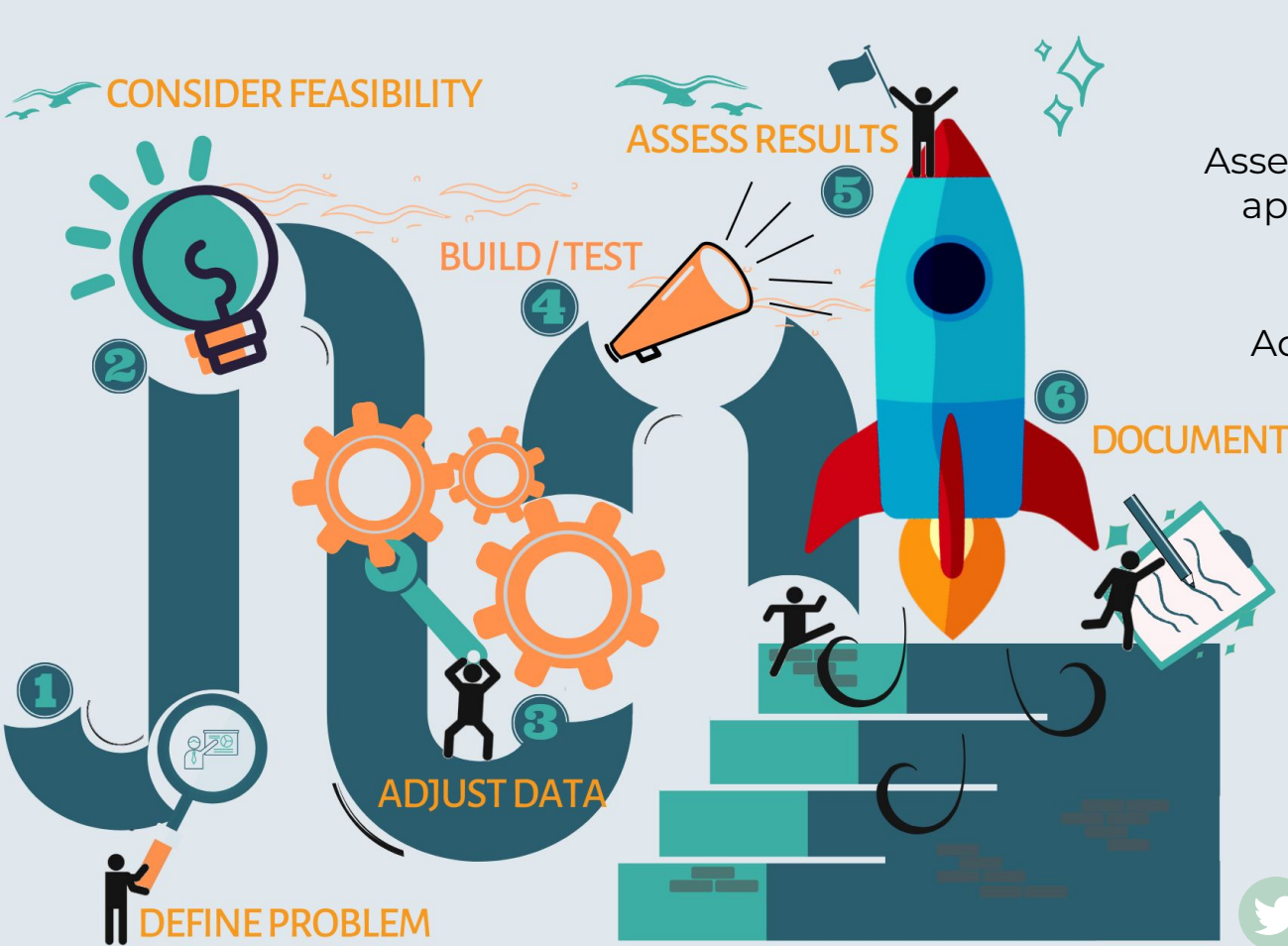
@lazarinastoy | #WTSFest

The background features a warm, orange-toned palette with large, flowing, organic shapes in white and light green. In the top right corner, there are thin, dark, wavy lines. In the bottom left, a cluster of small, dark, irregular dots is visible. In the bottom right, there are five vertical, dark, rectangular bars of varying heights.

**and when you  
encounter a task  
that matches  
some of the  
characteristics we  
discussed...**







Deconstruct each task you encounter

Assess whether ML is the correct approach to solve the problem

Adjust (prepare) or gather your data (if needed)

Build or test scripts, libraries and tools.

Assess results.  
Scrutinise the output.  
Compare, if possible.

Document the journey.  
Build your deliverable.  
@lazarinastoy | #WTSFest





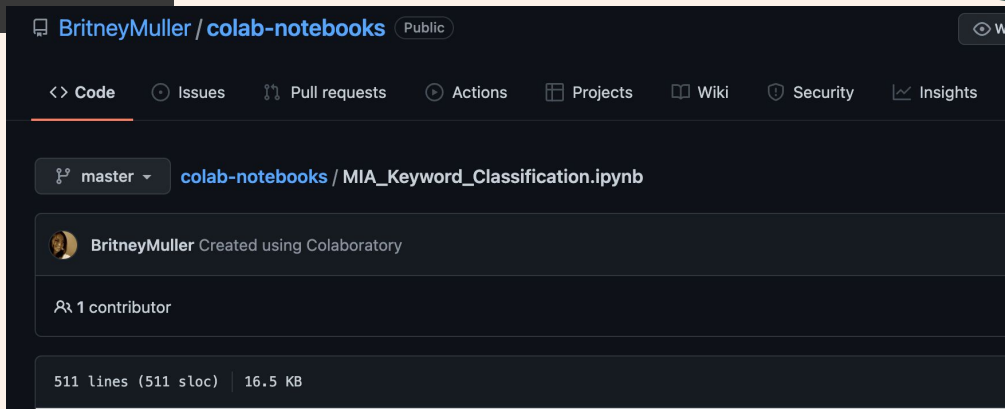
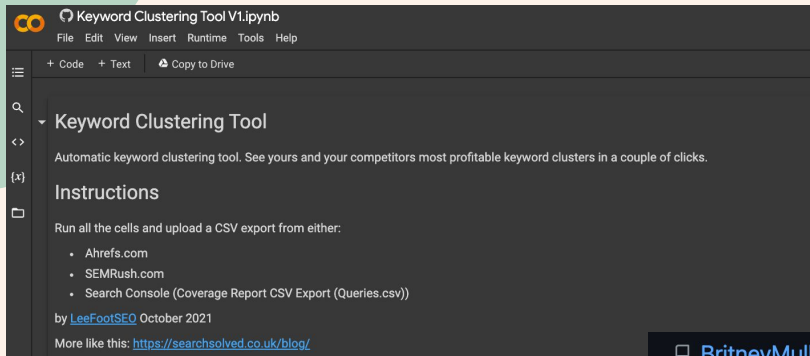
04.

**What is possible in a  
short time-frame?**

**Quick  
wins,  
no code**



# Keyword Clustering



@LeeFoot | @britneyMuller

# Entity Extraction

**Pick a choice**

- ☒ NER
- ☐ NER for URLs

### Named Entities

Entity labels

```

[PERSON, PERSON, LOCATION,
PERSON, PERSON, LOCATION,
PERSON, PERSON,
PERSON, PERSON,
PERSON, PERSON,
PERSON, PERSON,
PERSON, PERSON,
PERSON, PERSON]
  
```


**Enter Text**

[Sri] [Gupta] (born 24 September 1999 at 7:37 a.m.) resides in Purjidi, India. She is a Python developer and can speak fluently in english. She scored 89.8% in her senior secondary school and is now pursuing Bachelors in Technology from DAV Institute of Engineering and Technology, Jalandhar.

### Named Entities

[Sri] [Gupta] **PERSON**, (born 24 September 1999 at 7:37 a.m., **TIME**) resides in [Purjidi], **LOC**, [India], **CPE**. She is a [Python], **CPE** developer and can speak fluently in english. She scored **89.8%**, **PERCENT** in her senior secondary school and is now pursuing Bachelors in Technology from [DAV Institute of Engineering and Technology], **ORG**, [Jalandhar], **CPE**.

	text	label	start	end	start_char	end_char
0	SriGhi Gupta	PERSON	0	2	0	13
1	24 September 1999	DATE	4	-	9	20
2	7:37 a.m.	TIME	8	16	41	56
3	Purjidi	CPE	13	14	83	89
4	India	CPE	15	19	71	76
5	Python	ORG	20	21	87	93
6	english	LANGUAGE	27	28	130	137
7	89.8%	PERCENT	31	33	150	155
8	DAV Institute of Engin...	ORG	46	52	236	279
9	Jalandhar	CPE	53	54	261	266


 SERP Content Entity, Keyword and NLP Analysis.ipynb ☆  
 File Edit View Insert Runtime Tools Help Changes will not be saved

Srishti Gupta



# Topic Modelling

Run 50 iterations

Iterations: 0

Train with  25 topics

[0] government  
american peace  
america war  
made federal  
country nation  
states

[1] government  
national federal  
states american  
security war

Topic Documents

Topic Correlations

Time Series

Vocabulary

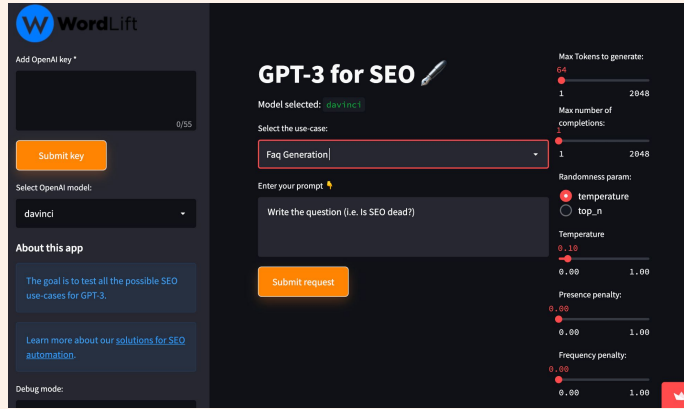
Downloads

Documents are sorted by their proportion of the currently selected topic, biased to prefer longer documents.

Use a different collection:

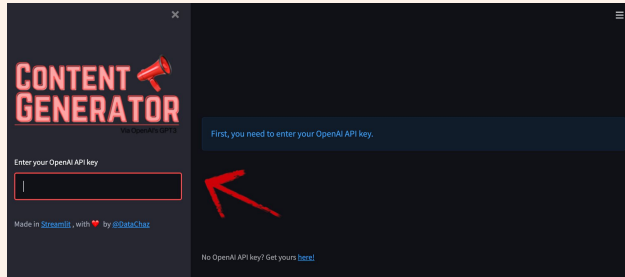
Documents  No file chosen

Stoplist  No file chosen



# GPT-3

- Content generation
  - Email writing
- Product Descriptions
  - FAQ Generation



@DataChaz | @cyberandy



# 05.

## **Best practice tips**

# Find a buddy.

Like-minded people or  
an ML-focused SEO tribe  
can be a mindset  
game-changer.



# Find the right tribe.



# Set reasonable expectations.

Managing expectations about what ML can and can't do will help you know when to apply it.



@lazarinastoy | #WTSFest



Don't let  
limiting beliefs  
keep you from  
experimenting



# collaborate. build relationships.

Ask for help when needed, developers or more experienced SEO builders would be happy to help you



@lazarinastoy | #WTSFest



# Test, test, test.

Test scripts every day. Dissect, analyse, **understand them**.

Find what you like and dislike about them.

Note best practices.

Building will become easier after.





Read the talk re-cap & drop any Qs  
in the comments 🔥:  
**[lazarinastoy.com/beginners-guide  
-to-machine-learning-for-seos/](https://lazarinastoy.com/beginners-guide-to-machine-learning-for-seos/)**



CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik and illustrations by Storyset

Check out the ML for SEOs collection  
on GitHub (scripts, notebooks, and  
tools):

**[github.com/lazarinastoy/beginners-  
guide-to-machine-learning-for-seos](https://github.com/lazarinastoy/beginners-guide-to-machine-learning-for-seos)**



CREDITS: This presentation template was created by Slidesgo, including icons  
by Flaticon, infographics & images by Freepik and illustrations by Storyset

