Tech talk

Wat zelfs experts in 2023 blijven fout doen

Als ik zie dat jullie niet meer gefocust zijn,
zal ik dit doen: klap, klap.

PLAY
Interessant.

Al die kennis...

SEO VOOR E-COMMERCE
Jan-Willem Bobbink

E-COMMERCE SEO FAILS
Steven Van Vessum

SITE MIGRATIES
Dirk Ceupens

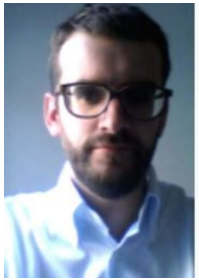
CONTENT INVENTARISATIE
Dave Lorrez

GOOGLE UPDATES AND THE FUTURE
Sander Tamaëla

SEO AT SCALE
Felix Peters

SEO CONSULTANCY & PROJECT MANAGMENT
Marloes Zweers

LOCAL SEO @ VDAB
Liesbet Devylder



LOCAL SEO @ RADISSON
Ken Derwael

SEO IN COMPLEXE B2B
Steven Van Duyse

CONTENT ALS SEO TROEF
Jente Kasprowski

Al die kennis...

. . .

Wat zou nog interessant kunnen zijn?

Tech talk

Wat zelfs experts in 2023 blijven fout doen

Disclaimer:

Ik maak ook fouten hé.

Wie?

Dev

SEOs

(ook ik)

PO/business

1

2

3

4

#1 Ze gebruiken robots.txt verkeerd

Drie fouten gezien.

De case van TikTok

Toen begon hij een heel lange uitleg.

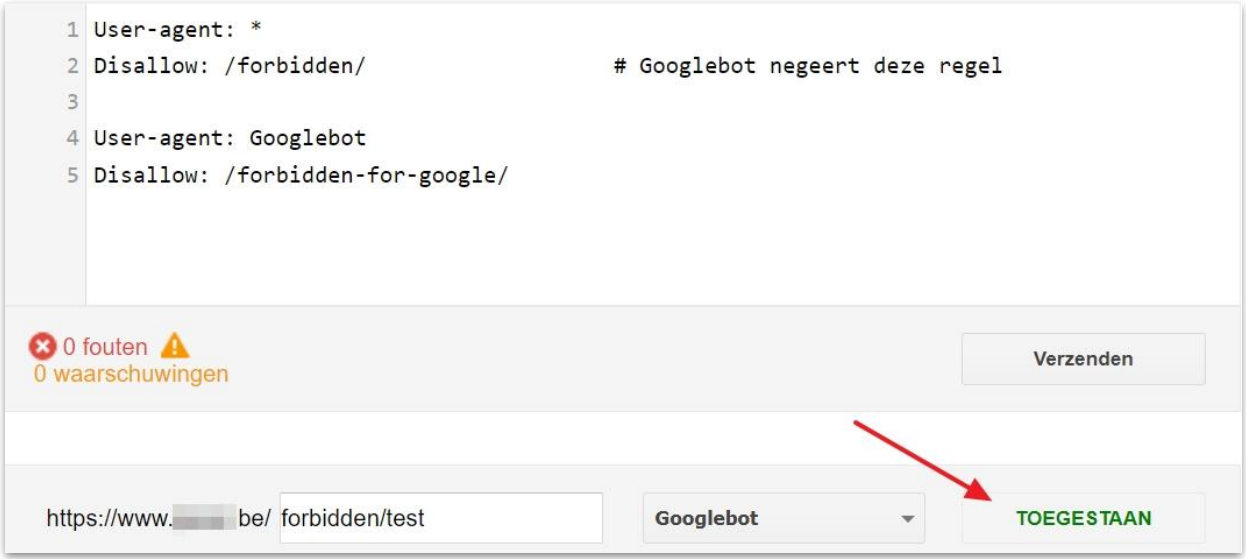
In feite komt het hier op neer

somesite.com/robots.txt

```
# Lezen Googlebot & Bingbot deze regel?
User-agent: *
Disallow: /forbidden/

# Dit bekijkt Googlebot
User-agent: Googlebot
Disallow: /secret/

# Dit bekijkt Bingbot
User-agent: Bingbot
Disallow: /confidential/
```

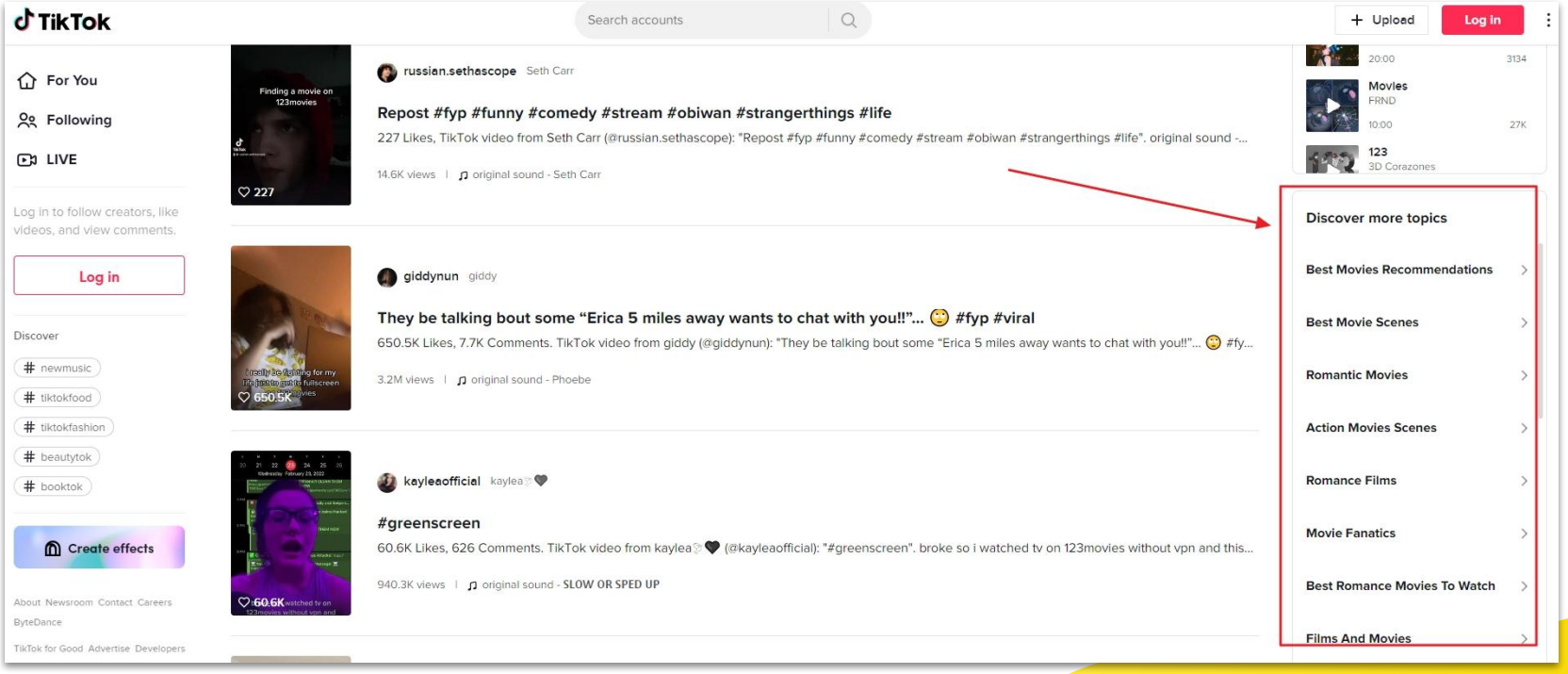
Regels worden niet gebundeld (geen overerving). Enkel de meest specifieke UA telt.

Als je me niet gelooft:



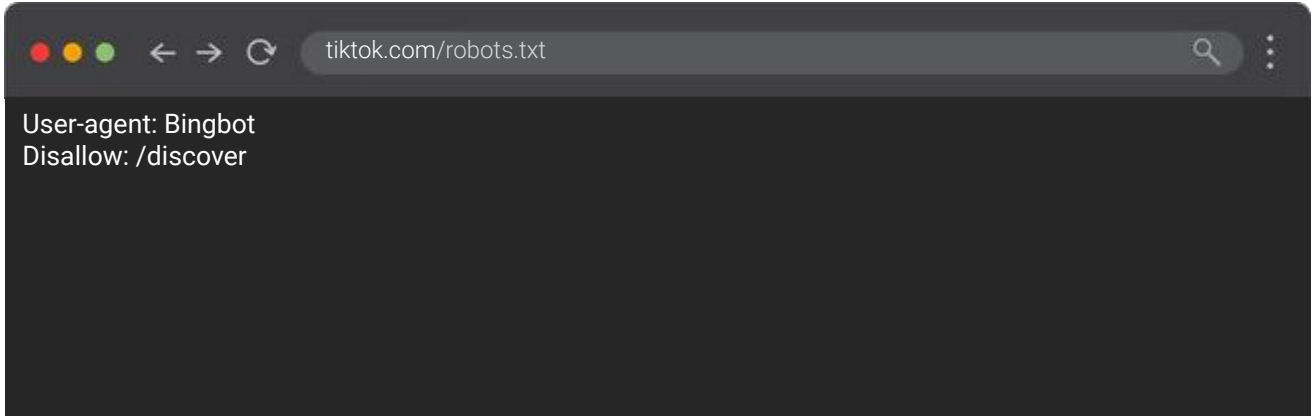
TikTok.com → wil niet dat Bing de /discover folder nog crawlt

Welke robots.txt code brief je naar de developer?



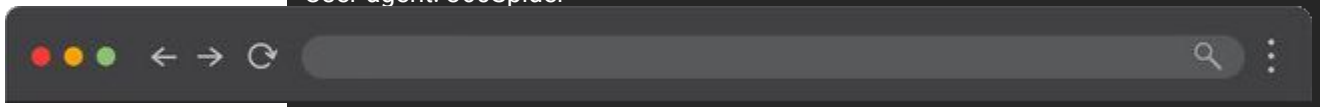
https://www.tiktok.com/discover/best-movies

TikTok.com


tiktok.com/robots.txt
User-agent: Bingbot
Disallow: /discover

TikTok.com

/robots.txt



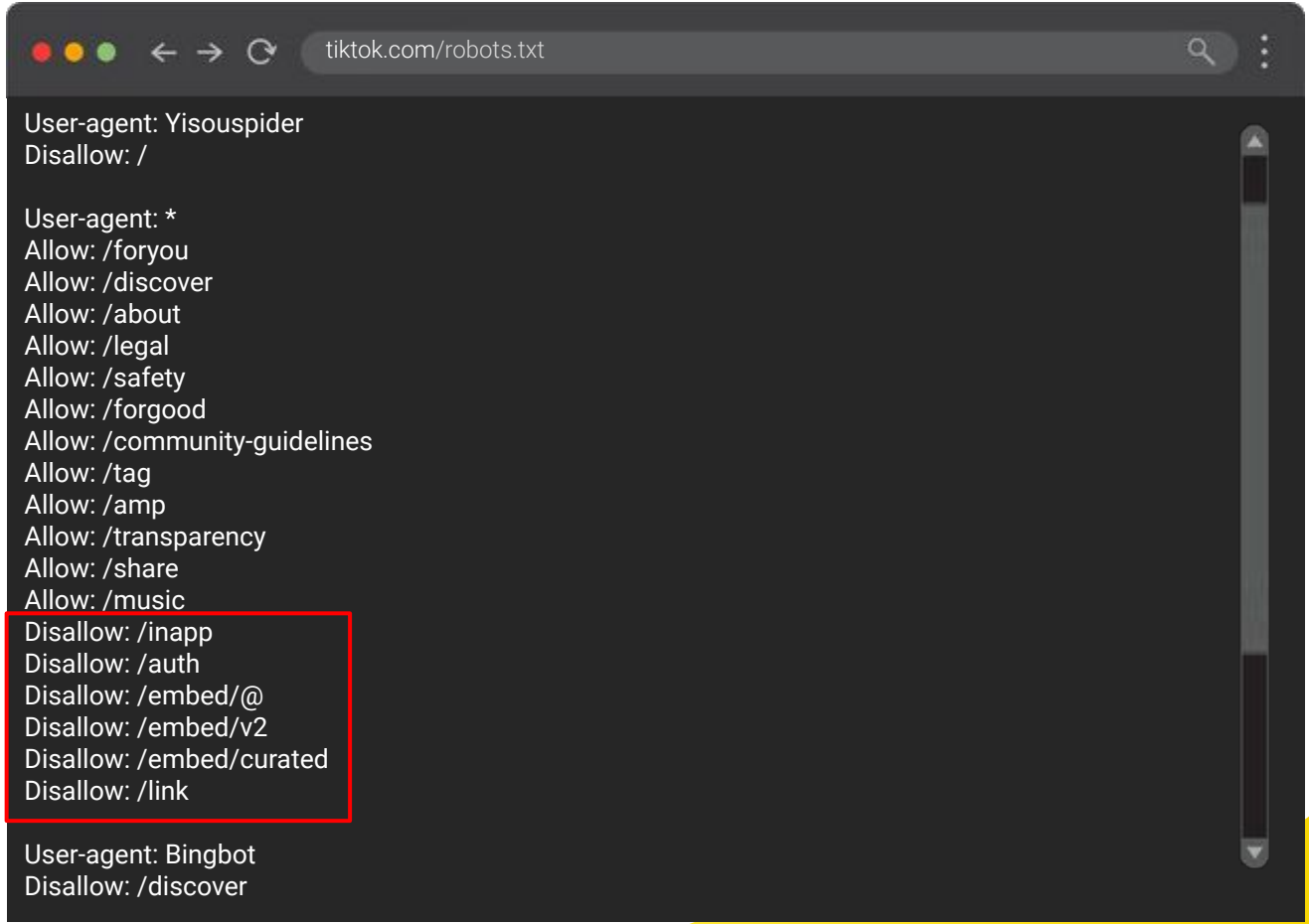
```
User-agent: Baiduspider
User-agent: 360Spider

User-agent: *
Allow: /foryou
Allow: /discover
Allow: /about
Allow: /legal
Allow: /safety
Allow: /forgood
Allow: /community-guidelines
Allow: /tag
Allow: /amp
Allow: /transparency
Allow: /share
Allow: /music
Disallow: /inapp
Disallow: /auth
Disallow: /embed/@
Disallow: /embed/v2
Disallow: /embed/curated
Disallow: /link
```

POV:

Dit was hun robots.txt.

TikTok.com

En dus, de developer copy-pastet uw code.



En ja, ook dat deden ze bij TikTok

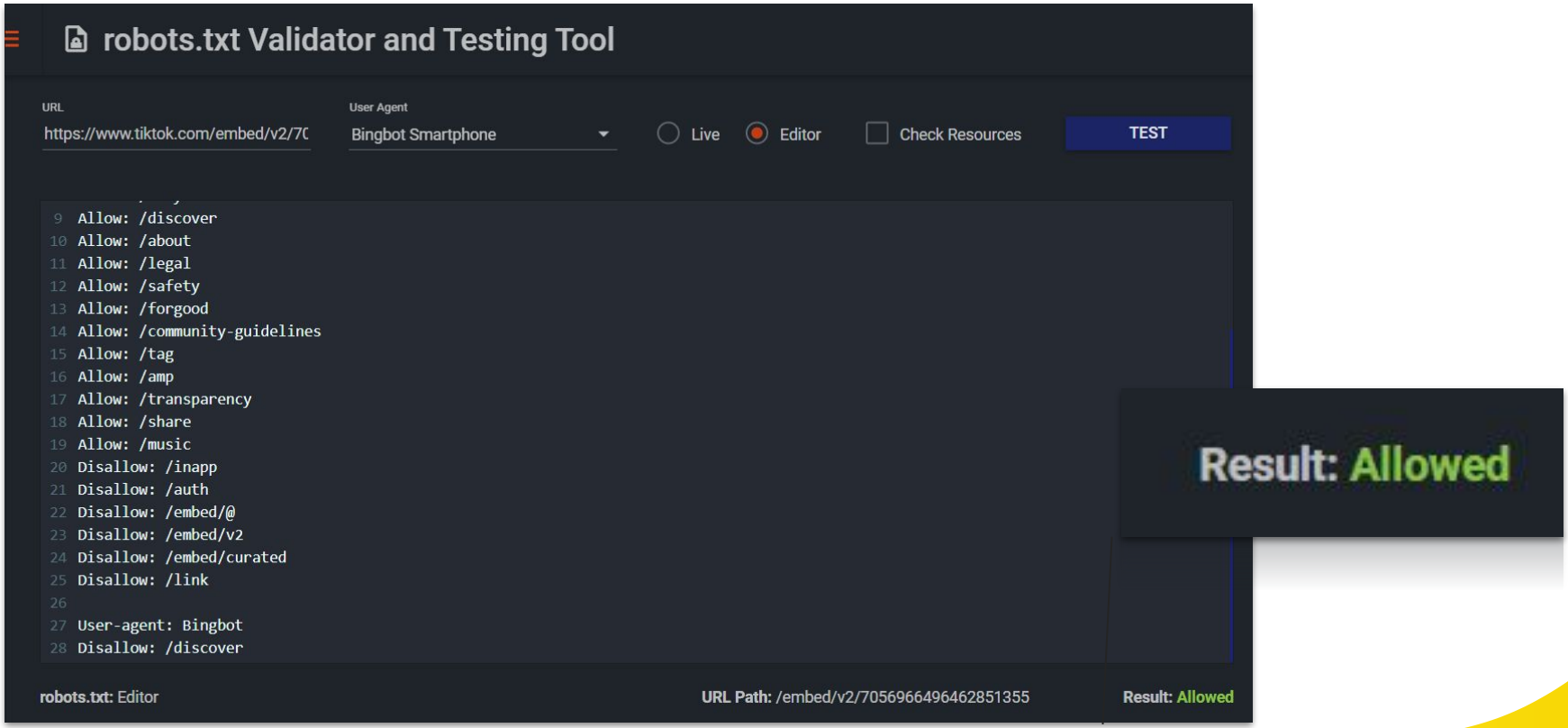
39 lines 0 Removals Copy all

```
User-agent: Baiduspider
User-agent: 360Spider
User-agent: Sogouspider
User-agent: Yisouspider
Disallow: /

User-agent: *
Allow: /foryou
Allow: /discover
Allow: /about
Allow: /legal
Allow: /safety
Allow: /forgood
Allow: /community-guidelines
Allow: /tag
Allow: /amp
Allow: /transparency
Allow: /share
Allow: /music
Disallow: /inapp
Disallow: /auth
Disallow: /embed/@
Disallow: /embed/v2
Disallow: /embed/curated


Sitemap: https://www.tiktok.com/tos/node/sitemaps/sitemap.xml
Sitemap: https://www.tiktok.com/tos/node/sitemaps/sitemap_tag_jpagency_index_20200509_-1553821641.xml
```

42 lines + 3 Additions Copy all

```
User-agent: Baiduspider
User-agent: 360Spider
User-agent: Sogouspider
User-agent: Yisouspider
Disallow: /

User-agent: *
Allow: /foryou
Allow: /discover
Allow: /about
Allow: /legal
Allow: /safety
Allow: /forgood
Allow: /community-guidelines
Allow: /tag
Allow: /amp
Allow: /transparency
Allow: /share
Allow: /music
Disallow: /inapp
Disallow: /auth
Disallow: /embed/@
Disallow: /embed/v2
Disallow: /embed/curated

User-agent: Bingbot
Disallow: /discover
Sitemap: https://www.tiktok.com/tos/node/sitemaps/sitemap.xml
Sitemap: https://www.tiktok.com/tos/node/sitemaps/sitemap_tag_jpagency_index_20200509_-1553821641.xml
```

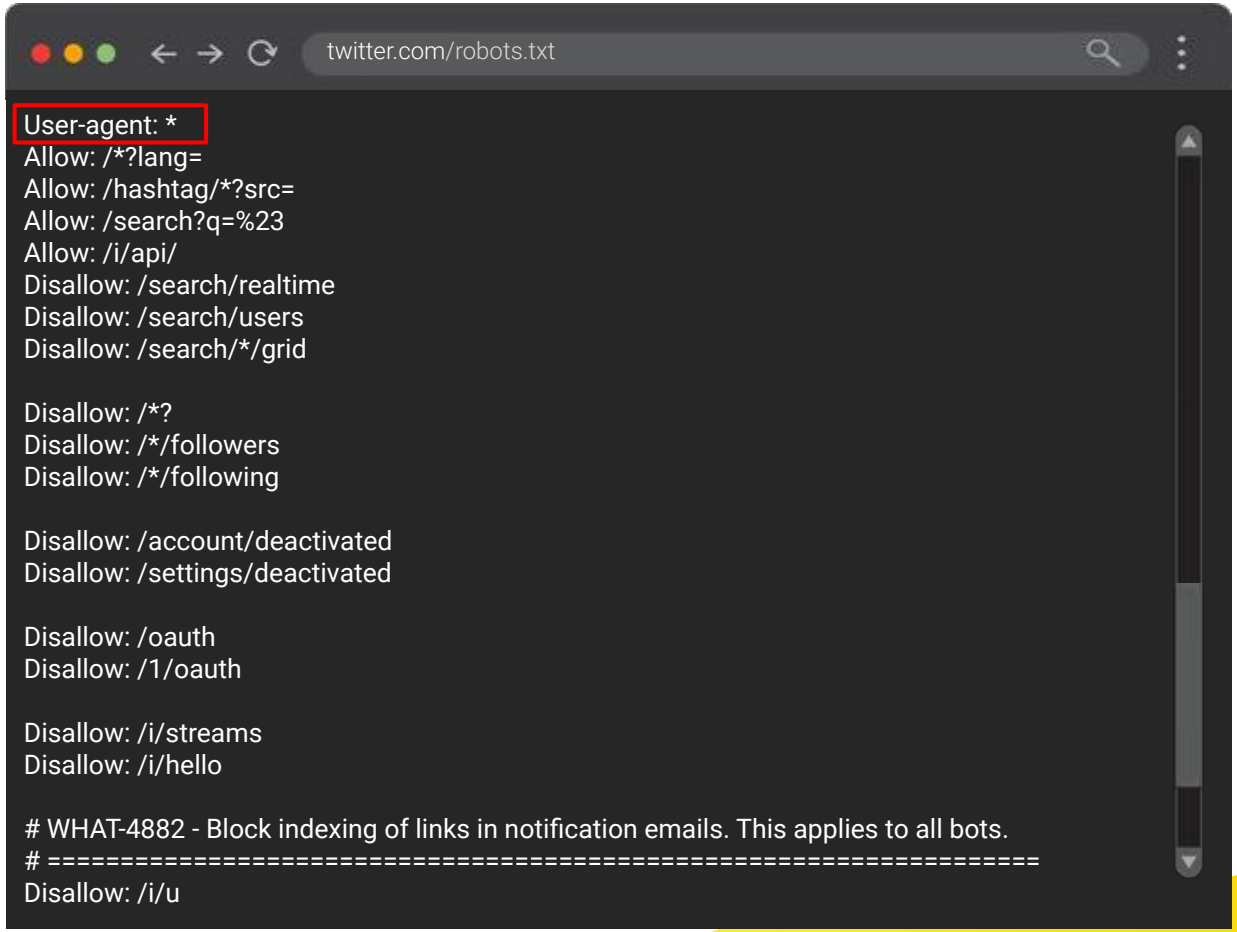
Het probleem:



https://www.tiktok.com/embed/v2/7056966496462851355

En doet juist hetzelfde.

De robots.txt van



We scrollen wat naar boven

twitter.com/robots.txt

```
# Bing Search Engine Robot
# ========================
User-Agent: bingbot
Allow: /?_escaped_fragment_

Allow: /*?lang=
Allow: /hashtag/*?src=
Allow: /search?q=%23
Disallow: /search/realtime
Disallow: /search/users
Disallow: /search/*/grid

Disallow: /*?
Disallow: /*/followers
Disallow: /*/following

Disallow: /account/deactivated
Disallow: /settings/deactivated

# Every bot that might possibly read and respect this file
# ========================================================
User-agent: *
Allow: /*?lang=
Allow: /hashtag/*?src=
Allow: /search?q=%23
```

Disallow: /oauth
Disallow: /1/oauth

Disallow: /i/streams
Disallow: /i/hello

Disallow: /i/u

We scrollen nog wat naar boven

twitter.com/robots.txt

```
# Google Search Engine Robot
# ==========================
User-agent: Googlebot
Allow: /?_escaped_fragment_

Allow: /*?lang=
Allow: /hashtag/*?src=
Allow: /search?q=%23
Allow: /i/api/
Disallow: /search/realtime
Disallow: /search/users
Disallow: /search/*/grid

Allow: /*?ref_src=
Allow: /*?src=
Disallow: /*?
Disallow: /*/followers
Disallow: /*/following

Disallow: /account/deactivated
Disallow: /settings/deactivated

# Yahoo! Search Engine Robot
# ==========================
User-Agent: Slurp
```

Disallow: /oauth
Disallow: /1/oauth

Disallow: /i/streams
Disallow: /i/hello

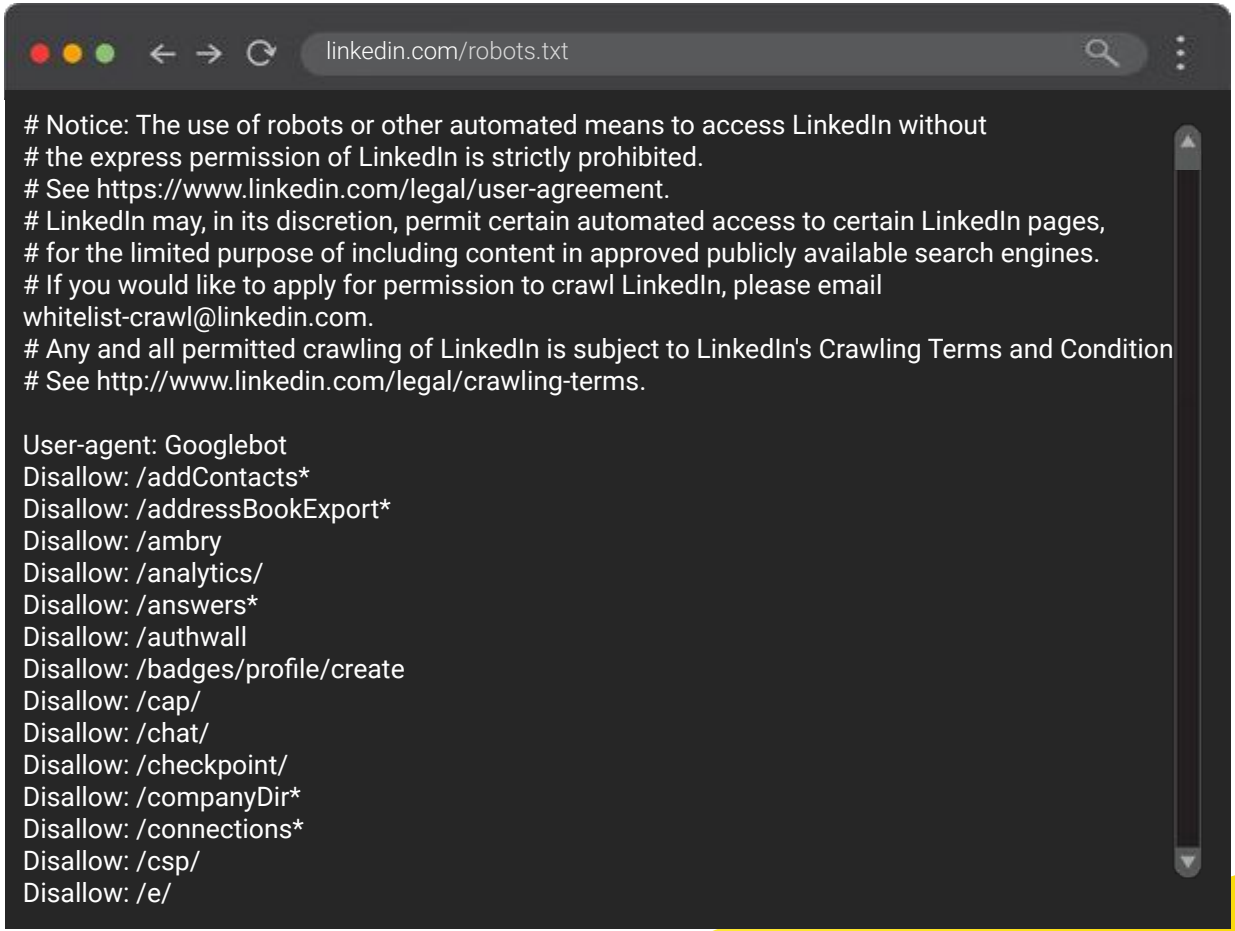
Disallow: /i/u

LinkedIn is dan weer chaos.

De robots.txt van LinkedIn

🤯

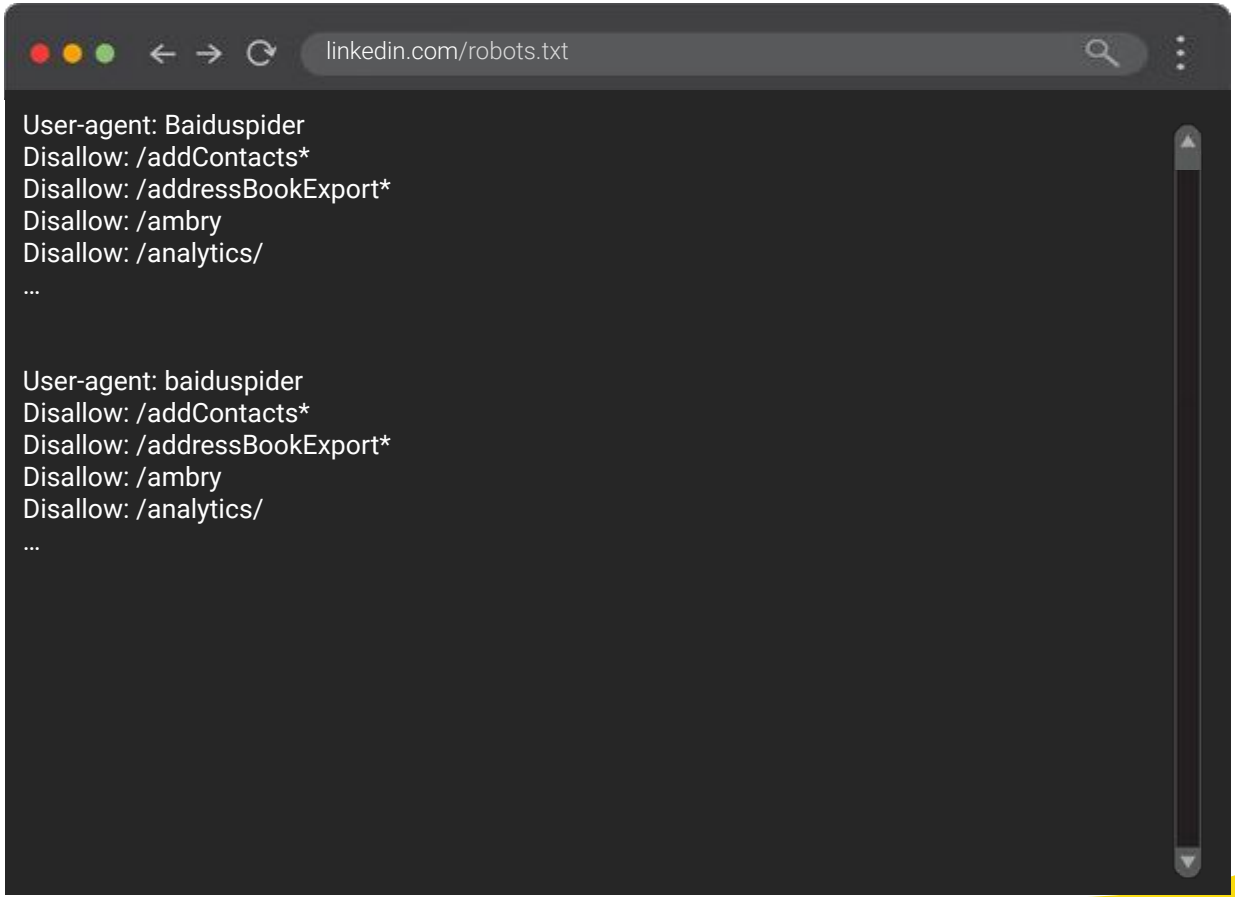
4102 lijnen tekst



Lichtjes chaotisch.

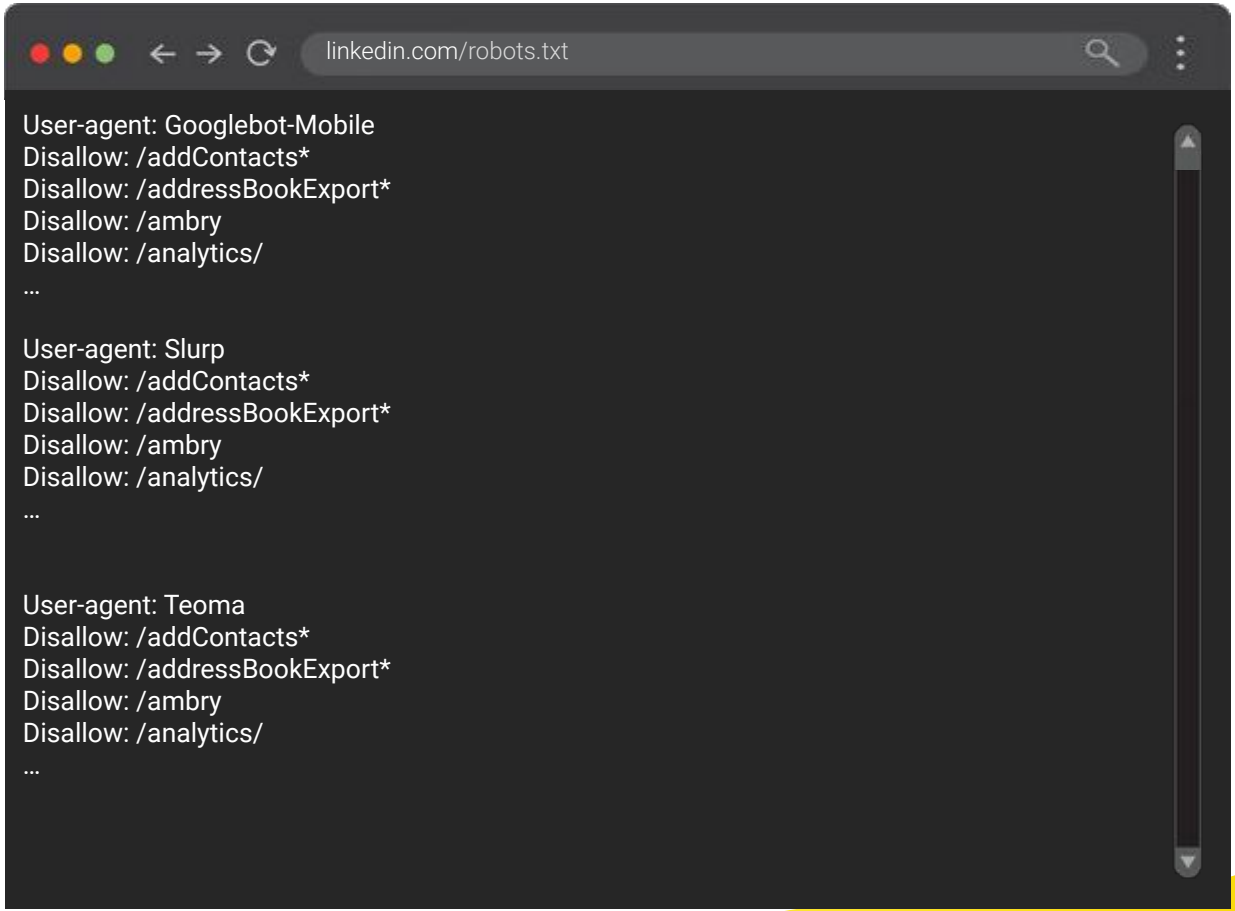
En wat opvalt:

eentje had ook volstaan.



En wat opvalt:

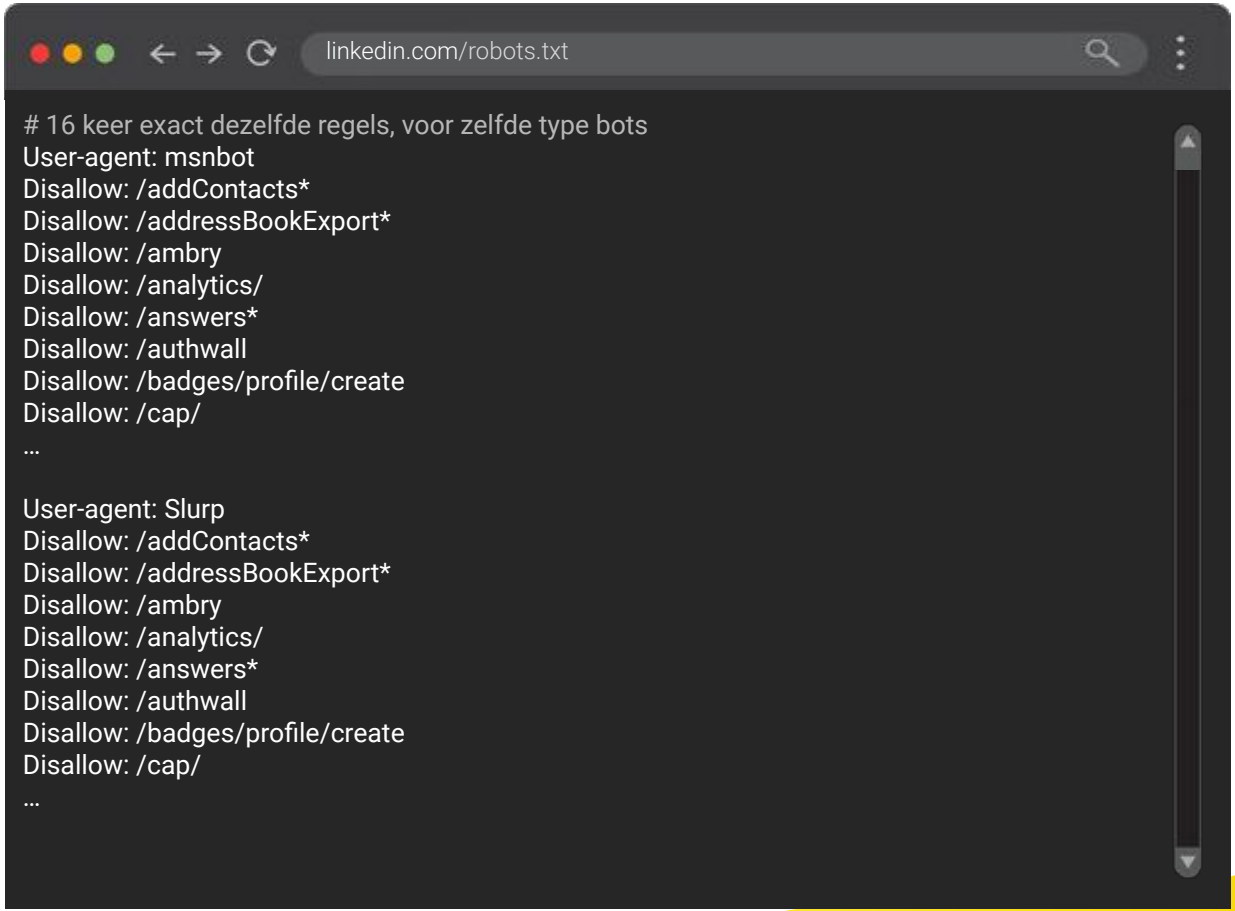
Outdated.



https://developers.google.com/search/docs/crawling-indexing/overview-google-crawlers

En wat opvalt:

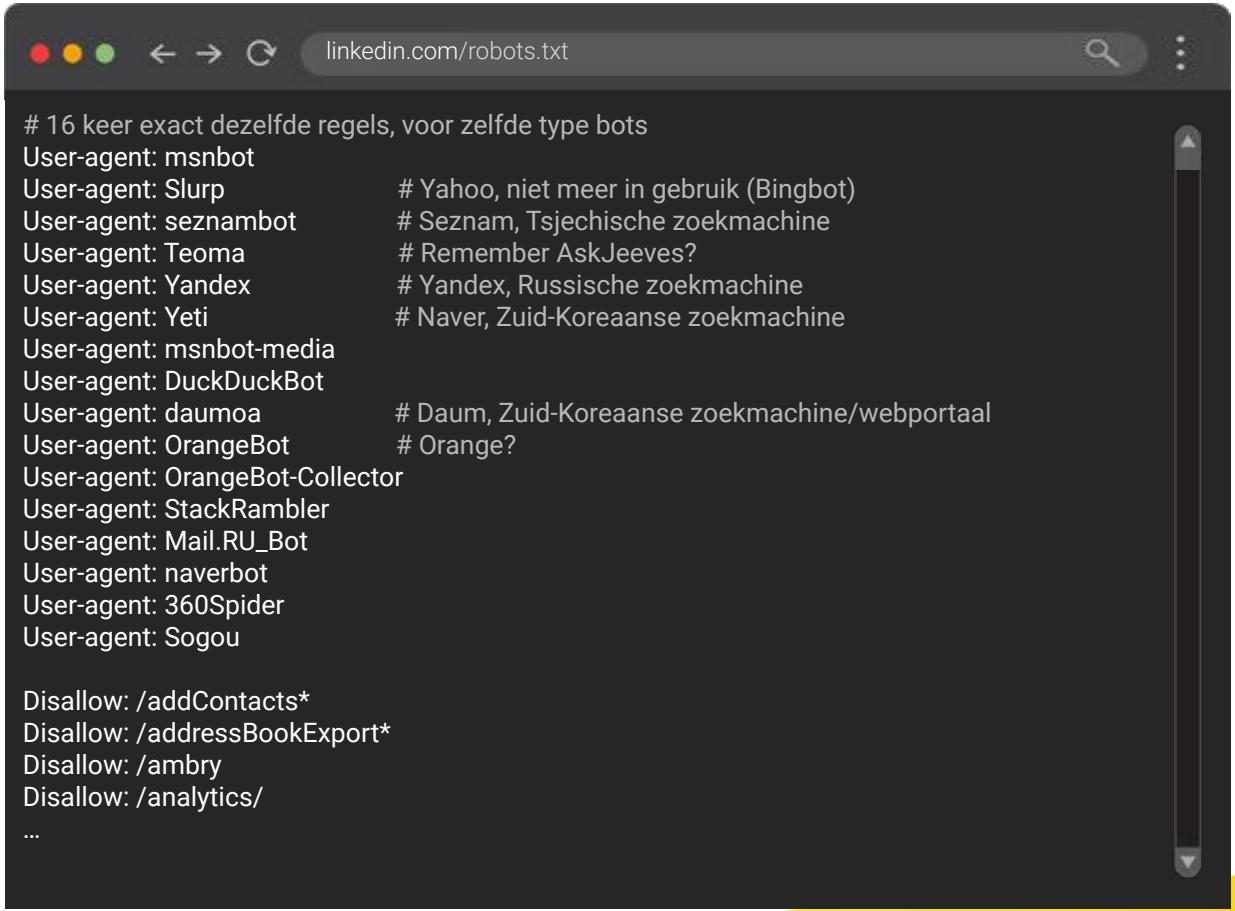
16 gelijkaardige UA's met **identiek** dezelfde regels.



En wat opvalt:

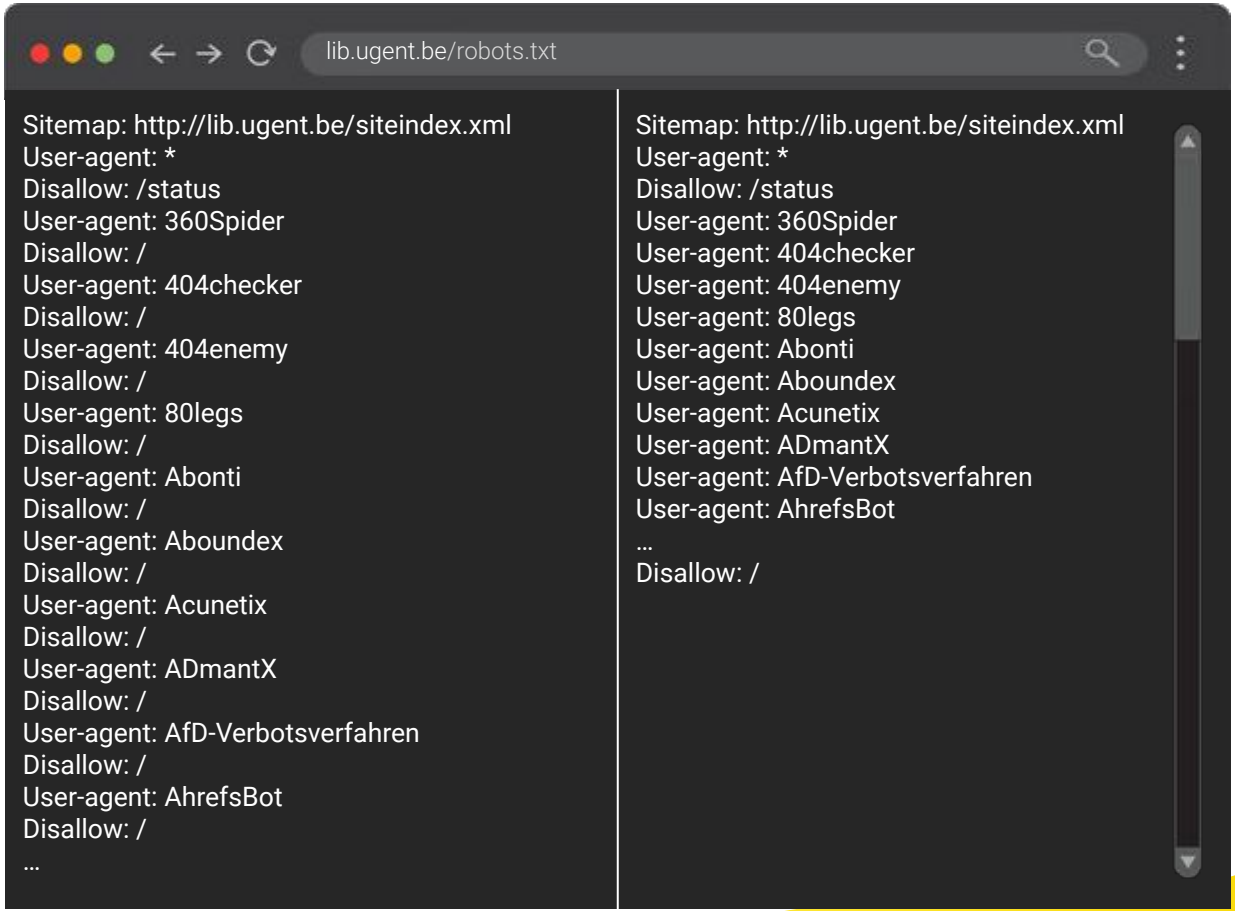
16 gelijkaardige UA's met **identiek** dezelfde regels.

1.710 (42%) lijnen uitgespaard.



Leuk weetje:

Dichter bij huis,
juist hetzelfde



Er is geen monitoring.

Herken je dit?

> Can you please add this code to the robots.txt?

DE MOL
Hier zitten we weer.

2 bezorgdheden voor tech SEO’ers:

1. (Wanneer) hebben ze het uitgevoerd?

2. Gaat er niks stuk bij een robots.txt aanpassing in de toekomst?

Wat als je dat zou monitoren met Slack alerts?

proxy.php

```php
<?php
header("Cache-Control: no-store, no-cache, must-revalidate, max-age=0");
if(isset($_GET['site'])){
  include_once 'inc/dbh.inc.php';

  // Check for errors
  if (mysqli_connect_errno()) {
      die('Connect Error (' . mysqli_connect_errno() . ') ' .
        mysqli_connect_error());
  }

  // Select the column you want to retrieve
  $column = 'site_url';

  // Write the SQL query to retrieve the column
  $query = "SELECT id FROM robotstxt_tbl WHERE $column = " . "'" . $_GET["
    site"] . "'" . "LIMIT 1";

  // Execute the query
  $result = mysqli_query($conn, $query);

  // Check if the query was successful
  if (!$result) {
    die('Failed to retrieve the data: ' . mysqli_error($conn));
  }
```

database.txt

robotstxt-alert 1

Sunday, April 16th

robotstxt-alert APP 10:03 AM
Your robots.txt has changed: https://www.orange.be/robots.txt

Wednesday, April 19th

robotstxt-alert APP 10:04 AM
Your robots.txt has changed: https://www[redacted]be/robots.txt

Friday, April 21st

robotstxt-alert APP 10:05 AM
Your robots.txt has changed: https://www.google.com/robots.txt

Thursday, April 27th

robotstxt-alert APP 10:09 AM
Your robots.txt has changed: https://www[redacted].be/robots.txt

Friday, April 28th

robotstxt-alert APP 10:03 AM
Your robots.txt has changed: https://www.netflix.com/robots.txt

Your robots.txt has changed: https://www.[redacted]com/robots.txt

Saturday, April 29th

robotstxt-alert APP 10:12 AM
Your robots.txt has changed: https://ebay.be/robots.txt

Yesterday

robotstxt-alert APP 10:07 AM
Your robots.txt has changed: https://www.facebook.com/robots.txt

In afwachting van eigen dashboard (work in progress)

Diffchecker.com

43 lines 0 Removals

```
User-agent: Baiduspider
User-agent: 360Spider
User-agent: Sogouspider
User-agent: Yisouspider
Disallow: /

User-agent: *
Allow: /foryou
Allow: /discover
Allow: /about
Allow: /legal
Allow: /safety
Allow: /forgood
Allow: /community-guidelines
Allow: /tag
Allow: /amp
Allow: /transparency
Allow: /share
Allow: /music
Disallow: /inapp
Disallow: /auth
Disallow: /embed/@
Disallow: /embed/v2
Disallow: /embed/curated

User-agent: Bingbot
Disallow: /discover
```

Copy all

44 lines + 1 Addition

```
User-agent: Baiduspider
User-agent: 360Spider
User-agent: Sogouspider
User-agent: Yisouspider
Disallow: /

User-agent: *
Allow: /foryou
Allow: /discover
Allow: /about
Allow: /legal
Allow: /safety
Allow: /forgood
Allow: /community-guidelines
Allow: /tag
Allow: /amp
Allow: /transparency
Allow: /share
Allow: /music
Disallow: /inapp
Disallow: /auth
Disallow: /embed/@
Disallow: /embed/v2
Disallow: /embed/curated
Disallow: /link

User-agent: Bingbot
Disallow: /discover
```

TikTok

PLAY
NIEUW
Dat hou ik geheim,
als jullie het niet erg vinden.

Geen codingkennis (of goesting)**?**

ContentKing for conductor

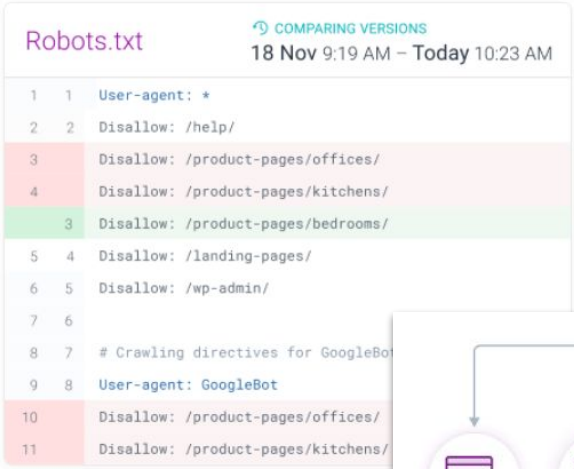
Features Solutions Case studies Start Trial

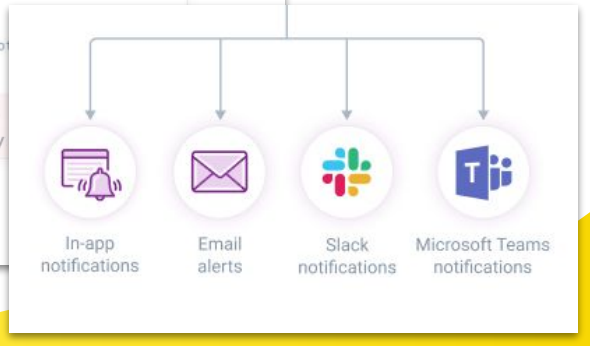
A time machine for your robots.txt file

A correctly implemented robots.txt is essential. A single mistake has the potential to drop your website completely from search engines.

That's why ContentKing immediately alerts you whenever your robots.txt file changes. You're also able to travel back in time to see and compare every single robots.txt revision.

Find answers, get context and gain insights to unlock new opportunities. Learn more about robots.txt monitoring.





Robots.txt: **de learnings**

Robots.txt: De learnings

1. Robots.txt doet niet aan overerving (enkel meest specifieke UA geldt)

2. Gooi altijd in de robots.txt tester. **ALTIJD!**

3. Bundel UA’s met dezelfde regels (leesbaarheid).

4. Monitoring & alerts zijn heel handig.

#2 Migraties: er loopt al eens wat fout.

Nee, het gaat niet hierover

- ~~302~~ → 301 redirect.

- robots.txt: Disallow: / van staging naar live

  (solution: browser paswoord)

Redirects met UTM’s

Redirects met UTM tags

- ?utm_medium=301&utm_source=website.com
- met canonical tag naar cleane URL

oude URL	nieuwe URL
/over-ons/contact	/contact?utm_medium=301&utm_source=website.com
/producten	/shop?utm_medium=301&utm_source=website.com
/p/gameboy-color	/p/gameboy-color-14522?utm_medium=301&utm_source=website.com
/nl/home	/nl?utm_medium=301&utm_source=website.com

Redirects met UTM tags

- **entries:** → attributie kapot
  - ~~google/organic~~ → website.com/301
  - ~~google/cpc~~ → website.com/301
  - ~~facebook.com/referral~~ → website.com/301

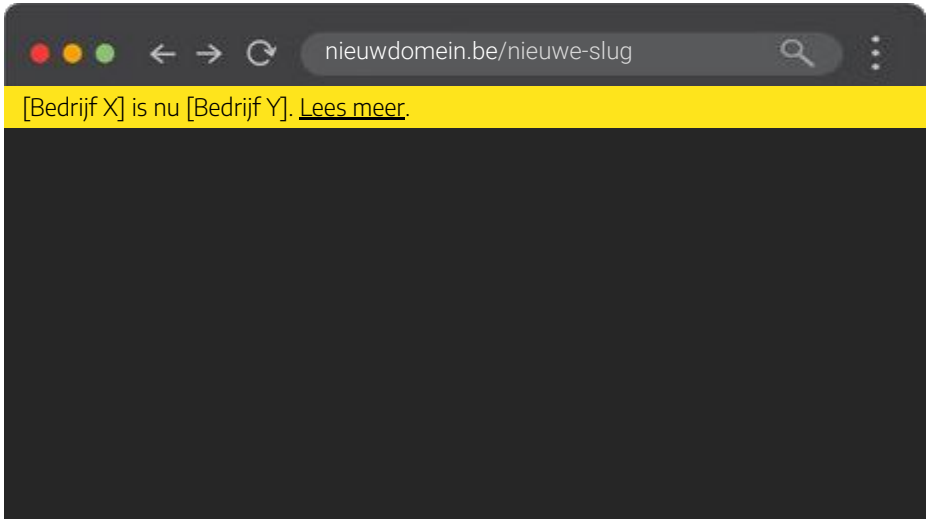
- **interne links naar oude URL:** redirect → nieuwe sessie gecreëerd (= 2 sessies)

Hoe dan wel?

- **Geen UTM tags**
- **Custom dimensie in GTM (“migratie_url=true”) instellen en meeschieten**

Redirects met UTM tags

“Maar we hebben dit nodig om een melding (banner) te tonen aan mensen die van oudewebsite.com komen.”



Redirects met UTM tags

“Maar we hebben dit nodig om een melding (banner) te tonen aan mensen die van oudewebsite.com komen.”

→ tussenredirect (1P cookie zetten) & uitlezen

```
12  # tussenredirect
13
14  RewriteCond "%{HTTP_HOST}" .*ouddomein\.be.*
15  RewriteRule .* https://nieuwdomein.be/redirect%{REQUEST_URI}
    [CO=blockcookie:true:%{HTTP_HOST}:600:/,R=301,L]
16
17  # Uw standaardredirect regels
18  ...
```

1. ouddomein.be/oude-slug
2. nieuwdomein.be/redirect/oude-slug (*plaatst cookie*)
3. nieuwdomein.be/nieuwe-slug

Geen staging opgezet.

Gij
PLAY
Ik zie geen hol.

Developer

Geen staging opgezet

“te veel werk” of “te grote database”

- Redirects kunnen niet getest worden
- Technische audit staging kan niet
- Veel assumpties
- **“Onze developers hebben reeds rekening gehouden met SEO tags.”**

Ik geloof er niks van.
- Ik geloof er ook niks van.

En dan gaan ze live...

De redirects zijn niet actief.

Echt error, gewoon error.

Voorbeeld van zulke assumpties

“De PDF’s blijven onveranderd.” → Ah, geen redirects nodig dus.

Nieuwe site: **PDF’s blijken niet overgezet.**

Voorbeeld van zulke assumpties

	A	B	C	D	J	L	M
1						GSC clicks	
2	Oude URL	Slug	Status nu?	Type	Nuttig	Clicks	Vertoningen
4			404	PDF	Ja	1576	9832
7			404	PDF	Ja	1394	18638
8			404	PDF	Ja	761	24186
10			404	PDF	Ja	599	30728
11			404	PDF	Ja	582	5255
12			404	PDF	Ja	576	21804
13			404	PDF	Ja	550	7104
14			404	PDF	Ja	544	15577
15			404	PDF	Ja	523	8320
16			404	PDF	Ja	491	12083
17			404	PDF	Ja	489	14693
18			404	PDF	Ja	464	5433
19			404	PDF	Ja	425	2750
20			404	PDF	Ja	388	14988
21			404	PDF	Ja	388	9058
22			404	PDF	Ja	369	7558
23			404	PDF	Ja	353	3298

ALWAYS GOOD
Niks aan te doen.

#3 Kan beter:

- communicatie
- dev relations
- buy-in krijgen

Altijd tekstuele briefings

Ticket loopt niet?

Case: prerender.io bleef regelmatig kapot gaan

- Meermaals gebriefd (ticket/mails)
- Vermoeden: Elke release ⇒ kapot
- Sleept aan (~ maanden)

Jan → PO → Lead dev → Front-end dev

Ticket loopt niet?

Case: prerender.io bleef regelmatig kapot gaan

En opnieuw:

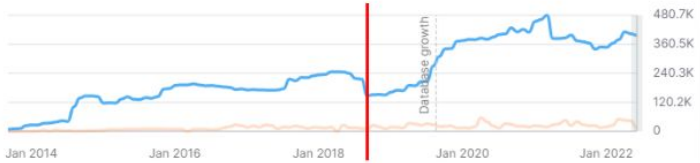
Jan → PO → Lead dev → Front-end dev

⇒ 1 week later definitief opgelost.

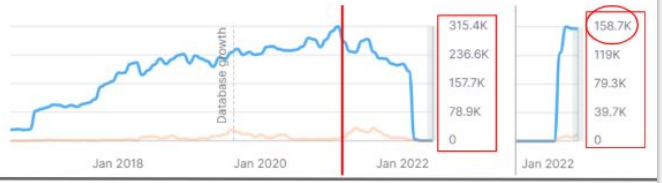
Nieuwe website plannen:
Hoe geraak je als SEO’er mee aan de tafel?

“Nieuw webplatform kiezen zonder nadenken over SEO? Dit is het risico.”

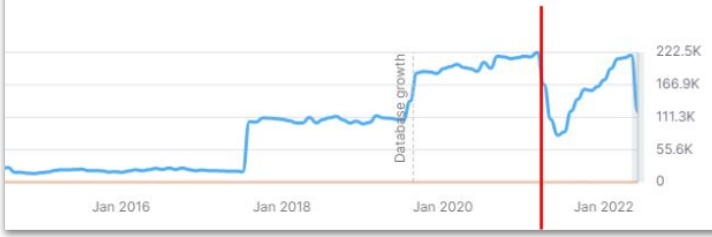
- **32% verlies** van organisch verkeer
- **81% verlies** van “top 3” posities in Google
- **Geraamde kost:** € 5,01 miljoen online omzet verloren
- Offline impact? Dev. kost? 12M vooraleer terug oud niveau



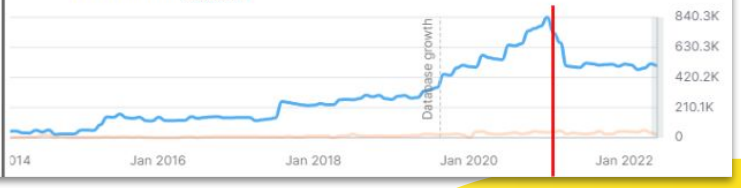
- **43% verlies** van organisch verkeer
- **53% verlies** van “top 3” posities in Google



- **50% verlies** van organisch verkeer
- **72% verlies** van “top 3” posities in Google



- **40% verlies** van organisch verkeer
- **38% verlies** van “top 3” posities in Google
- **XX% verlies organische omzet** (Confidentieel: cijfer op te vragen bij 29AC)



En jawel…

Niet veel later een mail:

> “De enterprise architecten zouden graag eens een meeting inplannen met je om te weten wat ze dienen te vragen aan de vendors voor SEO.”

NIEUW
THE END
Amai. Ik zou wel iets kunnen drinken.