

# SEO vs paywall

do it the right way



## Jakub Sawa

- Organic Growth Director at Wyborcza.pl
- 20 years of SEO experience at online businesses
- 10 years of SEO experience at online publishers (RASP)
- 3 years of SEO paywall experience at Wyborcza.pl
- 2 years as publisher of largest, weekly SEO newsletter in Poland

mail me: [ja@jakubsawa.pl](mailto:ja@jakubsawa.pl)

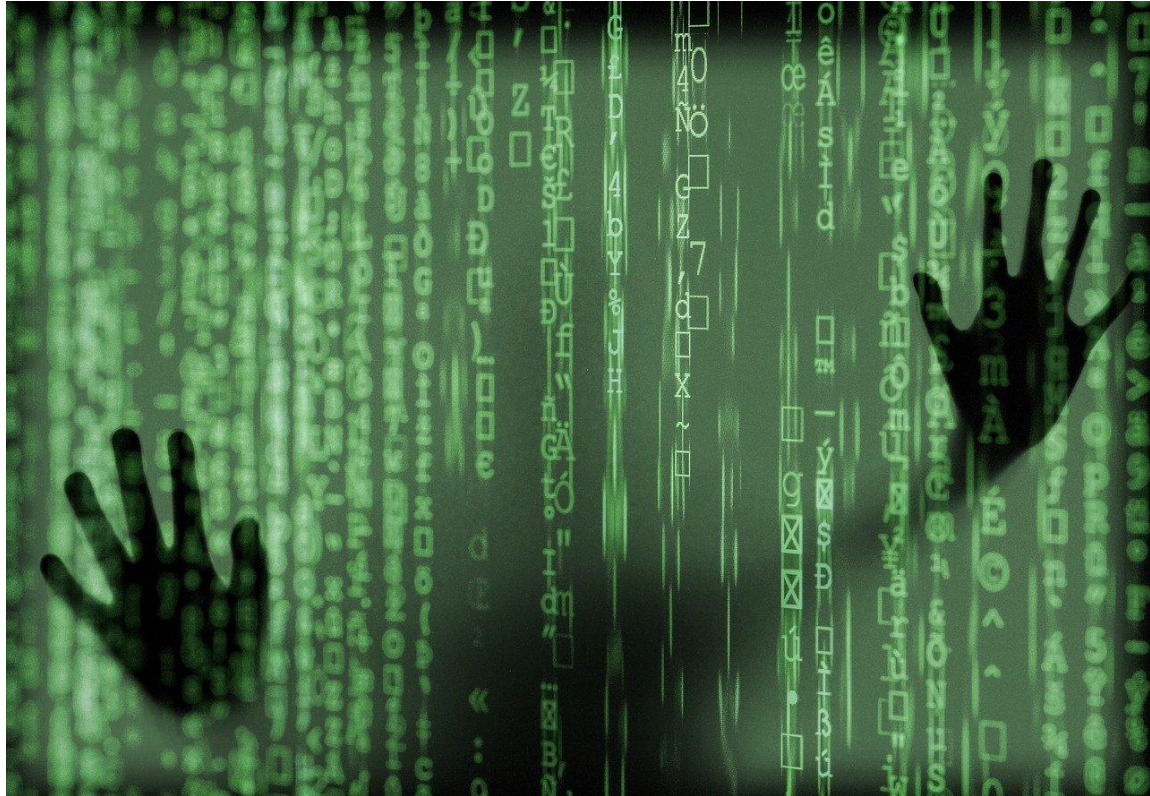
wyborcza.pl



# What is paywall?

---

“a method of **restricting access** to online content via a paid / free subscription”. A paywall acts as **a digital barrier**, and users can only access content behind it after they pay a fee or complete a subscription.



## Technical aspects of SEO and paywall

There are several paywall models from an SEO perspective. A hard paywall that works server-side or a soft paywall based on JS and CSS.

Paywall is a digital barrier for users - for Google and other search engines too.



# Does your paywall is Google friendly?

"Gierasimow zakładał, że równie istotne jak działania militarne są jedni

Wszystko

Grafika

Wiadomości

Zakupy

Mapy

Więcej

Narzędzia

Około 1 wyników (0,37 s)

Wszystkie słowa począwszy od "**chociaż**" zostały pominięte, gdyż zapytania nie mogą zawierać więcej niż 32 słów.

<https://wyborcza.pl> > 7,75398,28968903.gen-kraszewski-...

## Gen. Kraszewski: Putin zaanektuje część Ukrainy i użyje broni ...

6 godzin temu — Rozumiem, że Putinowi nie zależy szczególnie na życiu swoich żołnierzy, ale w wyniku wojny atomowej przez co najmniej pokolenie nie byłby w ...

The screenshot shows the Wyborcza.pl website interface. At the top, there is a search bar and navigation menu. The main article title is "Gen. Kraszewski: Putin zaanektuje część Ukrainy i użyje broni jądrowej? On chce straszyć Zachód". Below the title is a photo of General Kraszewski. The article text is partially visible, starting with "Rozumiem, że Putinowi nie zależy szczególnie na życiu swoich żołnierzy, ale w wyniku wojny atomowej przez co najmniej pokolenie nie byłby w stanie korzystać z ziem, które zdobył - mówi gen. dr Jarosław Kraszewski". A red dashed line indicates a paywall barrier. Below the barrier, there is an advertisement for "BOTANICAL CORE ZIELONA HERBATA BIO" featuring a woman holding a tea bag. The article text continues below the ad, mentioning "Gen. dr Jarosław Kraszewski, były dyrektor departamentu zwierzchnictwa i sił siłami zbrojnymi Biura Bezpieczeństwa Narodowego. Zaczajmy od tego, co obiecał Putin i jego współpracownicy, mówiąc o 'takturowej wojnie jądrowej'". At the bottom, there is another advertisement for "10" featuring a woman and a child.

paywall

Gierasimow zakładał, że równie istotne jak działania militarne są jednak działania w sferze informacyjnej. A Rosja nie tylko przegrany wojnę militarną, ale też wojnę informacyjną. Temu tymi wypracowani niebezpieczeństwo odkrywać i inicjatywę chociaż w tej ostatniej dziedzinie. To musi być sytuacja na strach.

# Recognize Google like a pro

You need a mechanism to let Googlebot into your site - Googlebot **can't log in** or **accept cookies**.

You need to **recognize Googlebot** and confirm that it is real Googlebot and not fake one.

**Tip:** You can recognize it by revDNS (simply checking if request came from Google networks).

<https://developers.google.com/search/docs/crawling-indexing/verifying-googlebot>



```
$ host 66.249.66.1  
1.66.249.66.in-addr.arpa domain name pointer crawl-66-249-66-1.googlebot.com.  
  
$ host crawl-66-249-66-1.googlebot.com  
crawl-66-249-66-1.googlebot.com has address 66.249.66.1
```

# “Cloaking” articles with paywall

---

Google needs to know that you are using a paywall on your website.

Showing different versions of an article to a user and a bot is called cloaking and is penalized by Google.

The only “legal” way is to mark your content as a paywall.

<https://developers.google.com/search/docs/appearance/structured-data/paywalled-content>

```
<html>
<head>
  <title>Article headline</title>
  <script type="application/ld+json">
    {
      "@context": "https://schema.org",
      "@type": "NewsArticle",
      "headline": "Article headline",
      "image": "https://example.org/thumbnail1.jpg",
      "datePublished": "2025-02-05T08:00:00+08:00",
      "dateModified": "2025-02-05T09:20:00+08:00",
      "author": {
        "@type": "Person",
        "name": "John Doe"
      },
      "description": "A most wonderful article",
      "isAccessibleForFree": "False",
      "hasPart":
        {
          "@type": "WebPageElement",
          "isAccessibleForFree": "False",
          "cssSelector": ".paywall"
        }
    }
  </script>
</head>
<body>
  <div class="non-paywall">
    Non-Paywalled Content
  </div>
  <div class="paywall">
    Paywalled Content
  </div>
</body>
</html>
```

# Structured data and JSON+ld

Use **structured data** and preferably as **ld+json** to mark paywalled content.

Are using **Accelerated Mobile Pages?** – implement a paywall module.

Implement `<meta name="robots" content="noarchive">`

Technical elements **are crucial** as part of your SEO strategy because, without them, it can cause a **negative impact** on your organic search performance.

**Tip:** Do you know that SEO specialists could read your closed AMP articles using Rich Results Tool – check your paywall here

<https://search.google.com/test/rich-results>

```
<html>
<head>
  <title>Article headline</title>
  <script type="application/ld+json">
    {
      "@context": "https://schema.org",
      "@type": "NewsArticle",
      "headline": "Article headline",
      "image": "https://example.org/thumbnaill1.jpg",
      "datePublished": "2025-02-05T08:00:00+08:00",
      "dateModified": "2025-02-05T09:20:00+08:00",
      "author": {
        "@type": "Person",
        "name": "John Doe"
      },
      "description": "A most wonderful article",
      "isAccessibleForFree": "False",
      "hasPart":
        {
          "@type": "WebPageElement",
          "isAccessibleForFree": "False",
          "cssSelector": ".paywall"
        }
    }
  </script>
</head>
<body>
  <div class="non-paywall">
    Non-Paywalled Content
  </div>
  <div class="paywall">
    Paywalled Content
  </div>
</body>
</html>
```

# SEO rules for paywalled website

---

1. Search engines needs access to your articles
2. If you have server-side paywall - you need to recognize SE bots
3. Structured data are important - beware of a cloaking
4. AMP has paywall module too
5. Use meta name noarchive

**SEO + paywall = ❤️**

# How to implement a paywall in the secure way

---

1. Make “test” version of your website – test everything there
2. Choose one “not so critical” folder or directory on your website – implement a paywall there – test everything (as I said before)
3. Wait 1-2 weeks for Google to crawl, render, index and rank your content – check if you can find your content at SERP’s
4. Implement a paywall at 50% of your website – check again, check mobile, check desktop
5. If you're 100% sure everything is fine – go full throttle
6. Monitor your visibility and traffic on daily basis

# Google is diversified

From a publisher's perspective, Google is composed of:

- **Google Universal Search**  
(“blue links” for evergreens) mobile and desktop
- **Google News and Top Stories**  
(for news) mobile and desktop
- **Google Discover**  
(for both) only mobile

**Different sources = different strategy!**

# Google Universal Search

Universal Search is a stable source of traffic, but difficult to get (competition) and requires a long-term strategy.

Mobile and desktop results could be different from each other and depending on the topic of the site have different shares. A service for accountants will have a large desktop share, a small mobile share and on weekends the traffic will be almost zero.

**User intent → “I am looking for an answer to a specific problem”.**

**Tip: Keep in mind a mobile first indexing!**



# Google News and Top Stories

Google News and Top Stories is a fast and less stable source of traffic. The strategy is to respond quickly to trends and write news stories.

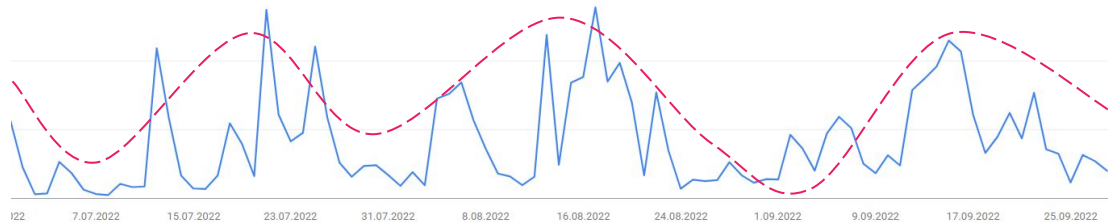
Mobile and desktop may differ slightly, but to a lesser extent than in Universal Search.

User intent → “I’m looking for the latest news”.

Tip: Publishing speed is essential.

The screenshot shows a Google search for "joe biden". The search bar is at the top, and the results are categorized under "Wszystko". The top result is "Joe Biden" with a sub-header "Prezydent Stanów Zjednoczonych". Below this, there are buttons for "W skrócie", "Wiadomości", and "Wykształcenie". The "Najważniejsze artykuły" section features two main articles: one from "wyborcza.pl" titled "Biden: Nigdy, nigdy, nigdy nie uznamy roszczeń Rosji do suwerennego terytorium Ukrainy" (4 godziny temu) and another from "TVP Info" titled "Biden mówi trzy razy „nigdy”. Chodzi o pseudoreferenda na Ukrainie" (15 godzin temu).

The screenshot shows a news aggregator page for "joe biden". The search bar at the top shows "joe biden" and the results are categorized under "Wszystko". The page displays "Okolo 298 000 000 wyników (0,70 s)". The "Najważniejsze wiadomości" section features several articles: "Biden: Nigdy, nigdy, nigdy nie uznamy roszczeń Rosji do suwerennego terytorium..." (4 godziny temu) from "wyborcza.pl", "Biden mówi trzy razy „nigdy”. Chodzi o pseudoreferenda na Ukrainie" (16 godzin temu) from "TVP Info", and "Biden pytany o roszczenia Rosji do terytorium Ukrainy: Pogwałcenie prawa. Nigdy..." (6 godzin temu) from "RMF 24". The "Inne wiadomości" section includes "Indie i Chiny przeciw Rosji? Biden ma plan" (21 godzin temu) from "TVP Info" and "Huragan Ian uderzył we Florydę. Joe Biden: Może kosztować życie wielu ofiar" (3 godziny temu) from "Polityka".



# Google Discover

Google Discover has been winning the hearts of publishers for the last few years. One article can earn a million visits and even more. This source is **the least controllable** and predictable, but yields the highest traffic volumes. It works **only on mobile**. Google is constantly developing and implementing more features.

Factors influencing Discover are the **speed of the website**, large, clear **photos (min. 1200px)** and specific topics. The algorithm is **based on CTRs**, the more clickable (an article) **title and thumbnail** is, the more strongly it is promoted by Google's algorithm.

User intent → “Entertain me”



Były kleryk: W seminarium wpaja się głównie posuszeństwo oraz przekonanie, że ksiądz to nadc...

Gazeta Wyborcza Wrocław... - 1 d



Andrzej Rysuje | NORD STREAM - Gazeta.pl

Wyborcza.pl - 19 g

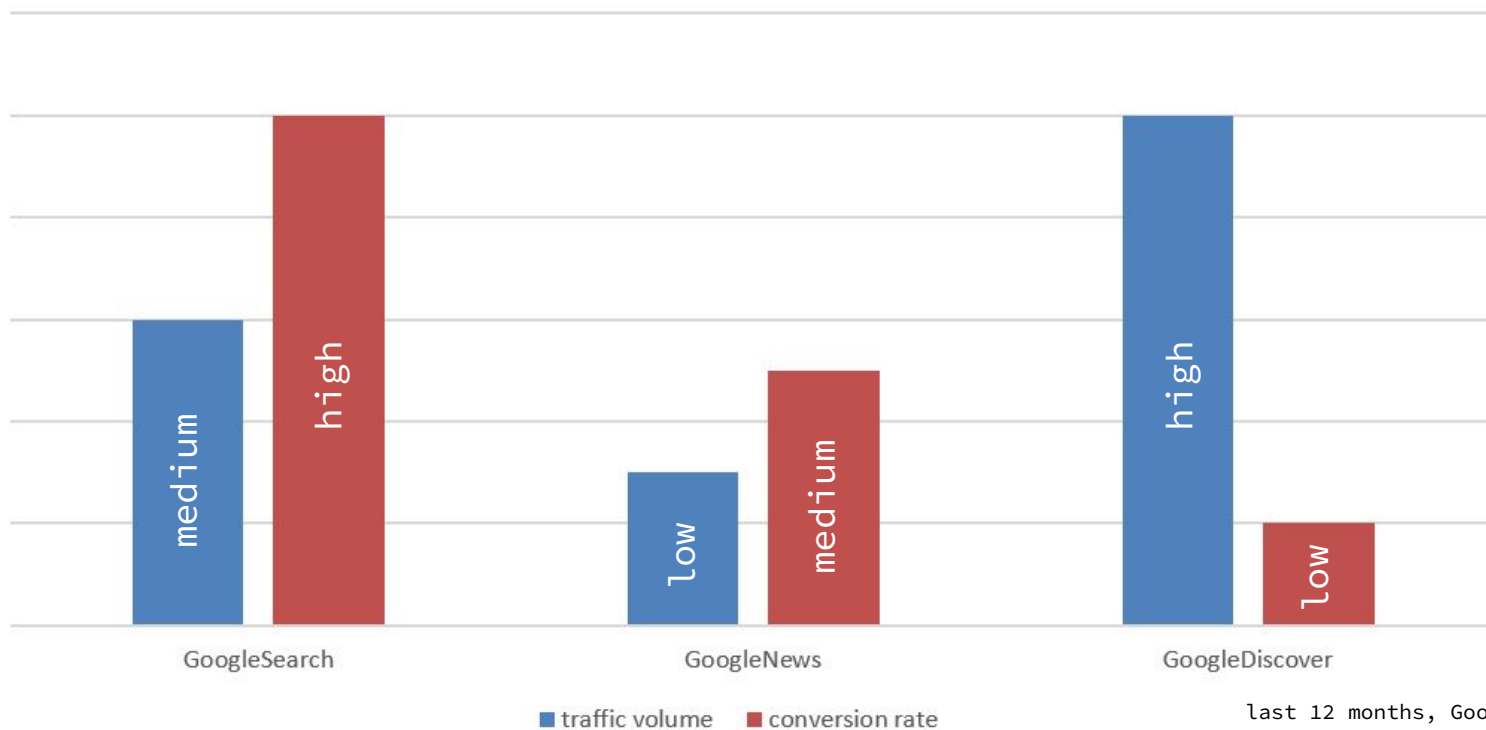


Przestali zatrudniać białych mężczyzn w imię różnorodności. Na początku szli w zaparte

HN PolishHR - 3 d

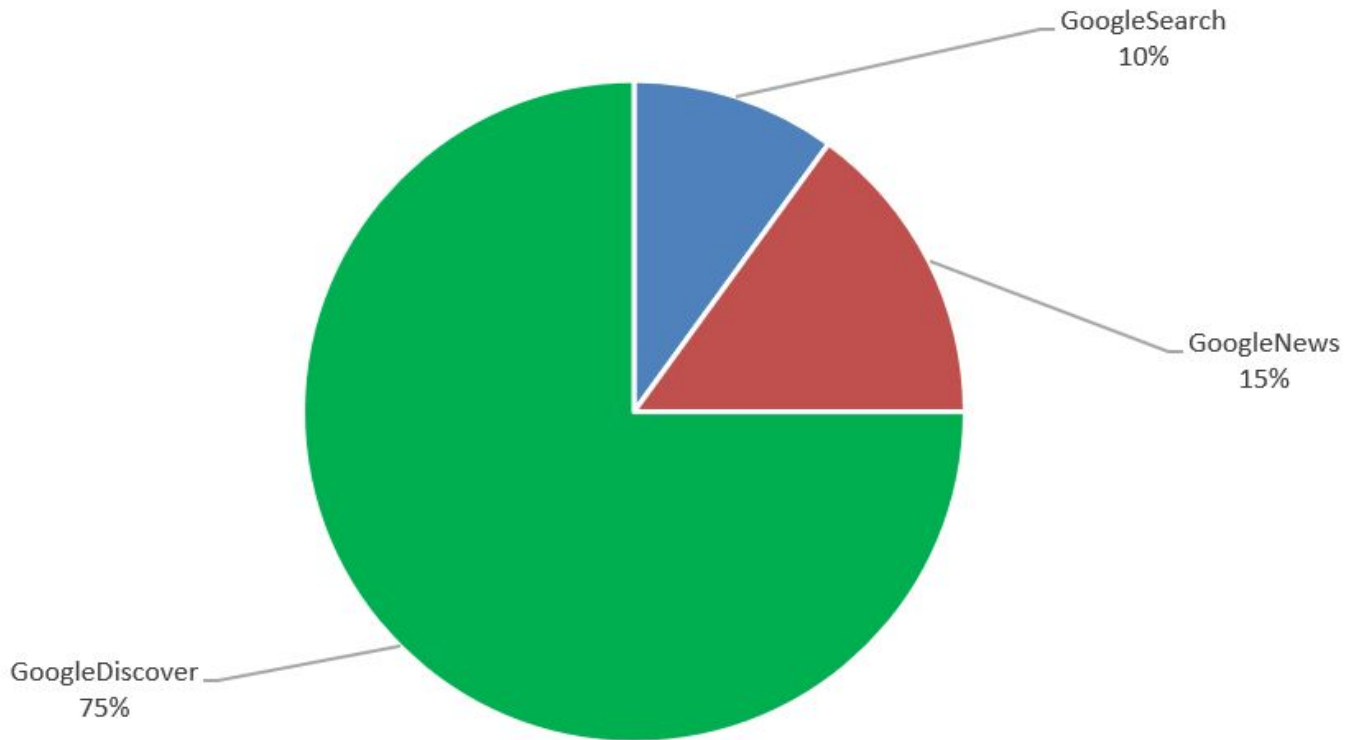
# Traffic volume vs. conv. rate (Google) at Wyborcza.pl

Traffic vs conv. rate

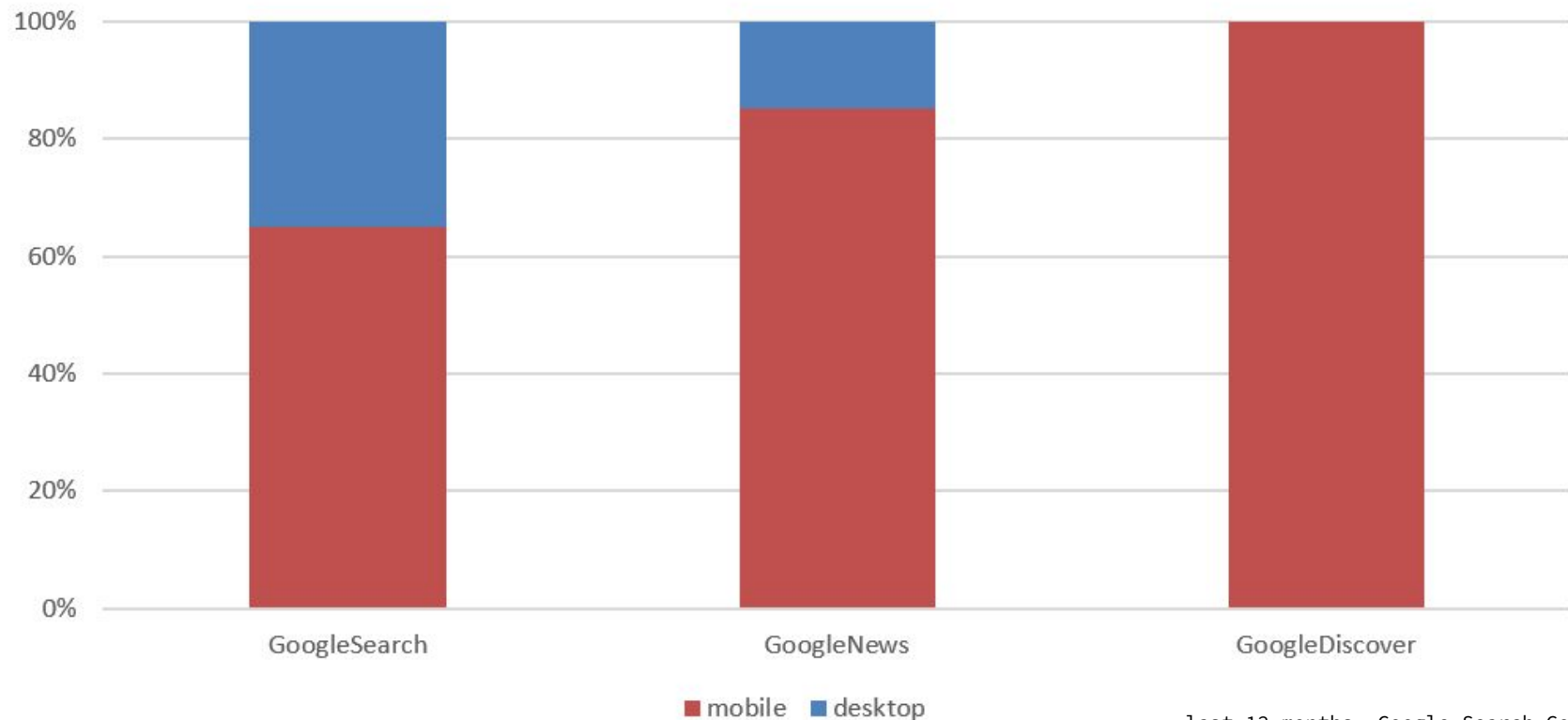


last 12 months, Google Search Console

# Google traffic sources average share (polish online publishers)



# Mobile vs Desktop (Google)



last 12 months, Google Search Console

# How we are working on SEO at Wyborcza.pl

— — —

We have a **dedicated SEO team** with **one technical SEO specialist** (all those revDNS, JS and IT stuff) and **two editorial SEO specialists**.

In our organization, the SEO team does not correct or optimize articles – we have approached this process differently than most publishers.

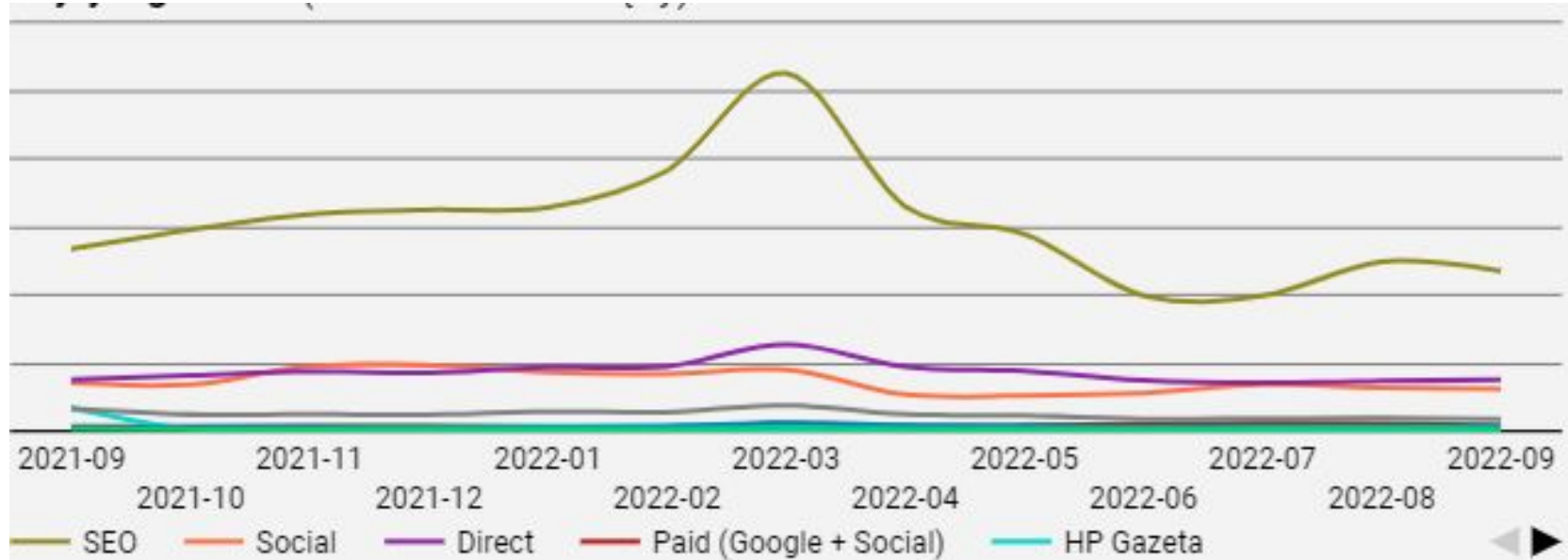
We placed a strong emphasis on **educating journalists and editors** (a few hundred people), but also managers. At least **once a year, EVERY member of the editorial team must** attend SEO training. During the year, we also work with them individually.

Every journalist, author, or editor do **basic content optimization** – taking care of the right title (including SEO title), headlines, photo descriptions, linking between texts, etc.

We educate our SEO team too, make reaserch&development and monitor all new features Google creates and how we can use them.



# Does it work at all?



# SEO success factors

- High-quality content (E-A-T)
- User intent is crucial
- Mobile-first approach
- Page experience (CoreWeb Vitals)
- Page speed (TTFB)
- SEO On-page optimization
- Internal links
- External links



# Dos and don'ts

---

- Ask your **SEO expert** to check a paywall system **before** you implement it (or check yours current solution) – you'll save **a lot of money**,
- Never publish newspaper content **without adaptation/optimization** for online needs (SEO),
- **Build processes**, not point action – improving 30 articles will not give you anything, but building a process to optimize each article will give measurable benefits,
- Always build a **few sources of traffic** – depending on the one is very risky (Google updates, Facebook updates).
- **Plan a strategy for SEO** at least for one year ahead, and do external SEO audits (regardless your in-house team or agency SEO).
- Remember that **you are writing for the humans**, not for Google's bot.

Thank you :)

contact: [ja@jakubsawa.pl](mailto:ja@jakubsawa.pl)