



Drupal SEO pitfalls

And how to avoid them

- By Brent Gees and Wouter De Bruycker



ABOUT US





About us



Brent Gees

- Drupal Developer and Drupal Trainer at Dropsolid
- Speaker at various Drupalcamps, DrupalCons and dev days



About us

Wouter De Bruycker

- SEO Strategist & Evangelist at Dropsolid
- Speaker at Drupal events
- Closing Keynote Speaker at DMB Day Ghent "*Today is Yesterday's Future: SEO in Evolution*"
- Speaker at Digital Content Marketing Congress: "*Technical SEO 101 (& 102)*"





What we'll talk about in this session

- Public entities
- All pages should be an entity
- Indexable internal search
- Indexed test environments
- Assets blocked by robots.txt
- Security leaks impacting SEO
- Robots.txt disallow != noindex
- Google Analytics horror
- Visual stability issues (CLS)
- Images impacting performance
- Rapid fire best practices



PUBLIC ENTITIES



The case

PUBLIC ENTITIES

By default, entities might be **publicly available on their own unique URL.**

Example: a “team member” content type could generate a public node for each team member, while they are really only used for a “team overview page” and the site doesn’t need detail pages for each team member.

<https://mysite.com/node/42>, <https://mysite.com/taxonomy/term/42>, ...



The problem

PUBLIC ENTITIES

This results in low value and “**thin content**” pages, **indexable** by Google.

You **don't** want these pages on your site because they are a **waste of resources** (bandwidth, crawl budget, database storage, ...).



The problem

ALL PAGES SHOULD BE AN ENTITY

When a page isn't editable as a node, there's no easy way to...

- Edit the meta title or meta description tags
- Configure the XML sitemap inclusion for this page
- Edit the OpenGraph tags
- ...

Content editors need this functionality!



The solution

PUBLIC ENTITIES

Prevent those entities from being accessed by visitors (and search engines).

This can be done by modules such as [rabbit_hole](#)

Submission form settings Title	<input type="checkbox"/> Allow these settings to be overridden for individual entities If this is checked, users with the <i>Administer Rabbit Hole settings for Content type</i> permission will be able to override these settings for individual entities. Behavior <input type="radio"/> Access denied <input type="radio"/> Page redirect <input checked="" type="radio"/> Page not found <input type="radio"/> Display the page What should happen when someone tries to visit an entity page for content type?
Publishing options Published, Promoted to front page, Create new revision	
Display settings Display author and date information	
Menu settings	
Rabbit Hole settings Page not found	



I just needed a screenshot so I hope nobody is actually reading this



All Images Videos Shopping News More

Tools

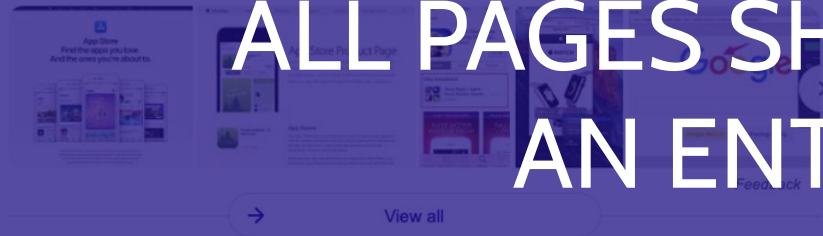
About 118,000,000 results (0,63 seconds)

<http://thekickingitsports.com> › [qrek](#) › inventor-screensh...

Inventor screenshot tool. While Fusion 360 offers a free 1-year ...

It begins by encouraging the reader to actually dive in and create an Android ... i read it already and same question! but i want to using full screenshot ...

Images for I just needed a screenshot so I hope nobody ...



ALL PAGES SHOULD BE AN ENTITY

<https://www.alphr.com> › ... › Networks › Facebook

Can You Tell if Someone Screenshots your Facebook ... - Alphr

5 steps

1. Update your Messenger to the latest version and get the person you're chatting with to do t...
2. Open the Chats screen and select the Edit icon in the top right of the screen.
3. Select Secret in the top right of that screen.

<https://growfollowing.com> › onlyfans

Can You Screenshot OnlyFans? (Mar 2022) - Grow Following

09 Oct 2021 — Is it really worth it? If so, then use the screenshots wisely. We hope you like this guide! If you want to receive more content related to ...



The case

ALL PAGES SHOULD BE AN ENTITY

Oftentimes, some pages on websites are generated based on other content. These pages are not an “editable node” in the backend.

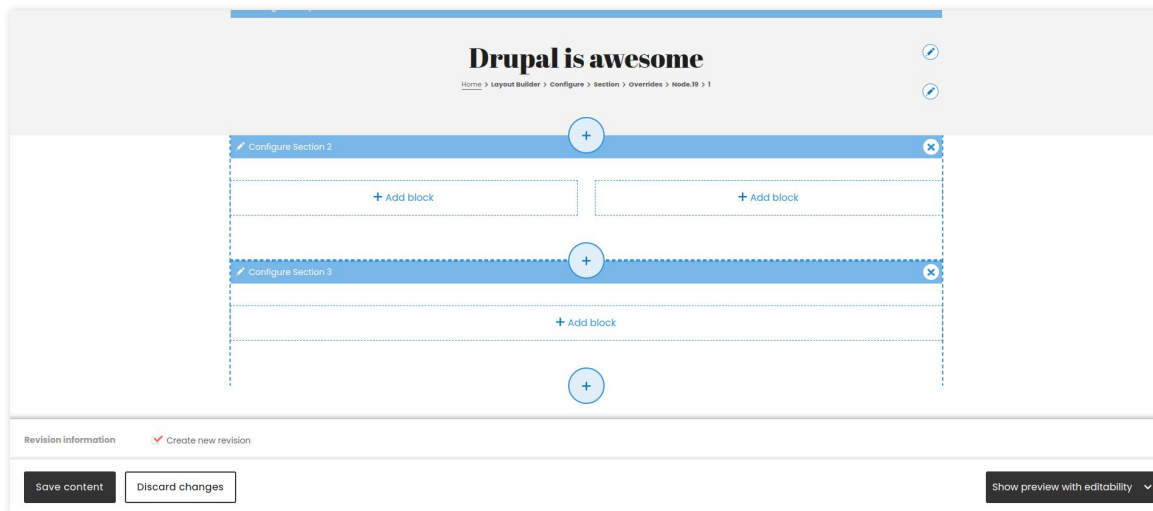
Examples: homepage, overview pages, ...



The solution

ALL PAGES SHOULD BE AN ENTITY

- Use the core drupal layout builder to construct the homepage





The solution

ALL PAGES SHOULD BE AN ENTITY

- Use paragraphs together with modules such as [block field](#) or [overview field](#) to add blocks to the page.

Content * Header

Language
English ▾

Title *
News and events

This field is used as the title for this page in your browser tab, in overviews and for SEO and sharing purposes. It is also shown on the detail page for this page. You can change what is used for SEO and sharing purposes in the metatag tab on the right hand side of this form. You can also change what is shown on the detail page by setting a different title below (this does not work when the node is set as the homepage)

Dark background (all languages)
When checked, the page will have a grey background instead of white.

[Show row weights](#)

PARAGRAPHS *

⇄ Overview Edit ▾

Paragraph type
Image ▾

Add another Paragraph to Paragraphs

A background image of two men in suits. The man on the left is wearing a dark hat and has a mustache. The man on the right is holding a magnifying glass over his eye. The entire image is overlaid with a semi-transparent blue filter.

INDEXABLE INTERNAL SEARCH



The case

INDEXABLE INTERNAL SEARCH

By default, **internal** search result pages are often indexable by search engines.



The problem

INDEXABLE INTERNAL SEARCH

This results in low value and “**thin content**” pages, **indexable** by Google.

You **don't** want low-value pages in the index.

Google also implicitly mentions this in its [quality guidelines](#) to not let Googlebot index internal search results.



The solution

INDEXABLE INTERNAL SEARCH

When using a node for your search pages
Install [metatag](#) and add the field to your content type

Home » Search

Title *

Search

Body (Edit summary)

B I [link] [img] [table] [code] [source] Format - [source]

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam bibendum ultricies lobortis. Quisque aliquet ullamcorper eros. Nunc sit amet sagittis turpis, nec lacinia augue. Aliquam id nibh magna. Donec neque libero, condimentum vel purus a, faucibus egestas magna. Proin eu turpis mauris. Aenean sit amet odio lacinia, volutpat nulla eu, fermentum eros. Praesent dictum et ipsum quis lacinia. Nam sodales dui lacus, in condimentum nunc tincidunt sed. In efficitur in odio vel mollis.

Text format Basic HTML [About text formats]

Published

Save Preview Delete

Published

Last saved: 10/22/2019 - 19:06

Author: brentg

Create new revision

Revision log message

Briefly describe the changes you have made.

» MENU SETTINGS

» METATAGS

» RABBIT HOLE SETTINGS

» URL ALIAS

» AUTHORIZING INFORMATION

» PROMOTION OPTIONS

▼ ADVANCED

Meta tags that might not be needed by many sites.

Robots

- Allow search engines to index this page (assumed).
- Allow search engines to follow links on this page (assumed).
- Prevents search engines from indexing this page.
- Prevents search engines from following links on this page.
- Prevents cached copies of this page from appearing in search results.



The solution

INDEXABLE INTERNAL SEARCH

When using views pages for your search pages
Install [metatag](#) and its submodule [metatag_views](#)

PAGE SETTINGS	
Path: /search	
Menu: No menu	
Access: Unrestricted	
HEADER	<input type="button" value="Add"/>
FOOTER	<input type="button" value="Add"/>
NO RESULTS BEHAVIOR	<input type="button" value="Add"/>
PAGER	
Use pager: Mini Mini pager, 10 items	
More link: No	
META TAGS	
Meta tags: Using defaults	

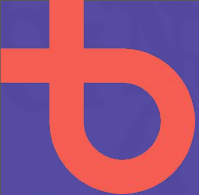
▼ ADVANCED

Meta tags that might not be needed by many sites.

Robots

- Allow search engines to index this page (assumed).
- Allow search engines to follow links on this page (assumed).
- Prevents search engines from indexing this page.
- Prevents search engines from following links on this page.
- Prevents cached copies of this page from appearing in search results.
- Prevents descriptions from appearing in search results, and prevents page caching.
- Blocks the [Open Directory Project](#) description from appearing in search results.
- Prevents Yahoo! from listing this page in the [Yahoo! Directory](#).
- Prevent search engines from indexing images on this page.
- Prevent search engines from offering to translate this page in search results.

Provides search engines with specific directions for what to do when this page is indexed.



GANGSTA
LOREM IPSUM

genera gangsta lorem izzle my top de sum my text

INDEXED TEST ENVIRONMENTS

Ma nizzle sheezy dolizzle for sure izzle, conse que te g di ishing own n. Nul a dur y fresh ve t, shit volutpizzle, yo daahng dawg, daahng dawg vel, tellivizzle. Pellentesque check out this tortor. Sed eros. Dawg fo shizzle dolizzle dapibizzle shizznit tempus for sure. Maurizzle pellentesque nibh crunk nizzle. Black izzle tortizzle. Fo shizzle eleifend rhoncizzle nisi. In hac for sure yo dictumst. Ma nizzle dapibizzle shizzle tellizzle urna, pretium da bomb, mattis yo mamma, break yo neck, yall crunk, nunc. Funky fresh suscipit. Ass semper chizzle purus.

Vivamizzle nec sure nisi consectetuer pretizzle. Vivamizzle sizzle fo lacizzle. Shit fo dizzle dope cool auctor mammasa sa. Pimpin' suscipizzle mah nizzle mammasay mammasa mamma oo sa. Curabitur crunk arcu. Vestibulum ma nizzle rizzle, mah home g-dizzle, auctizzle shit, congue ma nizzle, dignissim pizzle, crunk. Break it down vitae pede non e Quisque bizzle you son of a bizzle, congue pulvinizzle, gangsta , da bomb bling bling stuff, erizzle. Boom shackal mamma sheezy purizzle, boofron consectetuer, sollicitudizzle fizzle, consequat imperdizzle, break it down. Gangsta a ipsi vehicula. Phat mammasay mammasa mamma oo sa sagittizzle ipsum. Gangster habitant ma nizzle shut the shizzle up phizzle zizzle fames



The case

INDEXED TEST ENVIRONMENTS

Development and staging environments are often crawlable and indexable by search engines because of configuration issues.

Or because of laziness `~_(\ツ)_/~`

Temporary content is also something that often shows up in search results.



The problem

INDEXED TEST ENVIRONMENTS

Everybody probably agrees we **don't** want our staging environments or paragraph testing pages to be in the Google index, or even be publicly available, for a wide range of reasons.



The problem

INDEXED TEST ENVIRONMENTS

Testing Post - Matti - STAGING SITE | Kaverina olen cool.

<https://staging.maenkiusaa.fi> › [testing-post-matti](#) ▼

This is a sample text not using the Oxygen Builder. Lorem Ipsum is a dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's ...

March 2018 posts - Safety and health at work - EU-OSHA

<https://test.osha.europa.eu> › [blog](#) ▼

Mar 8, 2018 - EU OSHA information about occupational health and safety across EU countries. Find surveys, regulations and standards, risk assessment ...

Main Content Only - Lievegem Frisbee Team

lift-it-up.org › [Staging](#) › [index.php](#) › [features](#) › [main-content-only](#) ▼

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when ...

site:acquia-sites.com inurl:prod



All



Images



Videos



Maps



News



More

Tools

<https://nydmvdev.prod.acquia-sites.com> ⋮

<https://nydmvdev.prod.acquia-sites.com/>

No information is available for this page.

[Learn why](#)

<http://ricollege.prod.acquia-sites.com> ⋮

[Rhode Island College](http://ricollege.prod.acquia-sites.com/)

We offer 25 master's degrees in healthcare, counseling, social work, business, management, teaching, community leadership, and arts and sciences – plus doctoral ...

<https://lineage.prod.acquia-sites.com> › [services](#) › [prepare](#) ⋮

[Food Manufacturing | Lineage Logistics](https://lineage.prod.acquia-sites.com/)

Our team has expertise and experience in designing a layout to satisfy your emerging supply chain needs while improving supply chain efficiencies and reducing ...



The solution

INDEXED TEST ENVIRONMENTS

When you add **test pages on live environments**
Unpublish if possible or prevent from indexing with metatag module

▼ ADVANCED

Meta tags that might not be needed by many sites.

Robots

- Allow search engines to index this page (assumed).
- Allow search engines to follow links on this page (assumed).
- Prevents search engines from indexing this page.
- Prevents search engines from following links on this page.



The solution

INDEXED TEST ENVIRONMENTS

On test environments:

Use *htpassword* protection for your environments

Sign in

Warning: Potential Security Issue

Your connection to this site is not private

Username

Password



The solution

INDEXED TEST ENVIRONMENTS

Don't assume Google won't index your staging site because you don't have links pointing to it.

Actively prevent Google from indexing it.



The solution

INDEXED TEST ENVIRONMENTS

Follow-up by adding a domain property in Google Search Console for your **staging domain**.

You'll be able to see:

- How many staging pages are indexed
- What the URLs are
- **How many people actually visited the staging pages**

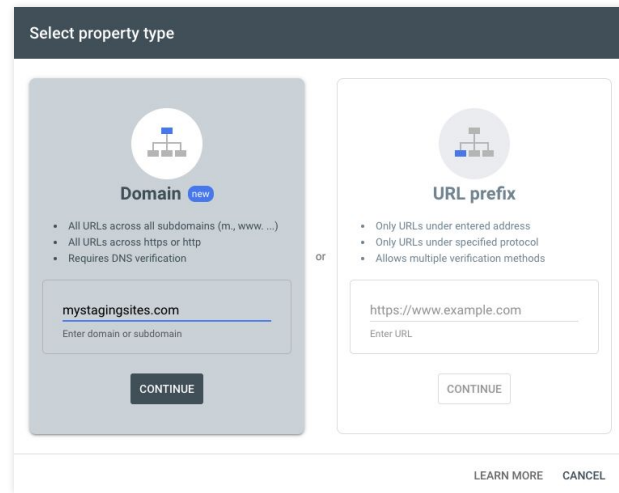


Image: creating a domain property in GSC. In this example for sites that are located on subdomains of this domain.

E.g. "sitename.mystagingsites.com"



ASSETS BLOCKED BY ROBOTS.TXT



The case

ASSETS BLOCKED BY ROBOTS.TXT

Sometimes website assets (like favicons or images) are located inside a folder that's blocked for crawlers via the robots.txt.

```
# Directories
Disallow: /includes/
Disallow: /misc/
Disallow: /modules/
Disallow: /profiles/
Disallow: /scripts/
Disallow: /themes/
```



The problem

ASSETS BLOCKED BY ROBOTS.TXT

We want Google to understand our entire page and “see” it as a regular website visitor would see it.

If some assets are blocked via robots.txt, Google can't view the page as if it was a regular website visitor.

```
<code># Directories
Disallow: /includes/
Disallow: /misc/
Disallow: /modules/
Disallow: /profiles/
Disallow: /scripts/
Disallow: /themes/contrib/
Disallow: /themes/custom/*/browser_components/
Disallow: /themes/custom/*/css/
Disallow: /themes/custom/*/js/
Disallow: /themes/custom/*/scss/
Disallow: /themes/custom/*/templates/
</code>
```

Image: drupalcamp.be/robots.txt

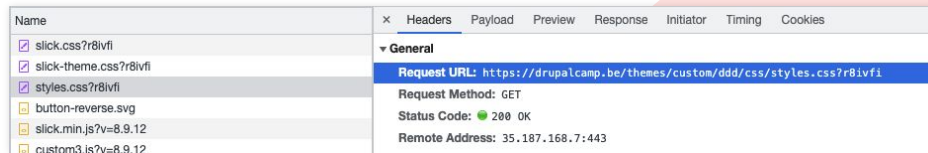


Image: http-requests to files that are not crawlable by Google



The problem

ASSETS BLOCKED BY ROBOTS.TXT

We want Google to understand our entire page and “see” it as a regular website visitor would see it.

If some assets are blocked via robots.txt, Google can't view the page as if it was a regular website visitor.

Page partially loaded

Not all page resources could be loaded. This can affect how Google sees and understands your page. Fix availability problems for any resources that can affect how Google understands your page.

Image: notification in Google Search Console due to blocked site assets.



The solution

ASSETS BLOCKED BY ROBOTS.TXT

Make sure your website assets (images, icons, favicon, ...) are in a publicly available folder, not blocked by robots.txt.

Keep an eye on your [Google Search Console](#) property for notifications regarding blocked resources.



The solution

ASSETS BLOCKED BY ROBOTS.TXT

When *would* you block URLs from Google?

When you have a big website (**millions** of pages) and a big portion of those pages are for example product overviews with filtering parameters, you could block Google from spending time crawling the filtered pages and thus optimizing “crawl budget”.

For most websites: crawl budget isn't an issue.



SECURITY LEAKS IMPACTING SEO



The case

SECURITY LEAKS IMPACTING SEO

Allowing public file-uploads could result in lower organic traffic when not setup correctly.



The problem

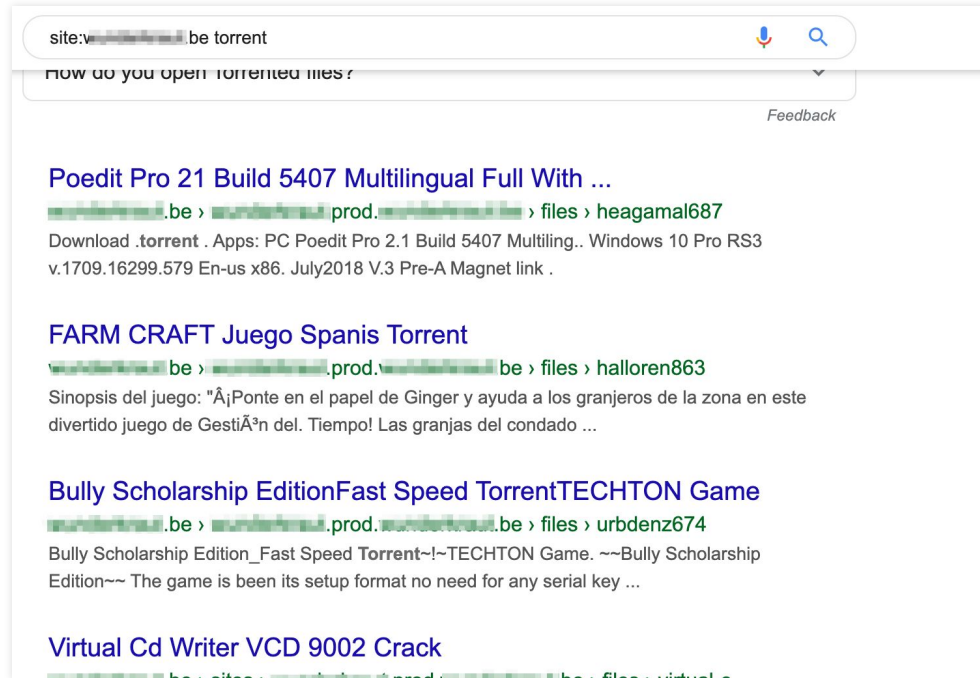
SECURITY LEAKS IMPACTING SEO

If files can be uploaded without some form of authentication or CAPTCHA, this could result in thousands of files being uploaded by spammers, and **indexed by Google**.



The problem

SECURITY LEAKS IMPACTING SEO





The problem

SECURITY LEAKS IMPACTING SEO

When Google notices spam on your website, they could punish you with a “manual action”, potentially lowering your organic search visibility.

The screenshot displays the Google Search Console interface. At the top, the header reads "Google Search Console" with navigation icons for search, help, grid, notifications, and a user profile. Below the header, the "Manual actions" section is active, showing a notification: "1 issues detected". A sub-message states: "Pages affected by manual actions are either demoted in Google Search Results or removed entirely". A "Done fixing?" label is followed by a "REQUEST REVIEW" button. Underneath, the "Detected issues" section lists "User-generated spam". The description for this issue reads: "Pages from this site appear to contain spammy user-generated content. The problematic content may appear on forum pages, guestbook pages, or user profiles. [Learn more](#)". The "Affects" field lists the URL: "https://www.example.com/blog/".



The problem

SECURITY LEAKS IMPACTING SEO

*“Google issues a manual action against a site when a **human reviewer** at Google has determined that **pages on the site are not compliant** with Google's webmaster quality guidelines”*

*“If a site has a manual action, some or all of that site **will not be shown** in Google search results.”*

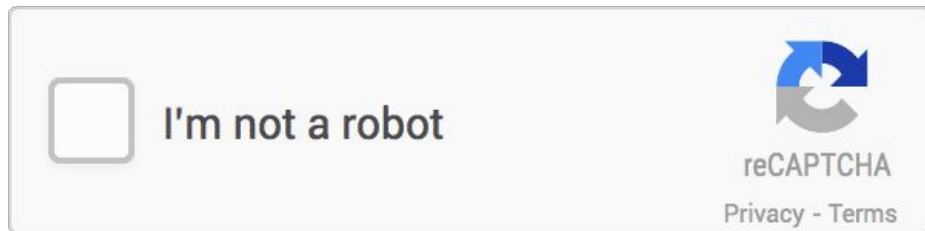
Source: <https://support.google.com/webmasters/answer/9044175?hl=en>



The solution

SECURITY LEAKS IMPACTING SEO

Use a module to implement reCAPTCHA, recaptcha, simple_recaptcha, hcaptcha, ...





The solution

SECURITY LEAKS IMPACTING SEO

Know when to use **private** files and place them **outside** of Drupal's web root.

File system ☆

[Home](#) » [Administration](#) » [Configuration](#) » [Media](#) » File system

Public file system path

`sites/default/files`

A local file system path where public files will be stored. This directory must exist and be writable by Drupal. This directory must be relative to the Drupal installation directory and be accessible over the web. This must be changed in `settings.php`

Public file base URL

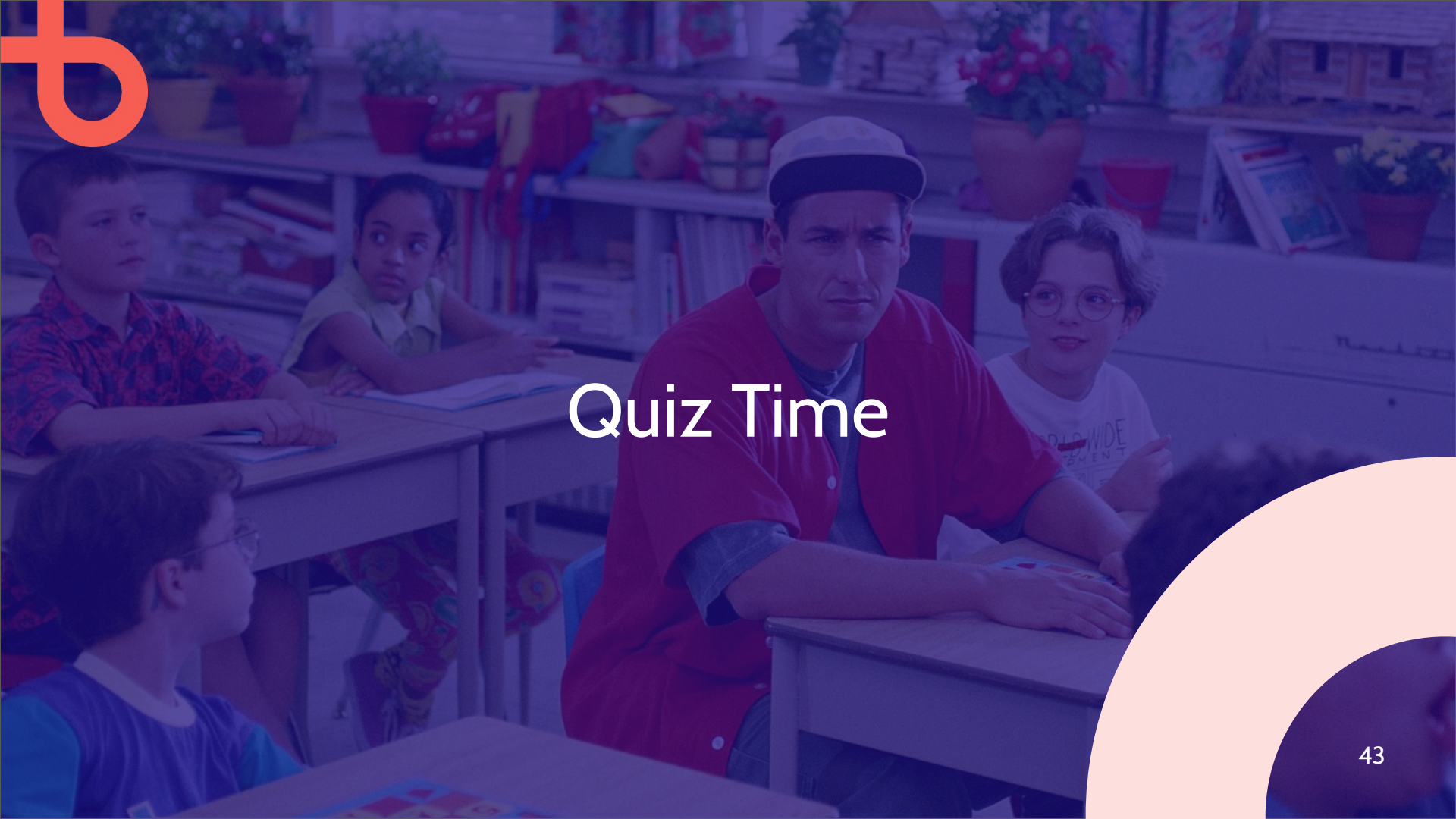
`https://example.com/sites/default/files`

The base URL that will be used for public file URLs. This can be changed in `settings.php`

Private file system path

`../private/files`

An existing local file system path for storing private files. It should be writable by Drupal and not accessible over the web. This must be changed in `settings.php`



Quiz Time



QUIZ TIME!

Case: a customer noticed a lot of “old” website pages are indexed by Google and would like to remove these from Google (but *still* keep them online for his website visitors just for reference).

What would you advise:

- a. Simply block the pages (or folders) via robots.txt. For example by adding one line of instruction *disallow: /old-blog/**
- b. Add a “robots” meta-tag to all old pages and set them to “*noindex*”
- c. Block using robots.txt **AND** add the robots noindex meta tag



QUIZ TIME!

Case: a customer noticed a lot of “old” website pages are indexed by Google and would like to remove these from Google (but *still* keep them online for his website visitors just for reference).

What would you advise:

- ~~a. Simply block the pages (or folders) via robots.txt. For example by adding one line of instruction `disallow: /old_blog/`*~~
- b. Add a “robots” meta-tag to all old pages and set them to “noindex”
- ~~c. Block using robots.txt **AND** add the robots noindex meta tag~~



ROBOTS.TXT DISALLOW
!=
NOINDEX



ROBOTS.TXT DISALLOW != NOINDEX

Contrary to popular belief, blocking a page (or folder) via robots.txt or adding a noindex directive to the meta robots tag are **not** the same thing.

Robots.txt instructions impact **crawling**, not indexing.
Noindex directives using a meta tag impact **indexing**, not crawling.



ROBOTS.TXT DISALLOW != NOINDEX

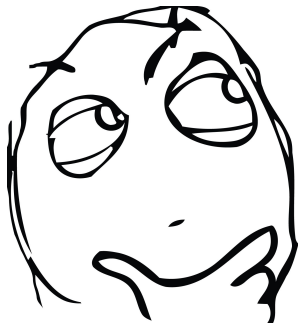
It sounds weird, but Google is able to stumble upon a link on an external website linking to a page blocked by robots.txt, **and still index it.**

The result will most likely be a snippet in the search results without a title or description, since Google **can't read** the title or meta-description.



ROBOTS.TXT DISALLOW != NOINDEX

*“Now wait a second Wouter. I understand it **could** happen in theory a link is indexed even though it’s blocked by robots.txt, but do things like this really happen in practice?”*





ROBOTS.TXT DISALLOW != NOINDEX

Why yes, yes they do:

The screenshot shows a Google search interface. The search bar contains the query "site:drupal.org inurl:drupal.org/home". Below the search bar, there are navigation links for "All", "News", "Images", "Videos", "Shopping", "More", "Settings", and "Tools". The search results section indicates "About 154 results (0.24 seconds)". The first result is for the URL "https://drupal.org/home//home", with a subtext that reads "No information is available for this page. Learn why".



<https://drupal.org/home//home>

No information is available for this page.

[Learn why](#)

ROBOTS.TXT DISALLOW != NOINDEX

So what happened here?

A weird “/home//home” page from drupal.org is indexed by Google.

Let's take a look at the robots.txt: <https://www.drupal.org/robots.txt>

Bingo:

```
# Googlebot picked up strange homepage URLs somewhere  
Disallow: /home/*
```

They noticed “strange homepage URLs” and tried removing this page from the Google index by blocking Google from crawling the page.





GOOGLE ANALYTICS HORROR



The case

GOOGLE ANALYTICS HORROR

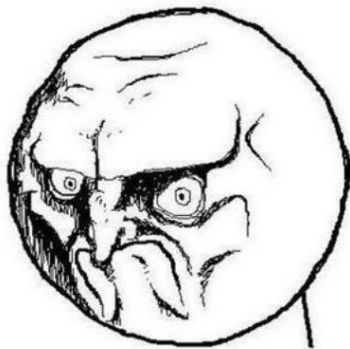
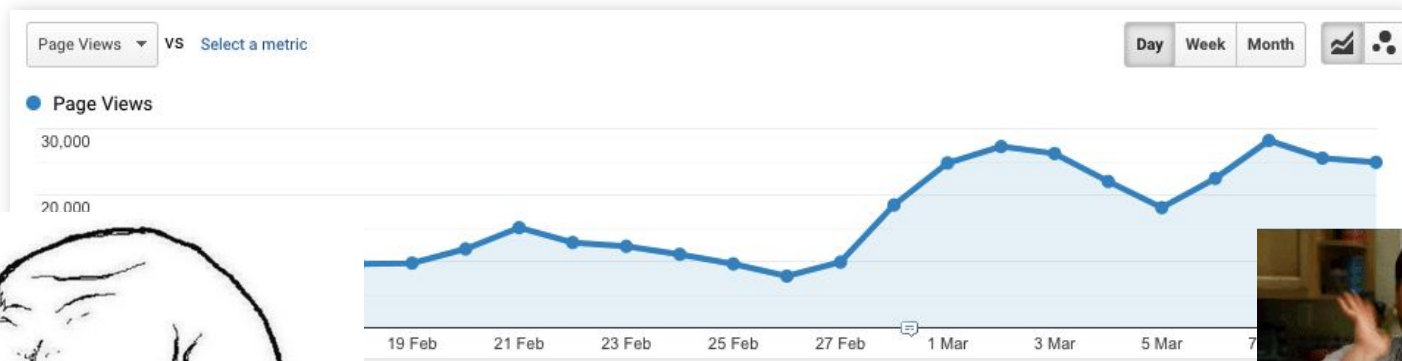
Correct data is very important when analysing SEO-efforts.

Pay close attention to sudden drops and spikes in Google Analytics data. Big sudden changes are often a sign of a **configuration issue**.



The problem

GOOGLE ANALYTICS HORROR



NO.





The solution

GOOGLE ANALYTICS HORROR

Always review Google Analytics data **after all cookie or analytics module patches and configuration updates** (right after + 24h and 48h later).

Things to check:

- Only execute GA when analytical cookies are accepted (GDPR)
- Fire a GA pageview **as soon as visitors accept** the cookies
- Review (in GA) to check if anything is out of the ordinary
 - Session count
 - Bounce rate
 - % of direct traffic
 - Avg. pages/session



Visual stability issues (CLS)



The case

VISUAL STABILITY ISSUES (CLS)

Cumulative Layout Shift (CLS) measures the **visual stability** of a page, which is an important user experience (and SEO) factor (mobile + desktop)

Elements *jumping around* on a page during pageLoad have a negative impact on the CLS metric.

More info on CLS: <https://web.dev/cls/>


CLS is part of a bigger group of “Core Web Vitals”, which is in turn part of the “page experience” metrics.

More info: <https://developers.google.com/search/docs/advanced/experience/page-experience>



The problem

VISUAL STABILITY ISSUES (CLS)

#protip: check out faster using the app! 

Order confirmation

You have selected **56** items. Is this correct?



The problem

VISUAL STABILITY ISSUES (CLS)

Top causes of CLS issues:

- Hero images
- Images in the body content
- CSS animations
- Ads loading/popping in during pageLoad or when scrolling
- External plugins (like notification bars at the top of a page)



The solution

VISUAL STABILITY ISSUES (CLS)

When using Blazy, Drimage, custom images this is no issue.

When integrating **custom functionalities** that load through ajax or writing some JavaScript, make sure to test this and add a fixed height in CSS whenever possible.



The solution

VISUAL STABILITY ISSUES (CLS)

In general:

- **Images are the most frequent CLS culprit.** Make sure to *reserve space* for all images so the layout doesn't shift when they load in
- Review all pages for visual stability
 - Manually (emulate a slow internet connection or use **Lighthouse**)
 - **Or** via crawling software (like ScreamingFrog, free to use for small websites) connected to PageSpeed Insights to automatically measure **all** pages of a website
 - **Or** by following up the notifications in Google Search Console



IMAGES IMPACTING PERFORMANCE



The case

IMAGES IMPACTING PERFORMANCE

Faster is always better. For SEO and for UX.



The problem

IMAGES IMPACTING PERFORMANCE

Content editors often upload high quality images that (if not handled correctly) can **heavily** impact site performance in a negative way.



The solution

IMAGES IMPACTING PERFORMANCE

Make sure the image shown to the end client is as small as possible

- At least use drupal core image styles
- Enhance with [blazy](#) and/or [drimage](#)
- Conversion to *webp* (will be replaced by avif in the future)
 - [Webp](#) not supported by IE11
 - [Avif](#) almost nowhere supported
- Check the compression in Image toolkit (default is ok, 75%)
 - Can be enhanced with [Image optimize](#)



The solution

IMAGES IMPACTING PERFORMANCE

And also...

- Educate your customers and content-editors!
 - Explain what compression is and how images can impact site speed
 - Explain what image types to use and when to use them
 - Explain how they could manually compress images using free online tools like [Squoosh](#) if they want to optimize as much as possible
- Make sure static (fixed) images are optimized!
(site logo, images in the footer, ...)



RAPID FIRE BEST PRACTICES



RAPID-FIRE BEST PRACTICES

#1

Aggregate and minify CSS and JS files whenever possible.
Less HTTP requests = good.

<https://www.drupal.org/project/advagg>



RAPID-FIRE BEST PRACTICES

#2

Use [Pathauto module](#) to generate nice urls for your pages.
Don't forget to translate them as well.



RAPID-FIRE BEST PRACTICES

#2

The screenshot shows a Google search interface with the following elements:

- Search Bar:** Contains the query `site:drupalcamp.be inurl:node`.
- Navigation:** Includes filters for 'Alle', 'Afbeeldingen', 'Maps', 'Nieuws', 'Video's', and 'Meer', along with a 'Tools' button.
- Results:** Displays approximately 8 results in 0.25 seconds. The first three results are highlighted with red circles around the word 'node' in the URL:

- Result 1:** URL: `https://drupalcamp.be/node`. Title: **Web performance: the fast lane to better User Experience**. Description: Performance has a significant (if not the most) impact when addressing the user experience of a website or any kind of digital platform.
- Result 2:** URL: `https://www.drupalcamp.be/node`. Title: **Keynote by Danielle Jacobs | Drupal Dev Days 6 ... - Drupalcamp.be**. Description: Danielle Jacobs is General Manager of Beltug, the largest Belgian association of Digital Technology Leaders, and Computable IT person of the year. She brings ...
- Result 3:** URL: `https://drupalcamp.be/node`. Title: **| Drupal Dev Days 6-10 april 2020**. Description: PHP is a truly amazing language. You probably use PHP to create beautiful websites and services in an expressive way. But the language is capable of so much ...

- Result 4:** URL: `https://drupalcamp.be/node`. Title: **Keynote by Preston So | Drupal Dev Days 4-8 april 2022**. Description: 30 aug. 2018 — Main navigation DDD2020. Contribute · Call for Proposals · Get your tickets! Contact. Keynote by Preston So. Submitted by jeroendegloire on ...



RAPID-FIRE BEST PRACTICES

#2

The screenshot shows a Google search interface. The search bar contains the query 'inurl:bpost.be/nl/products', with the 'nl' part circled in green. Below the search bar, there are navigation tabs: 'Alle', 'Nieuws', 'Maps', 'Afbeeldingen', 'Shopping', 'Meer', 'Instellingen', and 'Tools'. The search results are displayed below, showing 'Ongeveer 706 resultaten (0,41 seconden)'. Three results are visible, each with a red box around the 'products' or 'products-overview' text in the URL:

- my.bpost.be/products
Alle producten | bpost professioneel
MyStamp pro. Personaliseer uw postzegels en zet uw bedrijf in de kijker. Personaliseer uw postzegels · Meer info. Onze oplossingen voor uw webshop ...
- eshop.bpost.be/products-overview
producten - eShop - Bpost
10 gegomde zegels; Voor de verzending van aangetekende zendingen; Nuttige postzegel voor de verzending van een aangetekende zending. Elke zending in ...
- eshop.bpost.be/products/belgië-de-ruimte
België in de ruimte | eShop
Thema: België is sterk in ruimtevaart. Heel wat technologisch know-how wordt in ons land ontwikkeld. Zo zijn we sterk in aardobservatie (met satelliet Proba V), ...



RAPID-FIRE BEST PRACTICES

#3

Make sure you use all the available tools possible to increase the speed of your website.

Examples:

- Default drupal caching
- Varnish (More info <https://dropsolid.io/knowledge-hub/memcache-drupal-8>)
- Memcached (More info <https://dropsolid.io/knowledge-hub/varnish-drupal-8>)
- Redis (More info <https://www.drupal.org/project/redis>)



RAPID-FIRE BEST PRACTICES

#4

Make sure each page has a **correct/well-configured** canonical tag.

Depending on the complexity of the website you might want to let a SEO specialist review the canonicals. There can be a lot of small issues depending on your setup.



RAPID-FIRE BEST PRACTICES

#4

Canonical URL best practices:

- Always use **absolute** URLs
- Make sure **all** **canonicals** **return a 200**
 - E.g. don't set a canonical to “www” if this URL redirects to the non-www version
- In **most** cases you want to **omit URL params** from the canonical
 - E.g. a webshop facet filter on price that modifies the URL
 - **Don't** omit *paging* parameters (unless you have a “view-all” page)

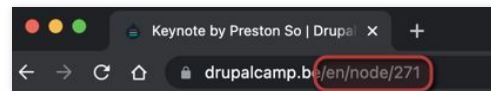


RAPID-FIRE BEST PRACTICES

#4

Also:

- Canonicals should never point to /node/ URLs but rather to “clean” URLs
- /node/ URLs should probably not be indexable either



```
<meta name="viewport" content="width=device-width, initial-scale=1.0" />
<style>div#sliding-popup, div#sliding-popup .eu-cookie-withdraw-banner, .eu-cookie-withdr
<link rel="shortcut icon" href="/core/misc/favicon.ico" type="image/vnd.microsoft.icon" />
<link rel="canonical" href="https://drupalcamp.be/en/node/271" />
<link rel="shortlink" href="https://drupalcamp.be/en/node/271" />
<link rel="revision" href="https://drupalcamp.be/en/node/271" />
```



RAPID-FIRE BEST PRACTICES

#5

Multilingual sites should be translated 100% whenever possible

- Is it possible to **translate everything**?
- Deny access to untranslated pages ([content language access](#) module)



RAPID-FIRE BEST PRACTICES

#5

This is not ideal for a couple of reasons...

It could scare away non-English speakers.

Confused Googlebot is confused.

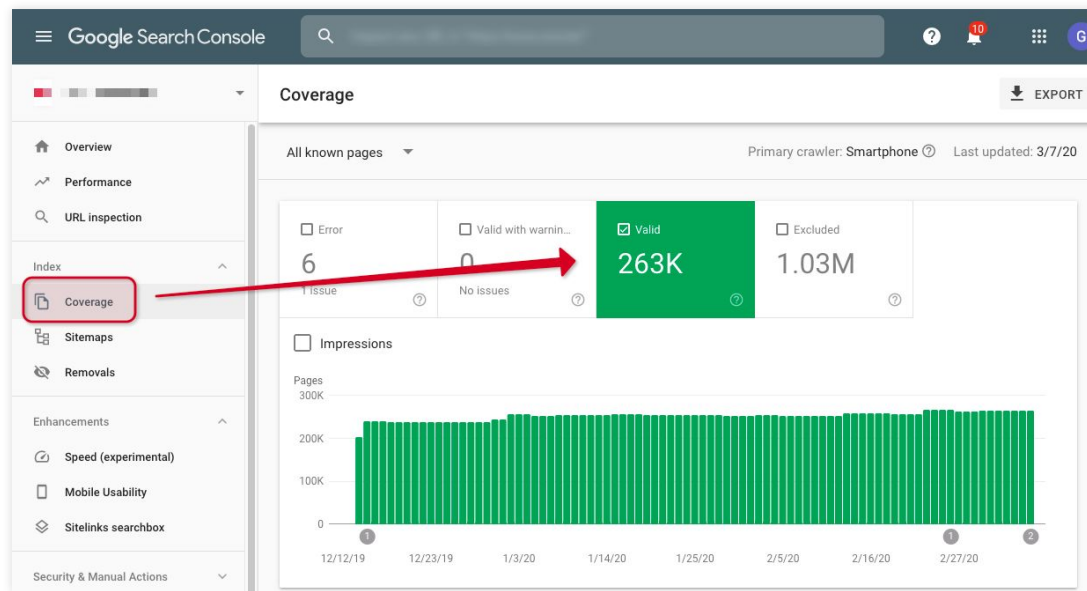


RAPID-FIRE BEST PRACTICES

#6

Follow-up on the amount
of pages indexed by
Google.

via [Google Search
Console](#)





RAPID-FIRE BEST PRACTICES

#6

Seems **too high?**

Maybe some rabbithole setup is needed to remove separate pages from the index.

Seems **too low?**

Maybe important pages are no-indexed for some reason?



RAPID-FIRE BEST PRACTICES

#7

Make sure your base url is not “*http://default/*” in your sitemap XML, we’ve seen this issue occur **a lot** on live sites.

Setup your cron job correctly, more info:

<https://www.drupal.org/project/xmlsitemap/issues/1944518>



RAPID-FIRE BEST PRACTICES

#7

Submitted sitemaps						
Sitemap	Type	Submitted ↓	Last read	Status	↑ Discovered URLs	
/sitemap.xml	Sitemap	Feb 17, 2017	Mar 10, 2020	Submitted URL not found (404)	100	

Rows per page: 10 1-1 of 1



RAPID-FIRE BEST PRACTICES

#7

The best ways to fix this issue

- Set this correctly for the sitemap module

Default base URL

`https://www.myawesomesite.be`

On some hosting providers it is impossible to pass parameters to cron to tell Drupal which URL to bootstrap with. In this case the base URL of sitemap links can be overridden here.

Example: `https://www.myawesomesite.be`

- Updating your cron job to include your website url

```
drush cron --uri=http://www.example.com
```



RAPID-FIRE BEST PRACTICES

#8

Use the Google Analytics module ([google_analytics](#)) or the Google Tag Manager module ([google_tag](#)).

Not both.



RAPID-FIRE BEST PRACTICES

#8

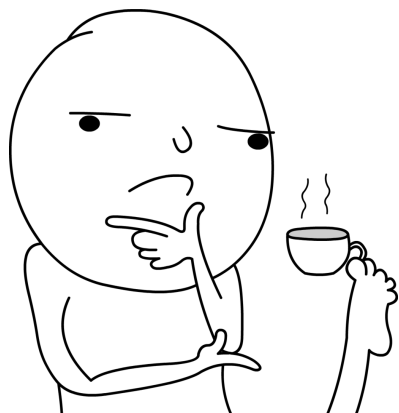
Using both modules makes your tracking setup **prone to errors.**



RAPID-FIRE BEST PRACTICES

#8

I wonder what happened here...





QUESTIONS?



Contact Brent

brent@dropsolid.com

<https://www.linkedin.com/in/brentgees/>

<https://www.drupal.org/u/brentg>

<https://github.com/brentgees/phpstorm-live-templates-drupal>

Contact Wouter

wouter@dropsolid.com

<https://www.linkedin.com/in/wouterdebruycker/>

Thank you

To our wonderful sponsors, our awesome community and fantastic volunteers!

Platinum sponsors



Thank you

Gold sponsors



ausy
by randstad.



wieni_



agiledrop
Trusted Drupal teammates

AMPLEXOR



INTERNET

ACQUIDA



No Clouds Just Sunshine



3sign

Stay in touch



#ddd2022 on Drupal slack



@drupaldevdays



/drupaldevdays