



AUDIT SEO



BuzzFeed



BuzzFeed



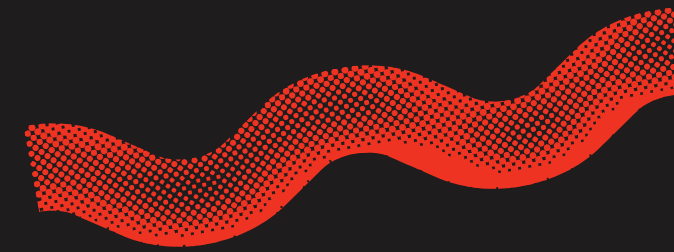
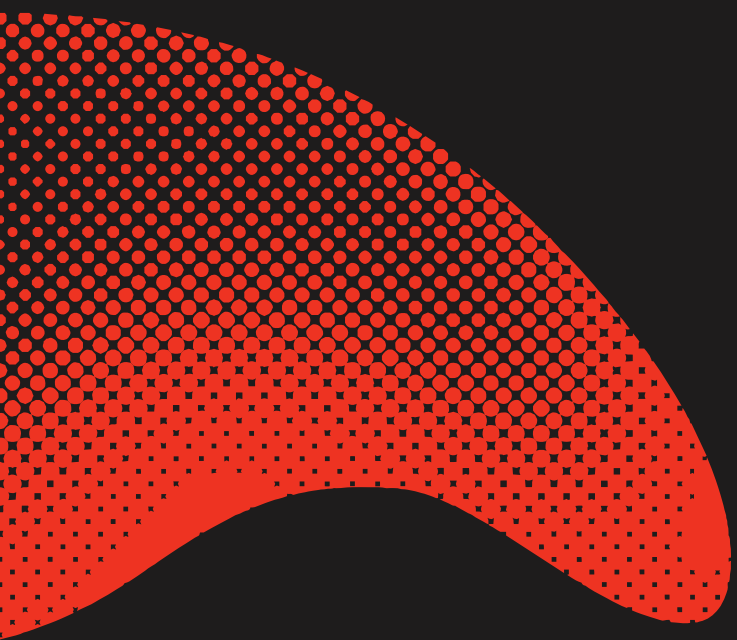
Nombre d'urls scrawl :

10 000


EXPORT SCREAMING FROG



AUDIT TECHNIQUE



Les codes de réponses



Les codes de réponses correspondent à un statut de la page à laquelle on souhaite accéder. En fonction de ce code, on peut facilement savoir si la page est fonctionnelle ou si elle rencontre des difficultés. En catégorisant les codes, cela permet d'identifier plus facilement les erreurs. La majorité des pages du site sont fonctionnelles, et renvoient à un code 200.



Code

301

302

307

308

Sur les **10 000** pages scrawl, une petite quantité sont en redirection **301**.

La page visée redirige automatiquement vers une nouvelle. Plusieurs pages sont également en code **302**, dont le principe est similaire à la **301**. La redirection est cependant censée être temporaire, car elle ne permet pas le transfert de popularité ce qui nuit au site à long terme.



Code

301

302

307

308

Il faudrait donc rebasculer les **302** en **301**, car ce ne sont visiblement pas des pages où la redirection temporaire est logique.

Certaines pages sont aussi en code **307** et **308**, qui permettent de masquer si la redirection est permanente ou temporaire. Il faudrait se pencher dessus pour vérifier que la redirection est bien logique vis-à-vis du besoin de la page.



Code 404

Plusieurs pages renvoient un code d'erreur 404. Les pages ne sont donc pas trouvées sur le serveur à l'url indiqué. C'est embêtant à la fois pour le seo d'avoir des liens morts, mais également pour l'expérience utilisateur (cela augmente potentiellement le taux de rebond).

Il y a deux possibilités pour résoudre le problème :

- supprimer tous les liens vers les pages en erreur 404 et les faire déréférencées par google via la search console
- effectuer une redirection du lien vers une page fonctionnelle et qui correspond à l'intention de recherche de l'ancienne page en erreur 404



Code 5xx

Les erreurs type **500** indiquent un souci entre la requête et le serveur, qui ne peut pas afficher la page demandée.

Les solutions sont multiples selon l'erreur exacte, mais il peut être nécessaire de contacter son hébergeur pour effectuer une maintenance ou une mise à jour du serveur, ou par exemple allouer plus d'espace de stockage ou de puissance au serveur si les pages à charger sont lourdes en ressources.



Robot.txt



De nombreuses pages sont bloquées par le fichier robot.txt. Ces pages ne sont donc pas scrawlées par les bots de google, qui ne peut pas les référencer. Cela ne pose pas de problème, puisque les ressources en question sont des éléments de vidéos miniatures accompagnant des articles, ou des éléments en provenance des réseaux sociaux. Il est donc normal de ne pas vouloir consacrer de son "budget crawl" à ces ressources.

Certificat

Https



→

Ce petit "s" rajouté à http devant une url signifie « sécurisé » en français. Attention, cela ne veut pas dire que le site est forcément sûr, mais qu'un certificat de sécurité est installé sur le site internet. Il permet de crypter les données qui s'échangent entre l'utilisateur qui effectue des requêtes et le serveur qui lui en renvoie en retour. Cependant, si le serveur est corrompu ou appartient à une personne mal intentionnée, les données et l'utilisateur restent en danger.

Certificat Https



Les pages en http ne sont pas appréciées de google car elles mettent automatiquement l'utilisateur en danger vis-à-vis de ses données. Cela a donc un impact sur le **SEO**. De plus, certains CMS génèrent une page en http et une autre en https ce qui crée du contenu dupliqué très mauvais également pour le SEO. Il faut donc rediriger les url en http vers la même en https ce qui est fait par BuzzFeed. Comme c'est une redirection permanente, il faut l'effectuer en **301** et non en **302** ce qui est également le cas. Autre souci, certaines pages sont en certificat https mais ont des ressources photos/vidéos en http. Il faut donc veiller à ce que l'ensemble des ressources du site soient bien sécurisées pour protéger les utilisateurs. C'est d'ailleurs également valable pour les formulaires sur le site de BuzzFeed.

Les balises sémantiques



Pour que le **SEO** fonctionne bien, il se base sur de nombreux mots clefs, chaque landing page visant un mot clef bien particulier. Ce mot clé doit avoir un champ sémantique présent sur la page, mais doit aussi apparaître en l'état. Il apparaît notamment dans l'**URL** de la page, la title dans la **SERP** Google, la meta description ainsi que le **H1** de la page internet.







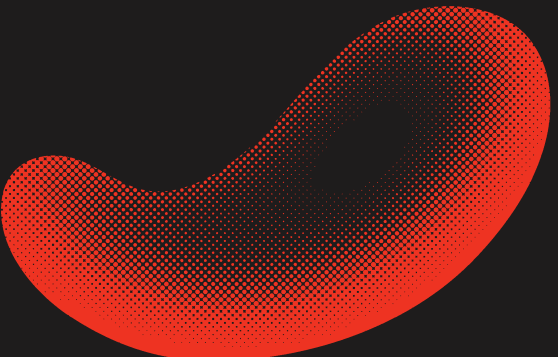
Les URLs



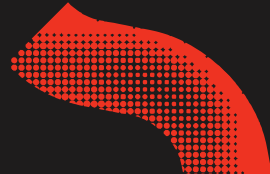
La grande majorité des pages scrawlées n'ont pas de problème vis-à-vis de leur **url**. Cependant, certaines sont trop longues. Ce n'est pas très graves en soit, mais cela limite l'expérience utilisateur puisqu'au-dessus d'une soixantaine de caractères google tronque l'**url** dans les résultats de recherche. Il est donc mieux de raccourcir les urls si possible. Il faut aussi faire attention à ce que l'url contienne le mot clef, ce qui n'est pas toujours le cas.



Les meta titles



Certaines pages ne disposent pas de titre. C'est à la fois mauvais pour le SEO, pour que Google analyse bien le mot clé ciblé, mais également pour l'expérience utilisateur. La seule chose à faire, c'est de prendre les pages concernées une par une et d'y ajouter une titre, ou de passer la page en no index dans le robot.txt si elle n'a pas d'intérêt à être référencée. Sur le site BuzzFeed, des formats sont récurrents mais avec des thèmes différents comme par exemple des quizz.




Les meta titles

Le souci, c'est que ces pages ont une balise title similaire et générique. C'est à la fois moins pertinent pour le ciblage de chaque mot clef, mais c'est aussi pénalisant car Google pourrait considérer ces pages comme des doublons surtout si elles contiennent peu de texte. Il faut donc modifier ces titres pour les rendre uniques et en adéquation avec le mot clé ciblé sur chaque page.

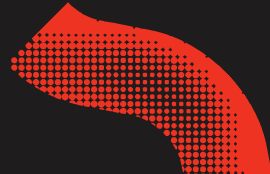
De même que les **urls**, la taille des titre ne convient pas. Beaucoup sont trop courtes, ce qui ne permet pas une idée sur le contenu de la page depuis la serp et rend le placement du mot clef plus compliqué, et d'autres sont trop longues ce qui tronque le texte et réduit la qualité de l'expérience utilisateur. Il faut donc augmenter/réduire la taille des titres pour atteindre environ **60/70** caractères.



Les meta descriptions



Les problèmes des meta descriptions sont similaires aux meta titles : certaines pages indexées n'ont pas la balise, il existe des pages avec un format similaire avec les mêmes description et les dimensions minimales et maximales ne sont pas respectées. Il faudrait donc rajouter celles qui manquent et modifier celles qui en ont besoin, mais également veiller à avoir des meta descriptions d'environ 150 caractères.





Les balises de niveau de titre H1 / H2


Les **Hx** correspondent à des niveaux de titre sur une page internet. Ils sont très importants pour hiérarchiser et organiser le site, à la fois visuellement pour l'utilisateur mais aussi pour Google. Le titre de **niveau 1**, la balise **H1**, doit être présent une et une seule fois sur la page, et il doit être positionné tout en haut de la page tout en contenant le mot clé ciblé.



Les balises de niveau de titre H1 / H2


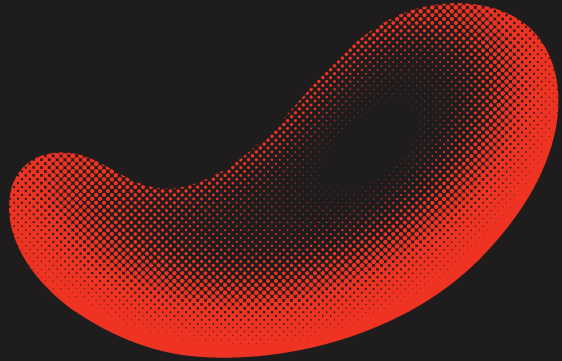
D'un point de vue technique, des balises **H1** sont manquantes, d'autres sont en doublons et d'autres encore sont trop longues. De plus, certaines pages ont plusieurs **H1** qu'il faut absolument supprimer. Il faut globalement apporter les mêmes correctifs que pour les meta descriptions et les titres des pages. Attention également, le problème est similaire aux **H2** et **H3**. Elles ne doivent pas non plus être négligées car la compréhension de la page par google et l'utilisateur est très importante.

Hreflang



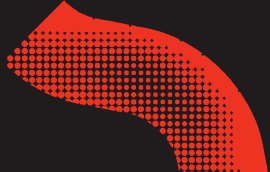
La balise Hreflang permet d'avoir deux pages au contenu identique mais traduite dans des langues différentes sans qu'elles soient considérées comme des doublons. Cela permet aussi de distinguer une langue dans deux localisations, par exemple le français en France et au Québec. L'url est alors adaptée avec un diminutif des deux informations dans le nom de domaine ou en fin d'url.

Hreflang



La majorité des pages de BuzzFeed ont cette balise, cependant les pages ne sont pas doublées dans différentes langues. La balise n'a donc pour le moment pas d'intérêt, et pourrait être supprimée.

Cette balise fonctionne par paire, sur la page originale et sur la traduite. Cependant, il ne faut qu'une fois la balise par page. Plusieurs pages du site BuzzFeed ont pourtant plusieurs fois la balise qui doit donc être retirée, ou alors l'une des deux pages n'a pas sa balise.



Les balises canonical


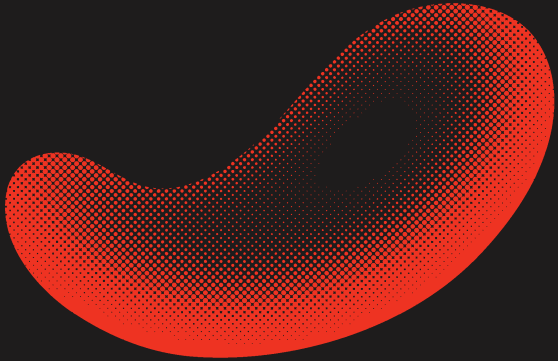


Les balises canonical permettent d'indiquer à Google que cette page fait référence, est calquée, sur une autre page du site ou d'un autre site. Cela permet par exemple de reprendre le contenu d'une autre page de son site sans être considéré comme un doublon. En auto-référencement, cela permet d'éviter les erreurs et considération de contenu dupliqué.

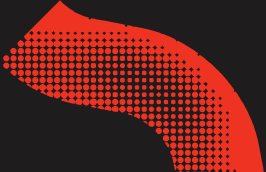
La grande majorité des pages sont, sur BuzzFeed, en auto-référencement. Cependant il y a quelques erreurs, avec par exemple des pages en boucle. La page A est canonisée vers la page B, en redirection **302/301** vers la page A. C'est une erreur pénalisante car cela peut perdre les crawlers et réduire le "budget crawl". Il faut donc changer les liens canoniques des pages concernées.



Contenu dupliqué



Certaines pages ont une url avec des paramètres. Ce sont des caractères ajoutés à la fin de l'url qui permettent par exemple les filtres lors de la recherche d'article entre deux dates ou sur tel sujet. Le problème, c'est que des pages du site pointent vers des urls avec des paramètres qui ne sont pas canonisées ou redirigées. Google peut donc considérer ces pages comme des doublons exacts et pénaliser le référencement de tout le site internet. La meilleure solution, c'est de veiller à ne pas inclure de lien avec des paramètres sur le site, et de corriger ceux déjà existants.

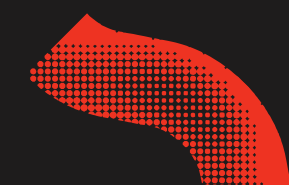





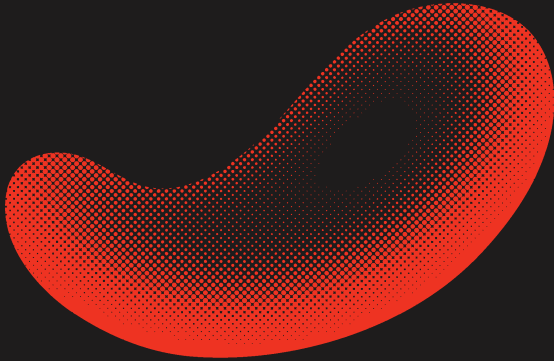
Les images

Pour apparaître dans Google image, les illustrations d'un site doivent être accompagnées d'une balise alt. Elle permet de rentrer un mot-clef sur lequel notre image pourrait correspondre à la requête. Elles sont absentes sur de nombreuses pages de BuzzFeed, ou présentes mais sans texte de renseigné.

Il faut donc les modifier une par une, et veiller à bien le renseigner pour les futurs images des articles. Attention également au mode de compression des images qui peuvent être lourdes et donc pénaliser la vitesse de chargement.

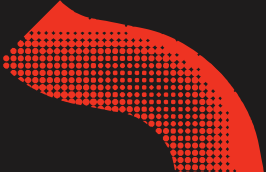


La pagination

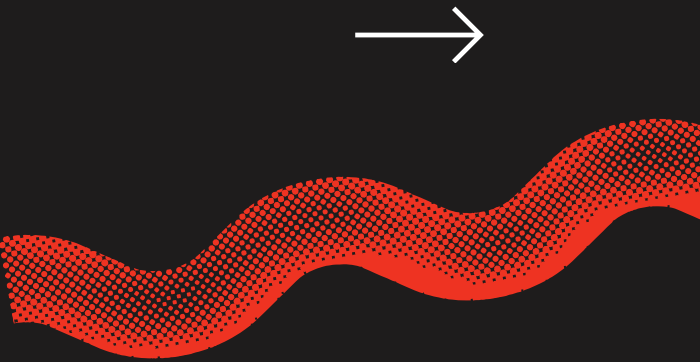


La pagination permet de créer une suite d'éléments avec des pages liées. Par exemple, il est possible de recenser tous les articles sur une même thématique. Ce lien est fait par la présence d'une balise, d'un attribut dans l'**URL** pour identifier la relation, et d'un lien url entre les deux pages.


Cependant plusieurs pages n'ont pas toutes ces conditions de validité ce qui peut perdre les crawlers de google. Il faut donc veiller à bien établir ces liens à la création des pages, et corriger celles qui ne les ont pas. Il faut d'ailleurs également veiller à ce que le lien réciproque de pagination ne soit pas entre des pages qui ne sont pas en **code 200**.



Structure du site

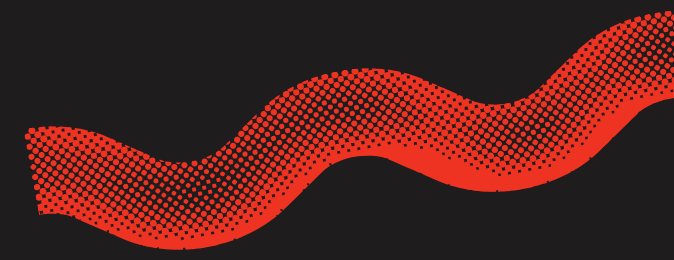
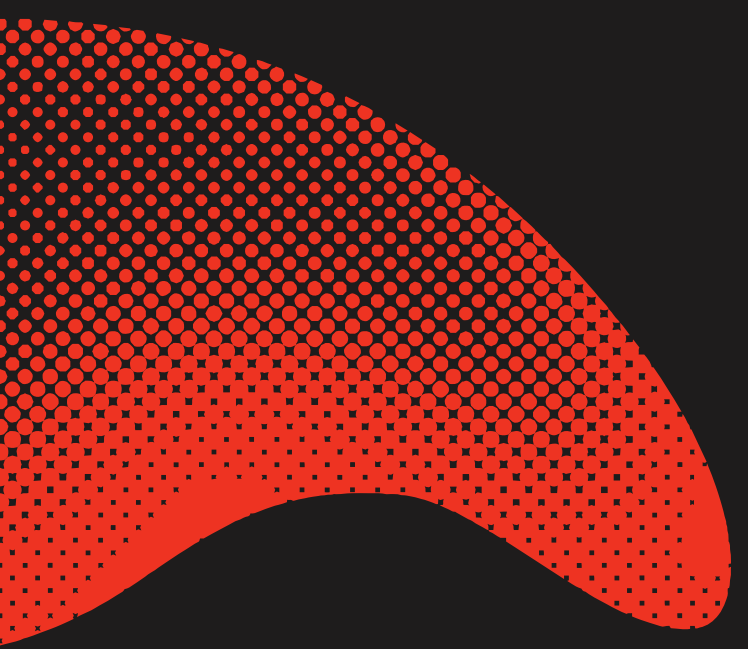


Pour qu'une page de site soit trouvable, elle doit être accessible en **3 clics** maximum. Un utilisateur n'ira sans doute pas plus loin pour chercher, ce qui peut être mauvais pour l'expérience utilisateur. Pour les bot google c'est également mauvais puisqu'ils perdent du temps sur des pages souvent inutiles, et c'est donc un point négatif sur l'ensemble du site. Elles sont rares mais il existe des pages, notamment d'archive, à plus de **3 niveaux** sur le site BuzzFeed. Il faudrait donc les retirer si elles ne sont pas utiles ou alors les réorganiser.





AUDIT UX

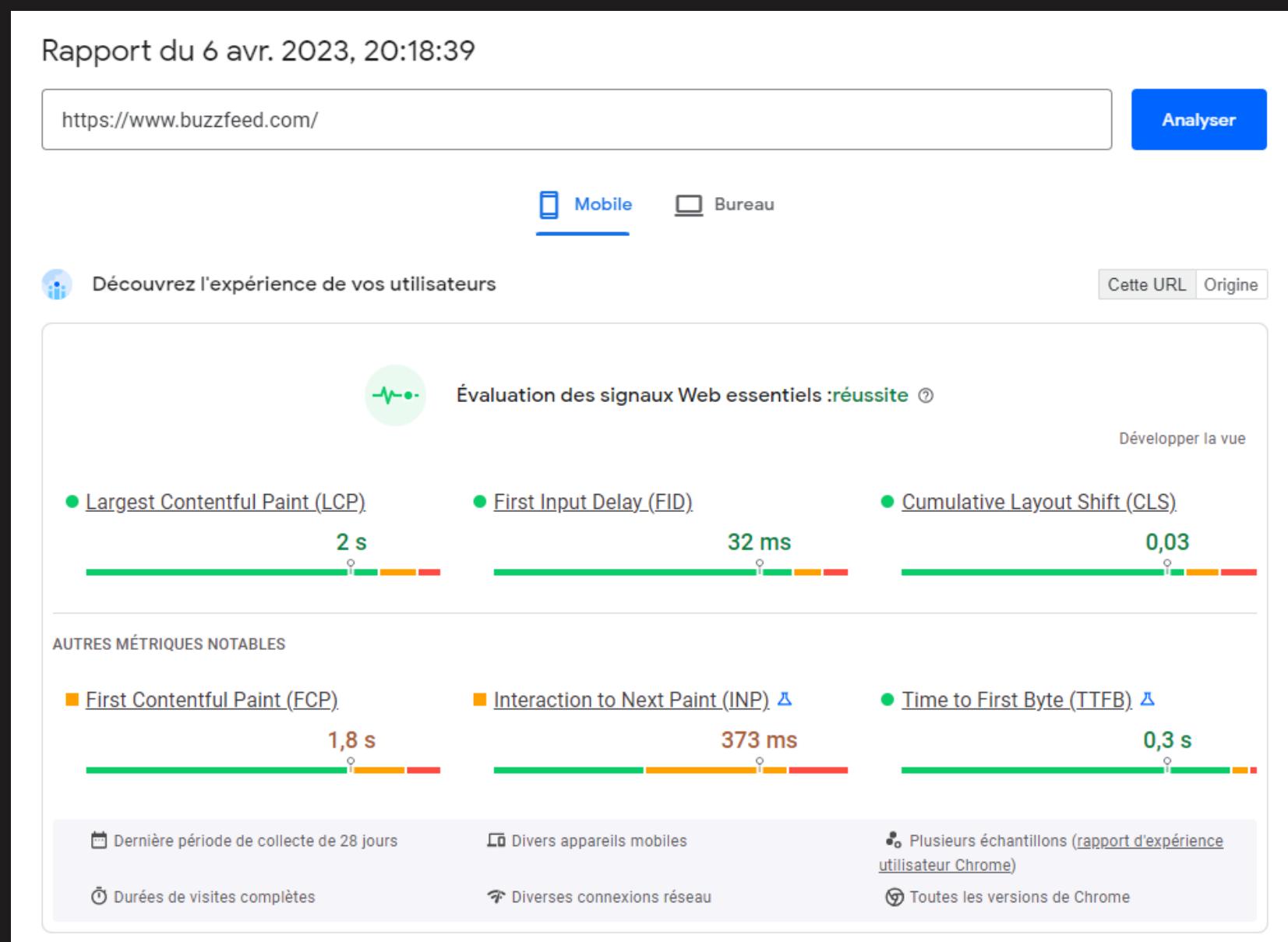


Signaux web



Landing page : Page d'accueil

Device : mobile



Depuis l'outil de Google Page Speed Insight, on peut mesurer la vitesse de chargement du site de BuzzFeed.

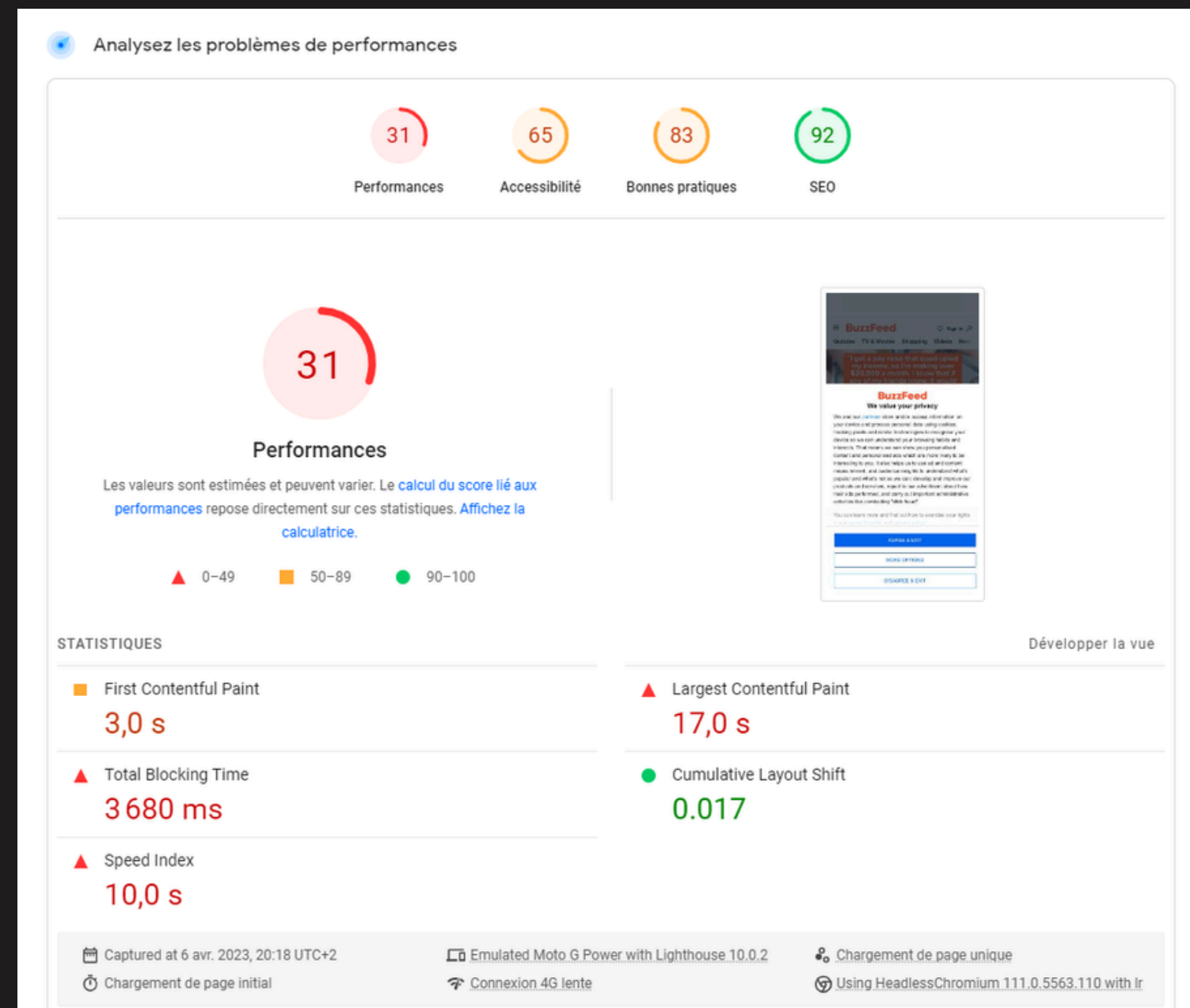
Le rapport Signaux Web essentiels indique les performances des pages en fonction de données d'utilisation. Globalement, la page d'accueil de BuzzFeed montre des résultats positifs sur mobile. Les utilisateurs sont moins susceptibles de quitter la page et/ou sortir du site. Le taux d'engagement pourrait donc être plus élevé.

Performances



Landing page : Page d'accueil

Device : mobile



Ici on peut voir que le score des performances sur mobile est de 31/100 et de 64 sur ordinateur. Ces score ne sont pas bon, surtout sur la version mobile. Afin de les améliorer, des optimisations sont nécessaires.

En suivant les recommandations de Google, il serait intéressant dans un premier temps **d'optimiser les images**. Une taille d'image trop importante peut limiter le temps de chargement du site. Il est donc utile de les compresser.

Ensuite, il faudrait **réduire le nombre de requêtes** : plus il y a de requêtes sur le site, plus il est lent.

Enfin, il est également possible **d'activer la mise en cache** : en activant la mise en cache, les utilisateurs pourront charger plus rapidement les pages qu'ils ont déjà visitées.

Diagnostic



Landing page : Home page

Device : mobile

The screenshot shows a diagnostic tool interface with two tabs: 'Mobile' (selected) and 'Bureau'. It displays a list of 'Opportunité' (Opportunities) and a 'DIAGNOSTIC' section.

Opportunité

Opportunité	Estimation des économies
▲ Réduisez les ressources JavaScript inutilisées	2,74s
■ Supprimez les modules en double dans les groupes JavaScript	0,45s
■ Différez le chargement des images hors écran	0,15s

Ces suggestions peuvent contribuer à charger votre page plus rapidement. En revanche, elles n'ont pas d'incidence directe sur le score lié aux performances.

DIAGNOSTIC

- ▲ Réduire l'impact du code tiers — Le thread principal a été bloqué par du code tiers pendant 1 650 ms
- ▲ La page n'utilise pas d'écouteurs d'événements passifs pour améliorer les performances de défilement
- ▲ Les éléments d'image ne possèdent pas de `width` ni de `height` explicites
- ▲ Réduisez le travail du thread principal — 12,7 s
- ▲ Évitez une taille excessive de DOM — 2 948 éléments
- ▲ Diffusez des éléments statiques grâce à des règles de cache efficaces — 56 ressources trouvées
- ▲ Réduisez le temps d'exécution de JavaScript — 9,2 s
- ▲ First Contentful Paint (3G) — 6261 ms
- Évitez de créer des chaînes de requêtes critiques — 2 chaînes trouvées
- Marques et mesures du temps utilisateur — 10 temps utilisateur
- Réduisez au maximum le nombre de requêtes et la taille des transferts — 114 requêtes • 2353 Kio
- Élément identifié comme "Largest Contentful Paint" — 1 élément trouvé
- Éviter les changements de mise en page importants — 5 éléments trouvés

Sur cette page, on peut voir le diagnostic intégral de la page avec 8 problèmes à résoudre.

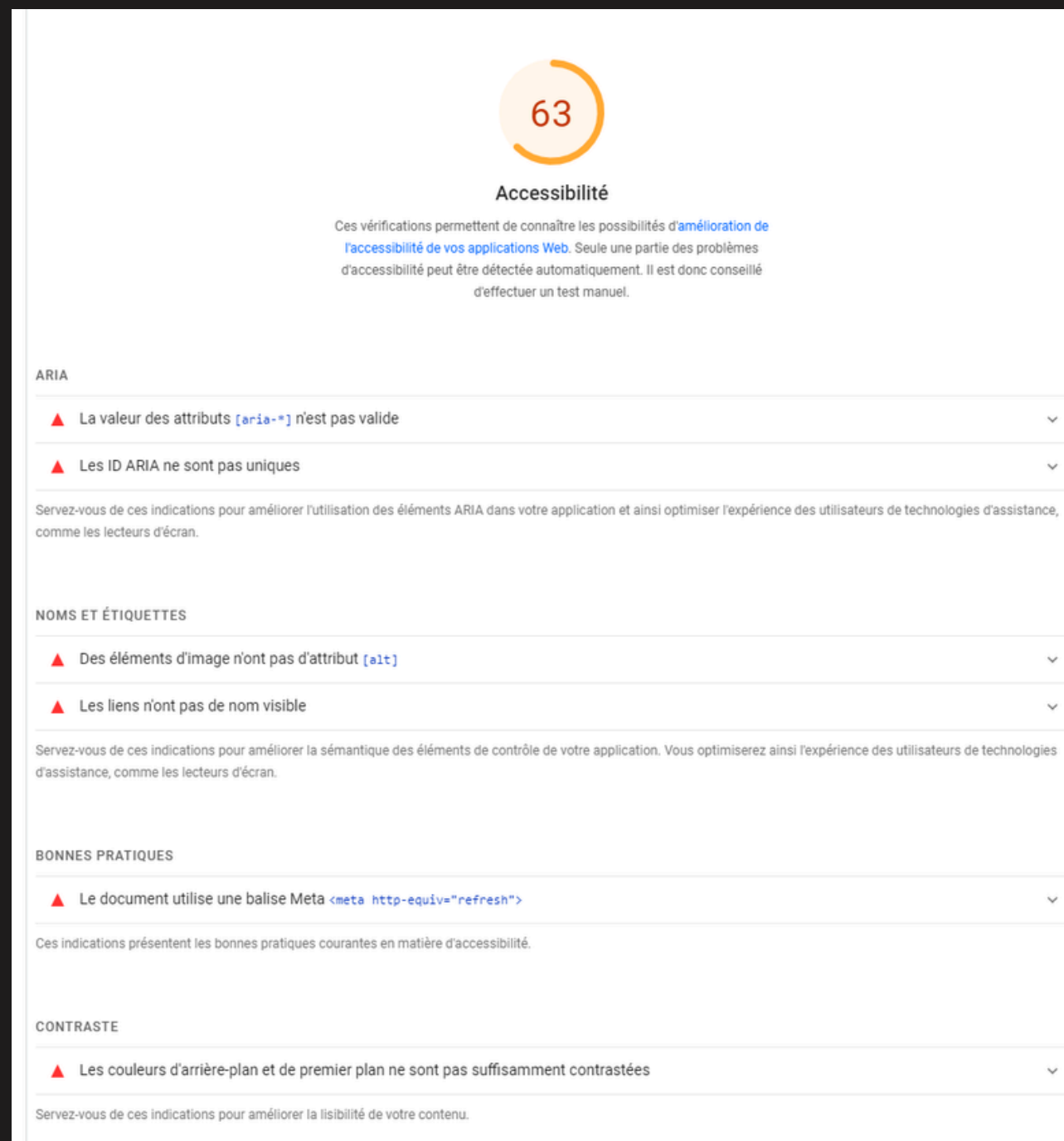
Par exemple, minimiser le code **CSS** et **JavaScript** qui permettrait de réduire la taille des fichiers et accélérer le temps de chargement.

Autre problème, les images n'ont pas de dimensions définies. Pour cela, il faudrait indiquer la largeur et la hauteur explicite sur les éléments d'image afin de réduire les décalages de mise en page et d'améliorer le **CSS**.

Performances

Landing page: Home page

Device : Desktop



63

Accessibilité

Ces vérifications permettent de connaître les possibilités d'amélioration de l'accessibilité de vos applications Web. Seule une partie des problèmes d'accessibilité peut être détectée automatiquement. Il est donc conseillé d'effectuer un test manuel.

ARIA

- ▲ La valeur des attributs [aria-*] n'est pas valide
- ▲ Les ID ARIA ne sont pas uniques

Servez-vous de ces indications pour améliorer l'utilisation des éléments ARIA dans votre application et ainsi optimiser l'expérience des utilisateurs de technologies d'assistance, comme les lecteurs d'écran.

NOMS ET ÉTIQUETTES

- ▲ Des éléments d'image n'ont pas d'attribut [alt]
- ▲ Les liens n'ont pas de nom visible

Servez-vous de ces indications pour améliorer la sémantique des éléments de contrôle de votre application. Vous optimiserez ainsi l'expérience des utilisateurs de technologies d'assistance, comme les lecteurs d'écran.

BONNES PRATIQUES

- ▲ Le document utilise une balise Meta <meta http-equiv="refresh">

Ces indications présentent les bonnes pratiques courantes en matière d'accessibilité.

CONTRASTE

- ▲ Les couleurs d'arrière-plan et de premier plan ne sont pas suffisamment contrastées

Servez-vous de ces indications pour améliorer la lisibilité de votre contenu.



Coté accessibilité du site, le score en desktop et en mobile n'est pas fameux non plus.

Voici quelques recommandation à mettre en place pour améliorer ce score :

- Utiliser des balises sémantiques : utiliser des balises HTML sémantiques appropriées pour structurer le contenu, tel que <header>, <nav>, <main>, <section>, <article>, <aside> et <footer>.
- Utiliser des descriptions alternatives aux images : les descriptions alternatives (ou "alt tags") aident les utilisateurs ayant des troubles de la vision à comprendre le contenu des images (également utile si l'image ne charge pas à cause de souci réseau/serveur).

Performances



Landing page: Home page

Device : Desktop

- Utiliser des couleurs contrastées : en s'assurant que le texte et les éléments du site ont suffisamment de contraste avec l'arrière-plan pour être facilement lisibles.
- Fournir des sous-titres et des transcriptions : en utilisant des vidéos ou des enregistrements audio, il faudrait fournir des sous-titres et des transcriptions pour les utilisateurs sourds ou malentendants.
- S'assurer que le clavier peut être utilisé pour naviguer : en s'assurant que le site peut être navigué à l'aide du clavier, car certains utilisateurs peuvent avoir des difficultés à utiliser une souris ou un trackpad.
- Tester le site avec des outils d'accessibilité : il existe de nombreux outils d'accessibilité gratuits en ligne pour tester un site et identifier les problèmes d'accessibilité.