How Google Discover REALLY Works

This is all based on the Google leak and tallies up with my experience of content that does well in Discover over time. I have pulled out what I think are the most prominent Discover proxies and grouped them into what seems like the appropriate workflow.

Like a disgraced BBC employee, thoughts are my own.

TL;DR

1. Your site needs to be seen as a **“trusted source”** with low SPAM, evaluated by proxies like **publisher trust score,** in order to be eligible.
2. Discover is driven by a six-part pipeline, using **good vs. bad clicks** (long dwell time vs. pogo-sticking) and **repeat visits** to continuously score and re-score content quality.
3. Fresh content gets an initial boost. Success hinges on a strong CTR and positive **early-stage engagement** (good clicks/shares from all channels count, not just Discover).
4. Content that aligns with a user’s interests is prioritised. To optimize, focus on your areas of topical authority, use a compelling headline(s), be entity-driven, and use large (1200px+) images.



Image Credit: Harry Clarkson-Bennett

I count 15 different proxies that Google uses to guide satiate the doomscrollers' desperate need for quality content in the Discover feed. It's not *that* different to how traditional Google search works.

But traditional search (a high-quality pull channel) is worlds apart from Discover. Audiences killing time on trains. At their in-laws. The toilet. Given they're part of the same ecosystem, they're bundled together into one monolithic entity.

And here's how it works.

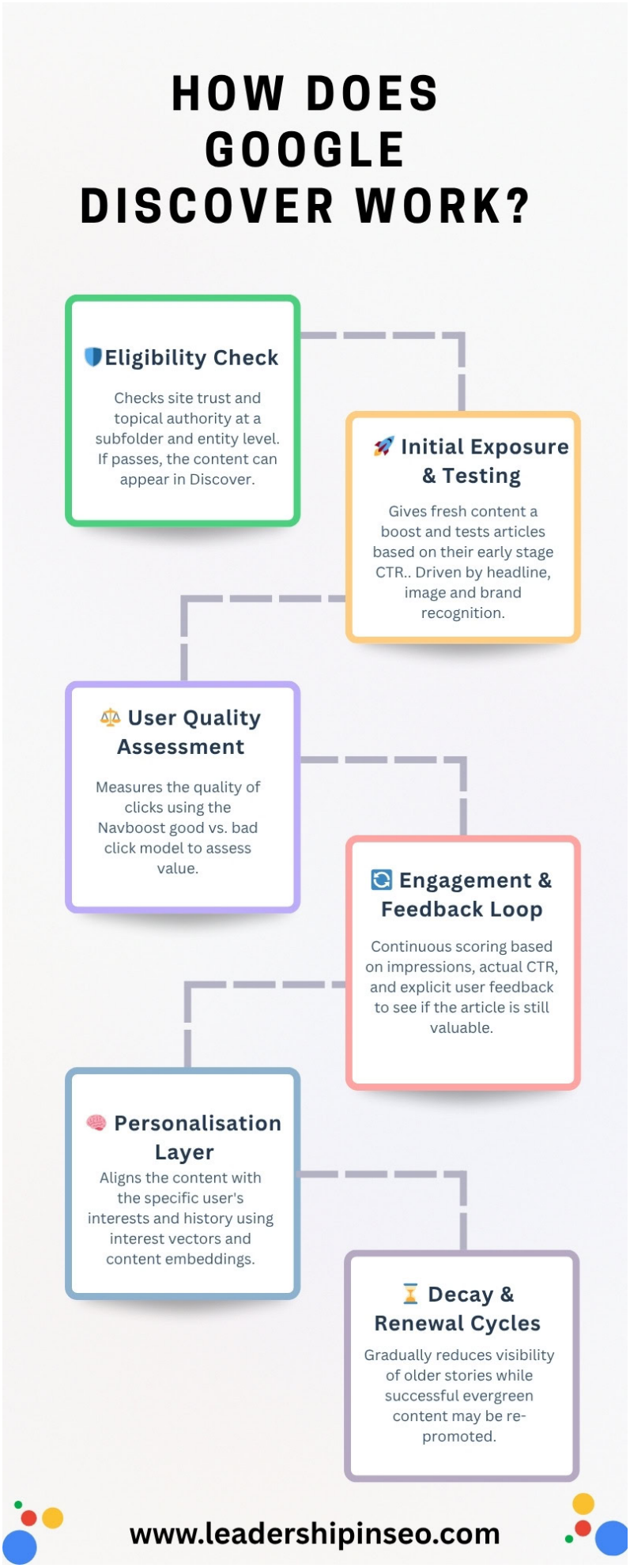


Image Credit: Harry

Clarkson-Bennett

Google’s Discover Guidelines

This section is boring, and Google’s guidelines around eligibility are exceptionally vague:

- Content is automatically eligible to appear in Discover if it is indexed by Google and meets Discover’s content policies.
- Any kind of dangerous, spammy, deceptive, or violent/vulgar content gets filtered out.

> *“…Discover makes use of many of the same signals and systems used by Search to determine what is… helpful, reliable, people-first content.”*

Then they give some solid, albeit beige advice around quality titles – clicky, not baity as John Shehata would say. Ensuring your featured image is at least 1200px wide and creating timely, value-added content.

But we can do better.

Discover’s Six-Part Content Pipeline

From cradle to grave, let’s review exactly how your content does or, in most cases, doesn’t appear in Discover. As always, remembering I have made these clusters up, albeit based on real Google proxies from the Google leak.

1. **Eligibility check and baseline filtering.**
2. **Initial exposure and testing.**
3. **User quality assessment.**
4. **Engagement and feedback loop.**
5. **Personalization layer.**

6. **Decay and renewal cycles.**

Eligibility And Baseline Filtering

For starters, your site has to be eligible for Google Discover. This means you are seen as a “trusted source” on the topic, and you have a low enough SPAM score that the threshold isn’t triggered.

There are three primary proxy scores to account for eligibility and baseline filtering:

- **is_discover_feed_eligible:** a Boolean feature that filters non-eligible pages.
- **publisher_trustScore:** a score that evaluates publisher reliability and reputation.
- **topicAuthority_discover:** a score that helps Discover identify trusted sources at the topic level.

The site’s reputation and topical authority are ranked for the topic at hand. These three metrics help evaluate whether your site is eligible to appear in Discover.

Initial Exposure And Testing

This is very much the freshness stage, where fresh content is given a temporary boost (because contemporary content is more likely to satiate a dopamine addicted mind).

- **freshnessBoost_discover:** provides a temporary boost for fresh content to keep the feed alive.
- **discover_clicks:** where early-stage article clicks are used as a predictor of popularity.
- **headlineClickModel_discover:** is a predictive CTR model based on the headline and image.

> *I would hypothesize that using a Bayesian style predictive model, Google applies learnings at a site and subfolder level to predict likely CTR. The more quality content you have*

> *published over time (presumably at a site, subfolder and author level), the more likely you are to feature.*
>
> *Because there is less ambiguity. A key feature of SEO now.*

User Quality Assessment

An article is ultimately judged by the quality of user engagement. Google uses the good and bad click style model from Navboost to establish what is and isn't working for users. Low CTR and/or pogo-sticking style behavior downgrades an article's chance of featuring.

Valuable content is decided by the good vs bad click ratio. Repeat visits are used to measure lasting satisfaction and re-rank top-performing content.

- **discover_blacklist_score:** Penalty for spam, misinformation, or clickbait.
- **goodClicks_discover:** Positive user interactions (long dwell time).
- **badClicks_discover:** Negative interactions (bounces, short dwell).
- **nav_boosted_discover_clicks:** Repeat or return engagement metric.

The quality of the article is then measured by its user engagement. As Discover is a personalized platform, this can be done accurately and at scale. Cohorts of users can be grouped together. People with the same general interests are served the content if, by the algorithm's standard, they should be interested.

But if the overly clicky or misleading title delivers poor engagement (dwell time and on-page interactions), then the article may be downgraded. Over time, this kind of practice can compound and nerf your site completely.

Tone/Technique	Headline
Outrage/Shock	**ROYAL SHOCKWAVE:** Andrew's Title GONE! You Won't **BELIEVE** the TWO Words the Queen Said When She Stripped Him!
Curiosity Gap (Immediate Secret)	The Secret Phone Call That **ENDED** Prince Andrew's Royal Life Forever (**Insiders Reveal Everything!**)
Curiosity Gap (Embarrassing Detail)	HE'S OUT! The One Tiny, Embarrassing Detail That Finally Forced the Queen to **Ban Prince Andrew for Good** (Click to See the Cringey Photo!)
Sense of Urgency/Drama	**THE ROYAL SCANDAL EXPLODES:** What Andrew Was Caught Doing **Seconds Before His Title Was Snatched!** (The Video Everyone is Sharing!)
Irrelevant Speculation	Andrew's Title: POOF! GONE! Find Out **Which Celebrity** Will Get His Duke Status Next (The Answer Will Stun You!)

Headlines like this are a one way ticket to devaluing your brand in the eyes of people and search engines (Image Credit: Harry Clarkson-Bennett)

> *Important to note that this click data doesn't have to come from Discover. Once an article is out in the ether – it's been published, shared on social, etc. – Chrome click data is stored and is applied to the algorithm.*
>
> *So, the more quality click data and shares you can generate early in an article's lifecycle (accounting for the importance of freshness), the better your chance of success on Discover. Treat it like* ***a viral platform. Make noise. Do marketing****.*

Engagement And Feedback Loop

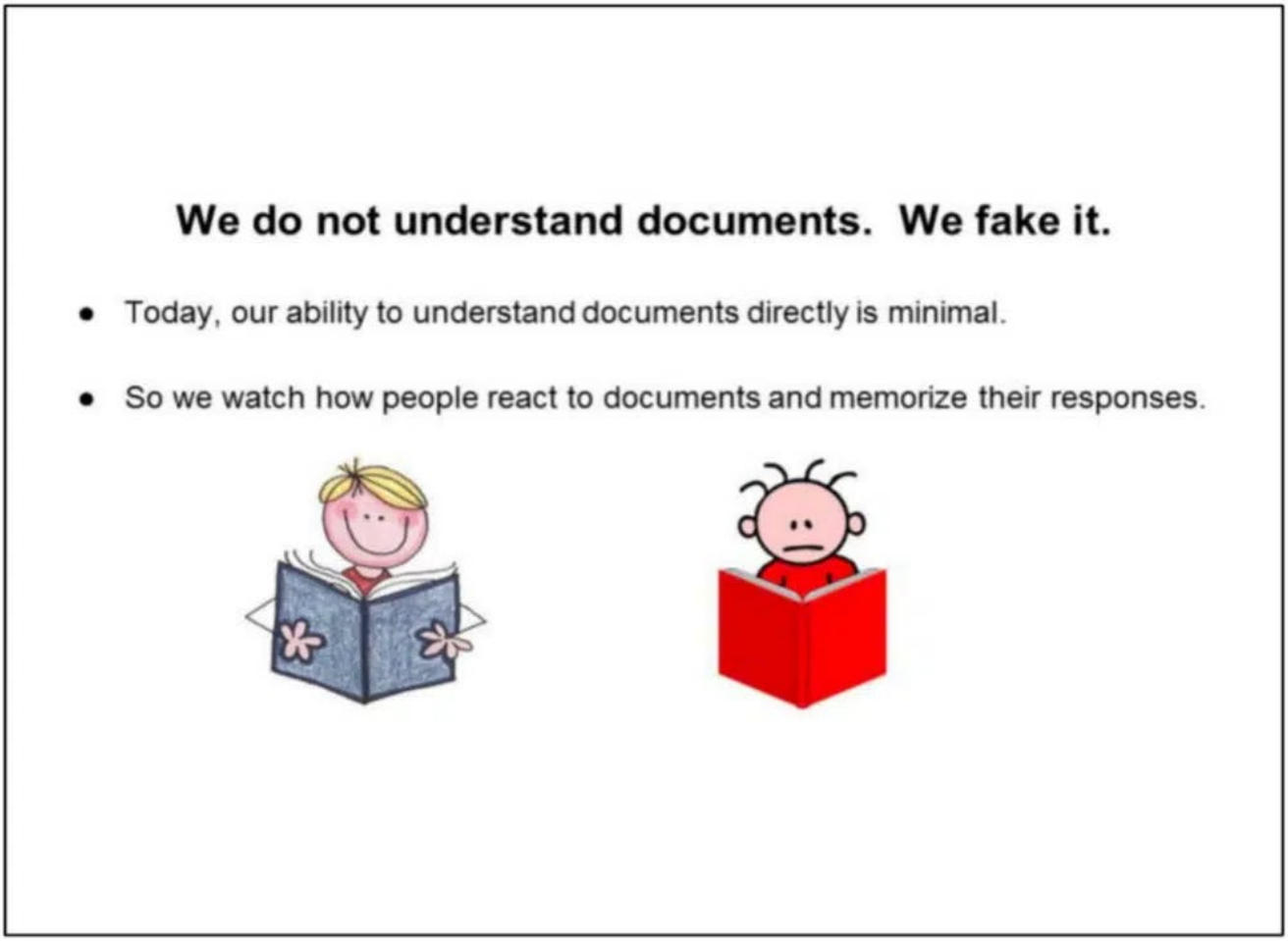
Once the article enters the proverbial fray, a scoring and rescoring loop begins. Continuous CTR, impressions, and explicit user feedback (like, hate, and "don't show me this again, please" style buttons) feed models like Navboost to refine what gets shown.

- **discover_impressions:** The number of times an article appears in a Discover feed.
- **discover_ctr:** Clicks divided by impressions. Impressions

and click data feed CTR modelling

- **discover_feedback_negative:** Specific user feedback, i.e., not interested suppresses content for individuals, groups, and on the platform as a whole.

These behavioral signals define an article's success. It lives or dies on relatively simple metrics. And the more you use it, the better it gets. Because it knows what you and your cohort are more likely to click and enjoy.



This is as true in Discover as it is in the main algorithm. Google admitted as such in the DoJ rulings. (Image Credit: Harry Clarkson-Bennett)

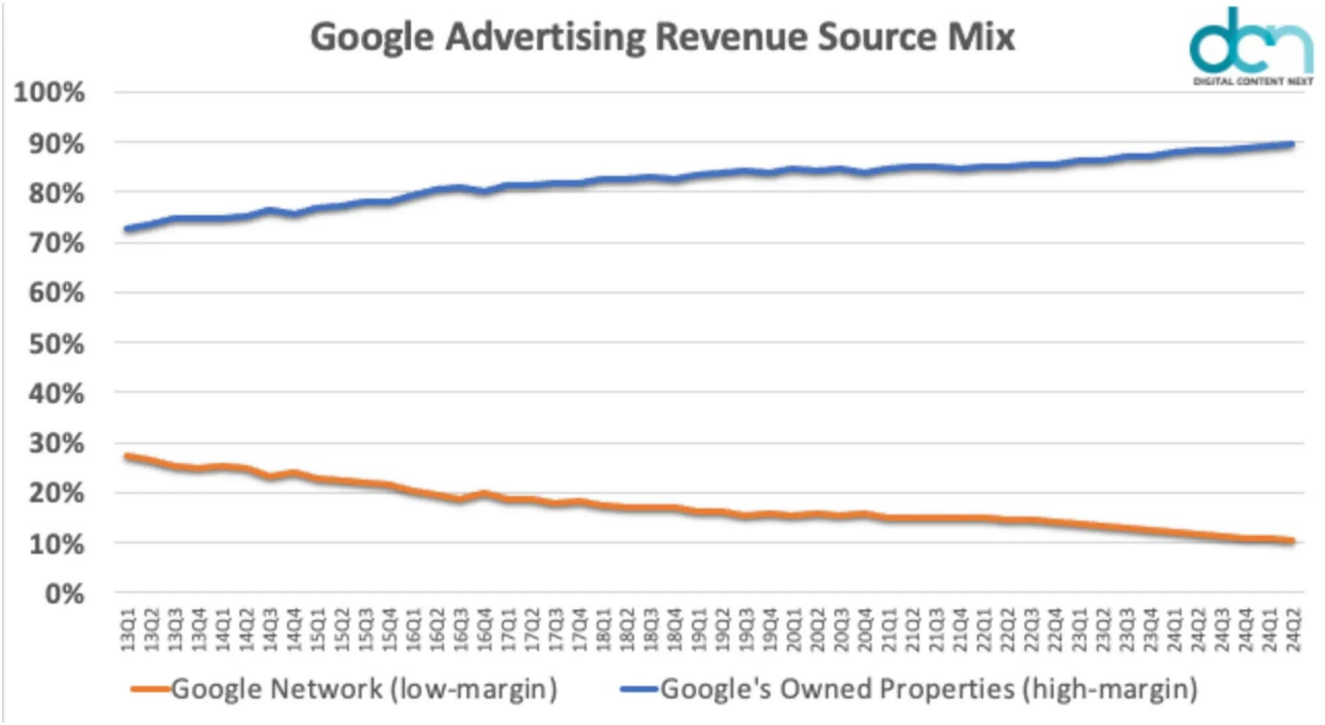
> *I imagine headline and image data are stored so that the algorithm can apply some rigorous standards to statistical modelling. Once it knows what types of headlines, images and articles perform best for specific cohorts, personalisation becomes effective faster.*

Personalization Layer

Google knows a lot about us. It's what its business is built on. It collects a lot of non-anonymized data (credit card details, passwords, contact details, etc.) alongside every conceivable interaction you have with webpages.

Discover takes personalization to the next level. I think it may offer an insight into how part of the SERP could look like in the future. A personalized cluster of articles, videos, and social posts designed to hook you in embedded somewhere alongside search results and AI Mode.

All of this is designed to keep you on Google's owned properties for longer. Because they make more money that way.



Hint: They want to keep you around because they make more money (Image Credit: Harry Clarkson-Bennett)

- **contentEmbeddings_discover:** Content embeddings determine how well the content aligns with the user's interests. This powers Discover's interest-matching engine.
- **personalization_vector_match:** This module dynamically

personalises the user's feed in real-time. It identifies similarity between content and user interest vectors.

Content that matches well with your personal and cohort's interest will be boosted into your feed.

You can see the site's you engage with frequently using the site engagement page in Chrome (from your toolbar: chrome://site-engagement/) and every stored interaction with histograms. This histogram data indirectly shows key interaction points you have with web pages, by measuring the browser's response and performance around those interactions.

It doesn't explicitly say *user A clicked X*, but logs the technical impact, i.e., how long did the browser spending processing said click or scroll.

Decay And Renewal Cycles

Discover boosts freshness because people are thirsty for it. By boosting fresh content, older or saturated stories naturally decay as the news cycle moves on and article engagement declines.

For successful stories, this is through market saturation.

- **freshnessDecay_timer:** This module measures recency decay after initial exposure, gradually reducing visibility to make way for fresher content.
- **content_staleness_penalty:** Outdated content or topics are given a lower priority once engagement starts to decline to keep the feed current.

Discover is Google's answer to a social network. None of us spend time in Google. It's not *fun*. I use the word fun loosely. It isn't designed to hook us in and ruin our attention spans with constant spiking of dopamine.

But Google Discover is clearly on the way to that. They want to make it a destination. Hence, all the recent changes where

you can “catch up” with creators and publishers you care about across multiple platforms.

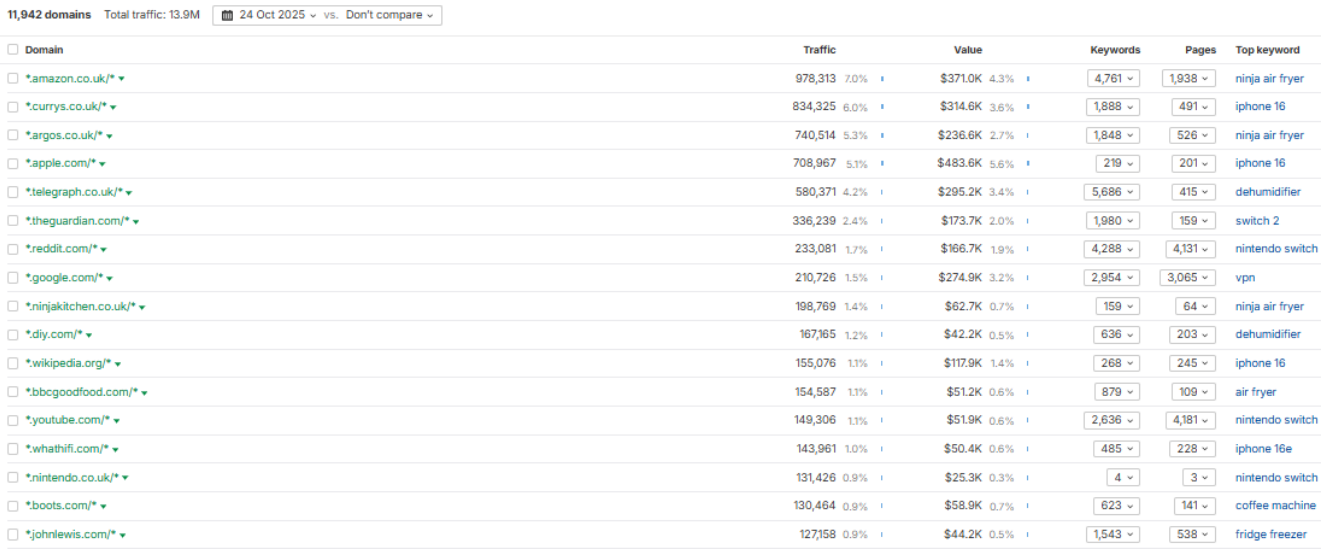
Videos, social posts, articles … the whole nine yards. I wish they’d stop summarizing literally everything with AI, however.

My 11-Step Workflow To Get The Most Out Of Google Discover

Follow basic principles and you will put yourself in good stead. Understand where your site is topically strong and focus your time on content that will drive value. Multiple ways you can do this.

If you don’t feature much in Discover, you can use your Search Console click and impressions data to identify areas where you generate the highest value. Where you are topically authoritative. I would do this at a subfolder and entity level (e.g., politics and Rachel Reeves or the Labor Party).

Also worth breaking this down in total and by article. Or you can use something like Ahrefs’ Traffic Share report to determine your share of voice via third-party data.



11,942 domains Total traffic: 13.9M 24 Oct 2025 vs. Don't compare

Domain	Traffic	Value	Keywords	Pages	Top keyword
.amazon.co.uk/	978,313 7.0%	$371.0K 4.3%	4,761	1,938	ninja air fryer
.currys.co.uk/	834,325 6.0%	$314.6K 3.6%	1,888	491	iphone 16
.argos.co.uk/	740,514 5.3%	$236.6K 2.7%	1,848	526	ninja air fryer
.apple.com/	708,967 5.1%	$483.6K 5.6%	219	201	iphone 16
.telegraph.co.uk/	580,371 4.2%	$295.2K 3.4%	5,686	415	dehumidifier
.theguardian.com/	336,239 2.4%	$173.7K 2.0%	1,980	159	switch 2
.reddit.com/	233,081 1.7%	$166.7K 1.9%	4,288	4,131	nintendo switch
.google.com/	210,726 1.5%	$274.9K 3.2%	2,954	3,065	vpn
.ninjakitchen.co.uk/	198,769 1.4%	$62.7K 0.7%	159	64	ninja air fryer
.diy.com/	167,165 1.2%	$42.2K 0.5%	636	203	dehumidifier
.wikipedia.org/	155,076 1.1%	$117.9K 1.4%	268	245	iphone 16
.bbcgoodfood.com/	154,587 1.1%	$51.2K 0.6%	879	109	air fryer
.youtube.com/	149,306 1.1%	$51.9K 0.6%	2,636	4,181	nintendo switch
.whathifi.com/	143,961 1.0%	$50.4K 0.6%	485	228	iphone 16e
.nintendo.co.uk/	131,426 0.9%	$25.3K 0.3%	4	3	nintendo switch
.boots.com/	130,464 0.9%	$58.9K 0.7%	623	141	coffee machine
.johnlewis.com/	127,158 0.9%	$44.2K 0.5%	1,543	538	fridge freezer

Essentially share of voice data (Image Credit: Harry Clarkson-Bennett)

Then really focus your time on a) areas where you’re already

authoritative and b) areas that drive value *for your audience*.

Assuming you're not focusing on NSFW content and you're vaguely eligible, here's what I would do:

1. Make sure you're meeting **basic image requirement**s. 1200 pixels wide as a minimum.
2. **Identify your areas of topical authority.** Where do you already rank effectively at a subfolder level? Is there a specific author who performs best? Try to build on your valuable content hubs with content that should drive extra value in this area.
3. Invest in **content that will drive real value** (links and engagement) in these areas. Do not chase clicks via Discover. It's a one-way ticket to clickbait city.
4. Make sure you're **plugged into the news cycle**. Being first has a huge impact on your news visibility in search. If you're not first on the scene, make sure you're adding something additional to the conversation. Be bold. Add value. Understand how news SEO really works.
5. **Be entity-driven**. In your headlines, first paragraph, subheadings, structured data, and image alt text. Your page should remove ambiguity. You need to make it incredibly clear who this page is about. A lack of clarity is partly why Google rewrites headlines.
6. **Use the Open Graph title.** The OG title is a headline that doesn't show on your page. Primarily designed for social media use, it is one of the most commonly picked up headlines in Discover. It can be jazzy. Curiosity led. Rich. Interesting. But still entity-focused.
7. Make sure you **share content likely to do well on Discover** across relevant push channels early in its lifecycle. It needs to outperform its predicted early-stage performance.*
8. Create **a good page experience**. Your page (and site) should be fast, secure, ad-lite, and memorable for the

right reasons.

9. Try to drive **quality onward journeys**. If you can treat users from Discover differently to your main site, think about how you would link effectively for them. Maybe you use a pop-up “we think you’ll like this next” section based on a user’s scroll depth of dwell time.
10. **Get the traffic to convert.** While Discover is a personalized feed, the standard scroller is not very engaged. So, focus on easier conversions like registrations (if you’re a subscriber first company) or advertising revenue et al.
11. Keep a record of your best performers. Evergreen content can be refreshed and repubbed year after year. It can still drive value.

**What I mean here is if your content is predicted to drive three shares and two links, if you share it on social and in newsletters and it drives seven shares and nine links, it is more likely to go viral.*

As such, the algorithm identifies it as ‘Discover-worthy.’

More Resources:

This was originally published on Leadership in SEO.

Featured Image: Roman Samborskyi/Shutterstock

https://www.searchenginejournal.com/how-google-discover-really-works/559304/