

brighton
SEO.

April
2022



How to Implement Machine Learning in Your Internal Linking Audit

LAZARINA STOY. | INTREPID DIGITAL



@LAZARINASTOY



SLIDESHARE.NET/LAZARINASTOYANOVA



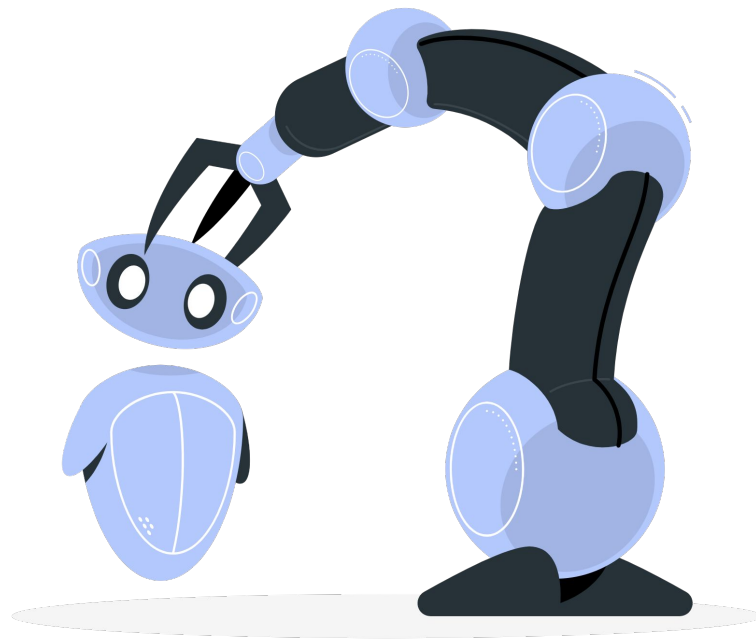
HELLO.

(THANK YOU FOR BEING HERE).

[\(IF YOU NEED A TRANSCRIPT OF THE VIDEO RECORDING OF MY TALK, CLICK HERE\)](#)



**BECAUSE
OF THE
NATURE OF
MY TALK...**



@LAZARINASTOY | #BRIGHTONSEO



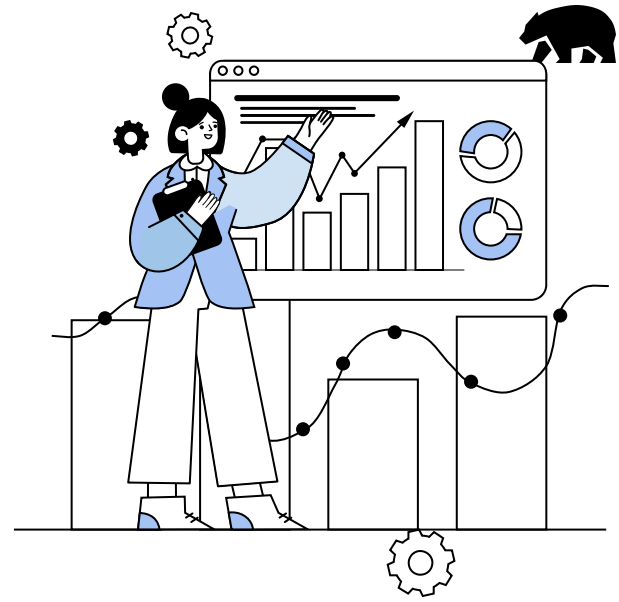
DISCLAIMER



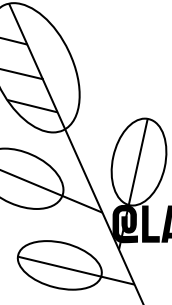
@LAZARINASTOY | #BRIGHTONSEO

I AM...

- ✓ SEO & Data Science Manager
@ **Intrepid Digital**
- ✓ Content-creator
- ✓ Growth & efficiency-obsessed




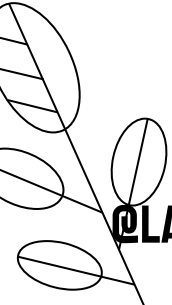
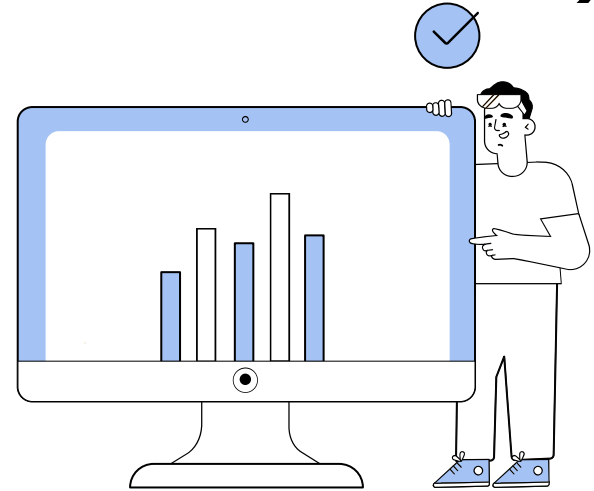
@LAZARINASTOY | #BRIGHTONSEO





I AM NOT...

- ✗ A developer
- ✗ A machine learning expert 
- ✗ An 'expert'
- ✗ A 'guru'





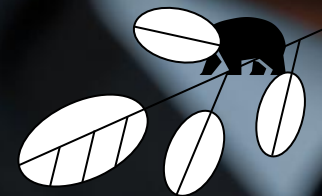
@LAZARINASTOY | #BRIGHTONSEO



**IN THE PAST 2 YEARS I'VE WORKED
AGENCY-SIDE
(WITH A BIT OF FREELANCE ON THE SIDE)**



@LAZARINASTOY | #BRIGHTONSEO



Quality

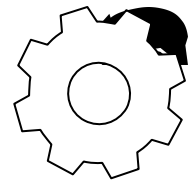
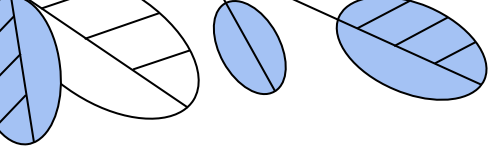
THE GOAL



THE CHALLENGE



THE OPPORTUNITY

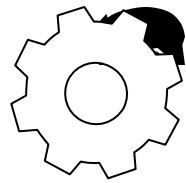
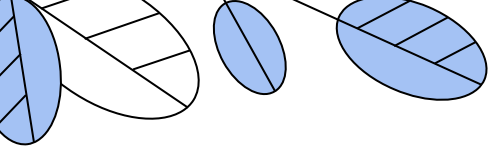


NO

**END-TO-END, CUSTOM-BUILT, GROUND-BREAKING,
NEW TECHNOLOGIES**

TODAY...

@LAZARINASTOY | #BRIGHTONSEO



ONLY A

TESTED CLAIM

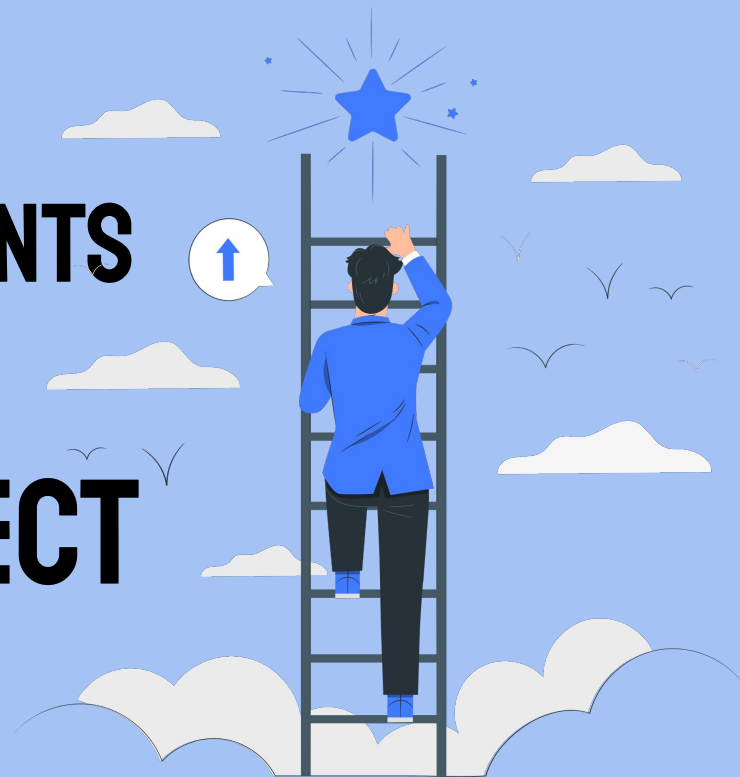
AND

SIMPLE PROMISE:

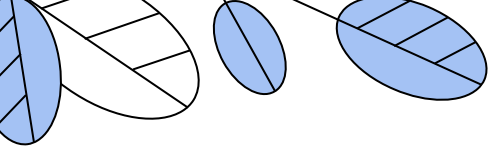
@LAZARINASTOY | #BRIGHTONSEO



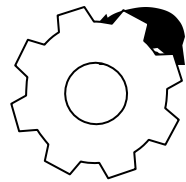
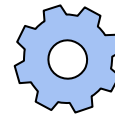
**ML-ENABLED
INCREMENTAL IMPROVEMENTS
CAN HAVE A
COMPOUNDING EFFECT
ON YOUR SEO IMPACT**



@LAZARINASTOY | #BRIGHTONSEO



IN JUST 20 MINUTES...



01

WHY

02

ANALYZE

03

CLUSTER

04

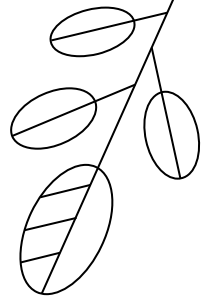
DISCOVER

05

LINK

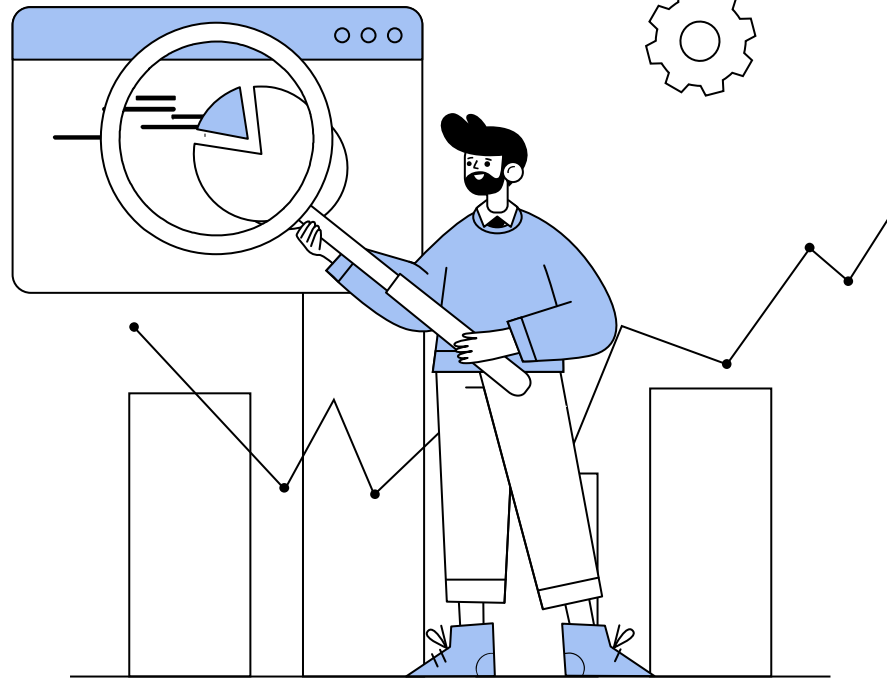
06

DELIVER



01

WHY BOTHER



@LAZARINASTOY | #BRIGHTONSEO



- Friends
- Projects
- Knowledge Graph
- Other

LET'S LOOK AT (RELEVANT) ALGORITHM HISTORY

@LAZARINASTOY | #BRIGHTONSEO



PAGERANK

Google's first algorithm relied (mostly) on links, **quality** and **quantity**.

How PageRank Works (A Simplified View)

PageRank is divided equally between the total number of links on a page.



Source: <https://ahrefs.com/blog/google-pagerank/>

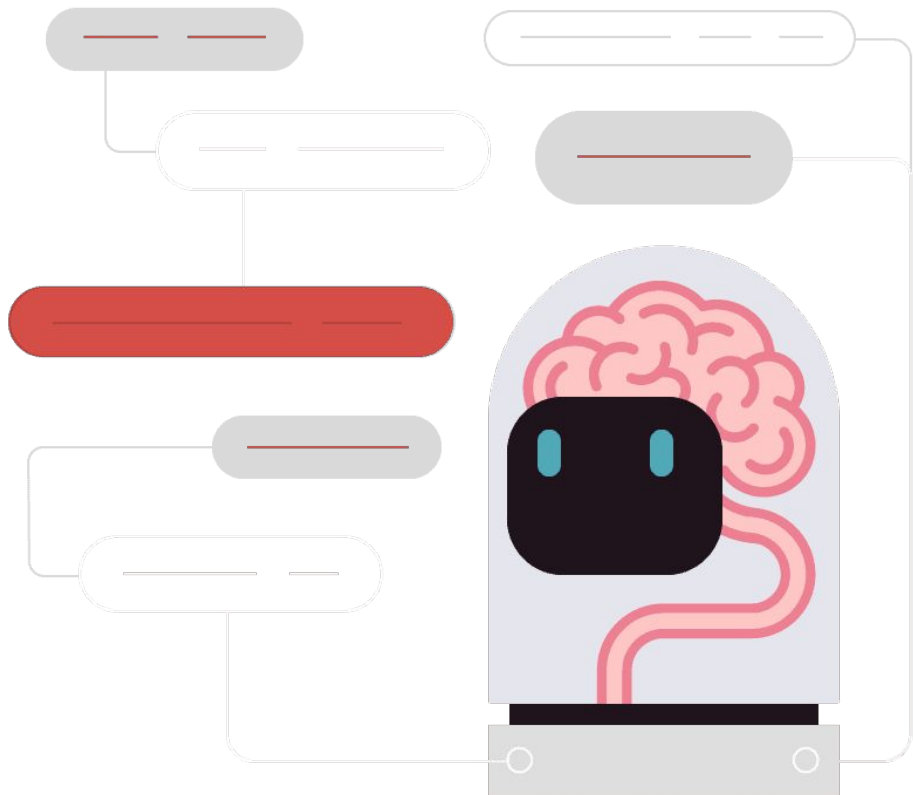
ahrefs

@LAZARINASTOY | #BRIGHTONSEO



RANKBRAIN

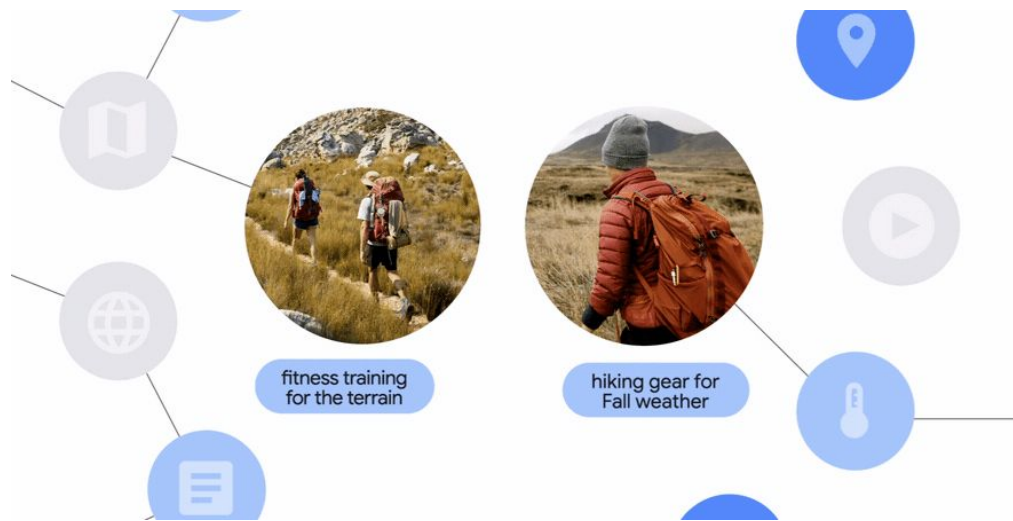
Focus on
intent-driven
structure and topic
comprehension.
Things over strings.

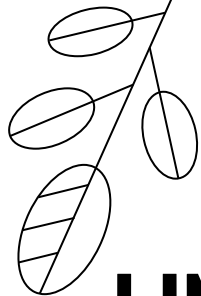




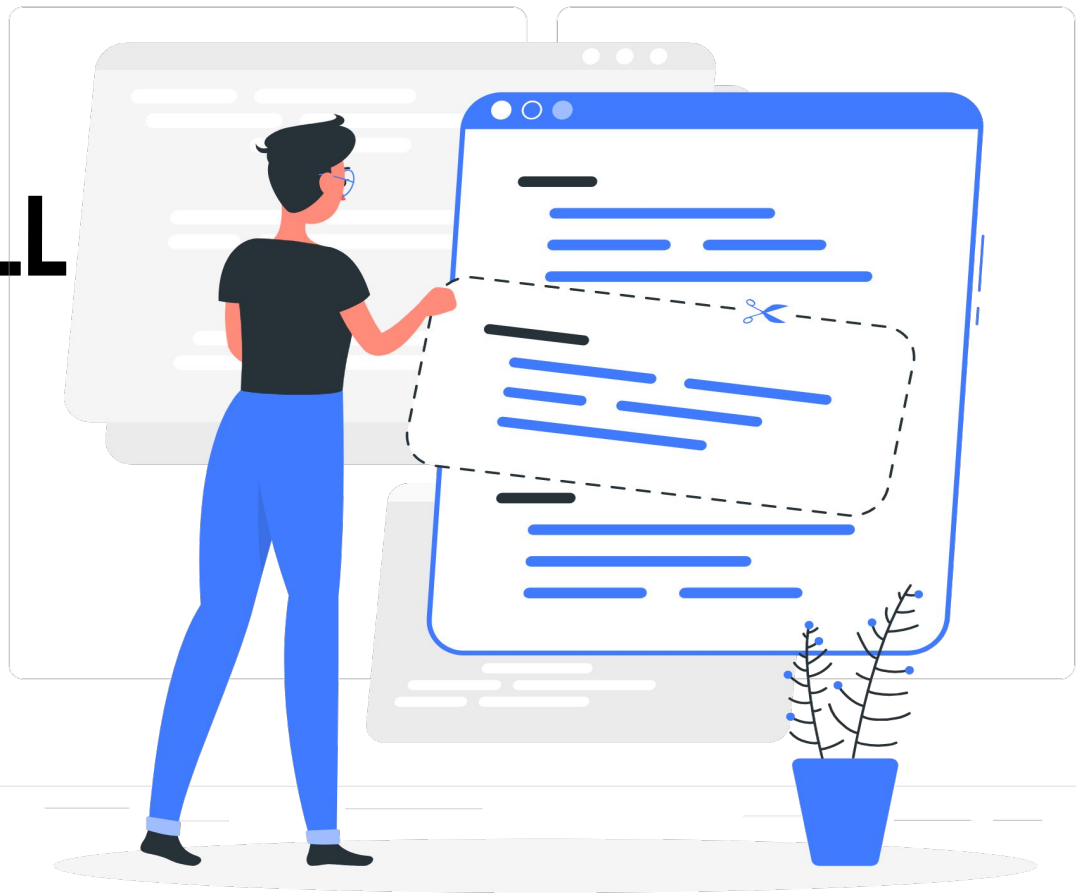
MUM

Surfacing subtopics,
semantic relationships,
multi-modal
comprehension





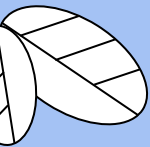
LINKS ARE STILL CRUCIAL FOR DISCOVERY (but...)



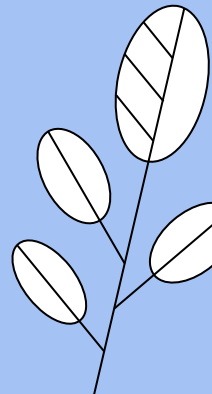
@LAZARINASTOY | #BRIGHTONSEO

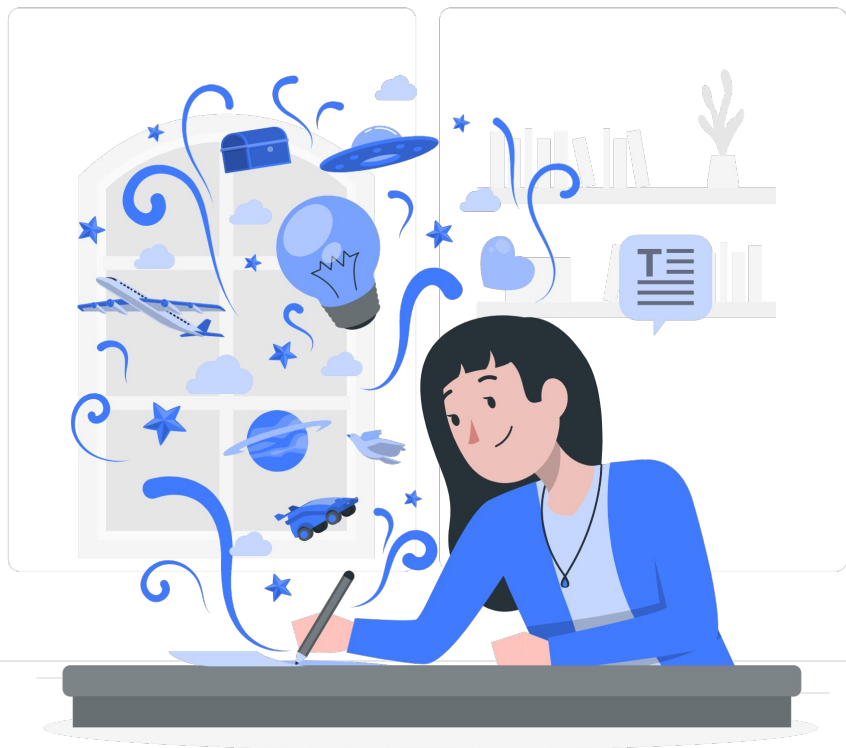


CONTENT IS KING.

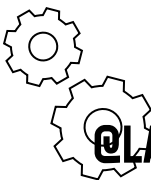


AZAPYASTOY | #BRIGHTONSEO

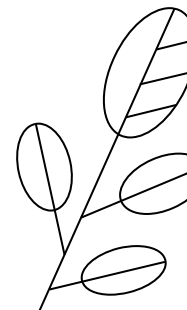


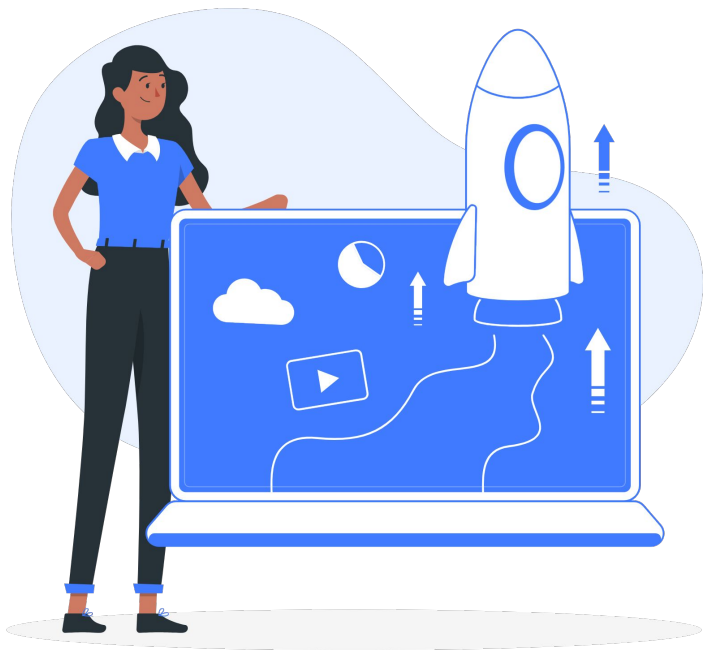


**WITHOUT GOOD CONTENT,
CONTEXT BECOMES
MEANINGLESS.**



@LAZARINASTOY | #BRIGHTONSEO

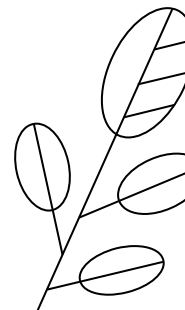




**IF THE QUALITY IS THERE,
CONTEXT BECOMES THE**

VESSEL

**TO PUSH THE CONTENT
FORWARD.**



@LAZARINASTOY | #BRIGHTONSEO



RELEVANT IN 2022?

48 MENTIONS OF INTERNAL LINKS.

17 FOR BACKLINKS.

@LAZARINASTOY | #BRIGHTONSEO



SEO in 2022

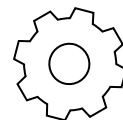
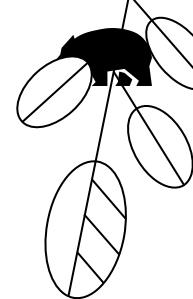


BY DAVID BAIN
AND
66 OF THE WORLD'S
LEADING SEO
EXPERTS

"It is hard to come across a resource these days that has input from so many brilliant and knowledgeable people in the SEO community. This is a must-read for anyone who is looking to take their SEO knowledge to the next level."

Joy Hawkins, Sterling Sky Inc

SEOIN2022.COM



IT'S WHAT YOU CAN CONTROL



HAZAR VASTOY | #BRIGHTONSEO





Aleyda Solis 🇺🇦 🟡🔵 @aleyda · Oct 12, 2020



What's the most frequent *high impact* but often overlooked Technical SEO issue that you see a lot across your projects happening again and again and try to fix asap?

In my case, it's likely poor **internal linking** of critical (priority to rank) pages! What's yours?



55



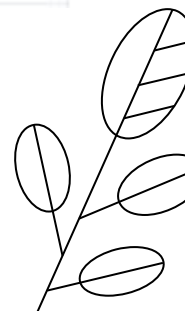
37



210



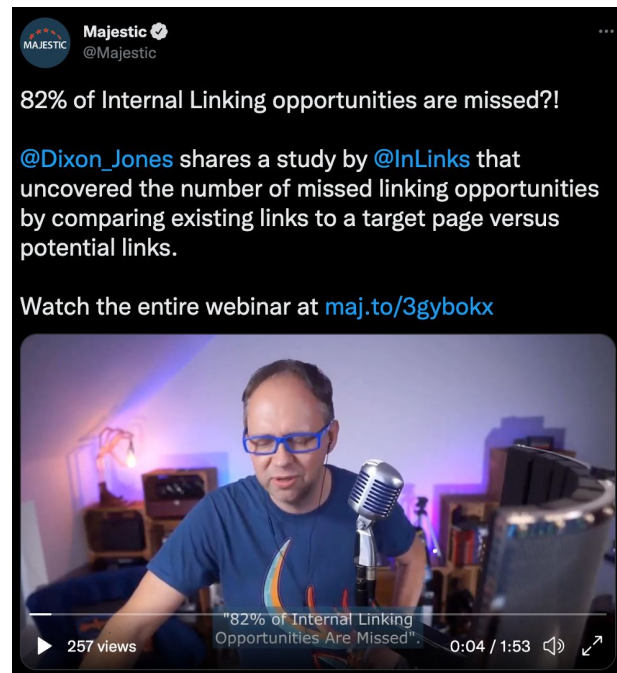
 @LAZARINASTOY | #BRIGHTONSEO





82 OF INLINK OPPORTUNITIES ARE MISSED

@LAZARINASTOY | #BRIGHTONSEO



A screenshot of a tweet from the account 'Majestic' (@Majestic). The tweet text reads: '82% of Internal Linking opportunities are missed?! @Dixon_Jones shares a study by @InLinks that uncovered the number of missed linking opportunities by comparing existing links to a target page versus potential links. Watch the entire webinar at maj.to/3gybokx'. Below the text is a video player showing a man with glasses and a blue t-shirt speaking into a vintage-style microphone. The video player interface includes a play button, '257 views', the title '"82% of Internal Linking Opportunities Are Missed"', a progress bar at '0:04 / 1:53', and volume and share icons.



John @JohnMu · J

Replying to @schachin
When I talk about site structure, I usually talk more about **internal linking**. Sometimes they overlap, and that's because everything else

3 8



John @JohnMu · Dec 8, 2020

Replying to @thinkofdave
Crawling and indexing is working as expected, nothing is suspended. The usual recommendations are to use a sitemap file and to have appropriate **internal linking**, so that we can recognize the importance of your new or updated pages.

2



John @JohnMu · Jun 25, 2017

Replying to @iPullRank @joeHall and @justinbriggs
Yes, if there's no content you wanted indexed there & your normal **internal linking** / crawlability is still fine.

2 1 2



John @JohnMu · Apr 3, 2021

Replying to @mohammadseo
Using a sitemap & also making it clear that these URLs are important within a website, through better **internal linking**, are good approaches. The problem with "pushing crawling" is that if Google doesn't want to crawl, they're not going to be shown in search anyway.

1 2



John @JohnMu · Apr 14, 2021

Replying to @maximilianarno
They're not crawled if they're blocked by robots.txt :-). They don't take away any of your site's crawl budget, but you might have better **internal linking** if you focused on crawlable & indexable URLs, so it can be good to check out.

2 1 3

Nov 12, 2021

...
... es that you need reprocessed quickly, I'd use the Google Search Console to resubmit them. Within a website, using the hreflang & link to what you really care about is also

20

B

seo

...
... /ling & indexing when we switch to a new site that doesn't have a clear **internal linking** strategy (but it's also something we

16

May 25, 2021

...
... website better -- and more crawling does not mean a sitemap file & **internal linking** can help us to understand, and which you consider important.

14



John @JohnMu ·

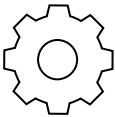
Replying to @Search_Bee
Sure! GA is based on what it crawls (usually :-)). However, you can set the preferred URLs by using **internal linking**.

3 2



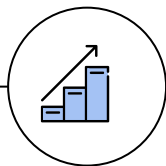
@LAZARINASTOY | #BRIGHTONSEO





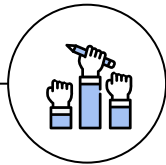
AN INTERNAL LINKING STRATEGY SHOULD AIM TO:

**ESTABLISH TOPIC
AUTHORITY**



**BOOST MONEY-MAKING
PAGES**

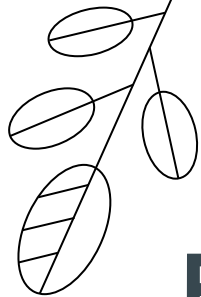
**ENRICH SEARCH
INTENT**



**BUILD SUBTOPIC
RELATIONSHIPS**



AND WHO DOESN'T NEED



**BETTER
STRUCTURE**



**IMPROVED
INDEXING**



**NAVIGATION
& UX**



**KEYWORDS
RANKED**

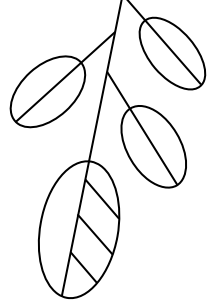


**TRAFFIC,
CTR**



**USER
ENGAGEMENT**





02

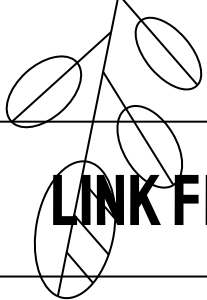
HOW TO ANALYZE INTERNAL LINKING STRUCTURE

@LAZARINASTOY | #BRIGHTONSEO



WHAT SHOULD AN INTERNAL LINK AUDIT ASSESS?

@LAZARINASTOY | #BRIGHTONSEO



DESCRIPTION

LINK FREQUENCY

- # inlinks / # unique inlinks per page
- Orphaned pages

**CRAWLING
PAGE DISCOVERY**

LINK QUALITY

- Link placement
- Source traffic
- Added link sections

**ENGAGEMENT
COMPREHENSION**

ANCHOR TEXT

- Type
- Length

**ENGAGEMENT
COMPREHENSION**

HIERARCHY

- Top pages
- Presence of topics, tags or categories
- Crawl depth

**ENGAGEMENT & UX
CRAWLING**

MISTAKES

- Broken links
- Nofollow tags
- UTM parameters

**CRAWLING
INDEXING
TRAFFIC**

MOST OF THIS CAN BE IDENTIFIED WITH A QUICK CRAWL & CLEVER DATA STUDIO TEMPLATE

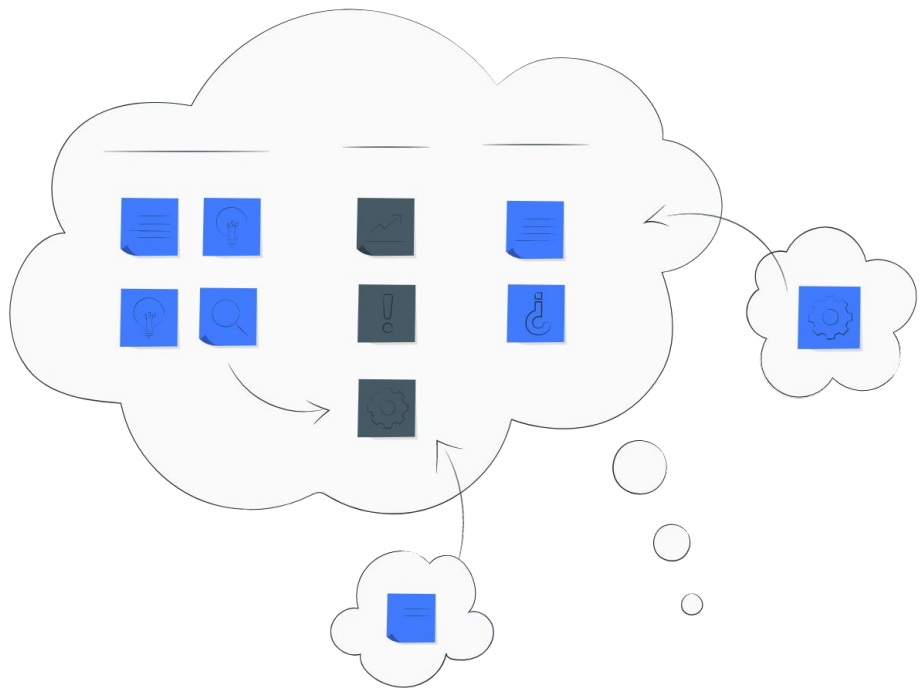
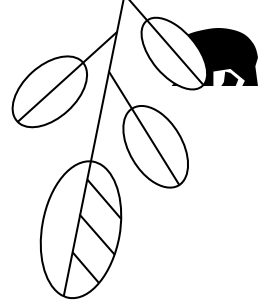
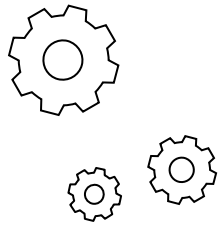


@LAZARINASTOY | #BRIGHTONSEO



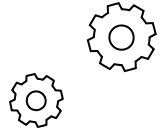
**BUT TO REALLY DIG DEEP YOU CAN
INCORPORATE MACHINE LEARNING
TO HELP YOU**

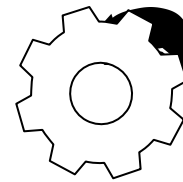
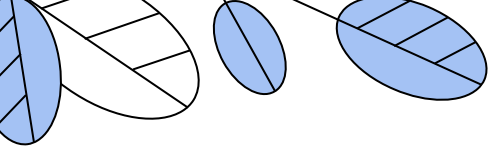
@LAZARINASTOY | #BRIGHTONSEO



CLASSIFY ANCHOR TEXT

@LAZARINASTOY | #BRIGHTONSEO



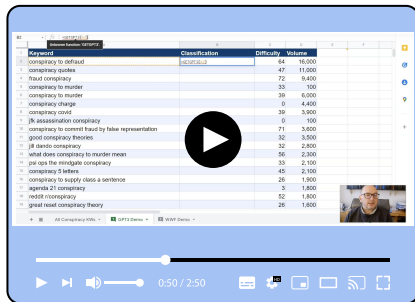


Classify Anchor Text N-Grams for Interlinking Insights with Python

[Home](#) / [Python](#) / [Automation](#) / [Classify Anchor Text N-Grams for Interlinking Insights with Python](#)



Greg Bernhardt

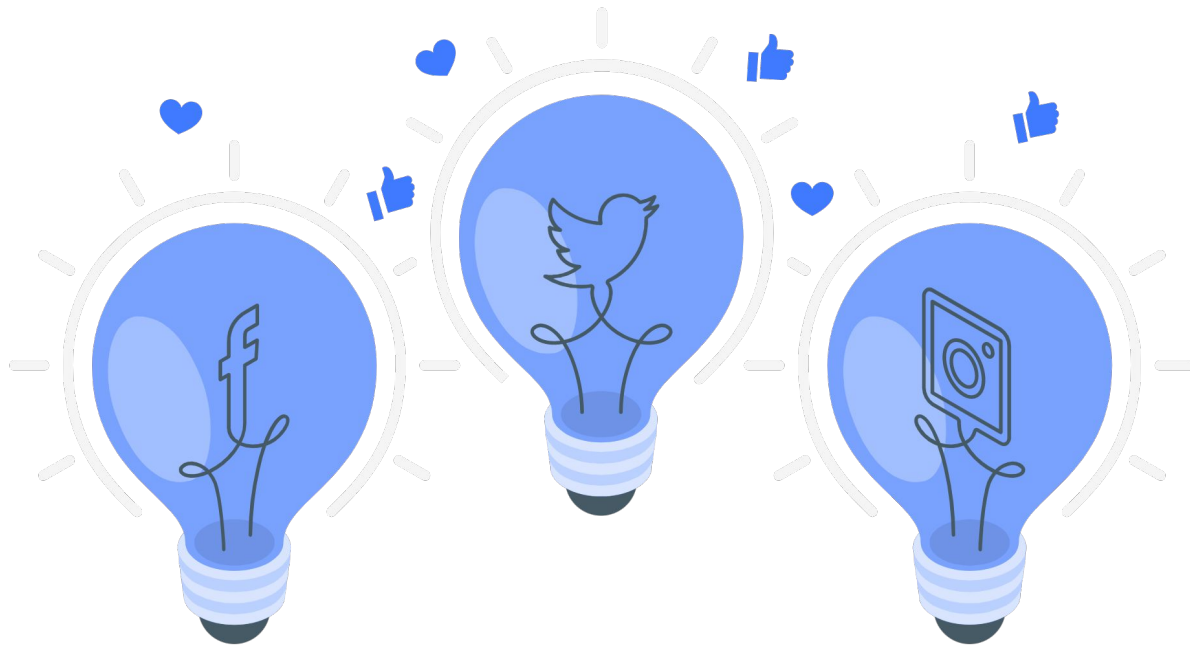


How to use GPT3 to classify Keywords



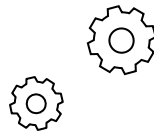
Richman SEO Training

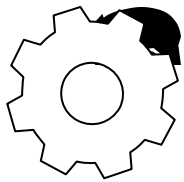
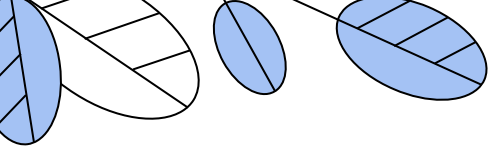
@LAZARINASTOY | #BRIGHTONSEO



IDENTIFY ENTITIES

@LAZARINASTOY | #BRIGHTONSEO





Analyze entities in a string

Inspect text for known entities (proper nouns such as public figures and landmarks), and return information about those entities.

[GO](#) [JAVA](#) [NODE.JS](#) [PHP](#) [PYTHON](#)

[View sample](#)

[View in documentation](#)

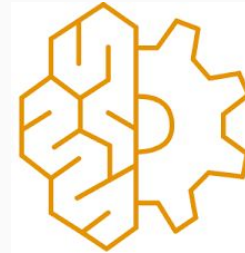
[Analyzing Entities](#)



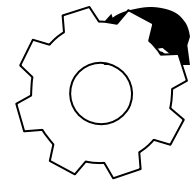
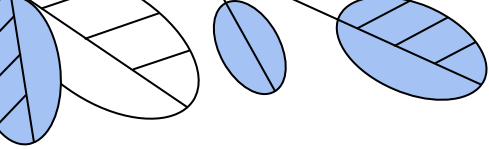
Sentiment Analysis & Entity Extraction with AWS Comprehend



Ram Vegiraju



@LAZARINASTOY | #BRIGHTONSEO



Person

p

Loc

l

Org

o

Event

e

Date

d

Other

z

Barack Hussein Obama II * (born August 4, 1961 *) is an American * attorney and

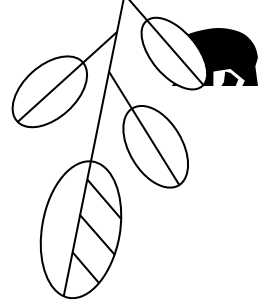
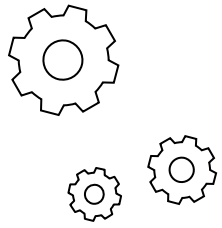
politician who served as the 44th President of the United States * from

January 20, 2009 *, to January 20, 2017 *. A member of the Democratic Party *, he

was the first African American * to serve as president. He was previously a

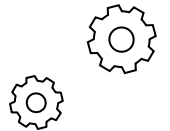
United States Senator * from Illinois * and a member of the Illinois State Senate *.

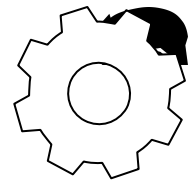
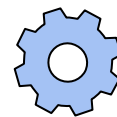
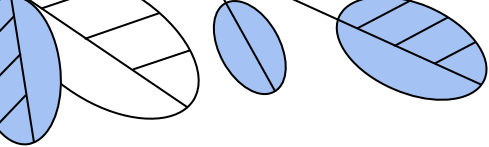
@LAZARINASTOY | #BRIGHTONSEO



ASSESS ENTITY RELATIONSHIPS

@LAZARINASTOY | #BRIGHTONSEO

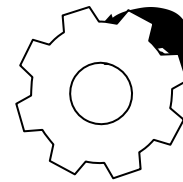
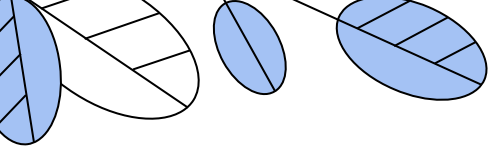




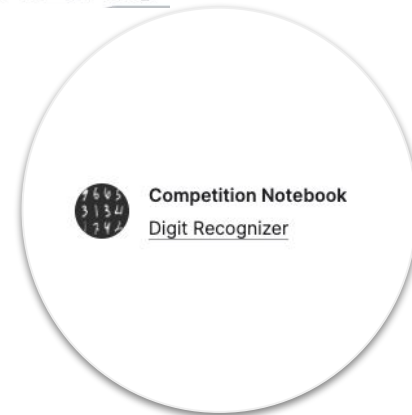
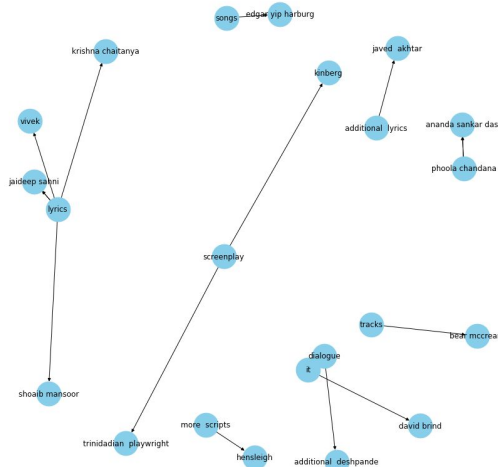
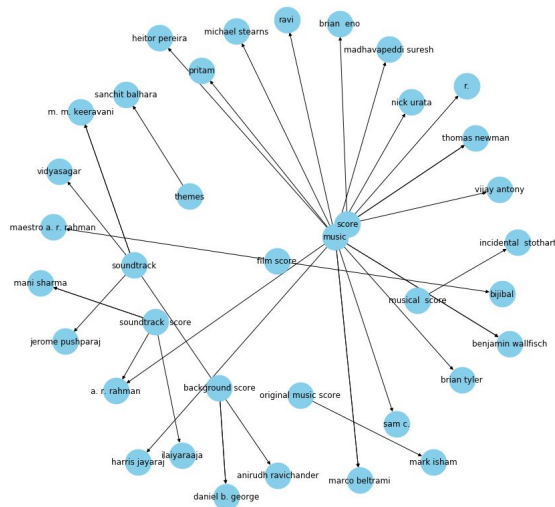
 neo4j

 KgBase

Google Knowledge
Graph Search API



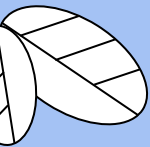
Knowledge Graph & NLP Tutorial-(BERT,spaCy,NLTK)



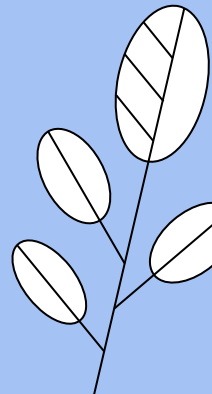
@LAZARINASTOY | #BRIGHTONSEO



MORE INSIGHTS



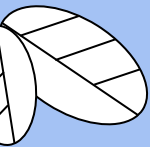
AZAPYASTOY | #BRIGHTONSEO



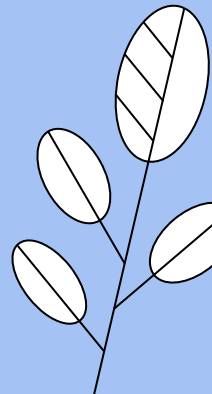


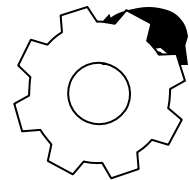
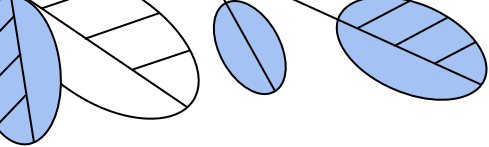
BETTER

RECOMMENDATIONS



AZAPYASTOY | #BRIGHTONSEO



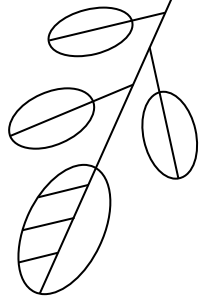


SEE YOUR CONTENT AND LINKS

THE WAY SEARCH ENGINES SEE IT

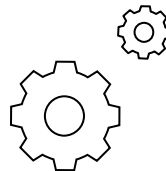
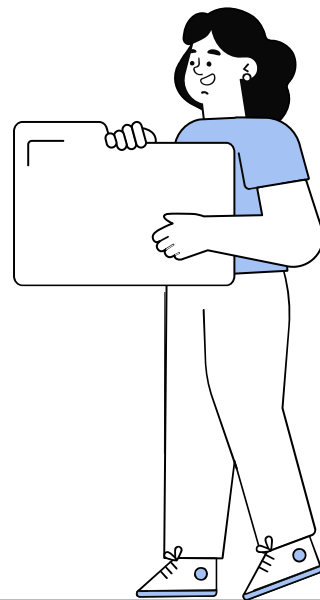
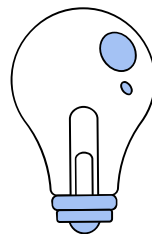
TO UNDERSTAND HOW TO IMPROVE IT.

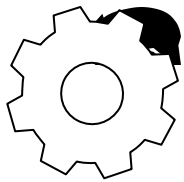
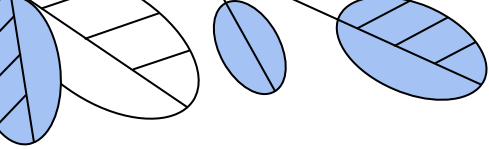
@LAZARINASTOY | #BRIGHTONSEO



03

CLUSTER CONTENT WITH ML

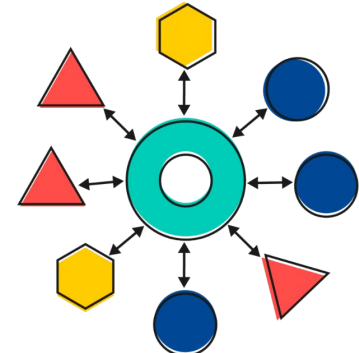
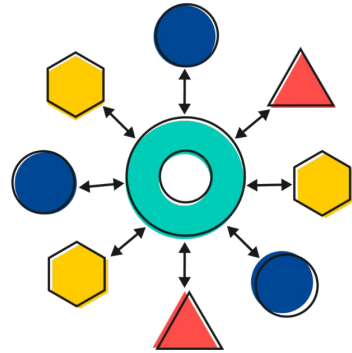




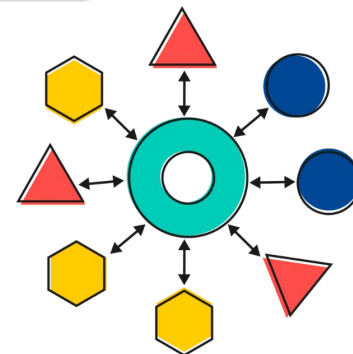
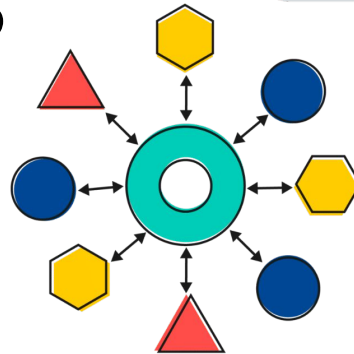
SITECARE

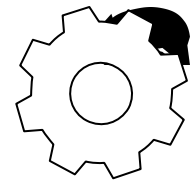
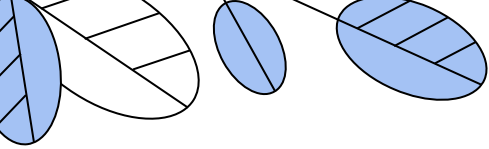
Content Clusters

- Cluster Content
- Internal Linking
- Pillar Page

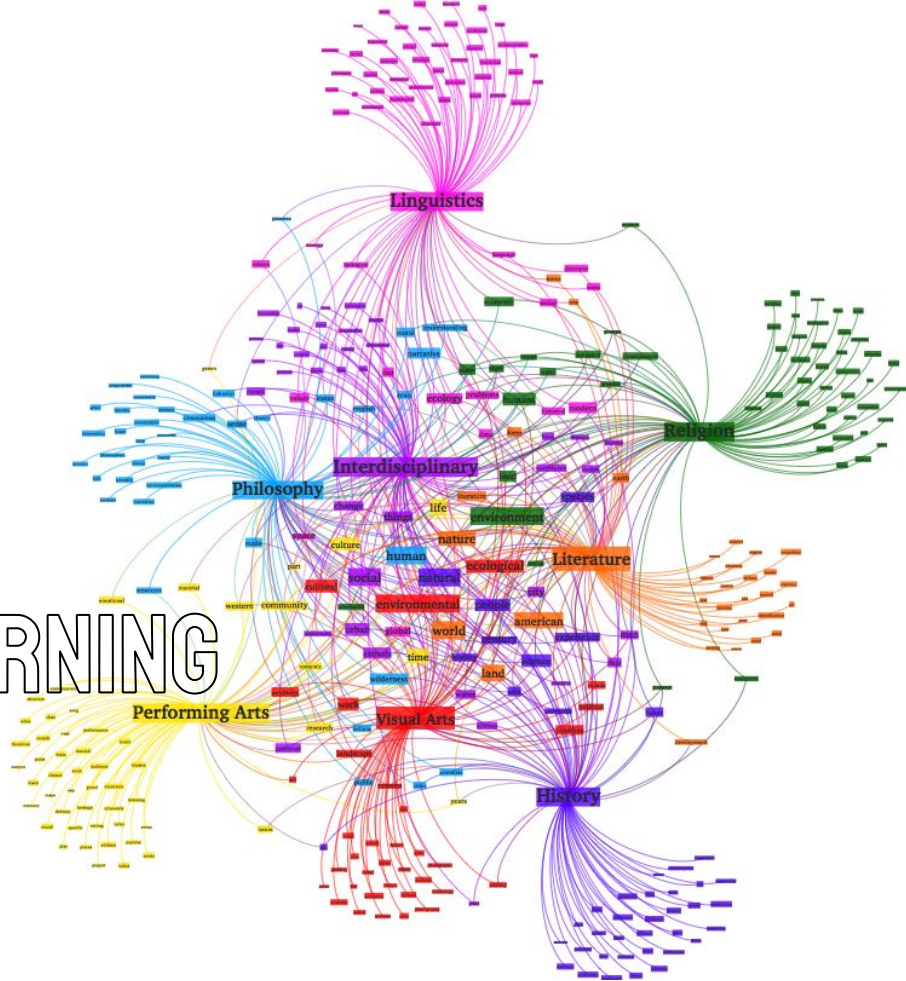


CONTENT CLUSTERS IN SEO





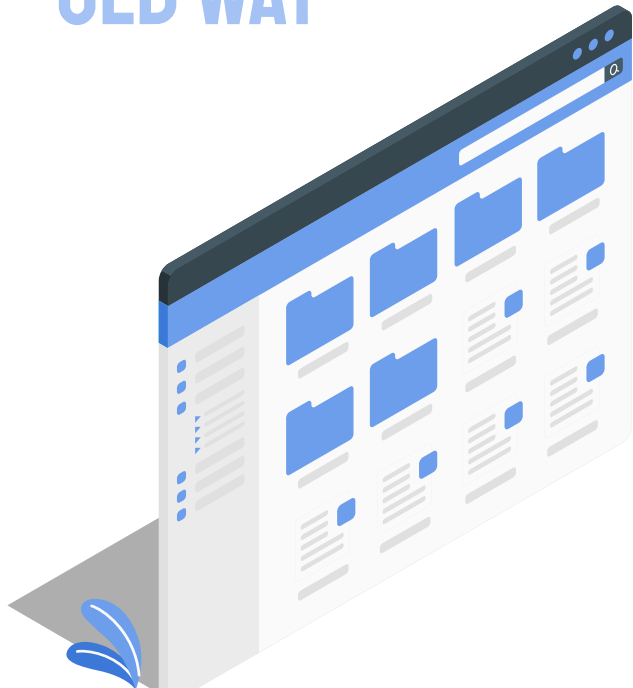
TOPIC MODELS IN MACHINE LEARNING



@LAZARINASTOY | #BRIGHTONSEO

THE FUNDAMENTALS ARE THE SAME

OLD WAY



NEW WAY

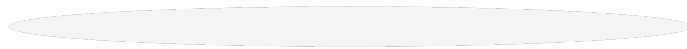


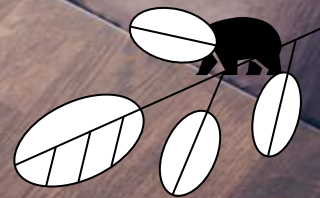
@LAZARINASTOY | #BRIGHTONSEO





HOW DID TOPIC MODELING COME ABOUT?



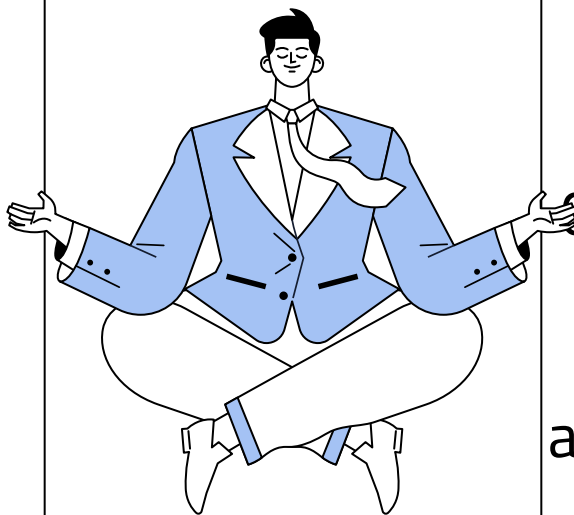


**ACCORDING TO THE AUTHOR
OF THE FIRST TOPIC MODELING ALGORITHM**



PROBLEM

Collective knowledge
increases
→ information
discovery gets more
difficult



SOLUTION

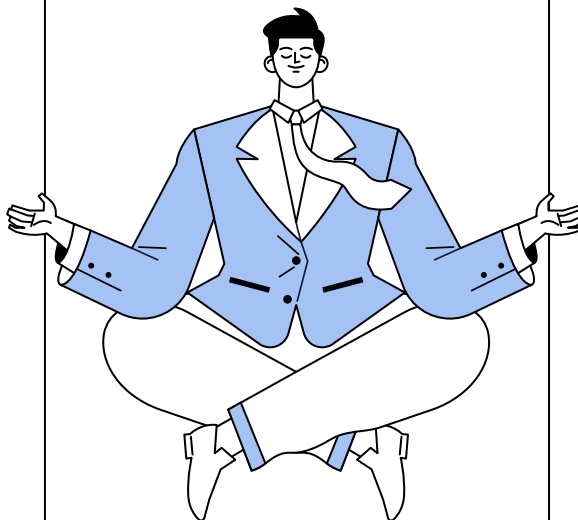


Computational tools to
help organize, search,
and understand vast
amounts of information



PROBLEM

Current systems of query & links are good, but can be **limiting**



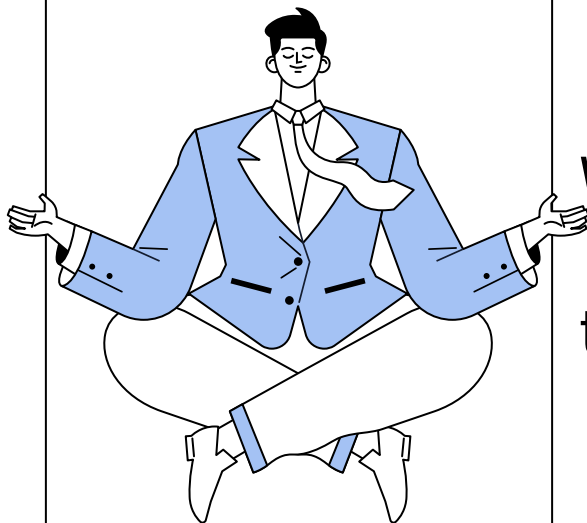
SOLUTION

We need systems that organise documents based on themes (topics) and subtopics.



PROBLEM

No ability to zoom in and out of topics of interest, highlighting patterns.



SOLUTION

We need topic models that operate under the assumption that a document has multiple topics.



Topics

gene	0.04
dna	0.02
genetic	0.01
...	

life	0.02
evolve	0.01
organism	0.01
...	

brain	0.04
neuron	0.02
nerve	0.01
...	

data	0.02
number	0.02
computer	0.01
...	

Documents

Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK—How many **genes** does an **organism** need to **survive**? Last week at the genome meeting here,* two genome researchers with radically different approaches presented complementary views of the basic genes needed for **life**. One research team, using **computer** analyses to compare known **genomes**, concluded that today's **organisms** can be sustained with just 250 genes, and that the earliest life forms required a mere 128 **genes**. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough. Although the numbers don't match precisely, these **predictions**

* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

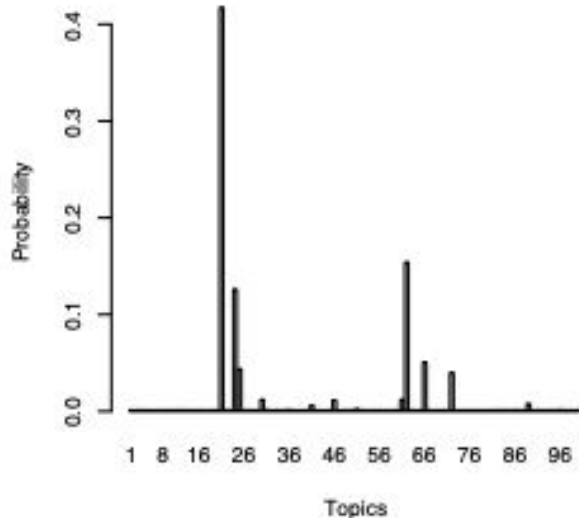
SCIENCE • VOL. 272 • 24 MAY 1996

Stripping down. Computer analysis yields an estimate of the minimum modern and ancient genomes.

Topic proportions and assignments



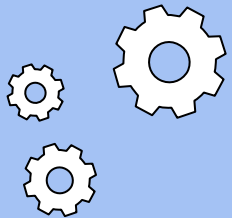
Introduction to Probabilistic Topic Models (Blei, 2012)



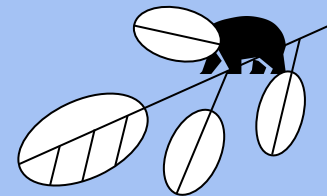
“Genetics”	“Evolution”	“Disease”	“Computers”
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations



[Introduction to Probabilistic Topic Models \(Blei, 2012\)](#)



“



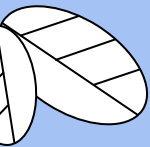
TOPIC MODELING IS

PATTERN RECOGNITION

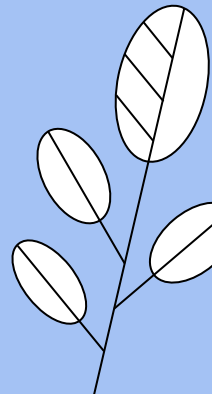
IN LARGE, TEXT-BASED CORPUSES OF DATA.



IT WAS MADE FOR US.



AZAPYASTOY | #BRIGHTONSEO



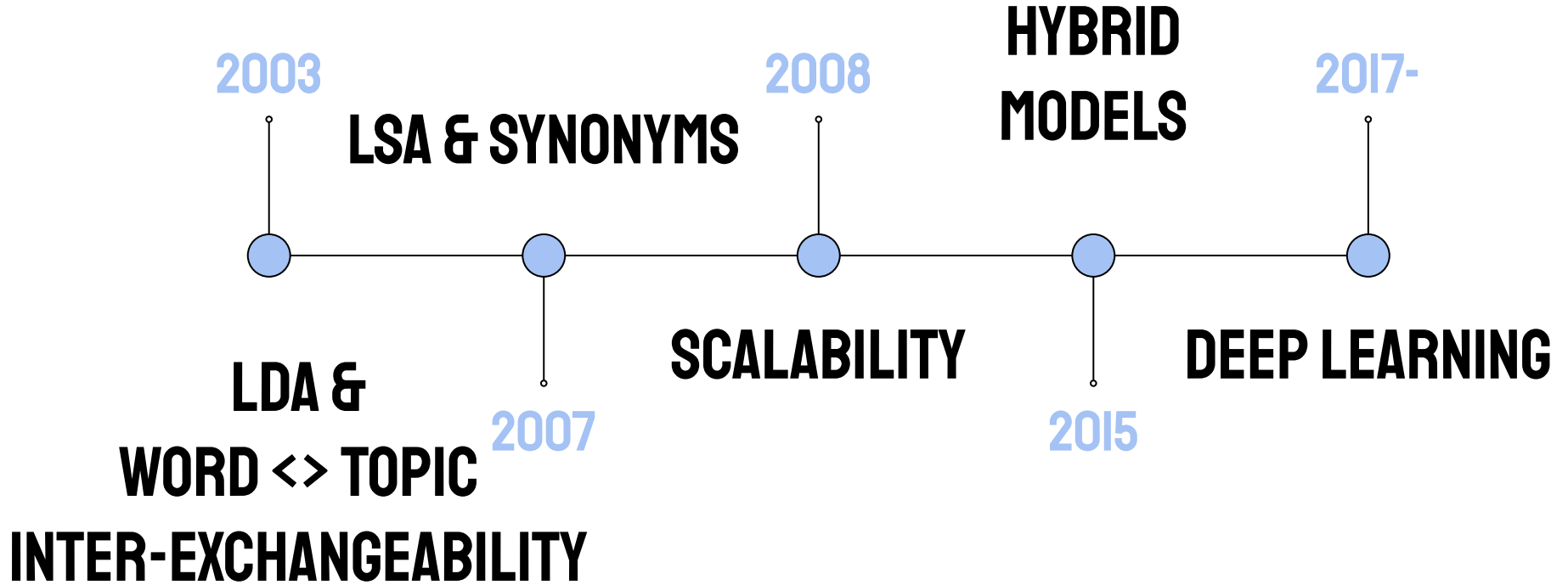


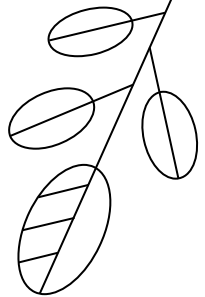
**IN THE PAST 20 YEARS,
A LOT HAS HAPPENED**

dison/Wabash



@LAZARINASTOY | #BRIGHTONSEO

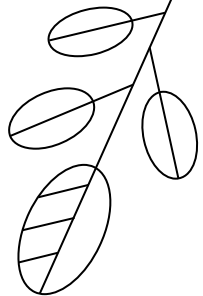




**BUT THE PROBLEM
WE FACE REMAINS
THE SAME**



@LAZARINASTOY | #BRIGHTONSEO



SO, WHY NOT TRY AN ML-BASED SOLUTION?

@LAZARINASTOY | #BRIGHTONSEO

TWO METHODS TO KICKSTART THIS JOURNEY

LDA

Topic Modeling in Python: Latent Dirichlet Allocation (LDA)



Shashank Kapadia

LSA

Latent Semantic Analysis using Python

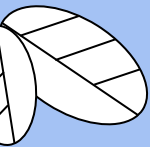


Avinash Navlani

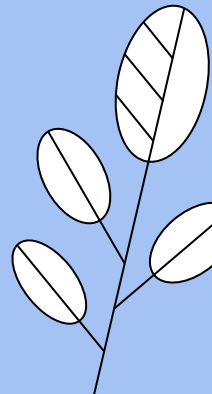
@LAZARINASTOY | #BRIGHTONSEO



OR, SKIP THE CODE AND JUMP STRAIGHT IN



AZAPYASTOY | #BRIGHTONSEO

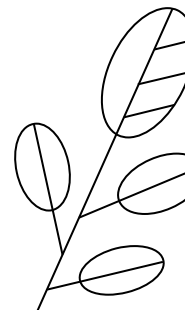




WATCH THE DETAILS LATER

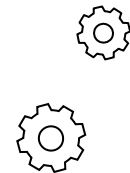
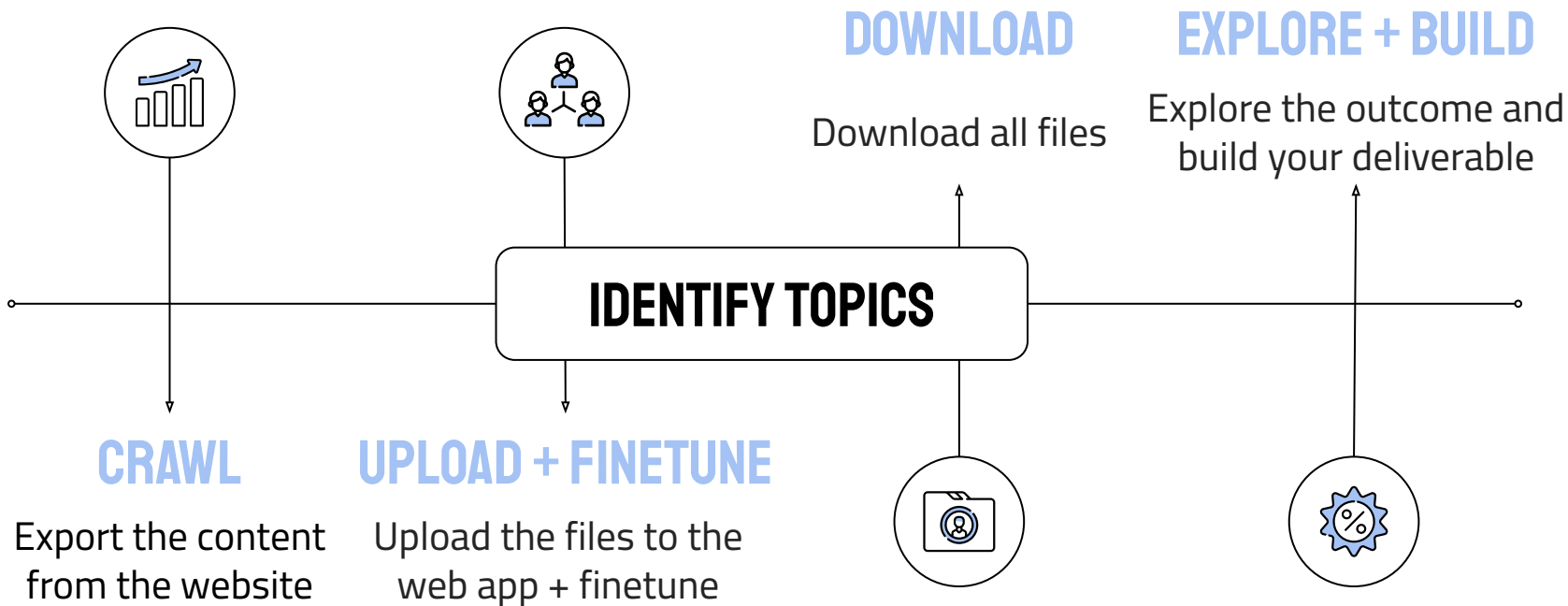
I've recorded a step-by-step tutorial on doing **topic modelling** using a no-code, publicly-available, web-based app using LDA.

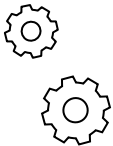
@LAZARINASTOY | #BRIGHTONSEO





BUT IN A NUTSHELL...





TOPIC MODELS

TOPIC MODELS

A	B	C	D	E	F	G	H	I	J	K
	health mental staff workplace home support group working eh students	corporate volunteering giving purpose social responsibility grawehr stéphanie	nonprofit volunteers time content media support volunteer nonprofits share form	alaya platform data user services users policy information conditions general	csr business social companies strategy responsibility corporate initiatives	volunteering program employees corporate giving programs matching benefits	people company time it's back feel start make mission that's	employees impact engage purpose activities make platform community	engagement work team employees good teams engaged make virtual	nonprofits season donors carmen amell nonprofit fundraising make strategy story
	0.00%	17.65%	-44.77%		-16.00%	-104.73%	-4.67%	9.29%	-5.99%	
	17.65%	0.00%	-21.13%	-88.64%	-8.61%	-62.86%	-36.37%	-23.42%	27.18%	6.80%
	-44.77%	-21.13%	0.00%	9.88%	-26.01%	-58.78%	12.52%	16.87%	-56.55%	70.50%
		-88.64%	9.88%	0.00%	-30.66%		-157.96%	-53.92%		
	-16.00%	-8.61%	-26.01%	-30.66%	0.00%	60.18%	54.64%	26.11%	-0.59%	-218.54%
	-104.73%	-62.86%	-58.78%		60.18%	0.00%	35.22%	30.17%	-0.59%	-77.01%
	-4.67%	-36.37%	12.52%	-157.96%	54.64%	35.22%	0.00%	28.20%	27.18%	-46.26%
	9.29%	-23.42%	16.87%	-53.92%	26.11%	30.17%	28.20%	0.00%	-1.53%	-40.31%
	-5.99%	27.18%	-56.55%		-0.59%	-0.59%	27.18%	-1.53%	0.00%	-29.35%
		6.80%	70.50%		-218.54%	-77.01%	-46.26%	-40.31%	-29.35%	0.00%

Topic to Topic Similarity ▾

Topic Modelling per Page ▾

Topic to Topic Similarity ▾

@LAZARINASTOY | #BRIGHTONSEO



PAGE INFO

TOPIC MODELS

Content export	Address	normalised title	[Blurred Content]											
			0.00%	0.00%	50.00%	0.00%	0.00%	50.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
			0.00%	0.00%	1.49%	0.37%	0.00%	20.07%	1.86%	33.09%	0.00%	0.00%	0.00%	0.00%
			1.33%	0.00%	2.00%	0.36%	5.74%	18.55%	6.20%	10.51%	7.64%	0.00%	0.00%	0.00%
			0.00%	0.00%	0.00%	0.00%	0.60%	14.83%	1.40%	38.48%	1.20%	0.00%	0.00%	0.00%
			0.00%	48.53%	0.74%	0.00%	0.00%	0.00%	0.00%	0.00%	8.09%	4.04%	0.00%	0.00%
			0.00%	0.00%	9.38%	0.00%	0.00%	0.00%	0.00%	3.13%	0.00%	53.13%	0.00%	0.00%
			0.00%	3.51%	15.59%	2.10%	0.86%	2.73%	6.63%	5.22%	0.00%	8.96%	0.00%	0.00%
			2.99%	0.48%	4.31%	1.08%	4.55%	1.20%	7.19%	6.23%	15.81%	0.96%	0.00%	0.00%
			6.50%	3.58%	3.17%	0.41%	6.81%	7.16%	7.98%	4.35%	15.05%	0.10%	0.00%	0.00%
			9.84%	7.81%	9.12%	1.74%	3.91%	0.00%	5.64%	9.41%	2.03%	0.00%	0.00%	0.00%
			0.09%	1.29%	0.76%	0.09%	25.68%	5.38%	8.90%	5.47%	4.09%	0.00%	0.00%	0.00%
			4.12%	1.17%	9.48%	1.58%	1.37%	2.68%	5.36%	4.53%	0.76%	15.52%	0.00%	0.00%
			1.40%	0.97%	7.97%	2.37%	3.47%	4.50%	5.36%	4.81%	1.58%	7.61%	0.00%	0.00%
			0.00%	2.53%	27.09%	2.61%	2.61%	0.00%	3.15%	6.45%	2.23%	8.21%	0.00%	0.00%
			0.00%	49.14%	0.00%	0.00%	0.00%	0.00%	0.69%	0.00%	9.62%	5.84%	0.00%	0.00%
			0.00%	5.18%	8.83%	6.91%	7.49%	1.92%	4.41%	15.36%	0.58%	0.58%	0.00%	0.00%
			0.00%	0.00%	43.24%	0.00%	0.00%	0.00%	2.70%	0.00%	0.00%	18.92%	0.00%	0.00%
			0.00%	19.63%	8.89%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	32.59%	0.00%	0.00%
			0.00%	62.77%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
			7.07%	51.52%	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	5.39%	1.68%	0.00%	0.00%
			4.34%	1.42%	5.54%	0.78%	2.35%	0.28%	10.73%	3.98%	0.57%	13.15%	0.00%	0.00%
			0.00%	54.29%	2.86%	2.86%	2.86%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
			1.93%	4.16%	12.01%	1.97%	6.22%	13.46%	4.72%	4.25%	2.19%	0.00%	0.00%	0.00%
			2.59%	4.25%	3.22%	0.00%	5.97%	3.14%	7.63%	12.74%	12.50%	0.39%	0.00%	0.00%
			2.03%	1.92%	1.05%	0.17%	23.55%	8.31%	14.42%	2.56%	3.43%	0.00%	0.00%	0.00%

+ [Menu] Search Intent Matching ▾ Topic to Topic Similarity ▾ **Topic Modelling per Page ▾** Other Linking Opportunities (3N ▾ Cluster 1 ▾ Cluster 2 ▾ Cluster 3 ▾

Topic Modelling per Page ▾

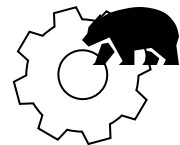




Selected Topic: 0 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:(2)

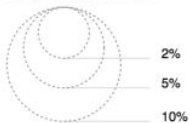
$\lambda = 1$



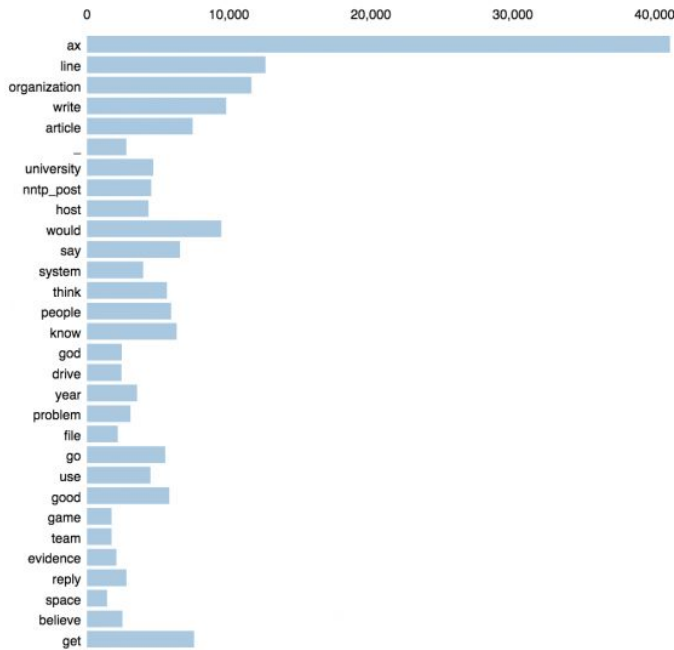
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Salient Terms¹



Overall term frequency

Estimated term frequency within the selected topic

1. $saliency(\text{term } w) = \text{frequency}(w) \cdot [\sum_t p(t|w) \cdot \log(p(t|w)/p(t))]$ for topics t ; see Chuang et. al (2012)

2. $relevance(\text{term } w | \text{topic } t) = \lambda \cdot p(w|t) + (1 - \lambda) \cdot p(w|t)/p(w)$; see Sievert & Shirley (2014)





**JUST LOOK AT
WHAT YOU
ACHIEVED ✨**



@LAZARINASTOY | #BRIGHTONSEO

SAVED A TON OF TIME

Took you **less than 30 minutes**. Hooray!



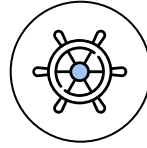
@LAZARINASTOY | #BRIGHTONSEO



BASELINE OVERVIEW

You've got a clear overview of the **main topic clusters** and their **relationship**





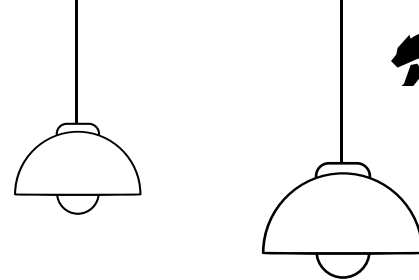
TESTED SOMETHING NEW

Scrutinize the output.

Tell your client how you have improved on it.

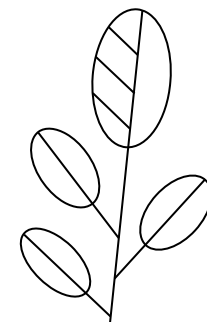
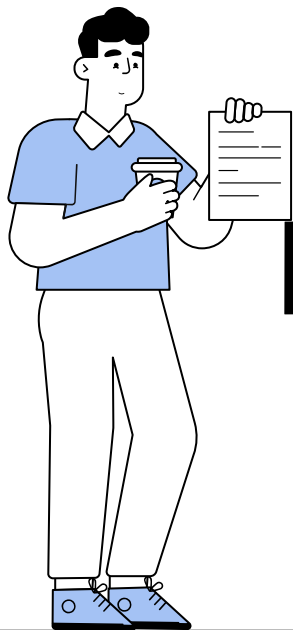
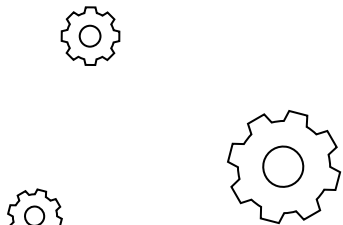


@LAZARINASTOY | #BRIGHTONSEO



04

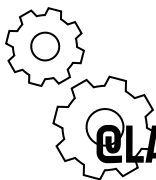
DISCOVER OTHER OPPORTUNITIES



@LAZARINASTOY | #BRIGHTONSEO



ENTER FUZZY MATCHING



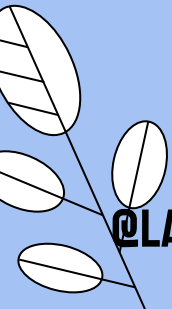
@LAZARINASTOY | #BRIGHTONSEO



A QUICK AND DIRTY WAY

FOR CALCULATING

SIMILARITY BETWEEN TWO STRINGS



@LAZARINASTOY | #BRIGHTONSEO

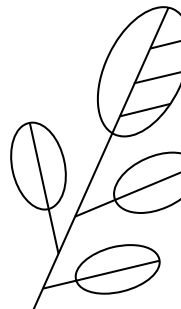


WATCH THE DETAILS LATER

I've recorded a step-by-step tutorial on using **fuzzy matching** for things like:

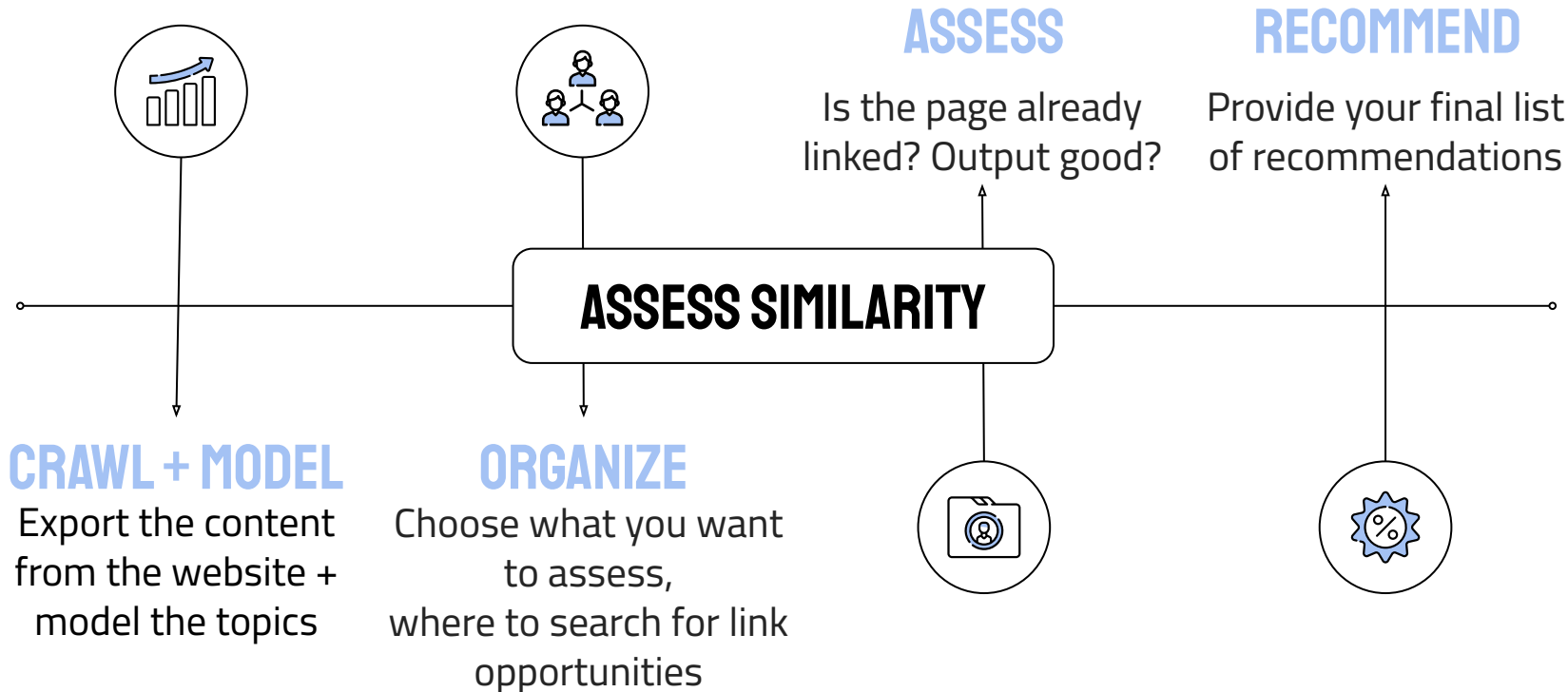
- Identifying link opportunities
- String Similarity Analysis
- Other cool things

 @LAZARINASTOY | #BRIGHTONSEO





BUT IN A NUTSHELL...





FROM /

TO

SIMILARITY SCORE

From	Address	To	Address	Similarity	Does a link already exist on this page to the target?
				72.80%	NO
				73.50%	NO
				68.70%	NO
				77.00%	NO
				90.30%	YES
				78.20%	YES
				71.00%	YES
				100.00%	YES
				78.20%	YES
				87.20%	YES
				93.40%	YES
				70.60%	YES
				83.50%	YES
				70.20%	YES
				69.90%	YES
				2.30%	NO
				30.90%	NO
				90.30%	YES
				62.70%	YES
				86.20%	YES
				38.20%	YES
				24.20%	YES

CHECK IF LINK EXISTS

```
=ARRAY_CONSTRAIN(ARRAYFORMULA(IFerror(IF(MATCH(B2&D2,'Copy of inlinks export'!B:B&'Copy of inlinks export'!C:C, 0), "YES"), "NO")), 1, 1)
```



Other Linking Opportunities (3N ▾)

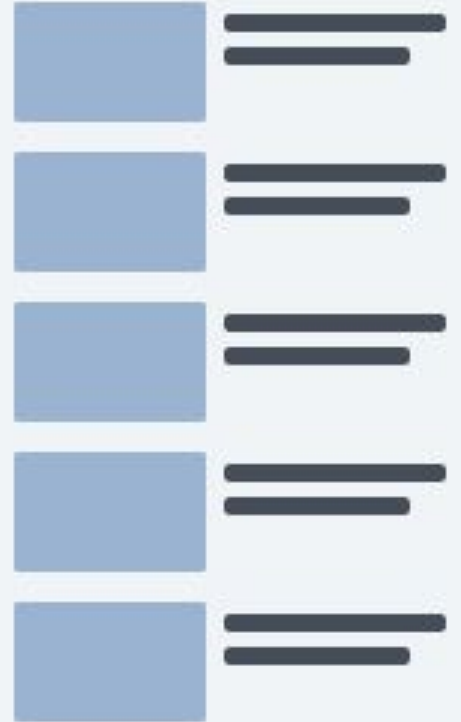


COVER YOUR BASES

@LAZARINASTOY | #BRIGHTONSEO



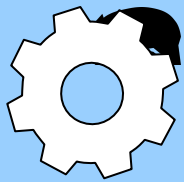
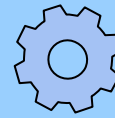
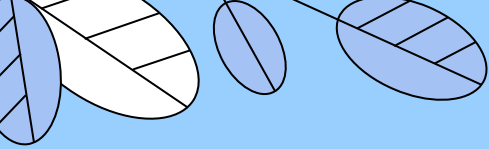
MOST POPULAR



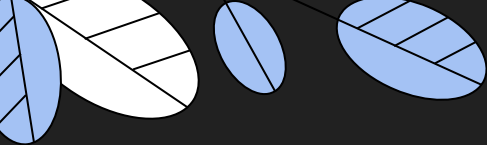
RELATED ARTICLES



@LAZARINASTOY | #BRIGHTONSEO



Home >> MY ACCOUNT >> PROFILE CHANGE PASSWORD



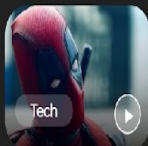
About Us



Phlox Is Your News, Entertainment, Music Fashion Website. We Provide You With The Latest Breaking News And Videos Straight From The Entertainment Industry



Popular Posts



Review Of Healthy Breakfast Meals For...

2 weeks ago | 30,451 views | 20:45



Review Of Healthy Breakfast Meals For...

2 weeks ago | 30,451 views | 20:45



Review Of Healthy Breakfast Meals For...

2 weeks ago | 30,451 views | 20:45

Most View



Review Of Healthy Breakfast...

2 weeks ago | 30,451 views | 20:45



Review Of Healthy Breakfast...

2 weeks ago | 30,451 views | 20:45



Review Of Healthy Breakfast...

1 month ago | 30,451 views | 20:45

Category

- Top Ads 15
- Gadgets 14
- Telephone 13
- Electronic 24
- Game Play 37
- Best Of 12
- Animations 19
- Series 26
- Cars 17



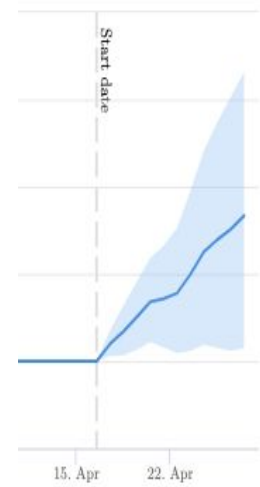
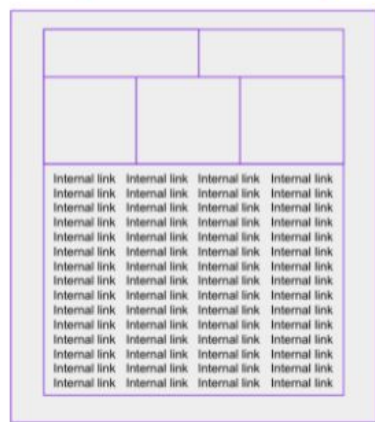
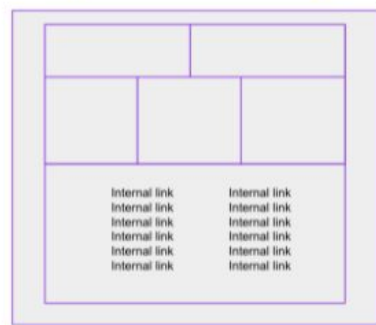
+10%

+5%

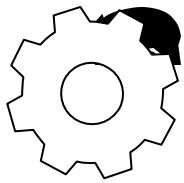
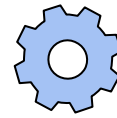
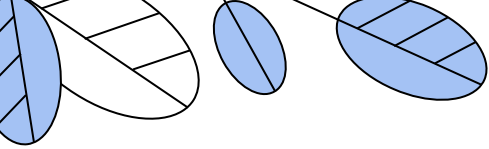
DESKTOP MOBILE

Control

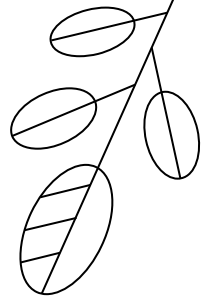
Variant



[Searchpivot Case Study](#)



~~RESEARCH MORE.~~



05

HOW TO FIND WHERE TO LINK





ML pages. The results can be seen in the **Custom** Tab. See our [User Guide](#) for examples.

nylon strings

Enter Regular Expression

Filter 3 **Contains** Enter Regular Expression

Filter 4 **Contains** Enter Regular Expression

Filter 5 **Contains** Enter Regular Expression

Filter 6 **Contains** Enter Regular Expression

Filter 7 **Contains** Enter Regular Expression

Filter 8 **Contains** Enter Regular Expression

Filter 9 **Contains** Enter Regular Expression

Filter 10 **Contains** Enter Regular Expression

Clear All Filters

OK

**FOR SMALL SITES /
INTERNAL LINK
INITIATIVES**

@LAZARINASTOY | #BRIGHTONSEO



The following query searches across all columns of the `Recipes` table for the value `toast` and returns the rows that contain this value.

```
SELECT * FROM Recipes WHERE CONTAINS_SUBSTR(Recipes, 'toast');
```

Breakfast	Lunch	Dinner
Potato pancakes	Toasted cheese sandwich	Beef stroganoff
Avocado toast	Tomato soup	Blueberry samon

**FOR REALLY BIG
SITES / INTERNAL
LINK INITIATIVES**

The following query searches the `Lunch` and `Dinner` columns of the `Recipe` table for the value `potato` and returns the row if either column contains this value.

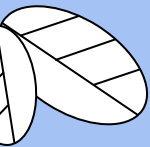
```
SELECT * FROM Recipes WHERE CONTAINS_SUBSTR((Lunch, Dinner), 'potato');
```

Breakfast	Lunch	Dinner
Blueberry pancakes	Egg salad sandwich	Potato dumplings
Corned beef hash	Lentil potato soup	Glazed ham

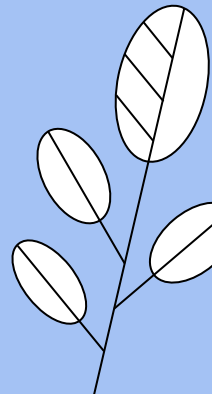
@LAZARINASTOY | #BRIGHTONSEO



THE LOGIC IS THE SAME



AZAPYASTOY | #BRIGHTONSEO





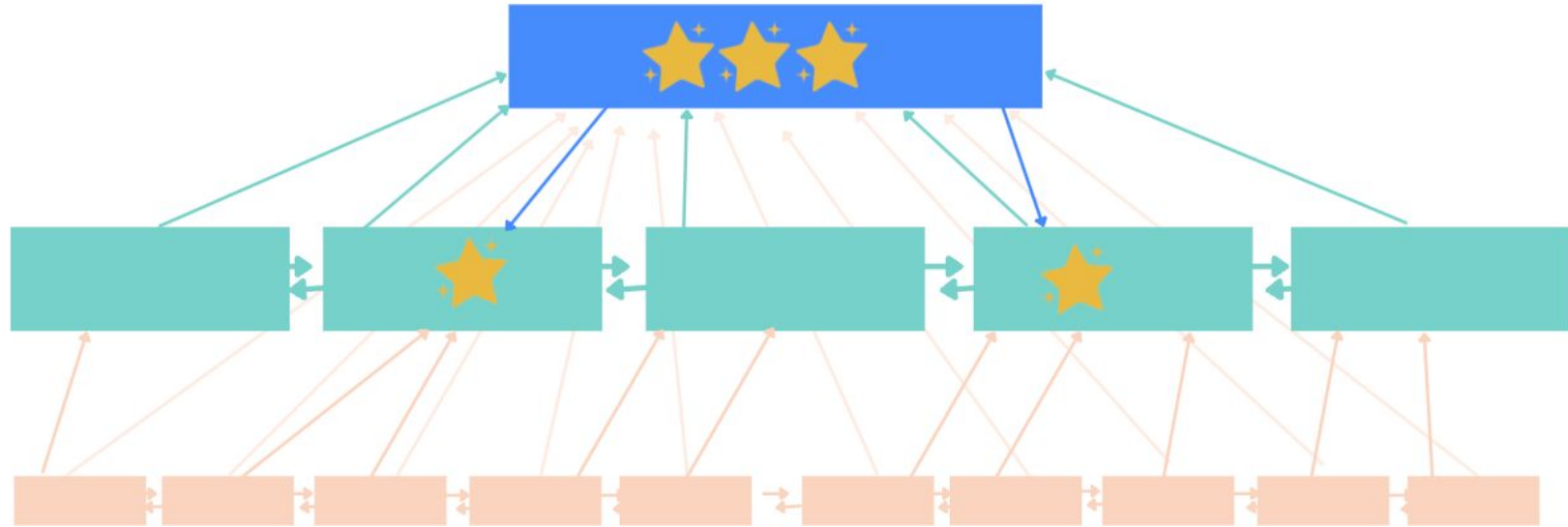


SEARCH FOR INTENT-BASED KEYWORD VARIATIONS

SEARCH FOR DIFFERENT ANCHOR TYPES

INCORPORATE ENTITY-BASED LINKING

THINK ABOUT MONEY-MAKING PAGES



Informational Intent Page

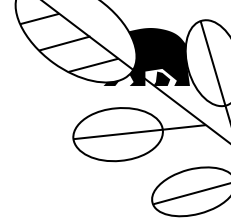
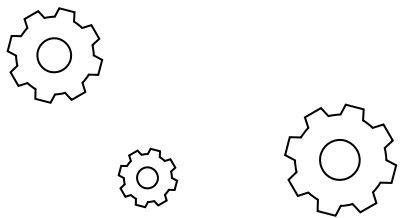
- 🔗 links to other informational content, based on similarity;
- 🔗 links to pillar page (money page, highest converting)
- 🔗 links to one other high intent page
- 🔗 receives link from one informational page, based on similarity

High Intent page

- 🔗 links to pillar page (money page, highest converting)
- 🔗 links to one other high intent pages
- 🔗 receives links from a few informational pages, based on similarity
- 🔗 receives a link from one high intent page, based on similarity

Pillar Page (🔥 conversion rate)

- 🔗 links to 1-2 high intent pages in the cluster (with the second and third highest conversion rates)
- 🔗 receives links from multiple informational pages
- 🔗 receives links from multiple high intent pages



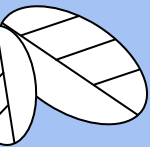
BUILD MEASURE LEARN

06

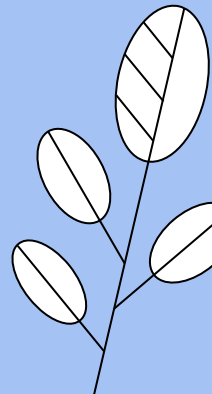
@LAZARINASTOY | #BRIGHTONSEO

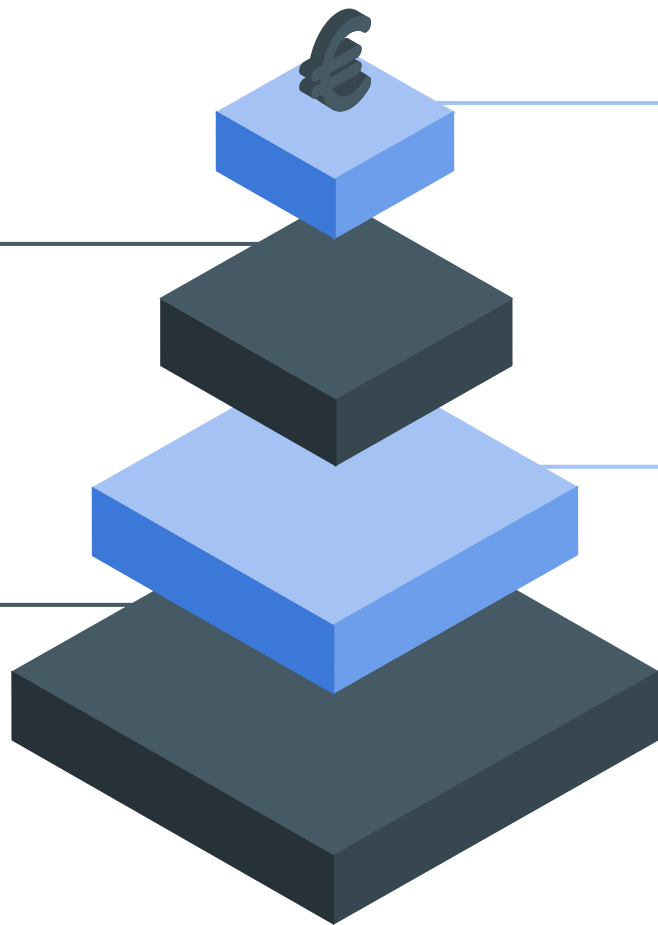
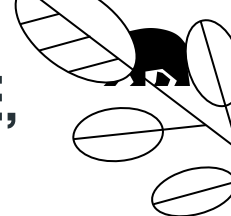


WHAT SHOULD YOU INCLUDE IN YOUR DELIVERABLE?



AZAPYASTOY | #BRIGHTONSEO





LINK OPPORTUNITIES

Highlight additional link opportunities, anchor strategies.

ANALYSIS

Present your analysis in a format that allows for benchmarking

PRIORITIZE, BUDGET

Talk about link implementation, timeline, tracking.

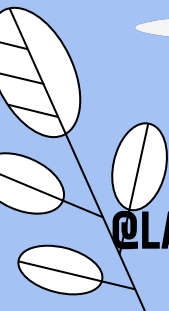
TOPIC MODELS

After finalising the categories and structure, create custom dimensions for the topics in GA/GDS/GSC





WHAT SHOULD YOU MEASURE?

@LAZARINASTOY | #BRIGHTONSEO



YOUR REPORTING SHOULD INCLUDE

	DESCRIPTION	
 LINK FREQUENCY	<ul style="list-style-type: none">• # inlinks / # unique inlinks per page• Orphaned pages	 CRAWLING PAGE DISCOVERY
LINK QUALITY	<ul style="list-style-type: none">• Link placement• Source traffic• Added link sections	ENGAGEMENT COMPREHENSION
ANCHOR TEXT	<ul style="list-style-type: none">• Type• Length	ENGAGEMENT COMPREHENSION
HIERARCHY	<ul style="list-style-type: none">• Top pages• Presence of topics, tags or categories• Crawl depth	ENGAGEMENT & UX CRAWLING
MISTAKES	<ul style="list-style-type: none">• Broken links• Nofollow tags• UTM parameters	CRAWLING INDEXING TRAFFIC

@LAZARINASTOY | #BRIGHTONSEO

BENCHMARKING

GROWTH METRICS

EXPECTED OUTCOMES

 **BETTER STRUCTURE**

 **IMPROVED INDEXING**

 **NAVIGATION & UX**

 **KEYWORDS RANKED**

 **TRAFFIC, CTR**

 **USER ENGAGEMENT**

@lazarinastoy | #brightonSEO | lazarinastoy.com | 

@LAZARINASTOY | #BRIGHTONSEO

FOR GROWTH REPORTING, TRACK



- Quicker Indexing
- # of pages indexed
- # of pages discovered
 - Organic traffic
- Inter-cluster user flow

TRAFFIC



- CTR
- Average time per session
 - Clicks
 - Interactions

ENGAGEMENT



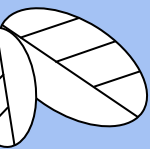
- # of ranking queries
- Length of queries
- Performance (visibility; engagement, rankings) split per topic cluster

QUERIES

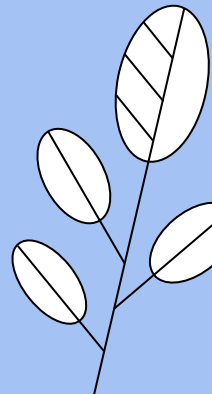


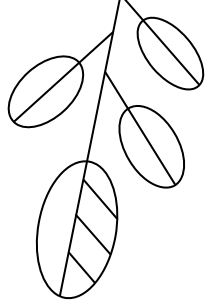


HOWEVER,



AZAPYASTOY | #BRIGHTONSEO





AS WILL CRITCHLOW EXPLAINS...

And even, potentially, on other pages that the **original page** already **linked to**

The diagram illustrates a hierarchical structure of pages. At the top is a light green box. Below it are three boxes: a purple box on the left, and two light green boxes on the right. Arrows point from the top box to each of these three boxes. Below these three boxes are five boxes: a blue box on the far left, a blue box, a pink box, a light green box, and a light green box on the far right. Arrows point from the purple box to the blue, pink, and light green boxes. Arrows point from the two light green boxes above to the pink, light green, and light green boxes below. A play button icon is overlaid on the purple box. A circular inset image of a man's face is positioned to the left of the diagram, with blue arrows pointing from his face to the blue and pink boxes in the bottom row.

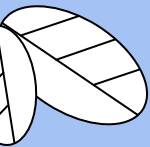
0:50 / 2:50

@LAZARINASTOY | #BRIGHTONSEO

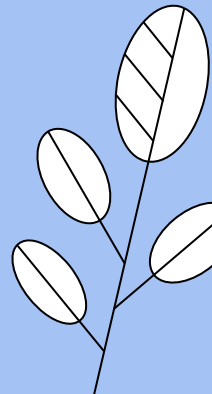


INTERNAL LINKING EFFECT MEASUREMENT IS

HARD



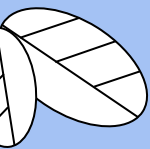
AZAPYASTOY | #BRIGHTONSEO



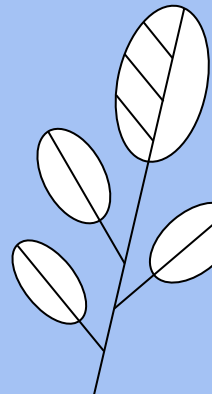


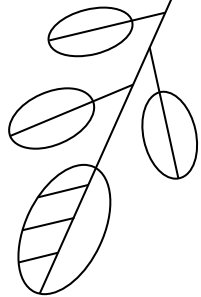
OUR INTUITION ABOUT LINK STRUCTURES IS

POOR



AZAR VASTOY | #BRIGHTONSEO





IN OTHER WORDS...

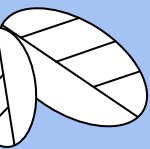
@LAZARINASTOY | #BRIGHTONSEO



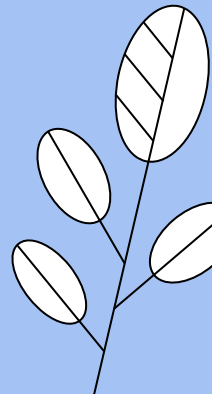
FAILING

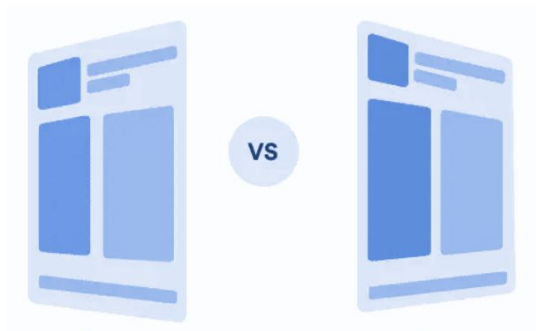
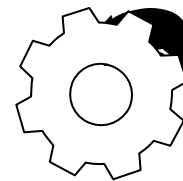
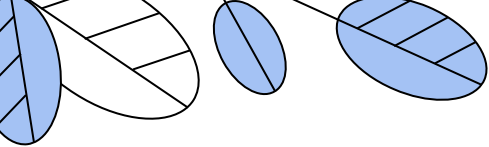
IS PART OF

THE PROCESS.



AZAPASTOY | #BRIGHTONSEO





TESTING

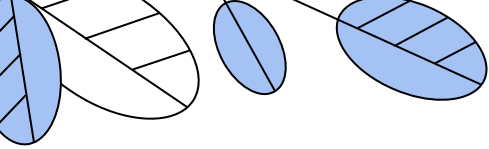


EXPERIMENTS



LEARNING

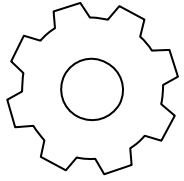
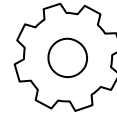
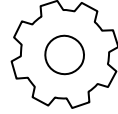
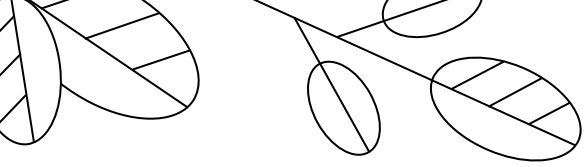
@LAZARINASTOY | #BRIGHTONSEO



ARE THE WAY TO GO

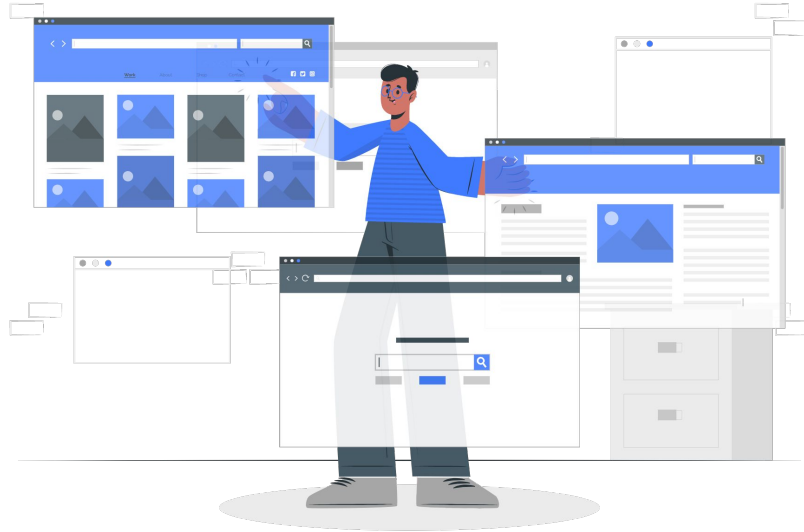


@LAZARINASTOY | #BRIGHTONSEO



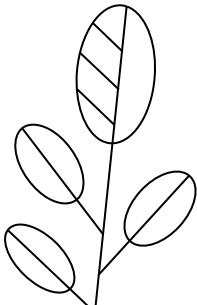
BEFORE I GO...

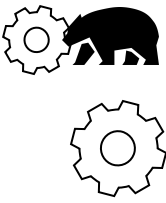
@LAZARINASTOY | #BRIGHTONSEO



**THIS DECK,
REFERENCED WORK
OF OTHER COOL PEOPLE, AND A
TALK WRITE-UP
WILL BE AVAILABLE**

@LAZARINASTOY | #BRIGHTONSEO



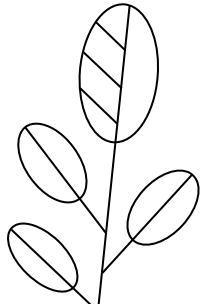


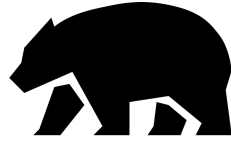
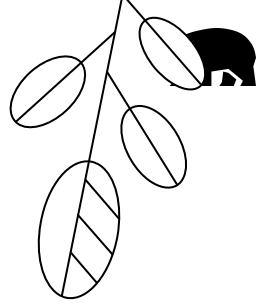
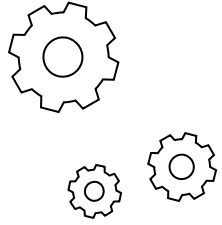
HERE.



<https://lazarinastoy.com/how-to-incorporate-machine-learning-in-internal-linking-audits/>

@LAZARINASTOY | #BRIGHTONSEO



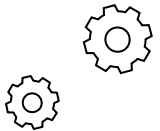


HIRING ALERT

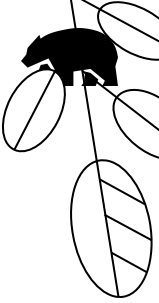
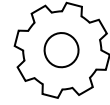
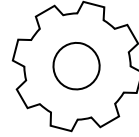
We are a pretty awesome team on the lookout for talent.

Search for **Intrepid Digital** to work with some exciting US & global brands.

@LAZARINASTOY | #BRIGHTONSEO



THANK YOU.



lazarinastoy.com

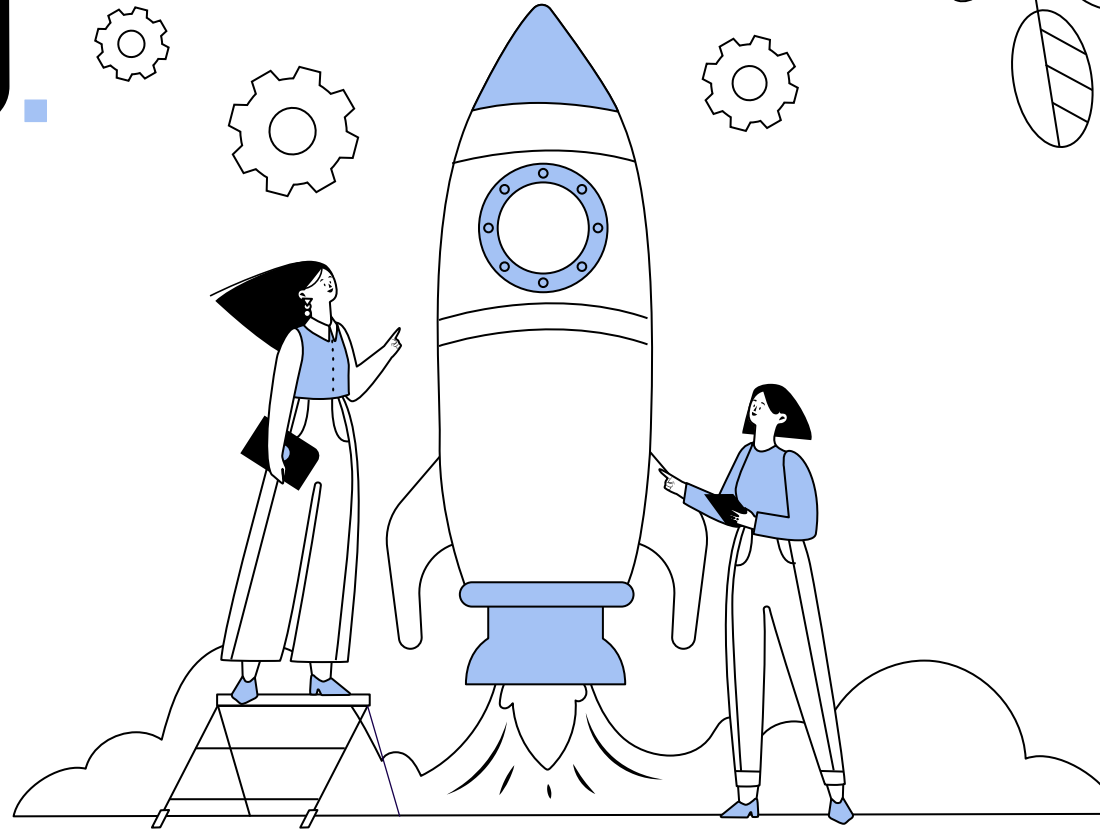


@lazarinastoy



/in/lazarinastoy/

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), and infographics & images by [Freepik](#)



@LAZARINASTOY | #BRIGHTONSEO



Intrepid Digital