L'analyse de logs pour tous.

SEOLYZER

SEO Good Vibes

Hey!

Olivier PAPON, dans le SEO depuis 2005.

Créateur de la solution **Seolyzer**

DECATHLON ELLE AIRFRANCE Calvin Klein JUST EAT

@Seolyzer_io

SEOLYZER SEO Good Vibes

Quand quelqu’un
entend “logs”.

D'où viennent les logs ?

Code source
= un hit

1 Police = un hit

1 Fichier CSS = un hit

1 Image = un hit

1 Fichier Javascript = un hit

Un hit = une demande envoyée au serveur

Astuce : dans votre navigateur, visualisez toutes les ressources chargées par une page avec F12 puis onglet "Network"

SEOLYZER

Le fichier de logs

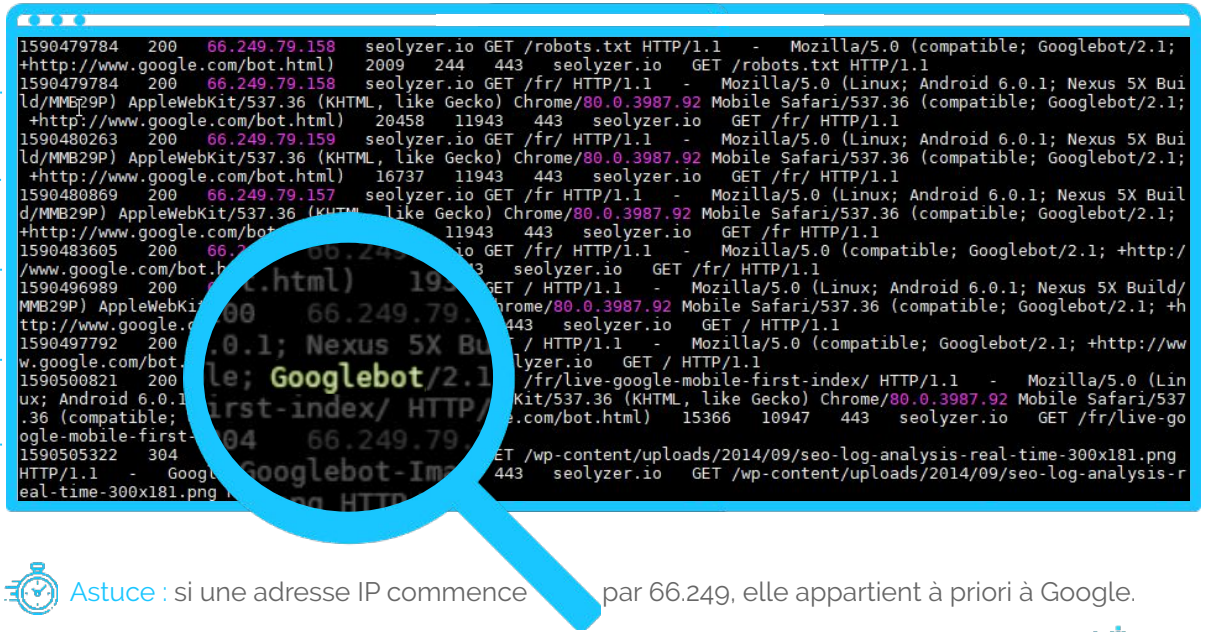
1 Code source page

1 Police

1 Fichier CSS

1 Image

1 Fichier Javascript



Astuce : si une adresse IP commence par 66.249, elle appartient à priori à Google.

SEOLYZER

Que contiennent les logs ?

Chaque ligne contient

- timestamp
- client-ip.
- url
- status code
- user-agent
- referrer
- weight (bytes)
- response time (optionnel)
- protocole ou port (optionnel)

Où récupérer les logs ?

Cela dépend de votre infrastructure / hébergement

CDN

ou Load Balancer

CLOUDFLARE

Akamai

aws

Live | Fichiers | API

Mutualisé

o2switch

OVH

PlanetHoster

Live | Fichiers

Dédié

APACHE HTTP SERVER PROJECT

NGINX

VARNISH CACHE

Live | Fichiers

SEOLYZER

SEO Good Vibes

Pourquoi utiliser
les logs ?

Log files are so underrated, so much good information in them.

John Mueller. Webmaster Trends Analyst @Google.

SEOLYZER

SEO Good Vibes

Google Search Console

A chacun sa vision...

Propriétaire du site

Le SEO et son crawler

Googlebot

Vision Business : catalogue...

Vision OnPage technique

Vision historique et globale

SEOLYZER

SEO Good Vibes

Crawler n’est pas
la vérité !

Crawler

Simulation

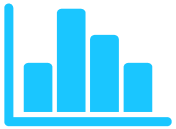
Instantané

A partir de la home page

Dépend de la configuration du crawler

Analyse de logs

Réalité



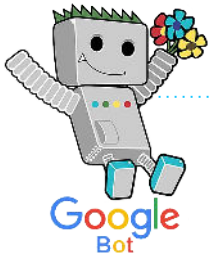
Historique

Live monitoring

SEOLYZER

SEO Good Vibes

Comprendre Googlebot



Site internet

Coût

130 000 milliards de pages

Electricité, serveurs, bande passante...

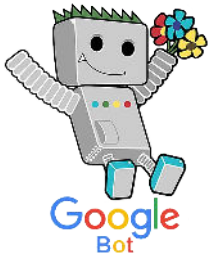
“Crawl budget”

SEOLYZER

SEO Good Vibes

Où est Googlebot ?

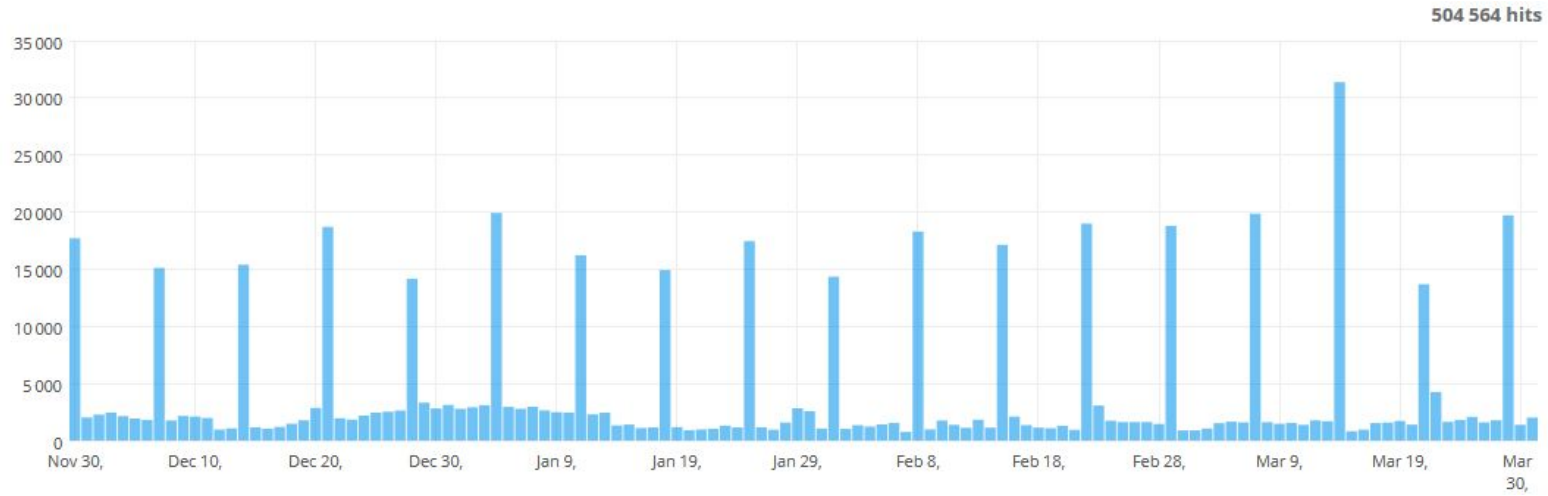
On résume !



Googlebot ♥ les sites où :

- Il ne perd pas de temps
- Les choses importantes sont proches de l'entrée
- Les rayons sont bien rangés
- Les rayons ne sont pas vides

Googlebot, c'est toi ?



SEOLYZER

SEO Good Vibes

Googlebot, c'est toi ?

Astuce : vérifiez les Reverse DNS des IP pour identifier GB

https://support.google.com/webmasters/answer/80553?hl=fr

SEOLYZER

Comment manipuler les logs ?

Vous oubliez !

Il y a des outils free[mium] pour ça.

Open Source

Watussi

ELK

Code modifiable

Freemium

Screamingfrog

1000 lignes

SEOLYZER

10000 pages

Fonctionnalités avancées

SEOLYZER

SEO Good Vibes

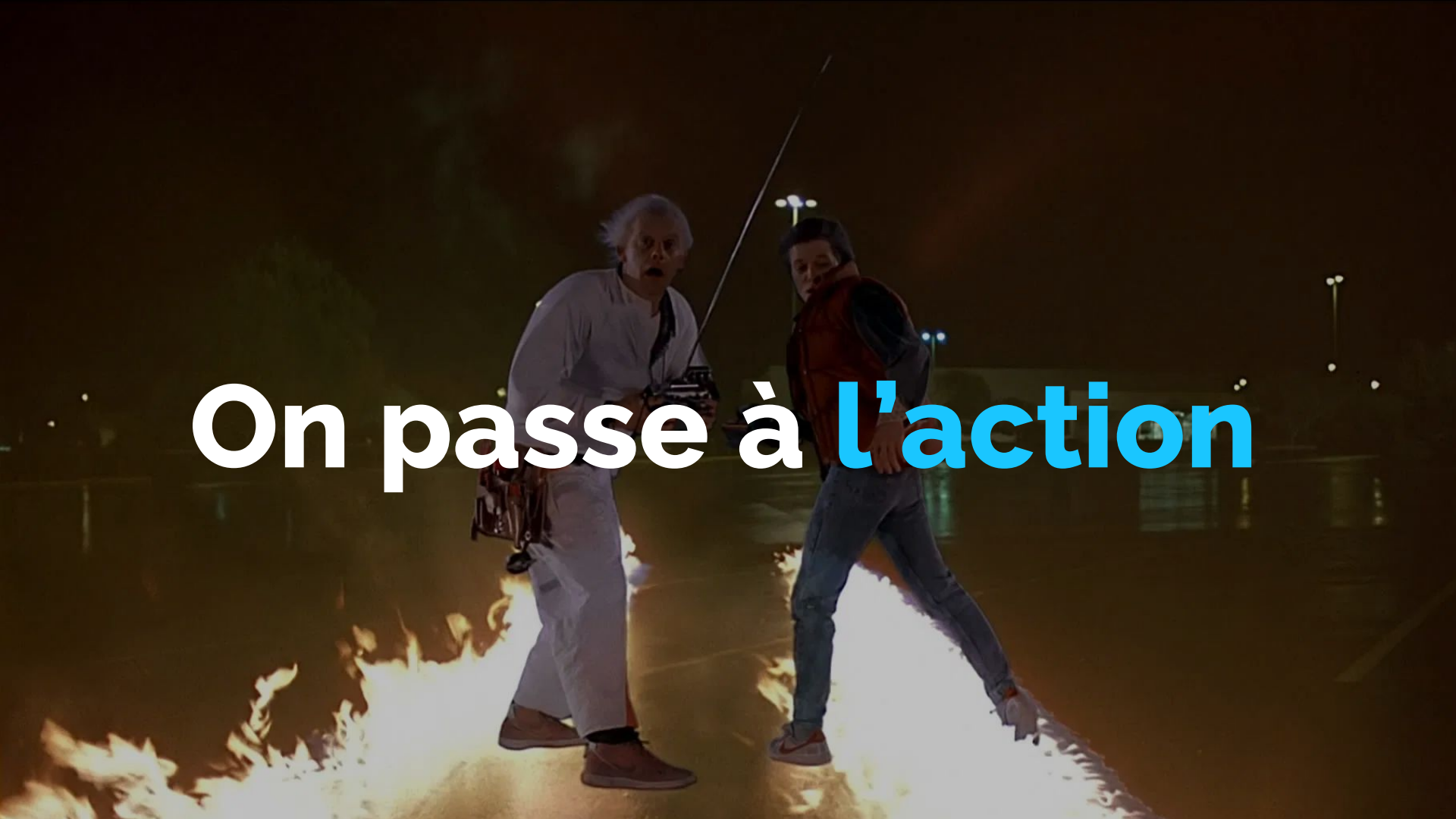
Pour une analyse pertinente du comportement de Googlebot

Catégorisez vos pages

Langue | Produit | Editorial | Template | ...

On peut catégoriser par l'URL ou par élément scrapé (Regex, Xpath, CSS)

SEOLYZER


On passe à l'action

On [résume|répète]
Avec l'analyse de logs vous...

Monitorez

1

- Volume de crawl
- Codes HTTP
- Performance par page
- Performance moyenne
- Poids des pages
- Nouvelles pages (temps de prise en compte)
- Pages les plus crawlées
- Visites SEO
- Robots.txt...

Contexte

- Routine quotidienne
- Refonte de site
- Migration
- Lot de redirection
- Upgrade de CMS...

Améliorez

2

- Erreurs 4XX (Liens perdus)
- Erreurs 5XX
- Pages lentes
- Pièges à robots
  - Recherche interne crawlable
  - Pagination infinie
  - Facettes
  - Scroll infini
  - Paramètres indésirables....

Pensez “business”

3

Googlebot a t-il la bonne vision de mon site ?

Et pour aller plus loin...

4

Croisement de données

- Crawl de Googlebot VS **Crawler SEO** [Pages orphelines]
- Volume de crawl Googlebot VS **Profondeur des pages**
- Volume de crawl Googlebot VS **Chiffre d'affaire**
- Volume de crawl Googlebot VS **Positions dans Google**
- Volume de crawl Googlebot VS **Liens externes**
- Volume de crawl Googlebot VS **Contenu**

Si vous avez la data, la seule limite est votre imagination.

SEOLYZER

SEO Good Vibes

Merci !

Des questions?

@Seolyzer_io olivier@seolyzer.io

SEOLYZER

SEO Good Vibes