

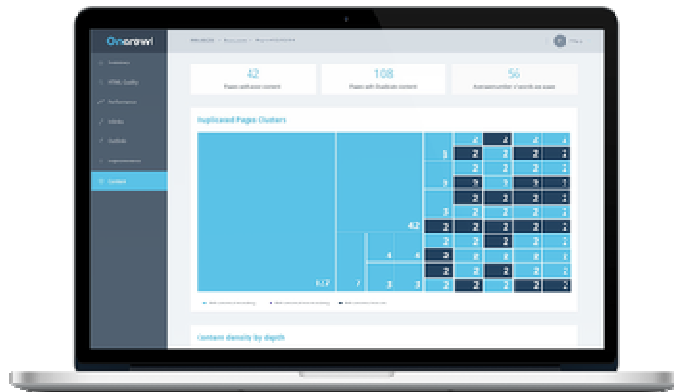
Mise à jour produit / Audit référencement naturel / Robot d'indexation SEO / Analyse de contenu

## Cogniteev lance sa nouvelle version d'OnCrawl, qui permet une analyse profonde du contenu pour les audits sémantiques SEO

*“Le contenu est l'un des plus importants facteurs d'évolution en termes de ranking, et une stratégie de contenu qualitatif devrait faire partie de n'importe quel plan marketing axé sur les moteurs de recherche.*

*Google pénalisant fortement les sites ayant un contenu pauvre ou dupliqué, il est essentiel de détecter, puis d'éliminer ou d'enrichir les pages qui paraissent similaires ou sans valeur ajoutée pour les moteurs de recherche. Malheureusement, il n'existe actuellement sur le marché aucun outil capable d'automatiser ce processus de détection et d'analyse pour des sites importants. Avec cette nouvelle version d'OnCrawl, nous avons souhaité mettre à disposition de ces derniers le premier explorateur sémantique SEO capable d'examiner le contenu de leurs pages en profondeur.”*

**François Goube, co-fondateur et CEO de Cogniteev**



**Bordeaux, le 8 juin 2015 - Trois mois après le lancement d'OnCrawl, Cogniteev présente aujourd'hui une version mise à jour de sa solution, dotée de nouvelles fonctionnalités sémantiques. OnCrawl est un outil puissant dédié à l'optimisation du référencement naturel de sites Web sur les moteurs de recherche. Il permet de détecter les erreurs de code HTML, de structure et de contenu, qui peuvent compromettre le positionnement du site sur les moteurs de recherche. Cette nouvelle version d'OnCrawl contient un module d'analyse de contenu qui détecte les pages similaires ou dupliquées, examine la densité de contenu et analyse la fréquence des principaux enchaînements de textes et mots-clés, plus connus sous le nom de *n-grams*.**

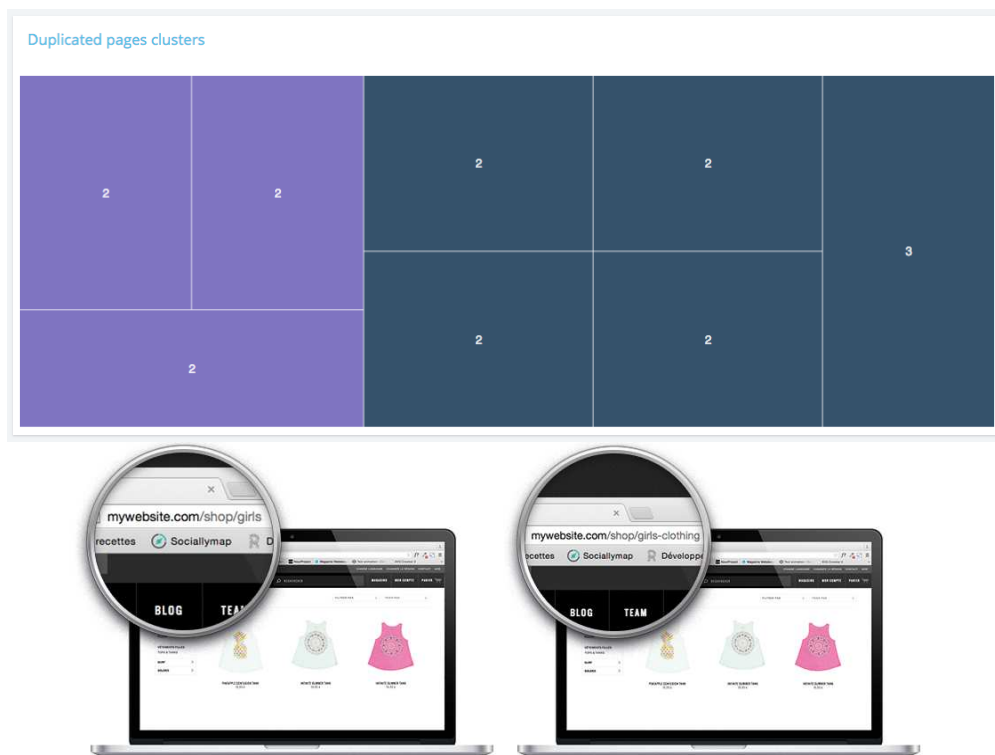
### L'expertise sémantique au service d'un contenu web unique et d'une stratégie de contenu équilibrée

OnCrawl explore l'intégralité du contenu d'un site, puis fournit une analyse complète tenant compte des éléments SEO suivants :

- Tags : titres, descriptions, tags H1, H2, etc...
- Performance : temps de chargement, poids...
- Architecture : structure interne, liens entrants et sortants, flux de popularité, etc...
- **Nouveau !** Contenu : duplications, nombre de mots, *n-grams*.

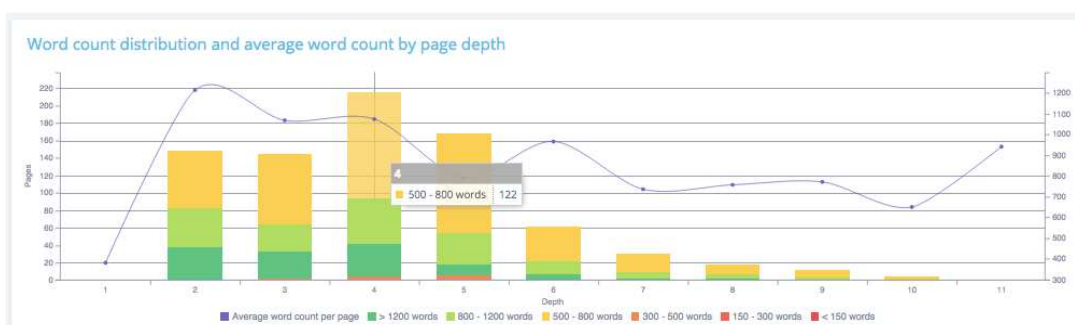
*“Nous avons décidé d'incorporer des éléments en NLP (Natural Language Processing) à notre outil OnCrawl, afin de mettre notre expertise sémantique au service des stratégies SEO et contenu de nos clients. Nos robots d'indexation sont aujourd'hui capables d'analyser des corpus entiers de texte issus de sites, puis d'en extraire des informations fiables relatives aux suites de textes similaires, aux fréquences et à la distribution de mots-clés.”* explique Tanguy Moal, co-fondateur et CTO de l'entreprise.

### 1. Détection de groupes de pages avec du contenu similaire ou dupliqué



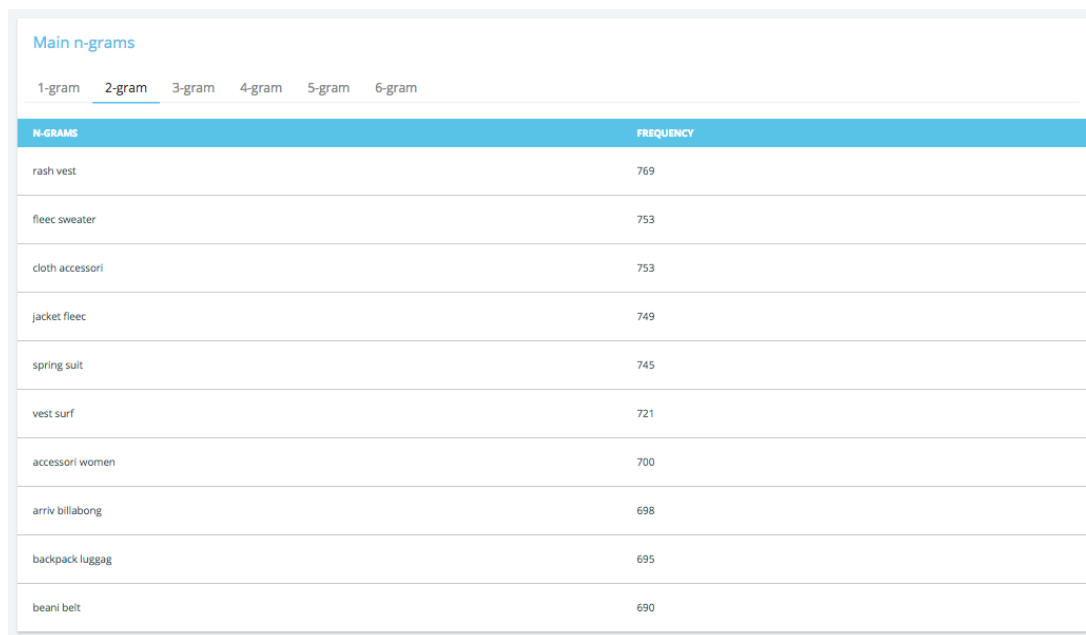
OnCrawl affiche les pages ayant du contenu dupliqué ou similaire dans des groupes. Les vendeurs e-commerce, par exemple, doivent souvent faire face à des problèmes de contenu dupliqué, dus à des options de filtres et de recherche, de structures catégorisées, de pagination, de versions imprimables, de site multi-langues, etc. Ces pages dupliquées ont un impact négatif sur le ranking sur les moteurs de recherche. Dorénavant, les utilisateurs peuvent facilement détecter et solutionner ces problèmes de duplication lorsqu'ils surviennent, par exemple en recalibrant leur système de management de contenu.

### 2. Analyse de la répartition du nombre de mots et moyenne du nombre de mots par page



OnCrawl offre dorénavant la possibilité d'analyser le nombre de mots par page, et donc par conséquent de détecter, puis d'éliminer ou d'enrichir les pages avec du faible contenu. En effet, concernant les articles, 800 mots ou plus sont recommandés, car les écrits de longue taille se positionnent mieux, apportant ainsi un meilleur taux de conversion et des prospects de qualité. A l'inverse, un contenu faible comportant moins de 150 mots est souvent exclu des résultats de recherche. Et lorsqu'il est pris en compte (en considérant d'autres critères comme le taux de rebond), il n'apporte pas une grande valeur ajoutée au visiteur.

### 3. Analyse des principaux *n-grams* (séquences de textes fréquents et mots-clés)



The screenshot shows a web interface titled 'Main n-grams'. At the top, there are tabs for '1-gram', '2-gram', '3-gram', '4-gram', '5-gram', and '6-gram', with '2-gram' selected. Below the tabs is a table with two columns: 'N-GRAMS' and 'FREQUENCY'. The table lists various n-grams and their corresponding frequencies.

N-GRAMS	FREQUENCY
rash vest	769
fleec sweater	753
cloth accessori	753
jacket fleec	749
spring suit	745
vest surf	721
accessori women	700
arriv billabong	698
backpack luggag	695
beani belt	690

OnCrawl est capable d'extraire les meilleurs mots-clés et séquences de textes trouvés sur un site. En d'autres termes, OnCrawl détecte pour quels mots-clés un site sera susceptible de bien se positionner. Cela permet à un manager SEO ou de contenu de trouver les failles, puis d'agir en conséquence en adaptant sa stratégie de contenu si nécessaire.

#### Réaliser un audit SEO en quelques étapes est à la portée de tous

OnCrawl est une solution basée sur le cloud, très simple à utiliser : il suffit d'indiquer une URL, de lancer l'audit, d'analyser et d'optimiser. Avec OnCrawl, Cogniteev souhaite rendre les audits SEO sur serveur accessibles aux petites et moyennes entreprises, à un prix modéré.

Jusqu'à maintenant, cette méthode était uniquement réservée à quelques entreprises de grande taille, disposant du budget et des ressources nécessaires pour développer leurs propres outils.

La solution OnCrawl est disponible à partir de 9,90 euros par mois, ce qui, comparé à d'autres solutions est une très bonne option !

#### A propos de Cogniteev :

Cogniteev est un éditeur de solutions Big Data. Tous les produits de Cogniteev partagent une brique technologique commune basée sur la recherche sémantique, le crawling et le Big Data, conçue pour extraire et agréger des données issues du web, et pour les structurer de manière intuitive. L'objectif de l'équipe rassemblée autour des co-fondateurs, Francois Goube et Tanguy Moal, est de simplifier, pour tous, l'accès aux informations dont ils ont besoin, et de rendre ses informations facilement utilisables. Cogniteev compte aujourd'hui 11 salariés et a levé récemment 500.000 euros auprès du fond de capital risque ACI et de Business Angels.

**Pour plus d'informations :** <http://www.cogniteev.com> ou <http://www.oncrawl.com>

#### CONTACT PRESSE :

Laurence Delval  
Email: [laurence.delvalrp@gmail.com](mailto:laurence.delvalrp@gmail.com)  
Tel. : +33 (0)6 82 52 98 47