

Les guidelines des Quality Raters de Google : qu'y apprend-on ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

En septembre dernier, le SEO français Paul Sanches a déniché sur le Web une version relativement récente des recommandations de Google destinées à ses "Quality Raters". Que nous apprennent-elles sur le fonctionnement de Google et sur les critères que le moteur de recherche juge pertinents pour positionner une page et qui sont ces "Quality Raters" ? Début de réponse avec l'analyse de ces "Guidelines"...

Avant-propos :

Vous trouverez à l'adresse suivante le fichier complet concerné par cet article, directement sur le site Internet SeoBlackOut : <http://www.seoblackout.com/2012/09/05/google-quality-raters-guidelines-2012/>.

Cette version date de Juin 2012, et elle ajoute quelques éléments supplémentaires par rapport à la version précédente de 2011.

Avant toute chose, il nous faut expliquer deux éléments.

Qu'est ce qu'un Quality Rater ?

Un Quality Raters est une personne employée par Google pour juger de la qualité d'un résultat du moteur de recherche.

Il utilise différents critères pour attribuer une note la plus objective possible à une page de manière générale, ou à un contenu vis-à-vis d'une requête précise. Les résultats de ces notes sont ensuite transmises aux développeurs de Google qui vont les prendre en compte pour affiner ou non l'algorithme global du moteur de recherche.

En soi, un Quality Rater ne peut donc pas influencer directement sur les résultats des moteurs de recherche. Une convergence de notation de certains types de contenus pourra cependant influencer les résultats en poussant les développeurs à modifier, améliorer et optimiser l'algorithme de classement. Pour rappel, Google le modifie entre 300 et 500 par an, avec un impact plus ou moins modéré sur l'ensemble des résultats.

Sources :

- <http://www.abondance.com/actualites/20120502-11372-mat-cutts-et-le-fonctionnement-des-quality-raters.html>
- <http://www.vincentabry.com/un-quality-rater-google-ne-peut-pas-descendre-un-site-dans-les-resultats-de-recherche-14506>

A quoi servent les Guidelines de Google ?

Pour pouvoir noter un contenu, les Quality Raters utilisent un guide transmis par le moteur de recherche pour pouvoir noter un contenu avec le plus de précisions possible : les "Guidelines".

Leur but est de donner des directives précises sur l'évaluation des contenus, afin de rendre le plus objectif possible chaque notation et pour réduire les écarts de subjectivité entre chaque Quality Rater.

Au delà de cela, les Guidelines de Google nous permettent de mieux comprendre les critères déterminants aux yeux du moteur dans le classement d'un site, critères qui sont déjà pris en

compte de manière plus ou moins importante dans l'algorithme de classement (ou qui seront amené à l'être).

Que nous apprennent les guidelines des Quality Raters ?

Nous allons donc passer en revue différents points déterminants abordés par ce guide. Le but est vraiment de pouvoir soi-même évaluer son contenu et de mieux comprendre pourquoi certaines pages peuvent être plus ou moins bien classées.

Il existe trois types de notation :

- La notation d'un résultat par rapport à une requête.
- La notation d'un résultat par rapport à une requête et à une localisation géographique.
- La notation de la qualité générale d'une page.

Nous allons donc voir point par point ces 3 éléments

La notation d'un résultat par rapport à une requête.

L'analyse du besoin

Pour comprendre comment Google positionne un contenu, il faut tout d'abord prendre en compte le fait que pour une même requête, les résultats peuvent différer d'un ordinateur à un autre.

Outre le fait que l'historique de navigation et que la connexion à un compte Google peuvent influencer les résultats, d'autres éléments sont pris en compte, y compris dans le système de notation des contenus des Quality Raters. Voici comment Google va décomposer la requête d'une internaute :

[Requête de l'internaute], langue (Localisation géographique).

Voici différents exemples de requêtes

- [Restaurant gastronomique], French (FR) : l'internaute recherche un restaurant gastronomique français, avec des contenus en Français.
- [Restaurant gastronomique], English (FR) : l'internaute recherche un restaurant gastronomique Français, avec dans des contenus en Anglais.
- [Car], English (GB) : un internaute anglais veut acheter des voiture, avec des contenus en Anglais.
- [voiture France], English (FR) : un internaute Français veut acheter des voiture en France, avec des contenus en Anglais.

Ces trois critères permettent de déterminer l'objectif de l'utilisateur (appelé *User Intent* dans les Guidelines). Le moteur de recherche analyse donc de base la requête de l'internaute, sa localisation géographique et la langue utilisée par celui-ci.

L'échelle de notation

Pour chaque requête, les Quality Raters doivent attribuer une note au contenu analysé. Cette note peut prendre 6 formes différentes :

1. *Vital* : c'est le résultat prioritaire pour la requête. Il sera presque à chaque fois positionné en premier.
2. *Useful* : la page est très utile par rapport à la requête.
3. *Relevant* : le contenu est pertinent.
4. *Slightly Relevant* : le résultat est potentiellement pertinent.
5. *Off-Topic or Useless* : le contenu ne sert à rien ou est hors-sujet.
6. *Unratable* : le Quality Rater n'a pas été capable de noter ce résultat.

En plus de ces critères, Google y ajoute d'autres critères complémentaires :

- *Not Spam* : le contenu n'est pas considéré comme du spam.

- *Maybe Spam* : le contenu est à la limite du spam.
- *Spam* : spam, donc non pertinent.
- *Porn* : contenu pour adulte.
- *Malicious* : contenu nuisible (virus, trojan, phishing, ...).

En tant que référenceur, votre but est donc de faire en sorte que vos contenus soient positionnés sur les meilleurs critères de cette grille.

L'objectif recherché

Pour pouvoir positionner un contenu selon cette liste de critères, il faut tout d'abord que le Quality Rater puisse déterminer le but réel de l'internaute. Pour cela, il doit pouvoir attribuer une interprétation à une requête. Là aussi, plusieurs classifications existent :

- *Dominant interpretation* : c'est l'objectif visé par la plupart des internautes ayant tapé ce résultat. Par exemple, la requête [Wikipédia] a pour interprétation dominante de se rendre sur le site de Wikipédia.
- *Common interpretations* : dans le cas où une interprétation dominante n'existe pas, il s'agit de celles qui sont les plus probables. Par exemple, la requête [wiki] peut avoir comme interprétations plausibles la recherche d'un outil pour créer un Wiki, ou la recherche d'un Wiki généraliste en ligne.
- *Minor interpretations* : il s'agit des interprétations peu fréquentes ou peu probables pour une requête donnée. Pour reprendre le 1er exemple, une interprétation mineure de [Wikipédia] est de consulter l'historique du service Wikipédia.

Rappelez-vous : même s'ils ne peuvent influencer directement sur les résultats, il s'agit des critères utilisés par le moteur de recherche pour adapter les résultats selon chaque requête.

Pour aider les Quality Raters à déterminer plus précisément l'interprétation, Google les classe selon trois type d'action : le *Do-Know-Go*.

- *Do* : l'utilisateur veut effectuer une action, comme par exemple acheter un produit, envoyer un email ou télécharger un logiciel.
- *Know* : l'utilisateur veut s'informer sur un élément.
- *Go* : l'utilisateur veut se rendre sur un site ou une page précise.

Certaines requêtes ne peuvent être classées qu'avec un seul de ces critères, mais d'autres peuvent être les trois à la fois. Pour reprendre un exemple concret des Guidelines avec la requête [Nikon digital cameras] :

- C'est une action (*Do*) pour acheter un Nikon.
- C'est une recherche d'informations (*Know*) pour connaître la marque Nikon.
- C'est une destination (*Go*) pour aller sur le site des caméras de Nikon.

La langue utilisée

Enfin, la langue du contenu aura également de l'importance. Une page en espagnol peut être pertinente pour l'utilisateur espagnol sans l'être pour l'internaute allemand. Une fois de plus, on peut le définir selon plusieurs critères :

- *Task Language* : le contenu est écrit dans la même langue que l'utilisateur.
- *Acceptable Language* : le contenu est écrit dans une langue acceptable pour l'utilisateur (par exemple le Catalan pour certaines parties d'Espagne).
- *English* : le contenu est en anglais.
- *Foreign Language* : le contenu utilise une autre langue que celui de l'utilisateur (et différente de l'anglais).
- *None of the Above* : il est impossible de déterminer la langue du contenu (erreurs, page indisponible, multiples langues, ...).

Retour sur l'échelle de notation

Il s'agit sans doute de la partie la plus important du guide, et celle d'ailleurs qui a subi le plus de modifications depuis la version de 2011 : l'échelle de notation. En utilisant la langue du

contenu, le but de l'internaute et sa localisation géographique, le Quality Rater peut attribuer une note objective à un contenu.

Voici en détail les différents classements :

Vital

Un contenu vital peut l'être pour deux raisons :

- Il s'agit de l'interprétation dominante de la requête de l'internaute, et le contenu correspond à l'objectif de l'utilisateur.
- Si la requête est une entité nommée (un lieu, une personne, une société...), le résultat ne sera vital que lorsqu'il s'agit du site ou de la page officielle.

Pour la requête [Abondance], le site <http://www.abondance.com> sera classé comme Vital.

Pour la requête [Olivier Andrieu], le résultat Vital sera la page de présentation suivante <http://docs.abondance.com/oa.html>, tandis que la page d'accueil sera classée *Useful*.

Le passage suivant est nouveau par rapport à l'édition 2011 des Guidelines : l'interprétation d'une requête "vital" dans le cadre d'une recherche sur une personne réelle. Pour ce type de requête, par exemple [Daniel Roch], l'attribution du caractère "vital" du contenu va dépendre de plusieurs points :

- la requête a une interprétation dominante qui est de trouver des informations sur la personne concernée ou de se rendre sur son site.
- la requête concerne une personne précise, donc sans homonymes possibles.
- la requête doit aboutir sur le site ou la page officielle de la personne concernée. Sans cela, il n'existera pas d'attribution "Vital" pour cette requête.

Ce que l'on apprend : Seuls les sites reconnus comme officiels peuvent s'assurer d'obtenir la première place sur une requête. Ils seront en conséquence beaucoup plus durs à "déloger" que tout autre type de contenu.

Cela peut aller plus loin. Par exemple, si vous créez une page avec un contenu très pertinent pour une célébrité (sans homonymes), que ce soit un site dédié ou une page Fan sur Facebook, il ne s'agira en aucun cas d'une réelle page officielle, et le résultat sera donc considéré au mieux comme *Useful*.

Pour ce qui est des homonymes, les Guidelines de Google donnent un exemple très pertinent qui explique pourquoi ils forcent les Quality Raters et le moteur de recherche à ne pas pouvoir classer un contenu comme "Vital". La requête [ADA] peut correspondre par exemple à :

- *American with Disabilities Act*
- *American Dental Association*
- *American Diabetes Association*

Comme il est impossible de déterminer le but réel de l'Internaute, chaque page officielle ne pourra être classée que "*Useful*".

Si plusieurs sites ou pages officiels appartiennent à la personne concernée par la requête, il est possible d'avoir plusieurs résultats classés comme "*Vital*" puisque tous concernent la bonne personne, avec un contenu officiel à chaque fois.

Ce que l'on apprend : Avoir un réseau de sites pour une marque, une société ou une personne permet de dominer plus facilement toute la première page des résultats.

Le concept même de page vitale va elle aussi dépendre de la langue du contenu et de la langue de l'utilisateur. Nous pouvons donc avoir ces valeurs :

- *Appropriate Vital* : c'est la page officielle dans la langue adaptée à l'internaute.
- *International Vital* : il peut s'agir d'une page officielle en anglais (donc adaptée pour l'international), ou d'une page permettant le choix de la langue : "*Choose Your Language*" ou "*Choose Your Location*".

- *Other Vital* : c'est une page vitale, mais dont la langue n'est pas forcément adaptée à l'internaute ayant fait la requête.

Là aussi, nous allons reprendre un exemple issu des Guidelines qui explique bien ce concept :

- [ikea], German (DE) : Appropriate Vital avec <http://www.ikea.com/de/>
- [ikea], German (DE) : Other Vital avec <http://www.ikea.com/fr/>
- [Ikea], English (US) : Appropriate Vital avec <http://www.ikea.com/us/en/>
- [Ikea], Chinese (CN) : International Vital avec <http://www.ikea.com/>

Useful

Comme expliqué précédemment, la notation "Useful" désigne une page qui répond dans la majorité des cas aux besoins de l'internaute. En d'autres termes, on peut trouver plusieurs contenus "Useful" pour une même requête. Il s'agit bien souvent de la meilleure classification possible.

Relevant

L'attribut "Relevant" désigne quant à lui des pages pertinentes pour certains utilisateurs, c'est-à-dire en prenant en compte un des buts potentiels d'une requête qui pourrait être interprétée de plusieurs manières différentes.

Google précise que les pages "Relevant" peuvent l'être à cause de plusieurs éléments différents (qui peuvent donc influencer l'algorithme actuel du moteur de recherche) :

- la date de publication et la fraîcheur du contenu.
- une source de contenu plus ou moins crédible (comprenez par là plus ou moins "récente" ou "populaire").
- la page ne couvre qu'un aspect de la requête.

Ce que l'on apprend : Gardez en tête le fait d'avoir un contenu qui couvre tous les aspects de vos requêtes cibles, et si possible avec une publication ou mise à jour récente de ces contenus.

Slightly Relevant

Cette fois-ci, il s'agit de pages peu utiles pour l'utilisateur. En d'autres termes, il s'agit de contenus annexes, peu précis, trop courts voire presque "hors-sujets".

Fait intéressant, Google indique également qu'un contenu mobile sera toujours considéré comme "Slightly Relevant" s'il existe une version normale du même contenu, preuve que Google considère à part les sites mobiles.

D'ailleurs, souvenez-vous que Google a récemment conseillé d'utiliser la technologie *Responsive* pour optimiser les visites pour les mobiles tout en gardant optimal son référencement naturel. Source :

<http://googlewebmastercentral.blogspot.ch/2012/06/recommendations-for-building-smartphone.html>

Ce que l'on apprend : Mettez-vous si possible au *Responsive Design* pour avoir une version mobile de votre site Internet. Dans le cas où ce n'est pas possible, suivez les guides de Google à partir du lien ci-dessus pour que le code source soit parfaitement compris par les robots du moteur de recherche.

Les interprétations

Notation et localisation

Un contenu doit parfois être jugé par rapport à la localisation de l'utilisateur. Si un contenu n'est pas pertinent par rapport au lieu de la requête (que ce soit un lieu inclus directement dans la requête, comme [restaurant Nantes] ou dans la position géographique de l'utilisateur),

Google conseille aux Quality Raters de considérer le contenu comme inutile : "*Off-Topic or Useless*".

Cela est donc en accord avec les résultats locaux de Google que l'on peut régulièrement apercevoir, et qui s'affichent pour une même requête pour certains utilisateurs et pas pour d'autres.

Ce que l'on apprend : La localisation de l'internaute ou la localisation précise d'une requête à un impact décisif dans les résultats du moteur de recherche.

Interprétations multiples

Si une requête peut être interprétée de plusieurs façons, voici comment Google gère son classement :

- si un contenu correspond à une interprétation mineure, la page doit avoir une note faible (*slightly relevant*).
- si un contenu correspond à une interprétation commune, la page doit être bien notée (*Relevant*).
- Si un contenu a une interprétation majeur, la page peut être notée *Useful*.
- Seules les pages une interprétation dominante peuvent être classées *Vital*.

Les fautes d'orthographe

Cette partie est très intéressante. En tant que référenceur, nous cherchons parfois à nous positionner sur une requête contenant une faute d'orthographe pour attirer du trafic supplémentaire.

Google indique aux Quality Raters qu'ils doivent juger un contenu en fonction du but recherché par l'internaute, et pas forcément en fonction de la manière dont il a décrit sa requête. En d'autres termes, le Quality Rater ne doit pas chercher à mieux noter la page qui utilise le mot clé avec la faute d'orthographe, mais bien celle qui correspond le mieux au besoin de l'utilisateur.

Ce que l'on apprend : Vouloir se positionner sur une requête mal rédigée ou mal orthographiée peut avoir un intérêt faible. Par exemple, la requête [fabook] renvoie bien vers le site officiel de Facebook, et non pas vers le domaine existant fabook.fr.

La date

Dans le cadre de contenus portant sur une thématique qui se renouvelle (congrès, compétitions, événements sportifs, salons, ...), Google considérera toujours les contenus récents comme plus pertinents, en partant du principe qu'un utilisateur va avoir tendance à rechercher en premier lieu des informations sur le prochain événement.

Attention cependant, si le prochain événement est situé dans un laps de temps très longs, Google considère que c'est la page de l'événement précédent qui reste la plus pertinente.

Ce que l'on apprend : renouvelez votre contenu, mettez-le à jour et créez de nouveaux contenus pour rester pertinent afin que la page positionnée apporte le contenu le plus récent et pertinent pour l'internaute (principe des QDF).

La pornographie

Il s'agit également d'une des parties des Guidelines qui a été mise à jour.

De base, tout contenu érotique ou pornographique sera balisé comme adulte ("*Porn*"). Ce n'est pas en soi une pénalité mais un classement à part de ce type de contenus pour ne pouvoir le proposer que dans certains cas (filtrage adulte ou non par exemple).

Dans le cas de requêtes explicites, ce classement est facile à appliquer, tout comme dans le cas de requêtes non explicites. Par exemple, avec la requête [voiture], l'utilisateur ne s'attend pas à trouver des photos de voitures avec des filles dénudées en train de les laver.

Mais certains requêtes peuvent parfois avoir une double interprétation. Par exemple, les requêtes suivantes :

- [seins]
- [fessée]
- [gay]
- ...

Dans ces cas de figure, le Quality Rater doit interpréter la requête de l'internaute pour pouvoir attribuer ou non l'attribut "Porn" au contenu.

Ce que l'on apprend : si un Quality Rater a du mal à correctement comprendre certaines requêtes, le moteur de recherche aura au minimum les mêmes difficultés. Le contenu que vous devez créer doit donc être explicite, et couvrir si possible tout l'aspect d'un même sujet. Si vous avez un site adulte, indiquez-le clairement comme tel pour ne laisser aucun doute et ne pas être considéré comme un contenu de spam.

La notation d'une URL en fonction d'une localisation géographique.

Nous venons de passer en revue les principaux points de la 1ère partie des Guidelines de Google. La seconde traite spécifiquement des requêtes ayant une localisation géographique précise, ou pour laquelle la localisation de l'utilisateur a une importance cruciale.

Une localisation explicite

Si l'utilisateur indique dans sa requête un lieu précis, la question ne se pose pas et le Quality Rater pourra facilement noter un contenu en fonction de celle-ci, par exemple avec la requête [concessionnaire Evry]. C'est ce qu'on appelle une "*Explicit Location*".

A l'inverse, les autres requêtes peuvent parfois avoir un objectif local, sans pour autant l'indiquer. Nous avons donc deux autres types de résultats :

- *Local Intent* : l'utilisateur cherche des résultats proches de lui, sans l'indiquer.
- *Non-local intent* : l'utilisateur ne cherche pas de résultats proches de lui.

Quelques exemples permettent de mieux comprendre :

- [Station essence Nantes] : *Explicit Location*, car l'internaute demande explicitement des informations sur les stations essence de Nantes.
- [Station essence] : *Local Intent*, car l'utilisateur cherche généralement une station essence proche de lui.
- [Législation station essence] : *Non-Local Intent*, car sa demande est généraliste.

Ce que l'on apprend : Se positionner sur une requête trop générique et potentiellement locale peut être peu pertinent. Non seulement le taux de transformation sera faible, mais la personnalisation des résultats fera apparaître beaucoup moins souvent le contenu.

Google considère qu'une localisation explicite sera toujours importante dans le classement, tandis qu'un "*Local Intent*" peut parfois avoir de l'importance : ce n'est pas systématique. Voici un exemple concret du guide :

- [pizza] : *Local Intent*, car l'utilisateur veut souvent manger des pizzas à proximité de chez lui.
- [hôtel] : *Non-Local Intent*, car l'utilisateur recherche généralement des hôtels en dehors de sa zone géographique.

La notation de la qualité d'une page.

Une notation du contenu

Au delà de la notation d'un contenu vis-à-vis d'une requête, il est également demandé aux Quality Raters de juger la qualité d'une page web de manière générale.

Il doit, comme le moteur de recherche, identifier et différencier le contenu principal d'une page et son contenu secondaire (footer, sidebar, header, ...). Pour chaque section, il doit attribuer une note.

Pour le contenu principal :

- *lowest* : très mauvais ;
- *low* : mauvais ;
- *medium* : moyen ;
- *high* : bon ;
- *highest* : très bon.

Pour le contenu secondaire :

- *no supplementary content* : pas de contenu secondaire ;
- *distracting/not helpful* : distrayant mais non pertinent ;
- *so-so* : moyen ;
- *helpful* : utile ;
- *very helpful* : très utile.

Pour le design et la mise en page :

- *misleading or deceptive* : décevant ou trompeur ;
- *poor* : mauvais ;
- *so-so* : moyen ;
- *good* : bon ;
- *excellent* : excellent.

En soi, on ne sait pas comment l'algorithme peut juger et analyser de tels critères, notamment pour la partie destinée à la mise en page. On peut cependant supposer que Google doit essayer de prendre en compte ces différents facteurs dans le positionnement d'un contenu ou d'un site dans son ensemble.

Ce que l'on apprend : utilisez ce schéma pour noter vous-même vos contenus. Vous pourrez ainsi améliorer vos pages pour vos utilisateurs (pour le taux de transformation), et par la même occasion pour les moteurs de recherche.

Les autres facteurs de notation

En complément, il est demandé de juger si le contenu analysé est en adéquation avec le reste du site. On pourrait par exemple avoir un excellent contenu dont la thématique est complètement hors-sujet par rapport à l'intégralité du site.

De même, voici plusieurs questions auxquelles doivent répondre les Quality Raters :

- Qui est le responsable légal du site et qui est le responsable du contenu analysé ? On peut en effet avoir un propriétaire pour le site mais plusieurs auteurs pour les contenus.
- Le site dispose t-il de suffisamment d'informations de contact ?
- Quelle est la réputation du site ?
- Est-ce que la page d'accueil du site est maintenue à jour ?

Ce que l'on apprend : un site Internet pertinent est reconnu dans son milieu, mis à jour régulièrement et l'on peut clairement identifier le créateur de chaque contenu. Un site qui répond à ces critères est donc généralement jugé plus pertinent par Google.

La lutte contre le spam

Le passage expliquant la lutte contre le spam et contre les contenus uniquement destinés à la publicité est, lui aussi, intéressant.

Google explique quelles sont ces techniques et comment les débusquer sur un site. Le moteur de recherche conseille ainsi d'utiliser les méthodes suivantes pour analyser le spam potentiel d'une page :

- Faire Ctrl-A pour sélectionner tout le texte et voir le texte caché ;
- Désactiver le CSS ;
- Désactiver le JavaScript ;
- Regarder le code source ;
- Regarder dans les zones où le regard n'irait pas naturellement ;
- Utiliser le Whois d'une page.

Ce que l'on apprend : Google ne parle à aucun moment de techniques de spam complexes, comme le cloaking ou les redirections 301. Reste à savoir si ces techniques peuvent être analysées par l'algorithme sans l'être par les Quality Raters.

Conclusion

Les Guidelines destinées aux Quality Raters sont très intéressantes pour comprendre les objectifs de Google quant à la qualité des résultats des moteurs de recherche. On y apprendra qu'un bon contenu :

- couvre tous les aspects de la requête de l'internaute ;
- est précis, pertinent et exhaustif ;
- est récent ;
- est en adéquation avec le reste du site ;
- est facilement associé à un auteur ;
- se trouve sur un site populaire, si possible sur l'un des sites officiels liés à la requête ;
- possède des informations de localisation géographique relatives à cette même requête ;
- dispose d'un design "convenable".

Les guidelines peuvent donc servir le graphiste, le rédacteur de contenus, le référenceur ou encore le développeur à analyser et noter leurs propres sites pour optimiser le contenu afin qu'il soit reconnu par l'utilisateur et le moteur de recherche.

Malheureusement, il reste à déterminer dans quelle mesure chacun de ces critères est implanté dans l'algorithme actuel du moteur de recherche, ou dans ses versions futures...

Daniel Roch, *Consultant WordPress, Référencement et Webmarketing chez SeoMix*
(<http://www.seomix.fr/>)