

# Vision forward

Tracing Image SEO from its roots  
to AI-enhanced horizons

**Roxana STINGU**

Head of Search and SEO

@roxanastingu

**alamy**



Amsterdam 2024



@roxanastingu | March 2024 | **alamy**

# Intro



Search fanatic



Tech SEO



AI enthusiast



01

Image Search

02

Best practices

03

Generative AI

04

What the future holds

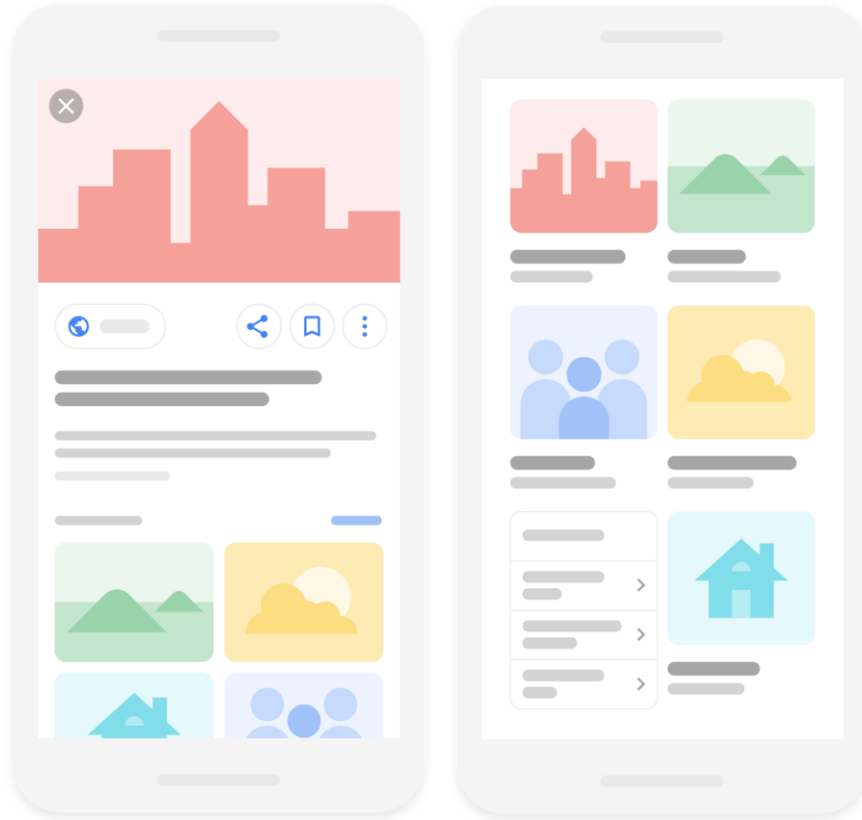
01

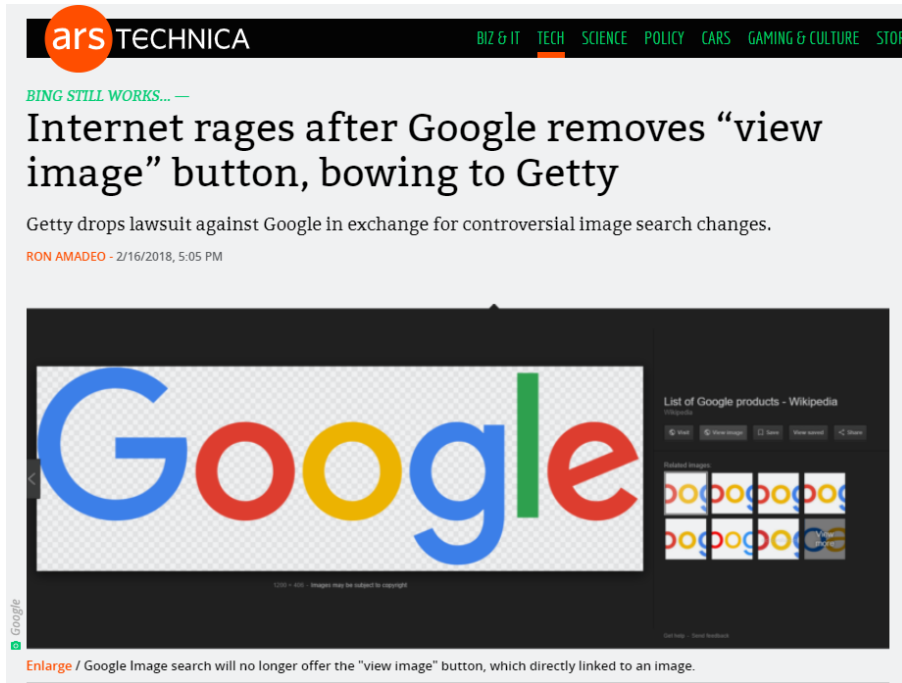
# Image Search

It all started with a green dress...



## New search facet for Images





At first there were only  
JPEGS

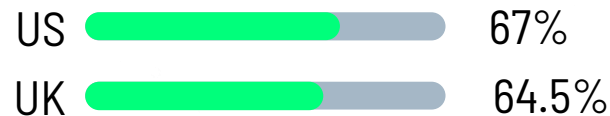
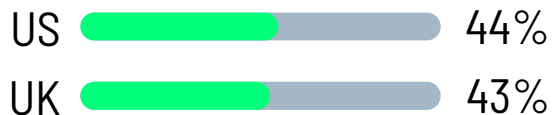
# Images are everywhere





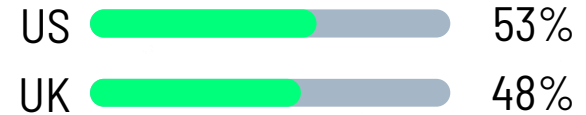
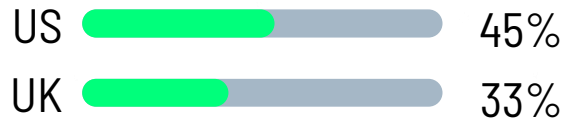
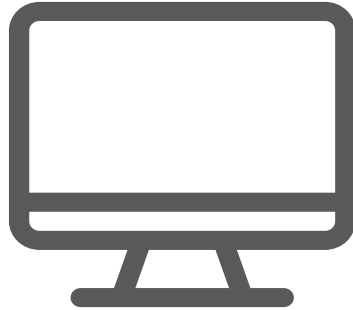
# Images in Google Web Search

Data from SEMRush Sensor on March 10th, 2024



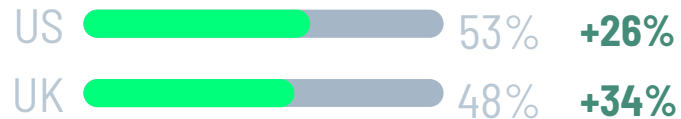
# Images in Google Web Search

Data from SEMRush Sensor on **May 2023**



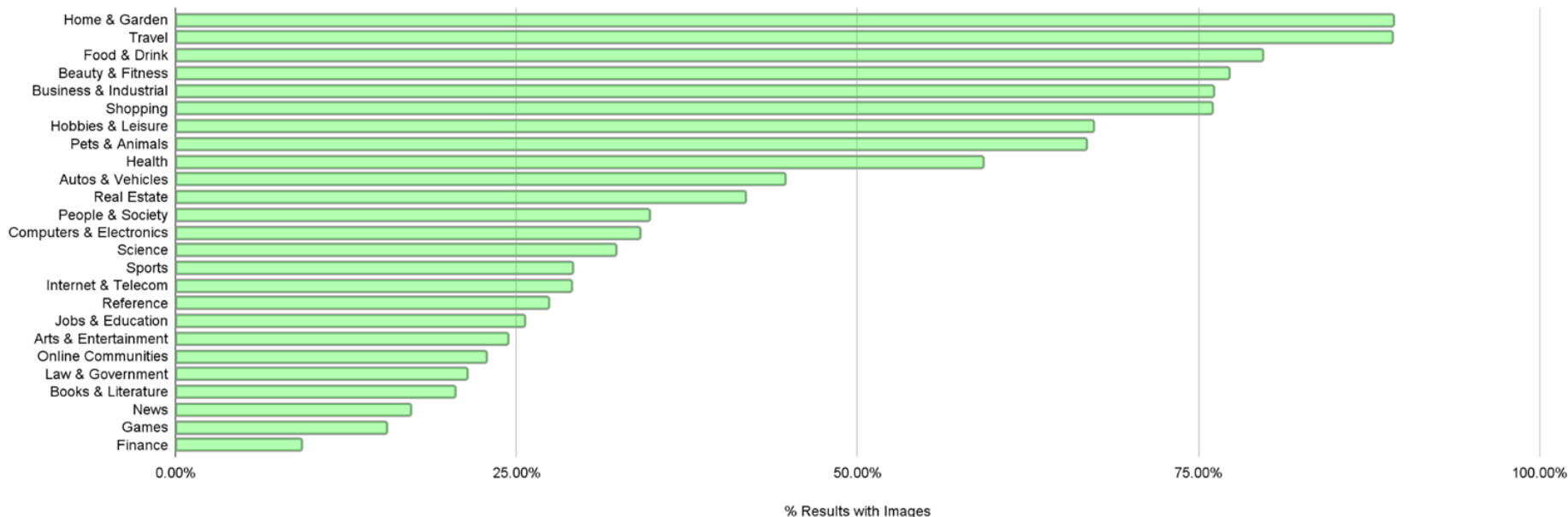
# Images in Google Web Search

Change in 10 months



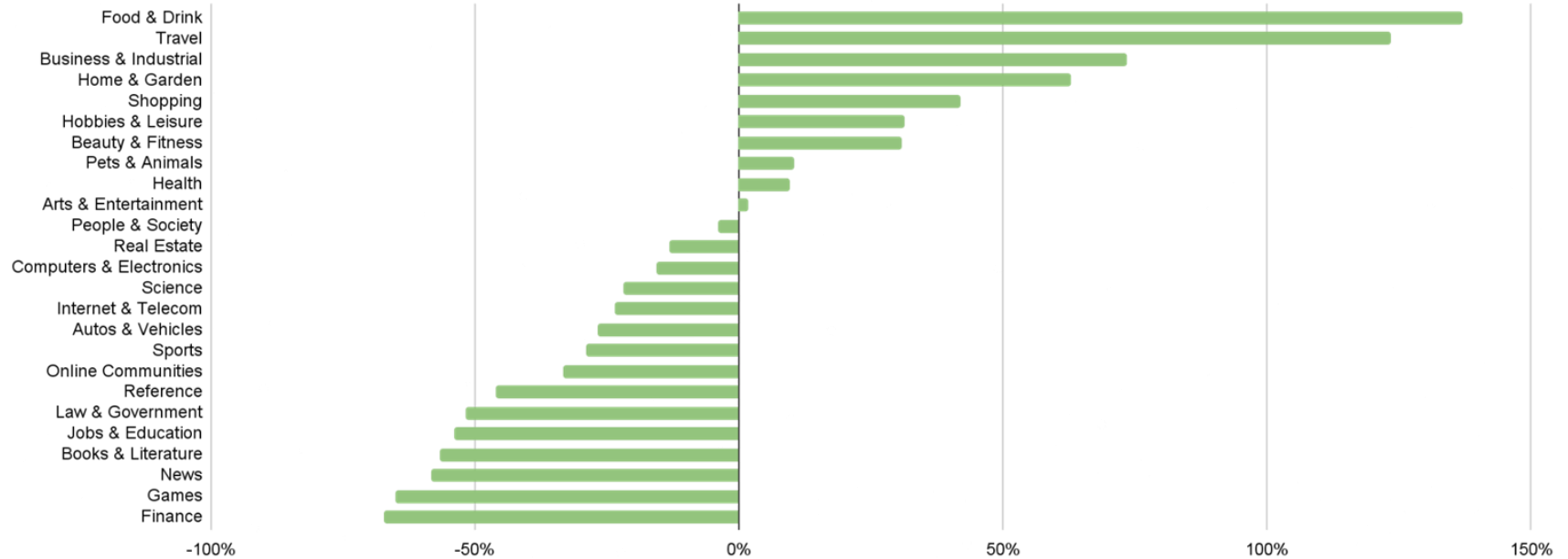
# % of results with Images in Web Search

SEMRush Sensor data, US Desktop, March 10th, 2024



# % growth by niche

SEMRush Sensor data, US Desktop, March 10th, 2024 Compared to May 2023

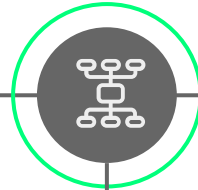
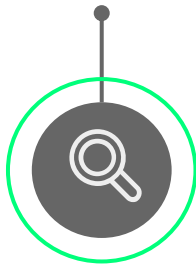


02

# Best Practices

DISCOVERY

INDEXING



CRAWLING

RANKING

# DISCOVERY







# Image sitemaps

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:image="http://www.google.com/schemas/sitemap-image/1.1">
  <url>
    <loc>https://www.example.com/sample1.html</loc>
    <image:image>
      <image:loc>https://www.example.com/image1.jpg</image:loc>
      <image:caption>Succinct sentence to describe the image</image:caption>
      <image:geo_location>Amsterdam, Netherlands</image:geo_location>
      <image:title>Image title</image:title>
      <image:license>Image Archive License, Standard Use</image:license>
    </image:image>
  </url>
</urlset>
```

Required

Deprecated

# HTML

DISCOVERY



**<img>**

```

```

**<div>**

```
<div style="background-image:url(image1.jpg)">Succinct visual  
description of the image</div>
```



# HTML: Responsive and image formats

**<picture><source>**

```
<picture>
  <source
    srcset="image-small.png 320w, image-
medium.png 800w,
image-large.png 1200w"
    sizes="(min-width: 60rem) 80vw,
(min-width: 40rem) 90vw, 100vw">
  
</picture>
```



# Lazy loading

## Browser level

```

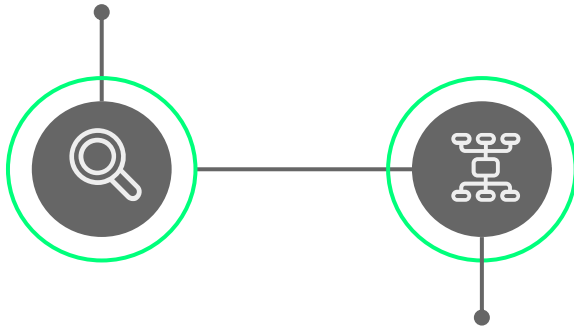
```

## Intersection observer

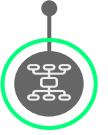
```

```

DISCOVERY

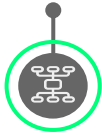


CRAWLING

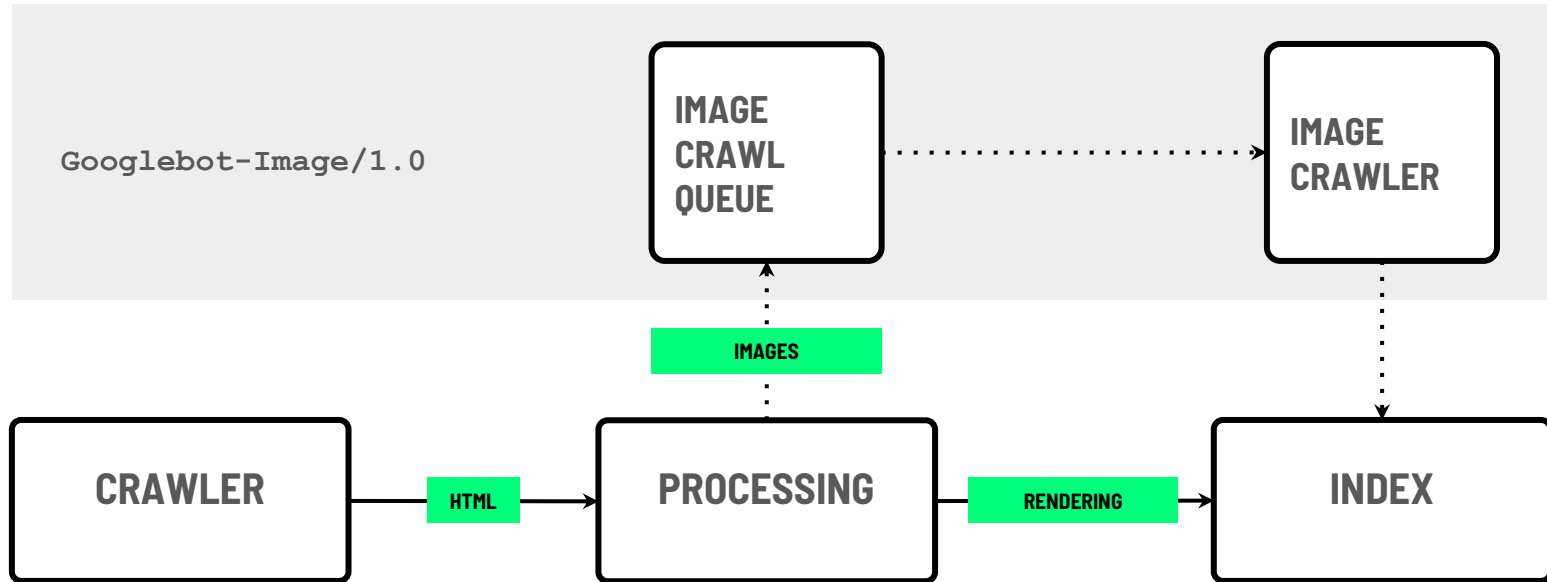


# The basics of crawling



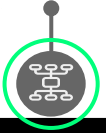


# In comes Googlebot-Image



# CDNs

CRAWLING



← Tweet



**Lily Ray** 🙄  
@lilyraynyc

...

What are the best practices around maintaining links, traffic & #SEO value for images when they are hosted on a CDN, since the CDN is on a subdomain or separate website? @googlewmc @JohnMu

9:35 PM · Mar 4, 2019

2 Retweets 9 Likes 1 Bookmark



**giovannimu (official) — #StaplerLife** @JohnMu · Mar 5, 2019

...

Replying to @lilyraynyc

It doesn't matter how you host the images, there's no SEO-bonus for having them on the same host/domain. However, since changing image URLs is "hard", I'd use your own hostname for the CDN (avoid [yoursite.cdncorp.com](#) and use [cdn.yoursite.com](#)), so you can change CDNs.

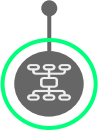
5

40

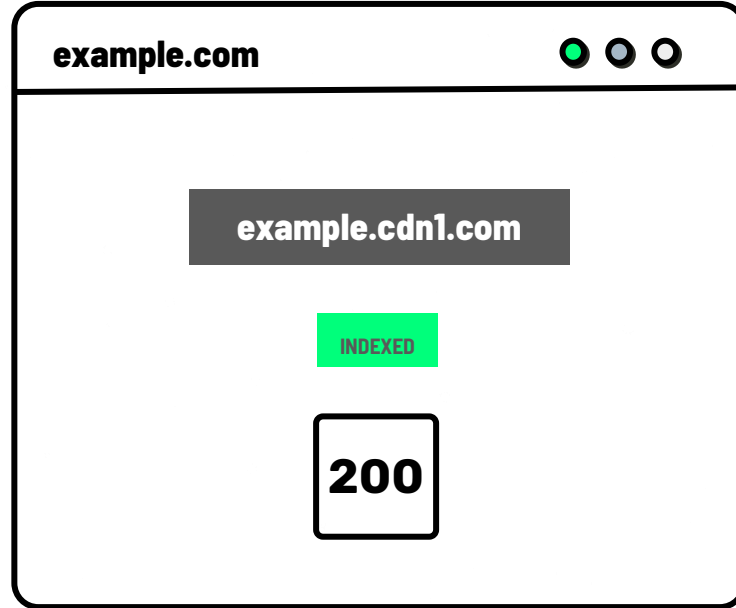
81

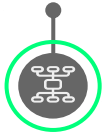




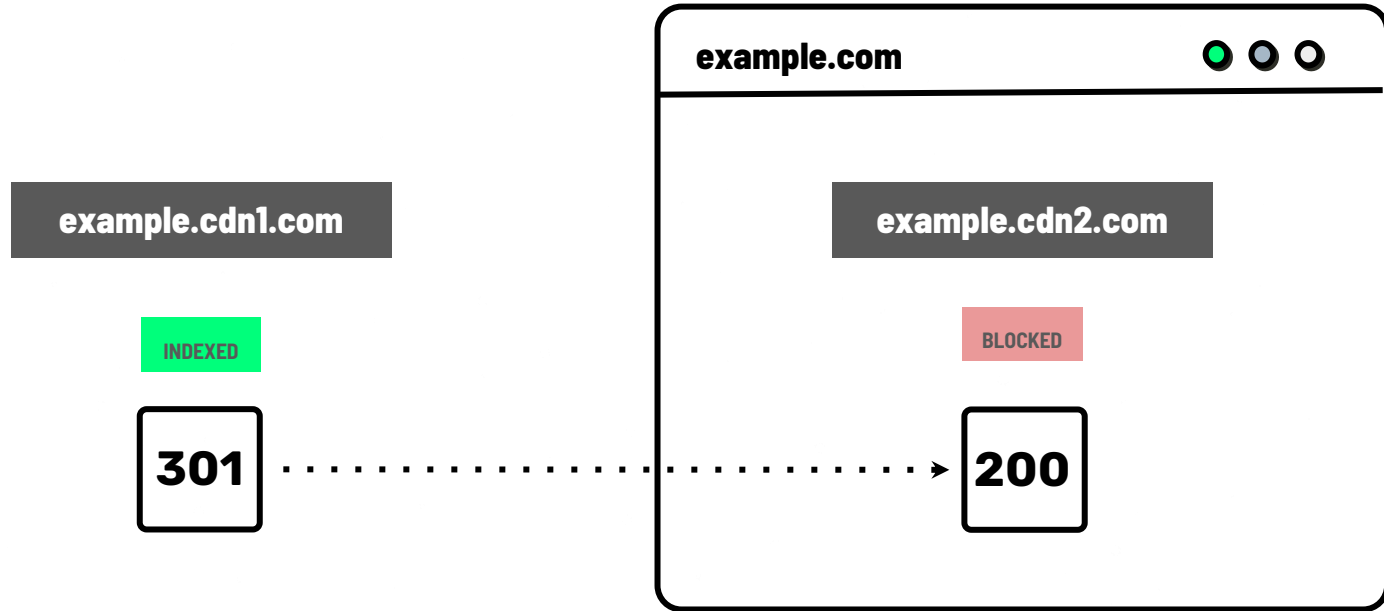


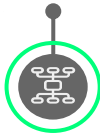
# CDNs + Redirects



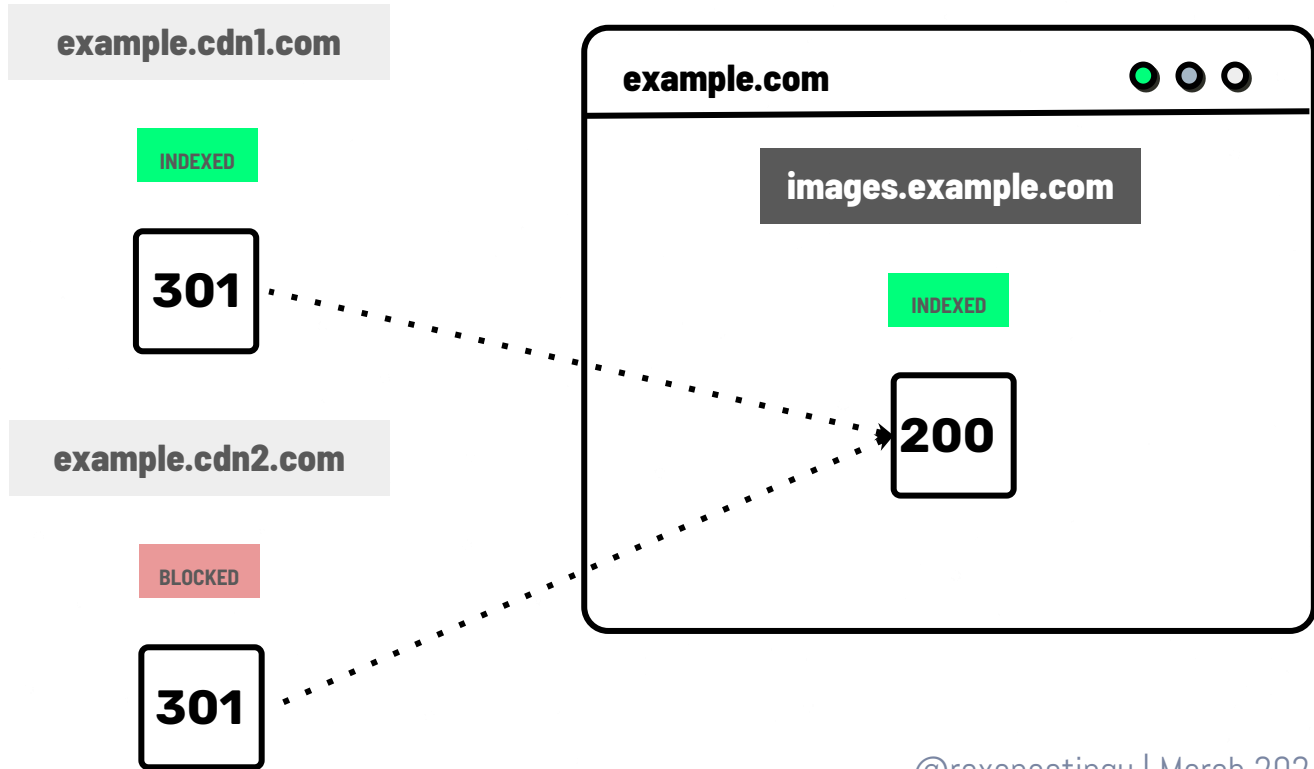


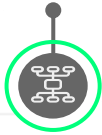
# CDNs + Redirects





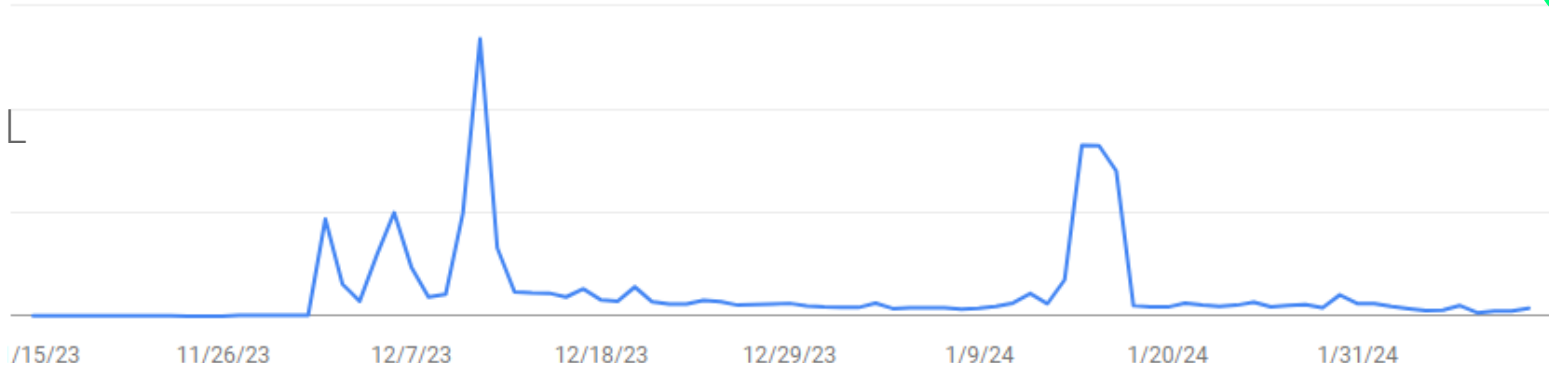
# CDNs + Redirects



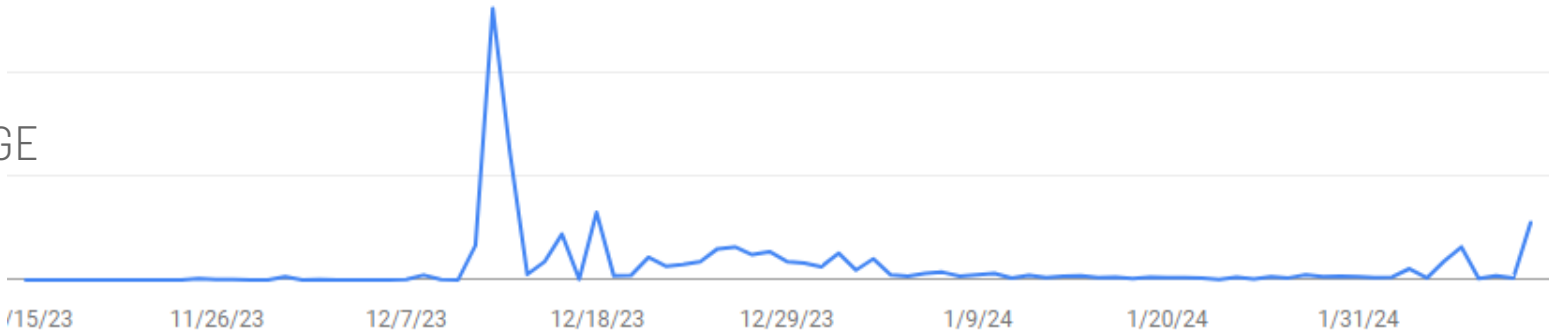


# Crawls by Filetype

HTML



IMAGE





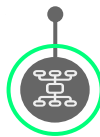
# Expect delayed results in Image Search

WEB  
SEARCH



IMAGE  
SEARCH





# Canonicalisation



I've set a canonical ( `.htaccess` ) header from an image to another one.

1

For instance:



```
site.com/image-small.jpg
```



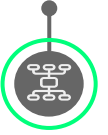
has a canonical header to:



```
site.com/image.jpg (Original image).
```

However the website only uses the `image-small.jpg` and because of that, `image.jpg` was never read by Google and is not indexed (only `image-small.jpg` is). After about a week, the image was removed from the results.

Is it normal for Google to remove credit from content because the original content is poorly indexed?



# No canonicalisation for images

1 Answer

Sorted by: Highest score (default) ▾



Google doesn't use rel=canonical headers for images, so you're probably just seeing normal fluctuations (there's no guarantee of indexing).

2



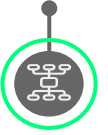
Share Improve this answer Follow

answered Mar 28, 2017 at 19:36



John Mueller

7,441 ● 1 ● 20 ● 26



# Advice still stands\*

answered Mar 28, 2017 at 19:36

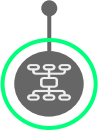


John Mueller

7,441 ● 1 ● 20 ● 26

\*Last checked on January 15th, 2024





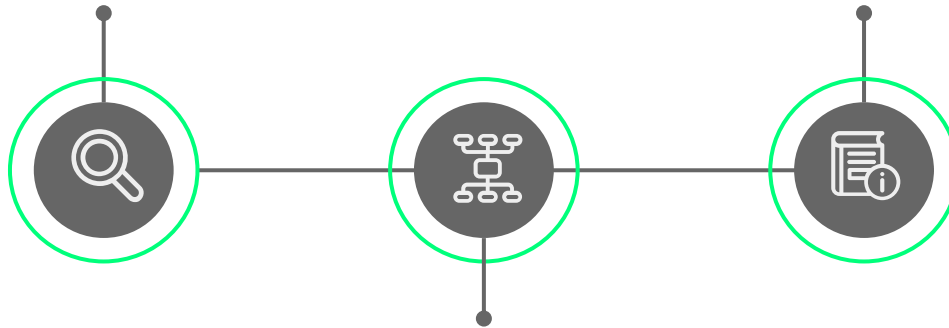
# Blocking crawls to images

**robots.txt**

```
User-agent: Googlebot-Image  
Disallow: /very-secret-images/
```

DISCOVERY

INDEXING



CRAWLING



# Indexable image formats



JPG



PNG



GIF



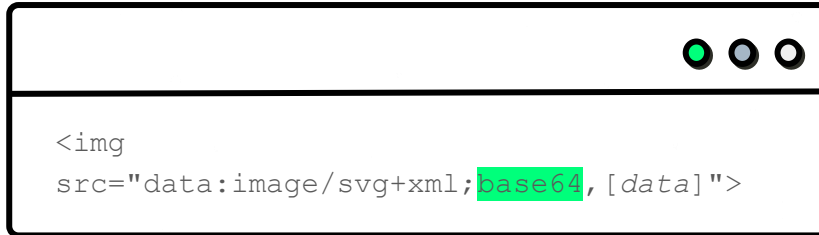
BMP



WEBP



SVG



# Image retrieval patent

**US8923626B1**  
United States

[Download PDF](#) [Find Prior Art](#) [Similar](#)

**Inventor:** Henrik C. Stewenius  
**Current Assignee:** Google LLC

**Worldwide applications**  
2012 · [US](#)

**Application US13/531,912 events** ©

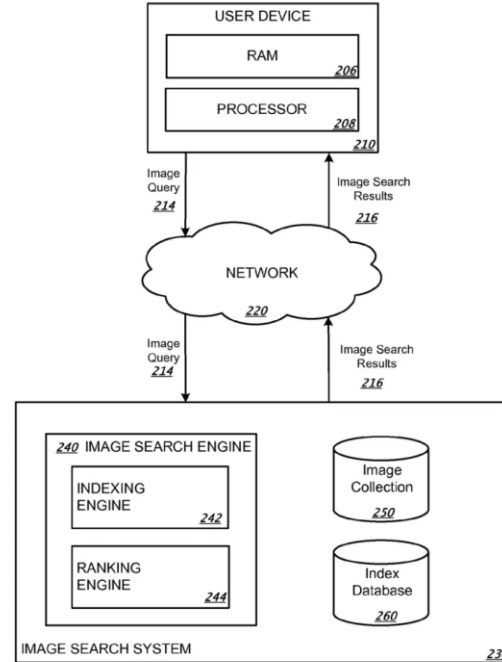
- 2012-06-25 • Application filed by Google LLC
- 2012-06-25 • Priority to US13/531,912
- 2012-07-31 • Assigned to GOOGLE INC. ©
- 2014-12-30 • Application granted
- 2014-12-30 • Publication of US8923626B1
- 2017-10-02 • Assigned to GOOGLE LLC ©

**Status** • Active

2033-06-22 • Adjusted expiration

**Info:** [Patent citations \(10\)](#), [Cited by \(6\)](#), [Legal events](#), [Similar documents](#), [Priority and Related Applications](#)

**External links:** [USPTO](#), [USPTO PatentCenter](#), [USPTO Assignment](#), [Espacenet](#), [Global Dossier](#), [Discuss](#)





# Textual information



WEB DOCUMENT

.....

Text around images  
Alt attribute values  
Image captions  
Anchor texts  
Metadata  
File name

.....



THE INDEX



# Keeping images out of the index

## Meta robots tag

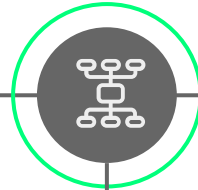
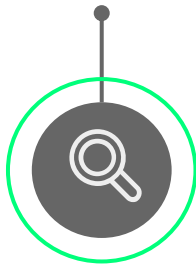
```
<meta name="robots"
content="noimageindex,
nofollow">
```

## X-robots tag

```
header("X-Robots-Tag:
noimageindex, nofollow", true);
```

DISCOVERY

INDEXING



CRAWLING

RANKING



# The basics of ranking for images



**PAGE WITHOUT IMAGE**

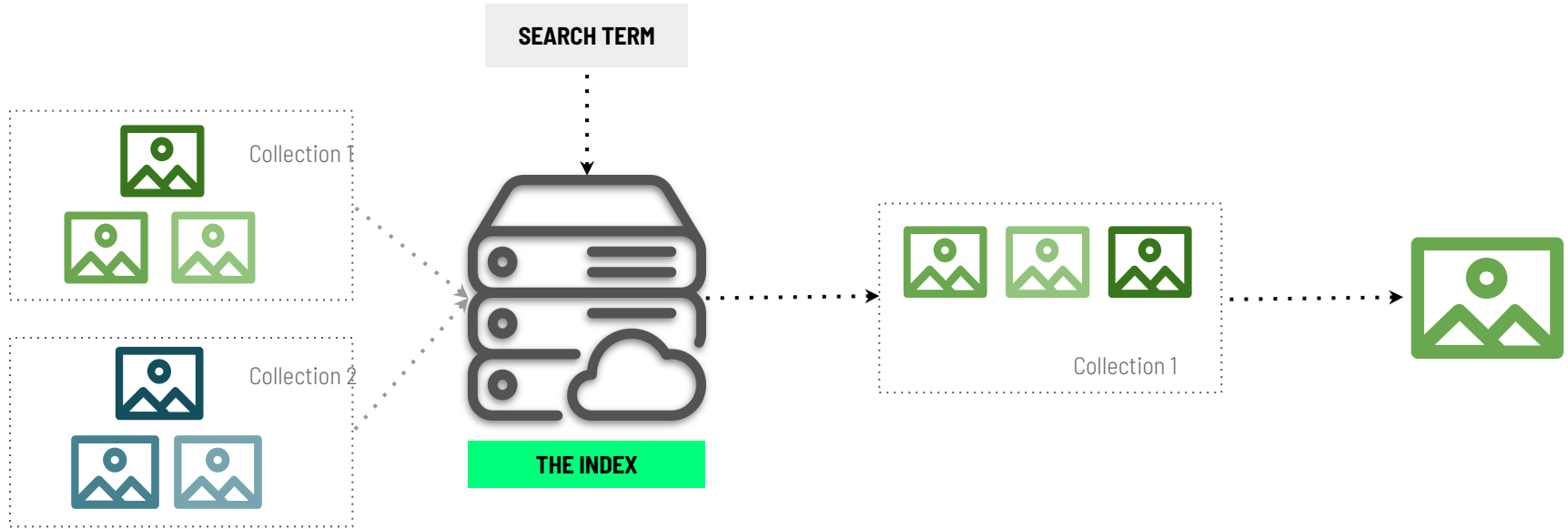


**IMAGE WITHOUT PAGE**



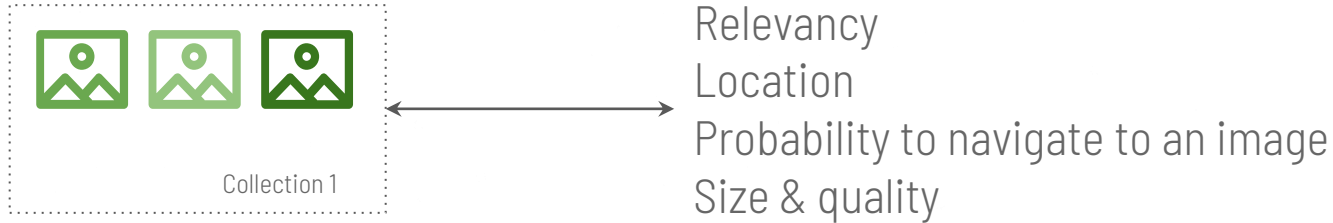
# How images are ranked

RANKING





# What can influence the order of images in a collection





# What is image size?

## File size

123 kb

## Image resolution

1200px wide

## Aspect ratio

16x9, 4x3, and 1x1



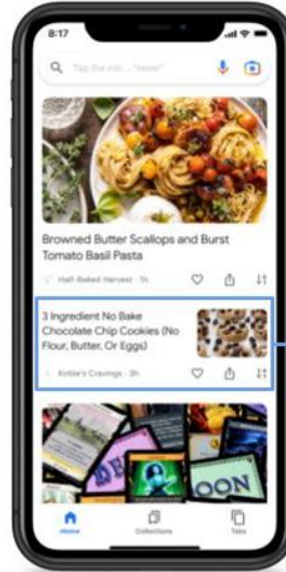
# Control image previews

## Meta robots tag

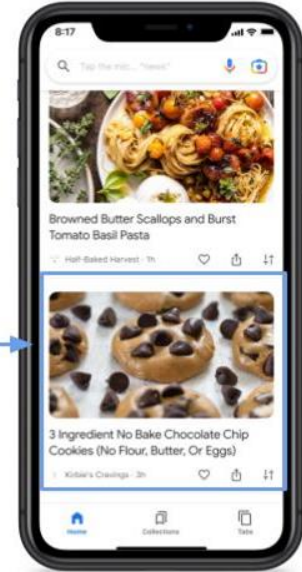
```
<meta name="robots" content="max-  
image-preview:large">
```

## Meta robots tag

```
<meta name="robots" content="max-  
image-preview:none">
```



Default thumbnail experience



Large image experience

# Does Google use EXIF?



2014

"The short answer is: [...] it is something that Google is able to parse out, and I think we do reserve the right to use it in ranking."

**Matt Cutts**

# Does Google use EXIF?



2014

"The short answer is: [...] it is something that Google is able to parse out, and I think we do reserve the right to use it in ranking."

**Matt Cutts**



2017

"This is still the same as before."

**John Mueller**

# Does Google use EXIF?



2014

"The short answer is: [...] it is something that Google is able to parse out, and I think we do reserve the right to use it in ranking."

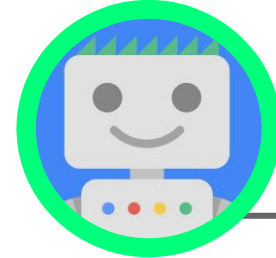
**Matt Cutts**



2017

"This is still the same as before."

**John Mueller**



2024

"To tell Google about your image metadata, **add structured data** or **IPTC photo metadata** to each image on your site."

**Google Search Central**

**Documentation - Image metadata in Google Images**

# Image Rights Metadata

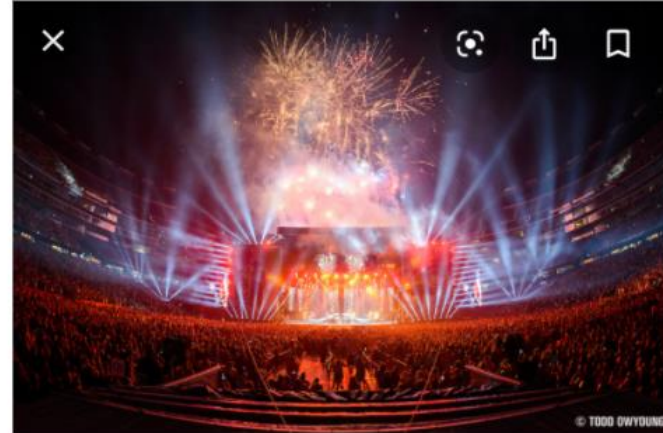
RANKING



Creator

Credit line

Copyright notice



**T/O** Todd Owyoun

Jason Aldean | New York Music  
Photographer Todd Owyoun

Visit

Creator: Todd Owyoun | Credit: Todd Owyoun  
Copyright: © Todd Owyoun  
Information extracted from IPTC Photo Metadata.

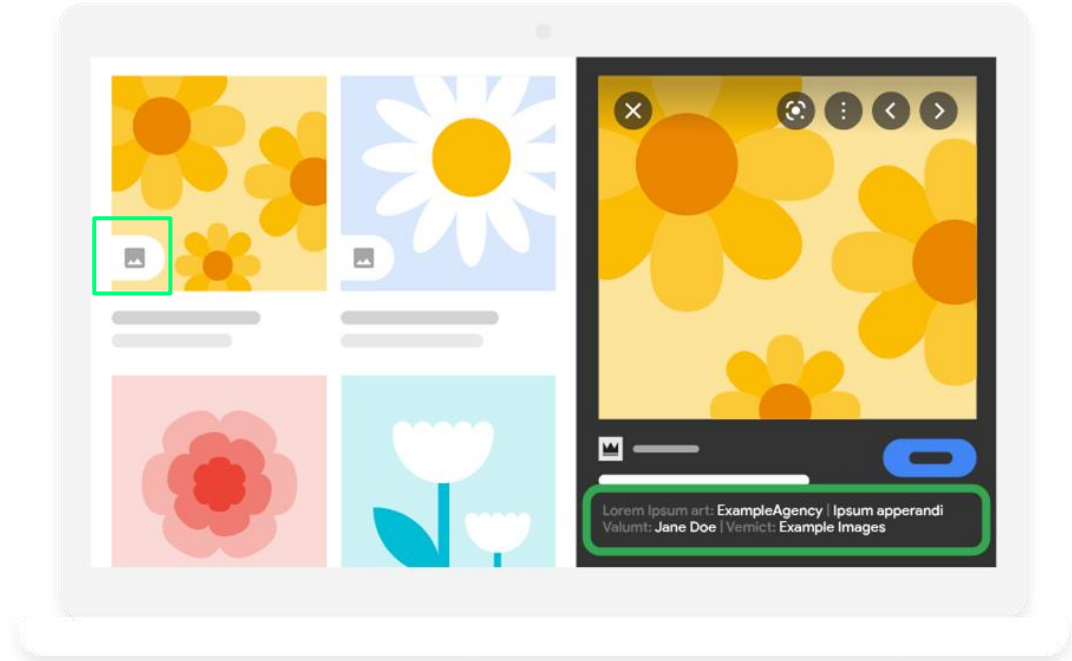




# Licensable Badge

license

acquireLicensePage





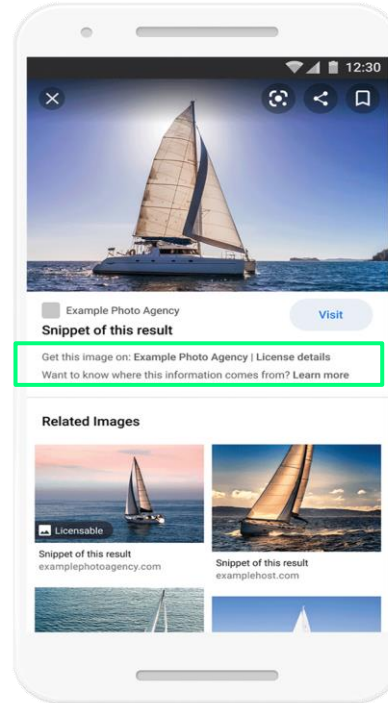
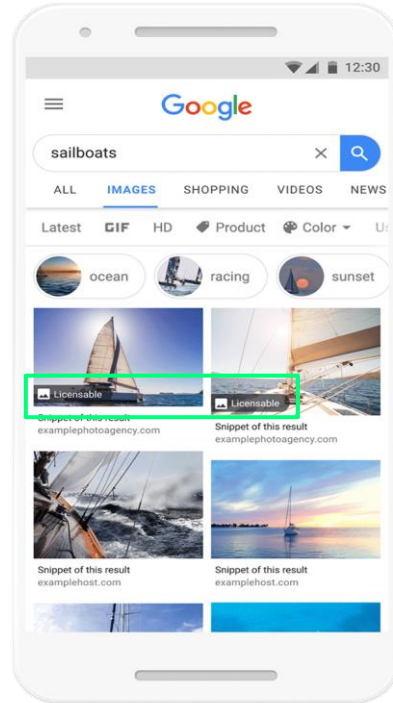
# Structured data v. IPTC



## Structured data

Page specific

Creates rich snippets for  
your pages that are  
ranking for a specific  
image



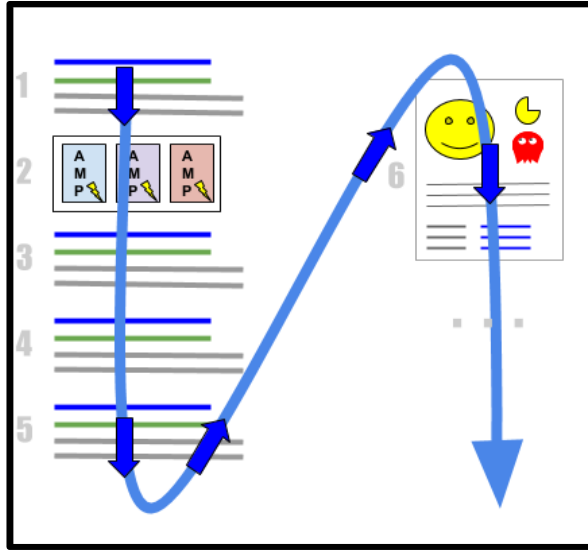
## IPTC metadata

Image specific

Creates rich snippets  
whenever the image  
shows up, no matter  
which page is ranking

# Web v. Image Search ranking

RANKING



Web Search

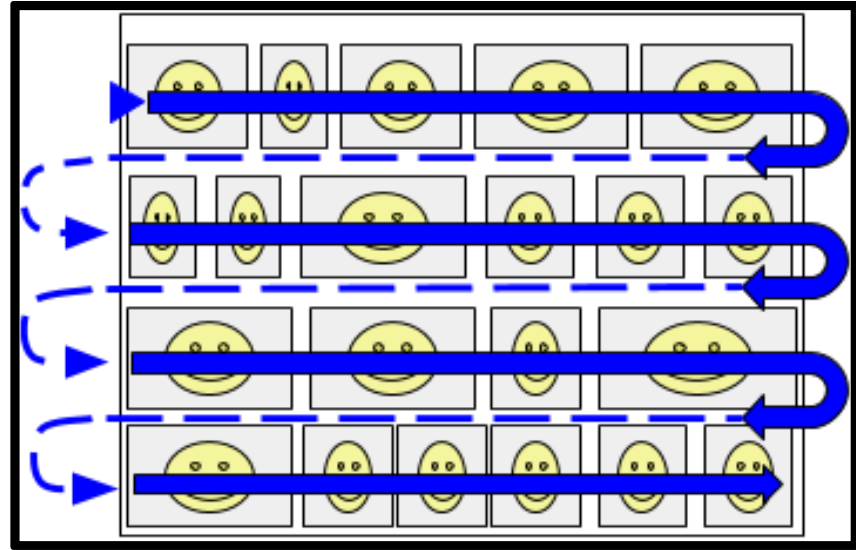
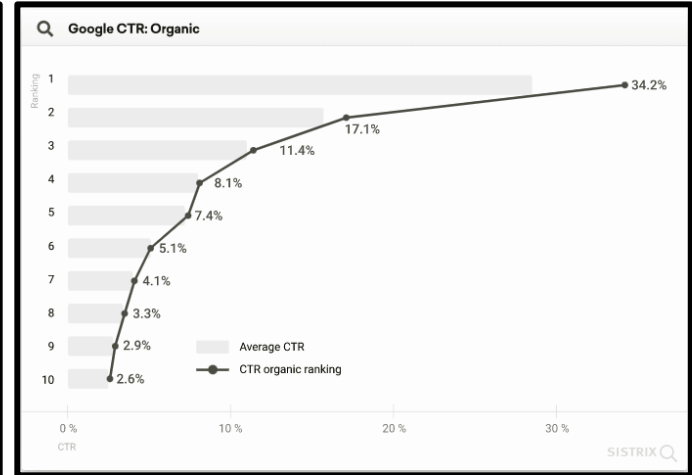
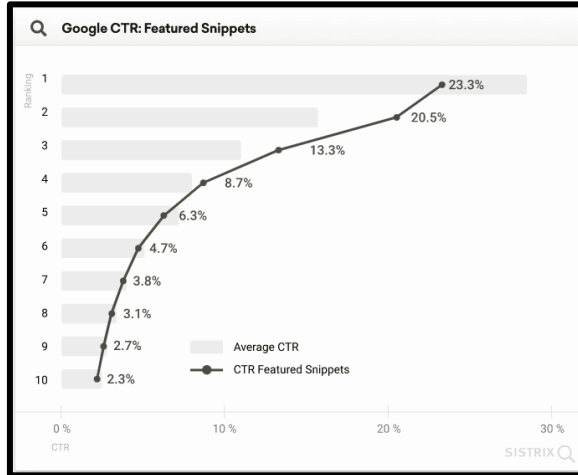
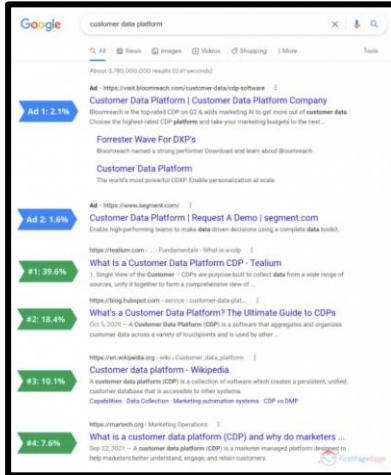


Image Search



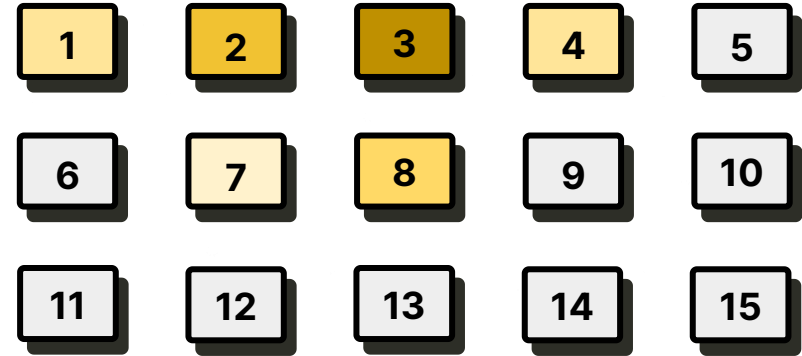
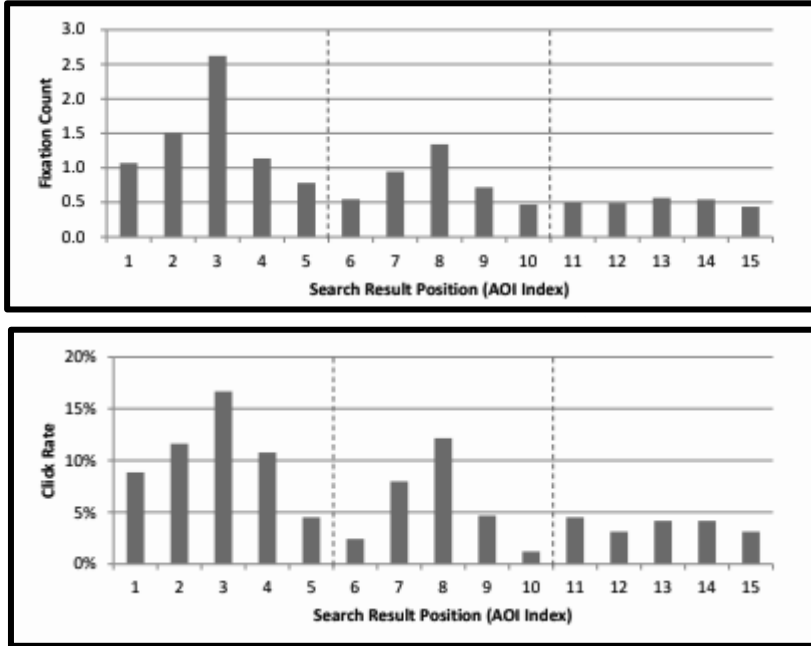
# Best position to rank for in Web Search



Web Search



# Eye tracking study for Image Search



An Eye-Tracking Study of User Behavior in Web Image Search  
 Wanxuan Lu and Yunde Jia Beijing Laboratory of Intelligent Information Technology,  
 School of Computer Science, Beijing Institute of Technology, Beijing 100081, China

@roxanastingu | March 2024 | **alamy**



# Spikes in impressions?



**Roxana Stingu**

@RoxanaStingu



Replying to [@bertiecharlton](#)

If you see keywords ranking 1st in Web Search with lots of impressions and virtually no clicks, it's most likely an image inside an image pack - clicks to these go to image search so check same keywords for clicks under Image Search rather than Web Search.

2:42 PM · Nov 5, 2021 · Twitter for Android

# Search Console - Web Search data cheat sheet

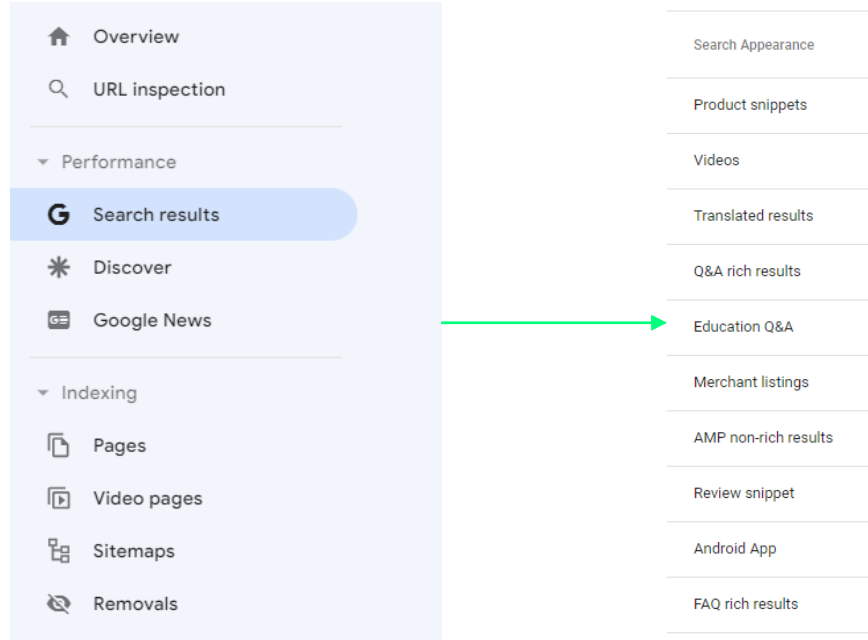
RANKING



		Impressions	Clicks	Average Position
Knowledge Panel	Clicking takes you to Image Search		 Query Refinement	Same as Knowledge Panel links, after blue links
Image Carousel	Clicking opens the Image Preview panel	 When Scrolled into view		Space it occupies
Image Pack	Clicking takes you to Image Search			Space it occupies
Image Thumbnails	Clicking takes you to the website			Space it occupies



# Search Console Search Appearance







# Get creative

Table 1  
Web Search - Site Impression

Your Website - Site Impression - Web Search

**Dimensions**

ABC	Query	GD
	Date	
+ Add dimension		

**Metrics**

AUT	Clicks
AUT	Impressions
AUT	Average Position
+ Add metric	

**Date range**

☒ Auto

**Available Fields**

🔍 Type to search

- ABC \_ Head or Tail Ter...
- ABC \_ Queries by numb...
- ABC \_ Counting Words
- ABC \_ Date
- ABC \_ Head or Tail
- ABC \_ Keywords By Inte...
- ABC \_ Queries containin...
- Country
- Date
- Device Category
- Google Property
- Query

1 condition



Table 2  
Image Search - Site Impression

Your Website - Site Impression - Image Sear...

**Dimensions**

ABC	Query	GD
	Date	
+ Add dimension		

**Metrics**

AUT	Clicks
AUT	Impressions
AUT	Average Position
+ Add metric	

**Date range**

☒ Auto

**Available Fields**

🔍 Type to search

- \_Date
- Country
- Date
- Device Category
- Google Property
- Query
- 123 Average Position
- 123 Clicks
- 123 Impressions
- 123 Site CTR



# Estimate what queries appear in image packs or knowledge graphs

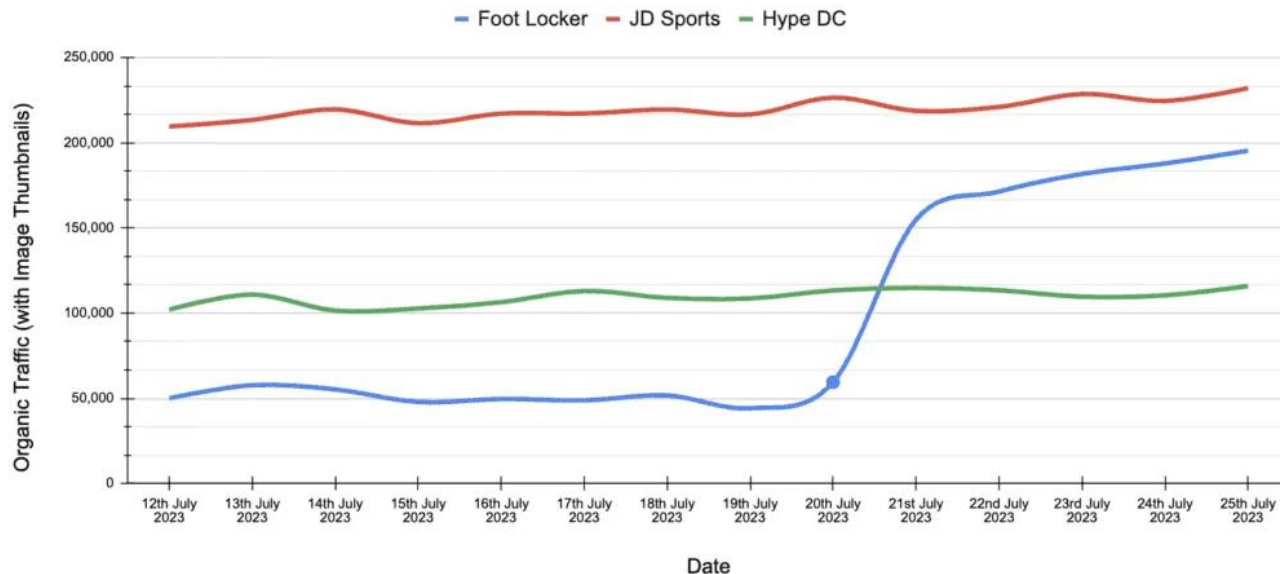
CASE

```
WHEN Query (Image Search)=Query (Web Search)
  AND Average Position (Web Search) =1
  AND Web Impressions > 100
  AND Web Clicks < (Web Impressions/100)
THEN "Image Pack or Image Carousel"
WHEN Query (Image Search)=Query (Web Search)
  AND Average Position (Web Search) >5
  AND Web Impressions > 100
  AND Web Clicks < (Web Impressions/100)
THEN "Knowledge Panel"
ELSE "Other"
```

END

# Are images worth the hassle?

Foot Locker Google image thumbnail organic traffic vs. competitors (Semrush index)



228% increase in estimated organic visits p/mo that now have image thumbnails

The change?  
Image dimensions update from 67x67 to 200x200.

03

# Generative AI

@roxanastingu | March 2024 |

alamy

# Is AI used for ranking?

**RankBrain**

**Neural Matching**

**BERT**

**MUM**

# Is AI used for ranking in Image Search?

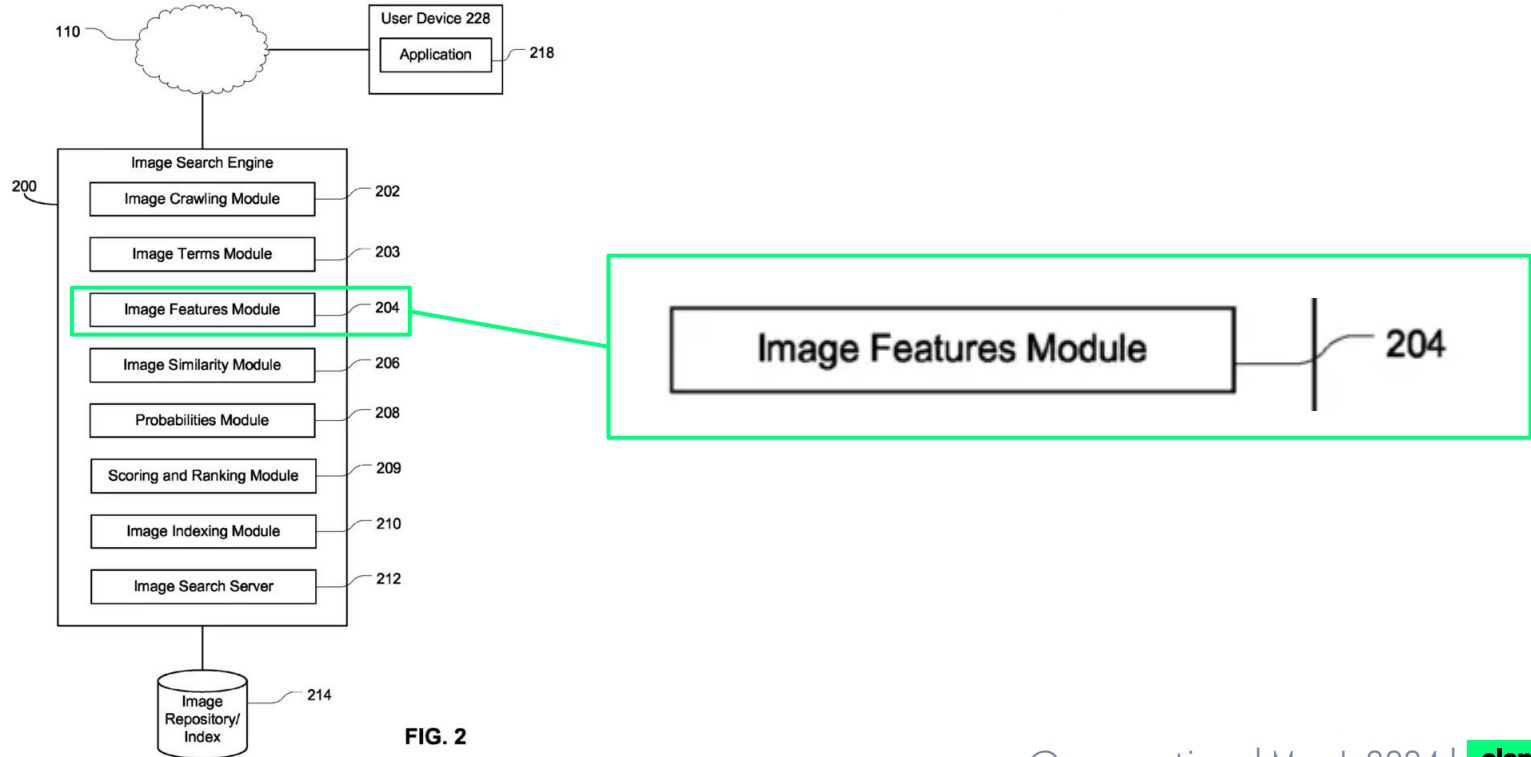
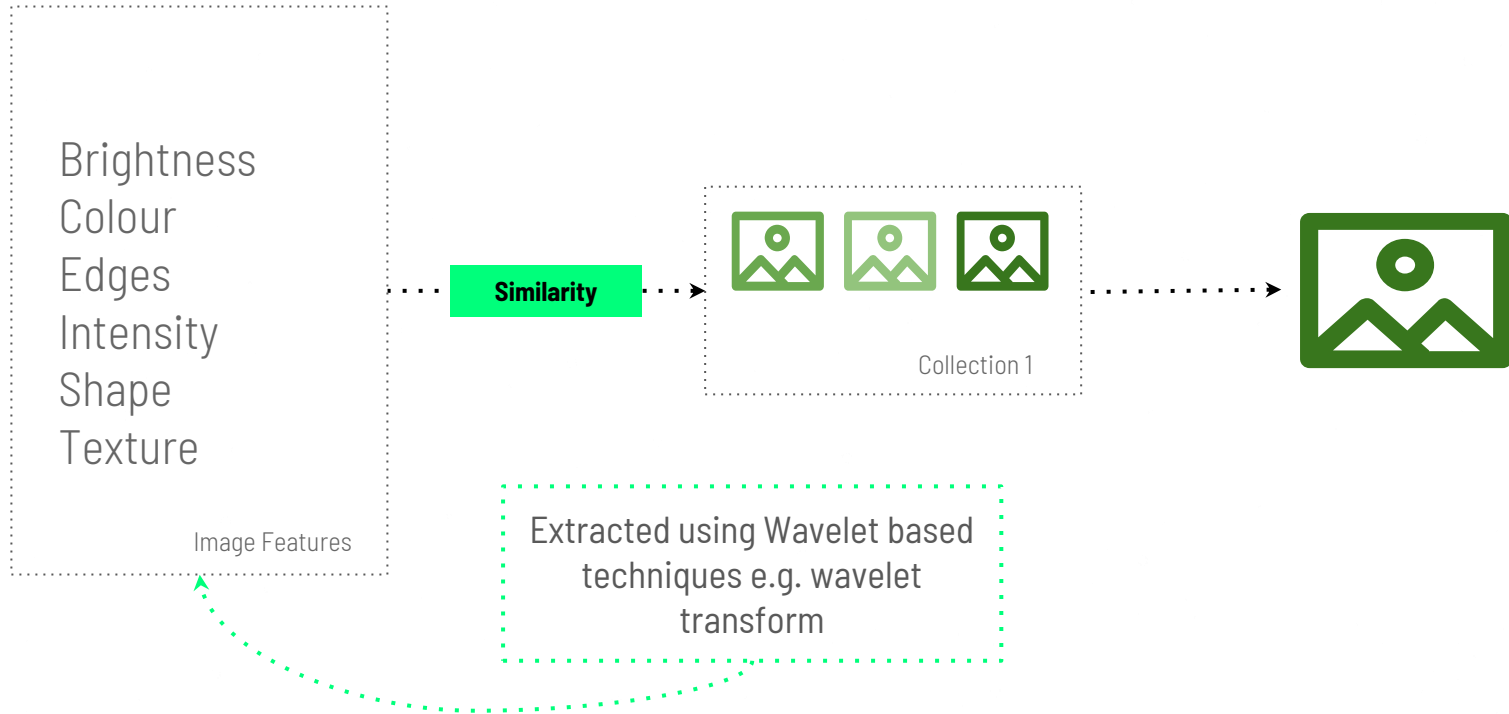


FIG. 2

# Image features



# Image patents made easy



**Bill  
Slawski**

seobythesea.com



**JC  
Chouinard**

jcchouinard.com



# Embeddings

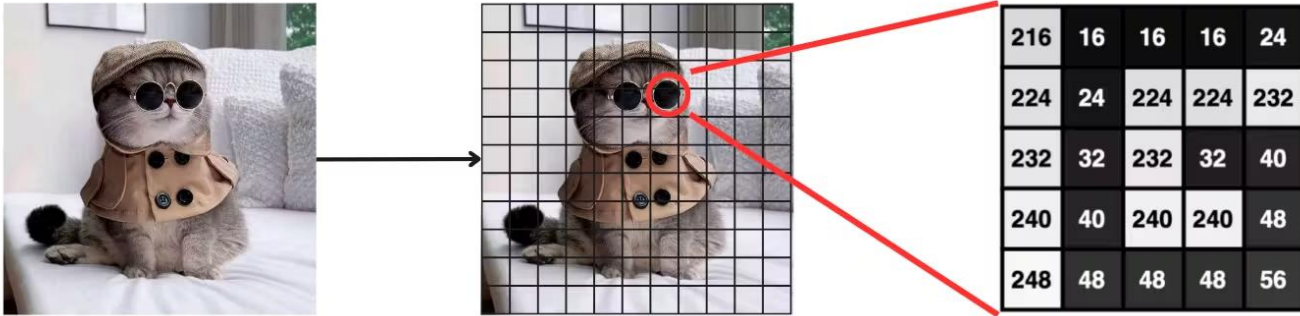
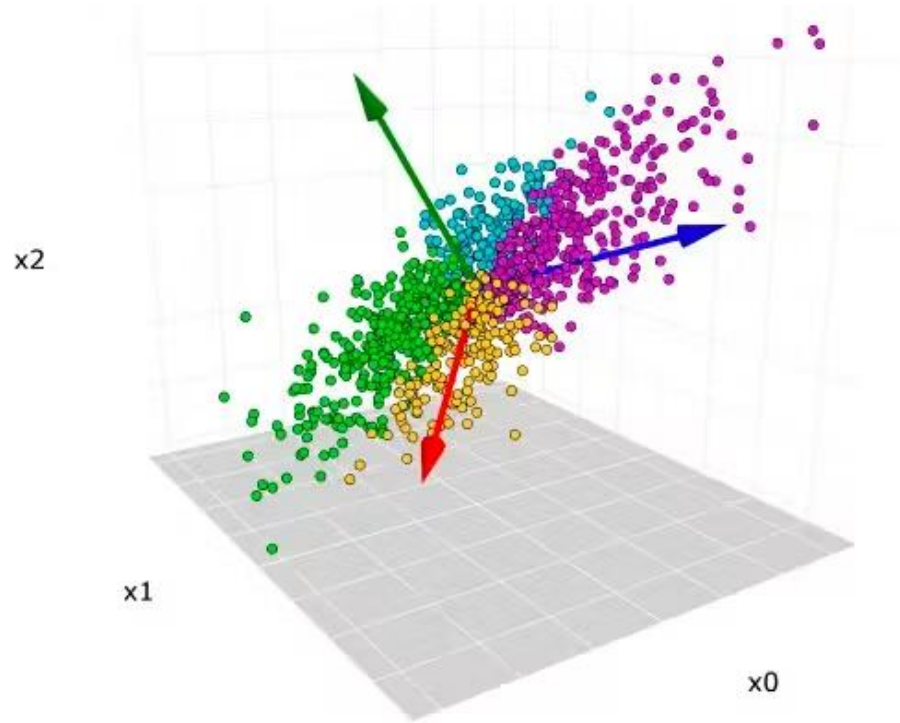
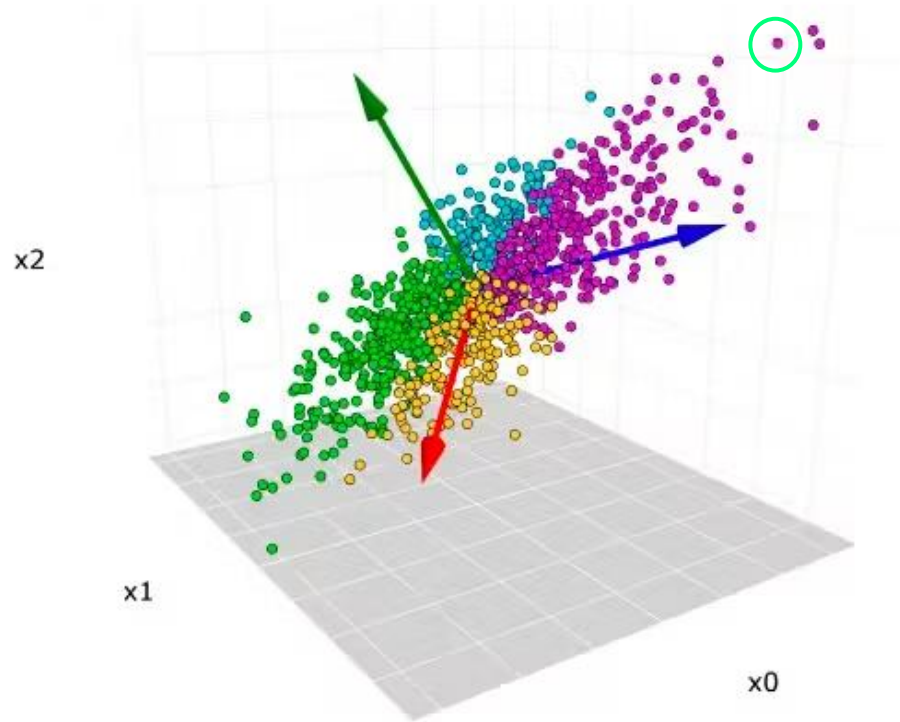


Image embeddings are numerical representations of images that capture their semantic meaning and visual features.

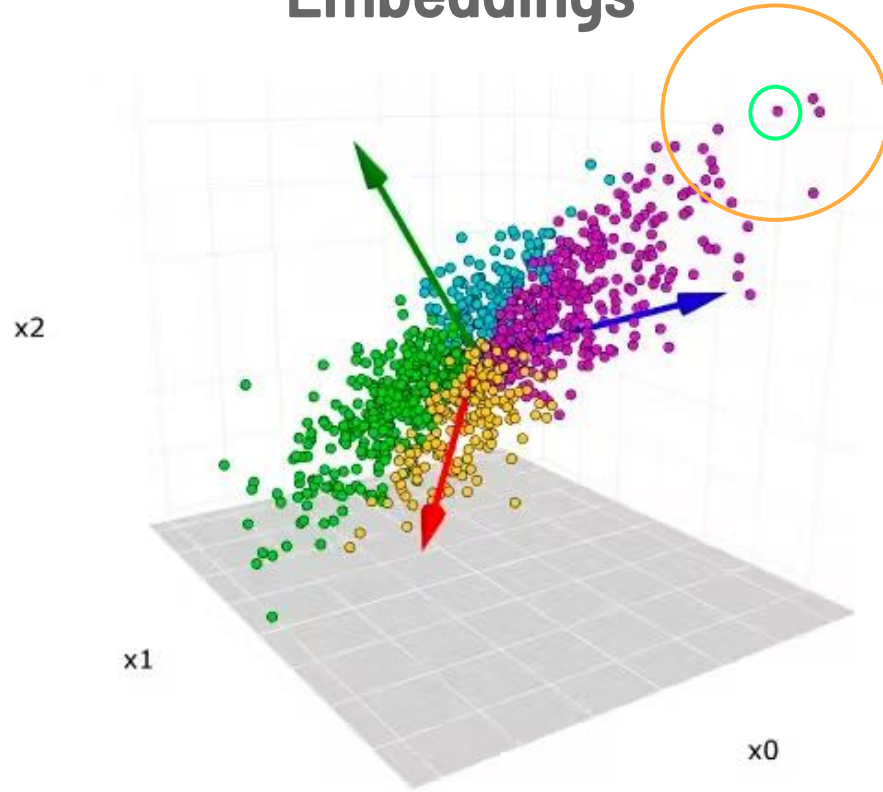
# Embeddings



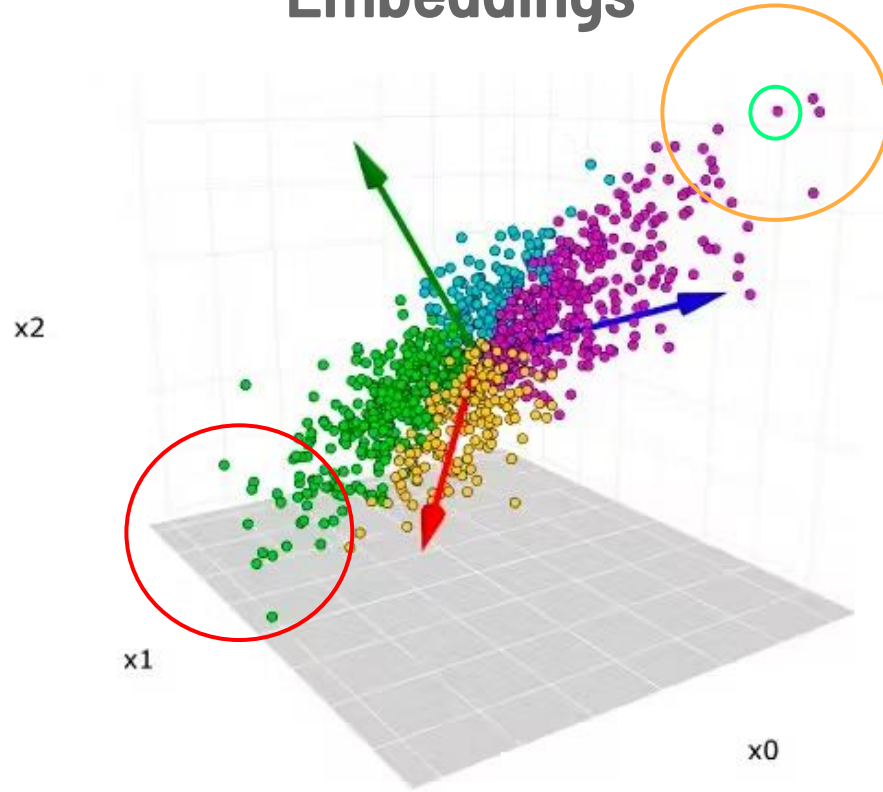
# Embeddings



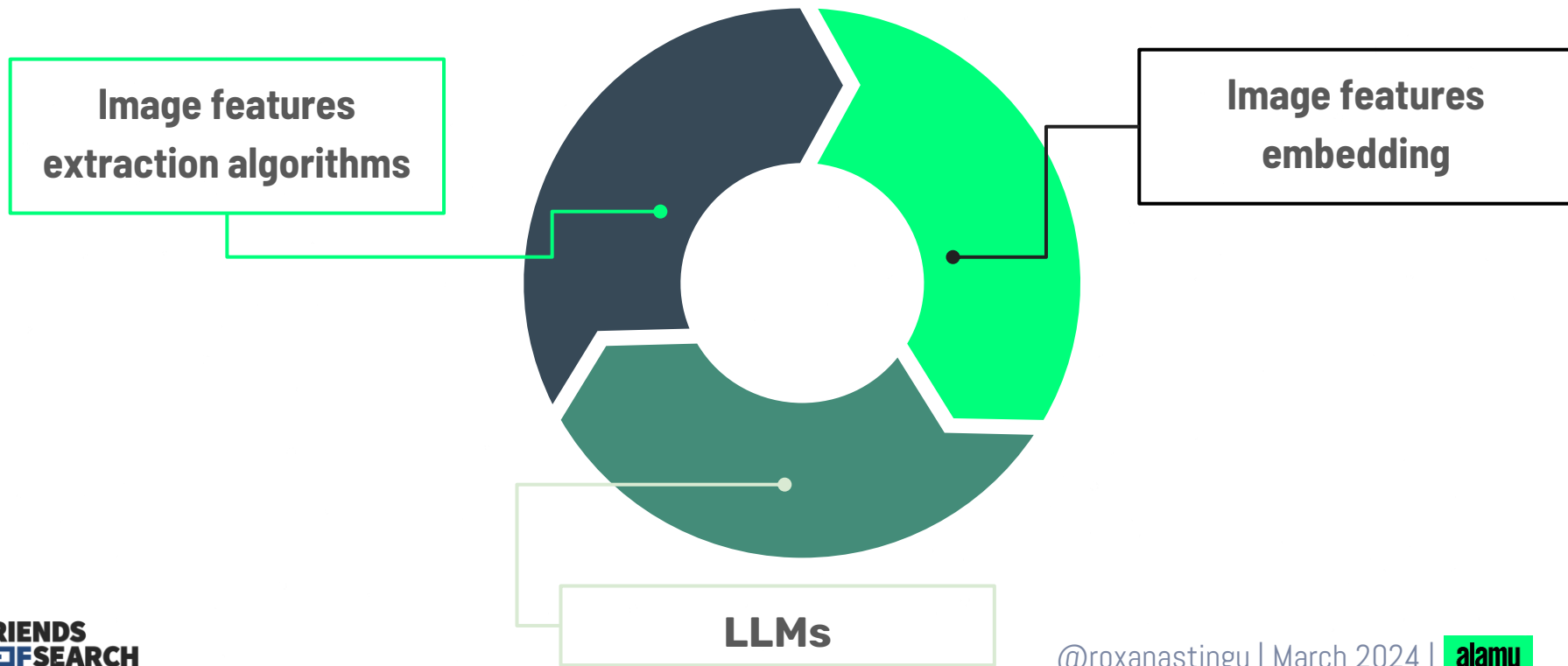
# Embeddings



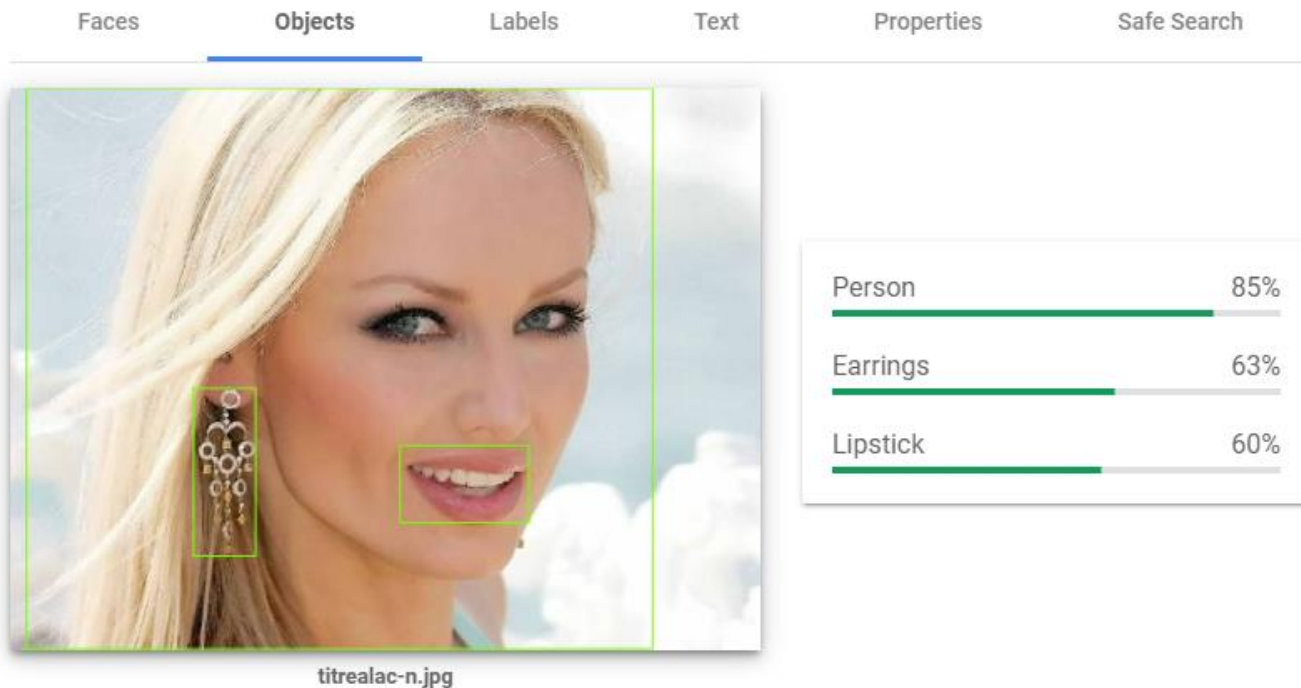
# Embeddings



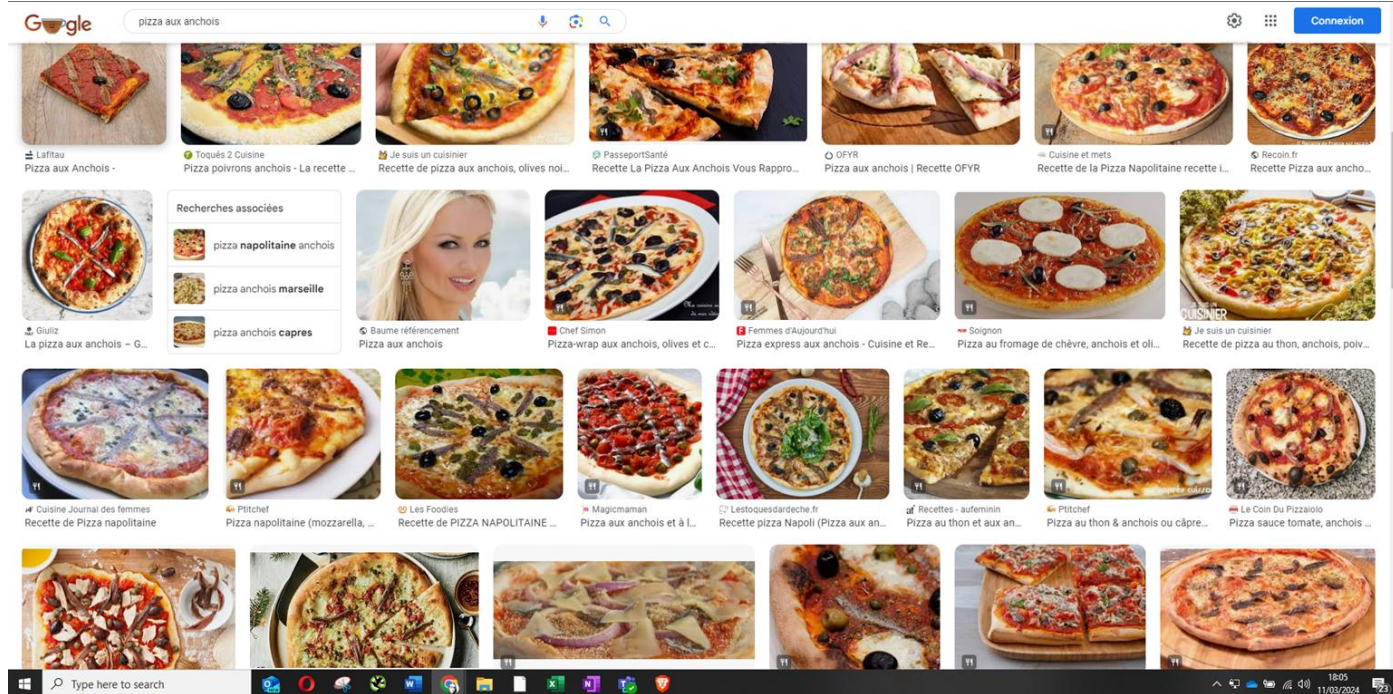
# It's all about balance



# Google can detect details in images



# Yet this woman is seen as pizza





# The greatest experiment

## Anchovy pizza

Posted on [June 20, 2011](#) by [gloq](#)

### What, how, but what is the relationship between seo and anchovy pizza?

We can't say that there really is a relationship, except that this article aims to do a test on Google image, and that we need "**anchovy pizza**" to gain strength. Or more seriously, we need it to see what we can do with this term in Google image.



"anchovy pizza"

### No image recognition for Google.

The sole objective of the test is to validate the fact that the image above will be positioned in GG images for the request "pizza with anchovies". There's no reason it shouldn't. For the moment, Google does not recognize images, but it takes into account the textual environment of these to determine the subject. And the environment here is much more **pizzeria** oriented than modeling.

Now, as the technical evolutions of image recognition go, it's a safe bet that the engines will one of these days be able to differentiate the pretty lady above from a vulgar pizza.

### What else can I say about this pizza?

We can say as for our test that this pizza recipe is not unanimous. Indeed, anchovies are welcomed as a delicacy by some, and as the worst for the taste buds by others.

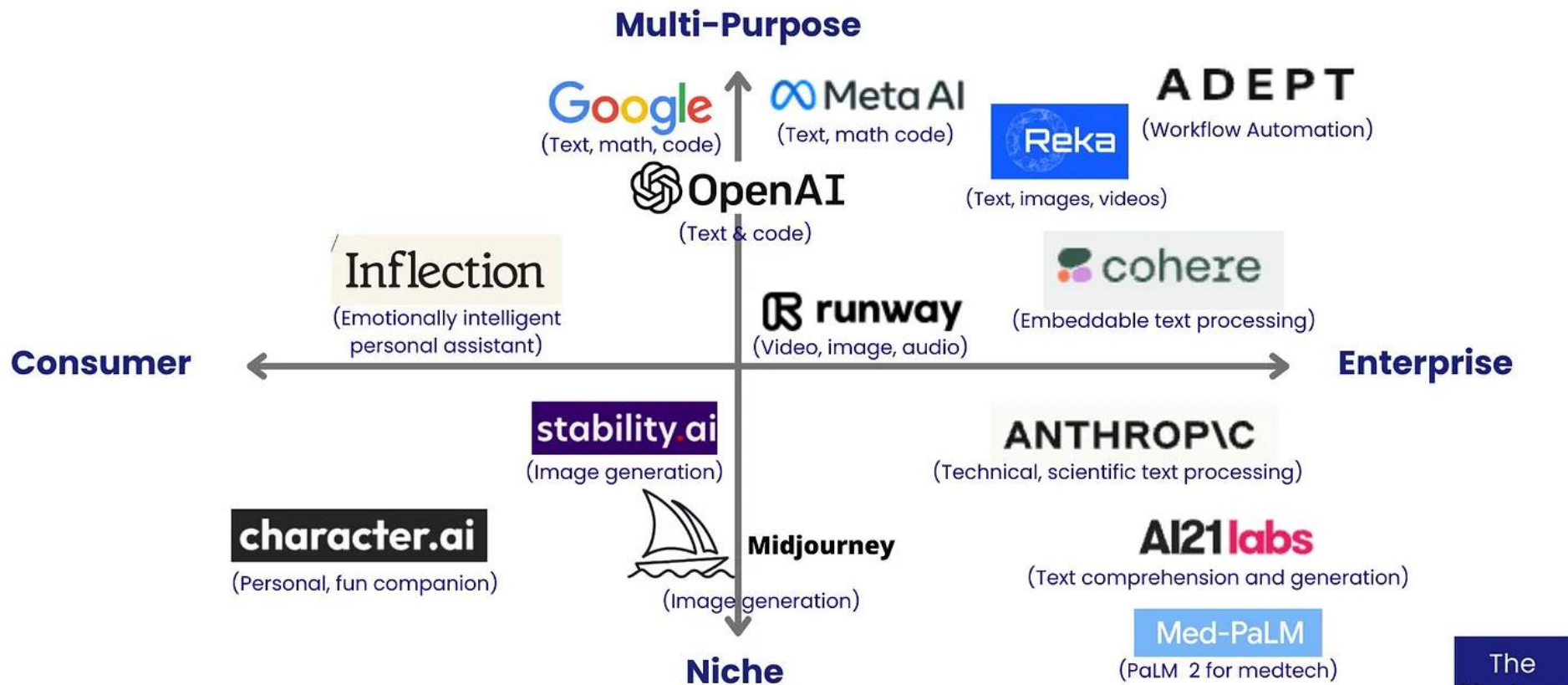
For the cranks who would like to know more about this subject, [go see here](#) , but overall, we don't care a bit, all we care about is seeing how our pizza will take to the king of engines.

### Anchovy pizza recipe

For the maniacs to whom we would have made hungry, it is necessary:

- tomatoes,
- **anchovies** (necessarily huhl)
- Garlic,
- pepper salt,
- a little olive oil ,

# Gen AI Companies with Foundational Models













# Getty Images sues AI art generator Stable Diffusion in the US for copyright infringement



# Who owns the copyright?

## author

The Copyright Designs & Patents Act 1988 states that, in the case of a computer-generated artwork, the 'author' (i.e. usually the first owner of copyright) will be "the person by whom the arrangements necessary for the creation of the work are undertaken". 14 Sept 2023



Strachan IP  
<https://strachanip.co.uk/copyright-in-ai-generated-art...>

UK Copyright in AI-Generated Artwork - Strachan IP

About featured snippets • Feedback



REUTERS® World Business Markets Sustainability Legal Breakingviews Technology Investigations

Litigation | Copyright | Technology | Intellectual Property | Litigation

## AI-generated art cannot receive copyrights, US court says

By Blake Brittain

August 21, 2023 7:29 PM GMT+1 · Updated 7 months ago



Firstly, there is the question of protection for AI-generated texts and images. In principle, only original works created by human authors are eligible for copyright protection, meaning AI-generated texts and images fall outside the scope of copyright law. 6 Oct 2023



Mondaq  
<https://www.mondaq.com/copyright/generative-ai-a...>

Netherlands: Generative AI And Copyright: The Story So Far

About featured snippets • Feedback

# Ethical foundation models

AI

## Shutterstock expands deal with OpenAI to build generative AI tools

Kyle Wiggers @kyle\_l\_wiggers / 6:52 PM GMT+1 • July 11, 2023

 Comment

## Adobe's Firefly: Ethical Generative AI to Revolutionize Content Creation



Roberto Cosentino · Follow


4 min read · Mar 21, 2023

## Ethical AI Image Initiative – Alamy Announces Industry Partnership

AI image generation technology opens up exciting creative possibilities for our customers, but it must be approached in the right way to protect the interests of Alamy content creators. As part of our commitment to exploring ethical opportunities within AI we are pleased to announce our participation in the 'The Fair Diffusion Program' - a ground-breaking generative AI initiative launched by [BRIA AI](#) in collaboration with [Alamy](#), [Getty Images](#), and [Envato](#). This is an alternative to the AI models already in existence that have been built on unlicensed content - instead it's designed to attribute and reward source material.



# Will AI-generated images rank?

**John Mueller** (He/Him) • 1st  
Senior Search Analyst / Search Relations team lead  
2mo •

I'm curious about your thoughts - is the use of AI-generated images on (non-art or AI-focused) websites equivalent to the use of stock photography?

My thinking ([#showerthought](#)) is if you care about a specific photo ("here's the suitcase we want to sell you"), you'll want to use a real photo. Real doesn't mean it's not digitally processed / cleaned / recolored / background removed with AI, but the basis is real, and it's meant to give the user a feel for something real.


On the other hand, if you just want to decorate your content, there's no real difference between stock photography / stock images and something initially AI-generated. They can be pleasing to the eye, make reading easier & more entertaining, but ultimately they're decoration.

Perhaps this is also a matter of the topic of the page. For some topics, people expect real images, and for others - perhaps most topics?, they don't mind either way. And to tie it over to SEO, my suspicion is that people are more likely to search visually for topics where they care about real images ("I want a small red suitcase, what are the options"), and not for pages that aren't expected to have real images.

So if you're wondering whether you should use AI-generated images on a page, it might be useful to ask yourself if you would use stock photography / imagery there.

Disclaimers: If you just want to save time & money by using your own AI-generated images, keep in mind that it's also easy to take a quick snapshot with your phone and call it "stock photography" -- it might not be as professional as you'd want on a business website though; making good things takes time & experience. None of this is SEO advice, there's no "ranking update" to pull from this post.

[#ai](#) [#midjourney](#) [#dalle](#) [#seo](#) [#photos](#) [#stockoptions](#) [#websitedesign](#)

 Zack Notes and 523 others      113 comments · 26 reposts

## ← Post



**Roxana Stingu**  
@RoxanaStingu

This is what I've been saying all this time - no need to fear stock (or now AI generated) images - you can't rank your own photos either if used out of context.

Thank you [@JohnMu](#) for touching on this subject.

[searchenginejournal.com/beyond-seo-joh...](https://searchenginejournal.com/beyond-seo-joh...)



**Roxana Stingu** @RoxanaStingu · Nov 17, 2023

This was a very important discussion for me that @Giridja made possible.

I don't think stock photos are great for SEO because I work for Alamy, I work for Alamy because I think stock photos can enrich any experience, SEO included.  
[x.com/Giridja/status...](#)

6:07 PM · Jan 9, 2024 · 260 Views

04

# What the future holds

# No one knows what the future holds



Garrett Sussman

@garrettsussman

WTF? @RoxanaStingu literally predicts Google two years in a row.

Stop it with this magic!

Do you work for Google? Lottery numbers?

Ridiculous.

[ipullrank.com/boring-generat...](http://ipullrank.com/boring-generat...)

ROXANA  
STINGU



WEBSITE: [Alamy](http://Alamy)

[@RoxanaStingu](https://twitter.com/RoxanaStingu)

## 🧙‍♀️ 2024 CONTENT AND SEO GENERATIVE AI PREDICTION

Following in Bing's footsteps, Google will also add text-to-image functionality in search through its Imagen model, as already announced.

Further to this, they might use it as part of search, instead of writing the query and seeing which images match, we might have the option to describe the image we want, have AI generate it and then use that to search for visually similar images. A combo of generative AI and Google Lens if you will.

ALT

# No, I can't guess the lottery numbers

ROXANA  
STINGU



WEBSITE: [Alamy](#)

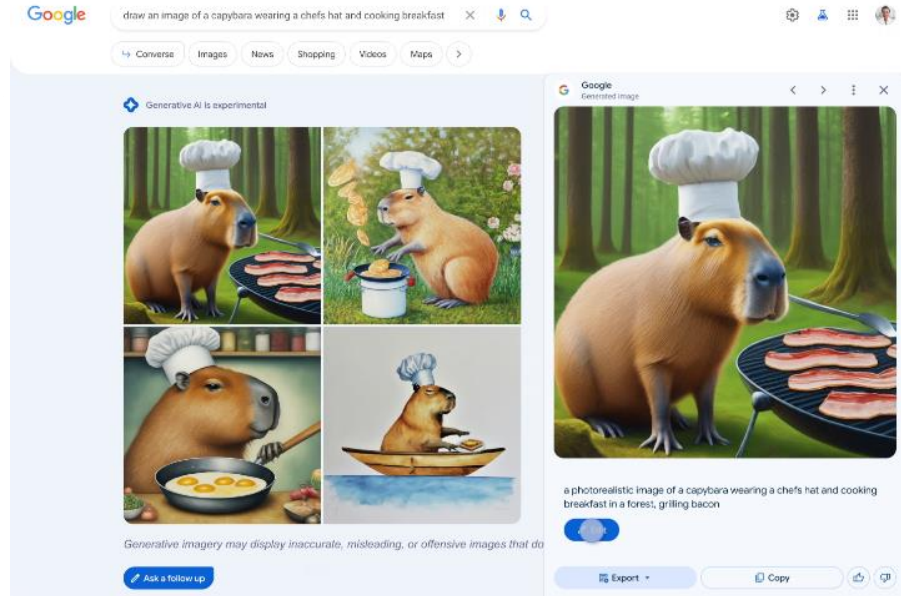
[@RoxanaStingu](#)

## 2024 CONTENT AND SEO GENERATIVE AI PREDICTION

Following in Bing's footsteps, Google will also add text-to-image functionality in search through its Imagen model, as already announced.

Further to this, they might use it as part of search. Instead of writing the query and seeing which images match, we might have the option to describe the image we want, have AI generate it and then use that to search for visually similar images. A combo of generative AI and Google Lens if you will.

I'm also expecting Google to create a series of image enhancement tools that will be free for Google users, ahem, as \*free\* as things are with Google in general (ads).



# My ridiculous prediction

ROXANA  
STINGU



WEBSITE: [Alamy](#)

 [@RoxanaStingu](#)

## **RIDICULOUS, OUT-OF-THE-BOX, SUPER-RANDOM GENERATIVE AI 2024 PREDICTION**

Upon introducing text-to-image functionality in Image Search, Google will try to monetise it by adding advertisements within the generated images.

When users search for an image based on a text description, Google will display relevant advertisements that are visually integrated into the generated image. For example, if a user searches for a “red sports car,” Google can overlay an ad for a local car dealership or related automotive products directly onto the image.





# MEET OUR TEAM



**ARNOLD DONALD**  
Q.C.



**AARON COOK Q.C.**  
Barrister



**HILDA MILLER Q.C.**  
Barrister



**GARY WEBB**  
Barrister



<https://thispersondoesnotexist.com/>

StyleGAN (Naras et al.)



# Provenance

Google DeepMind

SynthID



**FRIENDS**  
**SEARCH**

BY DORA lab

@roxanastingu | March 2024 | **alamy**

# Provenance

ARTIFICIAL INTELLIGENCE / TECH / OPENAI

## OpenAI is adding new watermarks to DALL-E 3



Image: OpenAI

/ OpenAI says watermarks in image metadata are not perfect, but they help build trust of digital information.

By [Emilia David](#), a reporter who covers AI. Prior to joining The Verge, she covered the intersection between technology, finance, and the economy.

Feb 6, 2024, 10:32 PM GMT



9 Comments (9 New)

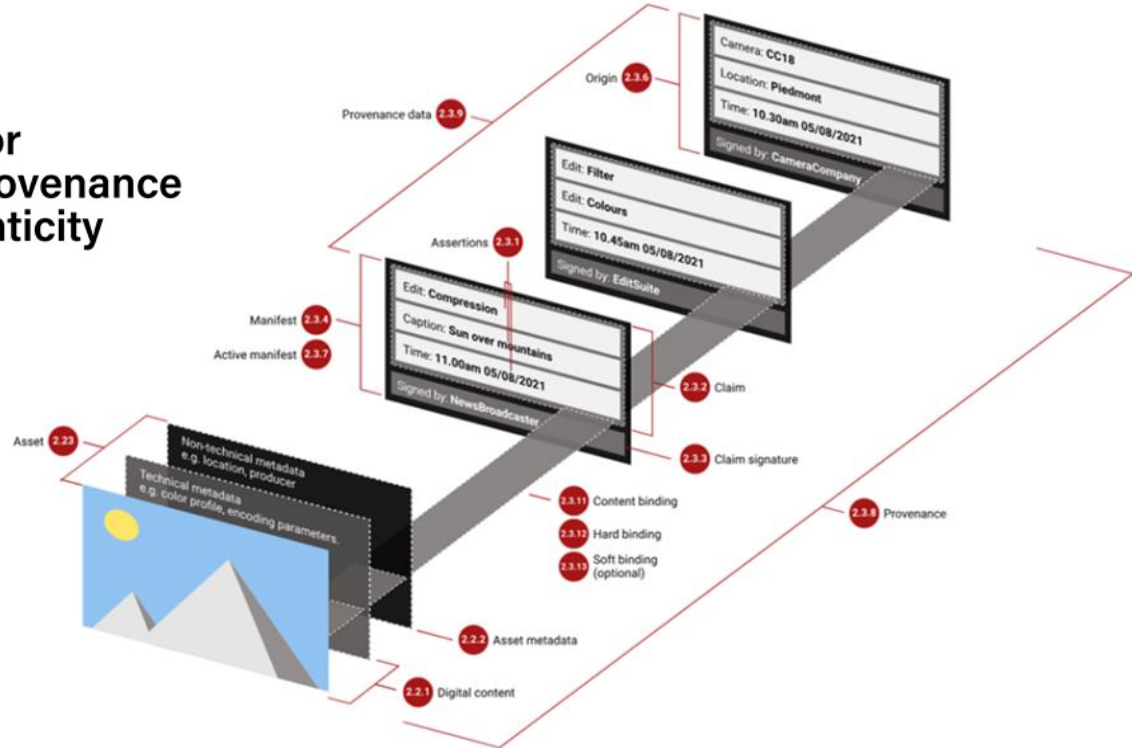
## Meta to fight AI-generated fake news with 'invisible watermarks'

Meta will make use of a deep-learning model to apply watermarks to images generated with its AI tool, which would be invisible to the human eye.

# One solution to rule them all?



**C2PA** Coalition for  
Content Provenance  
and Authenticity



# Thank you!