

La gestion des pages produits non disponibles ou épuisés sur une boutique e-commerce



Par Philippe Yonnet

| | | |
|------------------|-----------|----------------------|
| Domaine : | Recherche | Référencement |
| Niveau : | Pour tous | Avancé |

L'une des problématiques récurrentes sur un site e-commerce consiste à traiter les pages de produits qui ne sont plus disponibles, de façon temporaire ou définitive. Il existe plusieurs types de solutions disponibles pour gérer ce type de cas, selon la situation et le fait qu'on privilégie l'expérience utilisateur ou le SEO. Voici une vaste revue d'effectif des possibilités et des actions à mettre en place pour prévoir un maximum de variantes, plus ou moins complexes, qui peuvent survenir sur votre boutique...

Savoir comment gérer de manière optimale pour le référencement les produits devenus indisponibles ou qui sont épuisés est une préoccupation récurrente pour les e-marchands. Aux contraintes SEO se rajoutent celles concernant l'expérience utilisateur qui semblent (à tort) inconciliables.

En outre, les « bonnes pratiques » que l'on trouve sur Internet laissent souvent perplexes les webmasters car elles peuvent être contradictoires.

Dans cet article, nous allons commencer par mettre à plat les problèmes réels qu'il faut chercher à traiter. Puis nous allons analyser les différentes techniques à mettre en place pour aider les moteurs à comprendre comment vous voulez gérer votre catalogue. Et nous montrerons qu'il existe différentes approches possibles selon ce que vous placerez en tête de vos priorités : facilité d'implémentation, optimisation SEO ou l'expérience client.

Pourquoi faut-il un « traitement spécial » pour les pages produits non disponibles ou épuisées ?

Il existe trois cas de figure différents à prendre en compte :

- Le produit a disparu définitivement du catalogue ;
- Le produit est épuisé mais est en cours de réapprovisionnement ;
- Le produit est épuisé, et sa réapparition dans le catalogue est peu probable.

Le produit a disparu définitivement du catalogue

Un produit peut disparaître du catalogue pour différentes raisons : un changement de catalogue dû à la saisonnalité, le remplacement d'une référence, le déréférencement d'un produit ou de son fabricant, etc.

Dans tous les cas, la « disparition » d'un produit du catalogue se traduit en principe par :

- La disparition de la page produit correspondante sur le site web ;
- Un item de moins dans les pages de listing de produits.

La page produit n'existant plus, l'appel de son URL produira (par défaut) un code d'erreur 404.

Mais beaucoup de webmasters de sites e-commerce trouvent que renvoyer une « 404 » est une mauvaise solution. Ils défendent l'idée que délivrer dans ce cas une page d'usage 404 produit un « rebond systématique » : le « client » potentiel déçu ne poursuit pas sa visite et ira voir les sites de concurrents.

La perte des backlinks en cas de 404

Un référenceur fera remarquer que si la page recevait des backlinks (des liens provenant d'autres sites), la page 404 empêche ces backlinks de transmettre du jus de lien au site : ces backlinks sont « perdus ».

Le produit est épuisé mais est en cours de réapprovisionnement

Si le produit est épuisé, mais est susceptible d'être à nouveau disponible après une reconstitution des stocks, le problème devient un peu plus complexe encore :

- Faut-il sortir la page produit du catalogue ? Oui mais comment techniquement réintégrer la page dans quelques jours quand le produit sera réapprovisionné ? Pourra-t-on garder la même URL ? Car si ce n'est pas

possible, les URL du site vont changer sans arrêt ;

- Faut-il au contraire garder la page ? Mais dans ce cas, est-ce que l'on ne déçoit pas les clients potentiels ?

Le produit est épuisé, et sa réapparition dans le catalogue est peu probable

Ce cas semble être similaire au premier, mais il se gère différemment sur le plan technique. Car autant dans le premier cas, la « sortie » du catalogue résulte d'une décision du propriétaire, autant ici la décision de « sortir » le produit ou non est prise par l'applicatif web.

Quelques rappels techniques pour commencer...

Tout d'abord quel code renvoyer : 404 ou 410 ?

Rappelons qu'il existe deux codes http utilisables dans le cas de pages qui ont disparu de l'arborescence d'un site (ou que le serveur web ne sait pas renvoyer) :

- Le code 404 indique seulement que la « ressource est introuvable » ;
- Le code 410 que la ressource n'est plus disponible.

Pendant longtemps, Google a traité les URL renvoyant un code 410 comme les pages 404. Mais récemment, Google a décidé de traiter les deux codes d'une manière subtilement différente. Ce point a été confirmé par Matt Cutts dans une vidéo datant d'avril 2014.

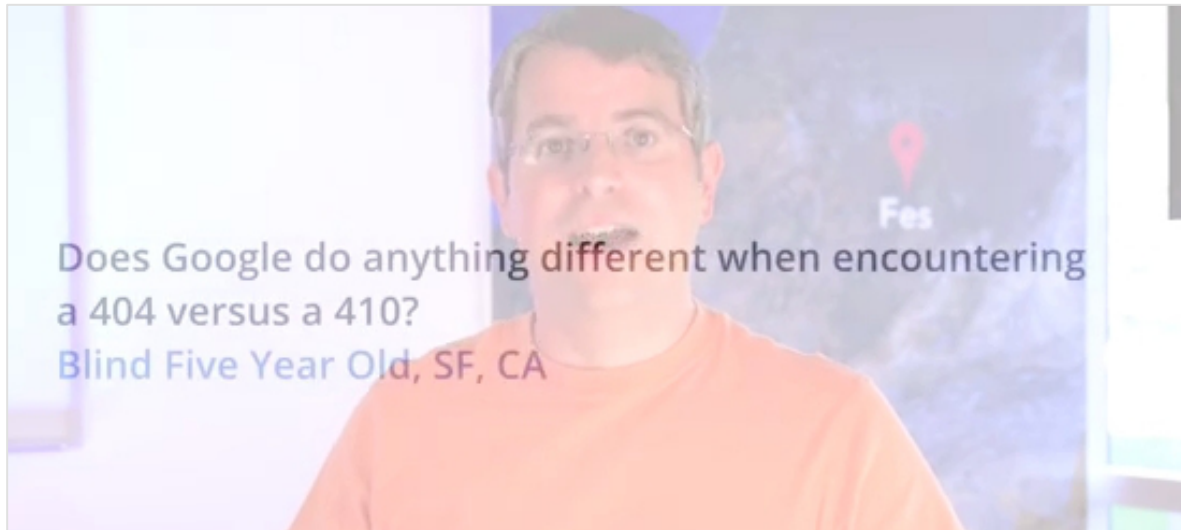


Fig.1. Une capture d'écran de la vidéo de Matt Cutts sur le sujet : la communication de Google a un peu « flotté » sur le sujet. Après avoir préconisé le recours au code 410 pour les pages qui n'existent plus, John Mueller a semé le trouble en indiquant plus tard qu'en fait, les deux codes étaient gérés de la même façon. Dans cette vidéo, Matt Cutts confirme qu'il y a bien une différence, mais une différence très subtile...

Si votre page renvoie une 410 : Google va prendre en compte immédiatement ce code, et « sortir » la page de son index. Si toutefois, le code 410 a été employé à tort, Google peut potentiellement recrawler la page dans les jours qui suivent, et la réintégrer dans l'index si elle répond à nouveau avec un code « 200 ».

Si votre page renvoie un code 404, Google va être plus circonspect, car beaucoup de problèmes techniques peuvent conduire un serveur web à renvoyer temporairement des erreurs 404 pour des pages qui réapparaîtront dans les jours qui suivent. Donc il s'écoulera environ 24h avant que Google décide de « sortir » la page de son index... Les URL renvoyant une 404 sont également recrawlées par la suite pour tester si elles n'ont pas réintégré l'arborescence du site.

Conclusion : si vous savez qu'un produit est devenu définitivement indisponible,

renvoyer un code 410 permettra de faire disparaître plus rapidement la page de l'index de Google. Mais fondamentalement, 404 et 410 sont gérées de manière très similaire par Google.

Les liens sur les pages en 404 transmettent-ils du linkjuice ?

Nous l'avons évoqué plus haut, les URL renvoyant un code d'erreur 404 sont censées ne plus transmettre le linkjuice en provenance des liens pointant vers l'ancienne page. Mais est-ce bien vrai ? Car de récentes rumeurs ont évoqué le fait que Google semblait suivre des liens qui ne figuraient pourtant que sur des pages renvoyant un code 404...

Mais rappelons ici que Google utilise des stratégies (parfois extrêmes) pour découvrir des liens (comme exploiter les URL non linkées - (pas de balise pour créer un « vrai » lien - découvertes dans des textes). Exploiter

l'information contenue sur des pages 404 pour découvrir des liens et en tenir compte pour transmettre du linkjuice est autre chose. John Muller a récemment confirmé qu'il était « douteux » que ces URL transmettent du PageRank et soient utilisées dans le calcul des « signaux » exploités par l'algorithme du moteur.

La meta « unavailable after » préconisée par Google

Google a annoncé en 2007 le support d'une valeur d'attribut nouvelle pour la balise meta robots : « unavailable-after ».

Cette balise permet d'indiquer à Google quand la page que Google vient de crawler devra être éliminée de l'index. La syntaxe type est la suivante (la date doit respecter le format défini par la norme RFC-850) :

```
<META NAME="GOOGLEBOT"  
CONTENT="unavailable_after: 25-  
Aug-2007 15:00:00 EST">
```

« unavailable-after » est également une valeur d'attribut utilisable dans une directive x-robots-tag. Par contre, elle n'a d'effet que sur l'index classique du moteur (donc pas sur Google News, Google Images, etc.).

Cette balise est intéressante à utiliser lors d'un changement radical de catalogue par exemple, pour favoriser la désindexation rapide des anciennes URL (notamment si le catalogue contient un nombre important de produits).

Par contre cette « meta » n'est supportée que par Google à ce jour, donc ne produira aucun effet sur le comportement de Bing ou de Yandex.

Les bonnes pratiques

Il n'existe pas une seule approche possible, selon qu'on privilégie l'expérience utilisateur, la facilité d'implémentation technique ou l'optimisation SEO.

Traiter les sorties de catalogue « définitives »

La page 404 ou 410 « intelligente »

Par exemple, si le webmaster (ou l'applicatif web) est en mesure de savoir quels produits de substitution proposer pour le produit qui a disparu du catalogue, alors on pourra créer une page dite « 404/410 intelligente ».

Si le produit de remplacement est unique, on peut afficher la fiche du produit de substitution à la place de l'ancienne, avec un message apparent indiquant quelque chose comme « *cette référence n'est plus disponible. Voici la fiche du nouveau produit équivalent* ». L'utilisateur voit donc une fiche produit presque normale, alors que pourtant cette URL renvoie un code 404 ou 410 : c'est la « smart 404 » ou « friendly 404 ».

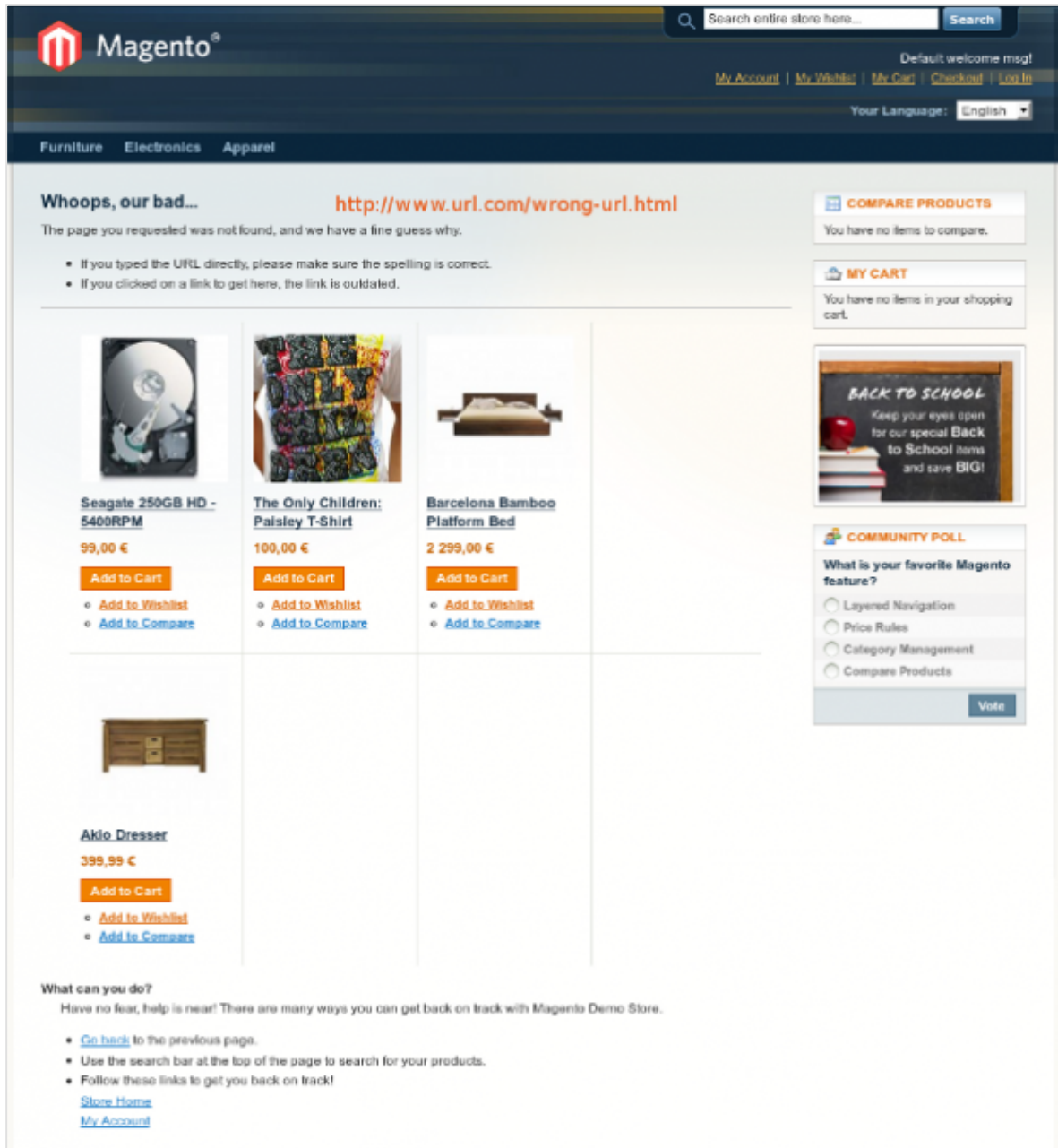


Fig.2. Un exemple de « 404 intelligente » fabriquée par le plugin pour Magento « friendly 404 ».

Si on doit suggérer plusieurs produits, alors la « 404 intelligente » aura une apparence différente. Le message devient « *ce produit n'est plus disponible. Voici d'autres produits équivalents* ». La même URL produit une page dont le contenu change par rapport à la fiche produit normale, et renvoie en parallèle un code d'erreur 404 ou 410.

Alors que la plupart des plateformes e-commerce ne savent pas gérer en natif la substitution d'une fiche produit à une autre, beaucoup savent gérer l'affichage de produits similaires ou appartenant à la même catégorie via une requête effectuée grâce au moteur de recherche interne. La seconde implémentation est donc plus facile et beaucoup plus fréquente, y compris si la substitution est univoque.

Notons une difficulté au passage : pour savoir vers quels produits de substitution renvoyer l'internaute, il faut que l'applicatif web se « souviene » des caractéristiques du produit disparu. Conclusion : si retirer le produit du catalogue conduit à effacer les données du produit, cette approche devient techniquement irréalisable. Il faut que l'on puisse simplement désactiver la fiche produit, au moins temporairement.

La page temporaire « produit indisponible » : une solution complexe et qui n'est pas dénuée d'inconvénients

La page « 404 intelligente » retourne vraiment un code 404 aux crawlers. Donc, même si d'un point de vue expérience utilisateur, ces pages sont plus qu'adaptées, ces pages ont deux inconvénients pour le SEO :

- Elles vont rapidement cesser de se positionner sur les mots clés liés à l'ancienne fiche produit ;
- Elles ne transmettent plus le linkjuice provenant des éventuels liens en provenance d'autres sites qui pointaient vers elles.

On peut donc imaginer de maintenir pendant quelque temps (quelques jours ? quelques semaines ?) une page avec le contenu décrit ci-dessus pour les 404 intelligentes, mais qui répondra temporairement un code 200.

C'est une pratique assez rare, car elle entraîne quelques complications :

- Si on affiche la fiche d'un produit de substitution, cette page est un doublon de la nouvelle fiche produit. Donc on affichera plutôt l'ancienne fiche produit, avec un message « *produit indisponible* » ... ce qui est décevant pour le client potentiel. Ce choix n'est

fait que quand on privilégie le SEO sur l'expérience utilisateur.

- Si on affiche une sélection de produits proches, ce n'est plus un doublon. Mais plusieurs pages de ce type peuvent être suffisamment proches pour être à nouveau considérées comme des doublons...
- Si on oublie le référencement de l'ancienne fiche produit, et qu'on veut juste que la page transmette du linkjuice, alors on peut utiliser une balise meta robots avec comme valeur d'attribut « noindex » ou une canonical pour éviter l'impact de cette génération de quasi doublons. Au bout de quelques semaines, la page sera redirigée selon l'une des méthodes évoquées plus loin. Mais dans la pratique, la solution la plus courante pour conserver le linkjuice est faire d'emblée une redirection 301 en cas d'appel de la fiche produit disparue, sans passer par une page temporaire.

La redirection 301 « intelligente »

Si on ne sait pas afficher une « 404 intelligente », on peut rediriger le visiteur vers la page produit de substitution (quand il est unique) ou la page de listing la plus proche (quand on doit proposer plusieurs produits de substitution).

Le recours aux redirections permet aussi de conserver une grande partie du linkjuice qui provenait des backlinks des produits devenus indisponibles.

Notons que cette redirection est permanente, et doit donc renvoyer un code 301 (et surtout pas un code 302, même si elle n'est maintenue en place que quelques jours).

A l'usage, cette solution peut être déroutante pour l'utilisateur, qui se voit

« téléporter » vers une autre page que celle qu'il attend. Il est donc important de faire apparaître les messages évoqués plus haut pour les « 404 intelligentes » sur la page produit ou la page listing finale. Ce qui demande un peu de développement car les plateformes e-commerce ne gèrent que rarement ce comportement.

Classiquement, l'URL vers lequel on redirige contiendra un paramètre complémentaire permettant le déclenchement de l'affichage du message. Le doublon d'URL est éliminé ensuite soit parce que le paramètre est derrière un caractère « # » et l'affichage du message traité en javascript », soit à l'aide d'une balise link rel=canonical.

La redirection 301 vers la catégorie parente

Si on ne sait pas identifier un ou plusieurs produits de substitution, on peut préférer rediriger vers la catégorie parente à laquelle l'ancien produit appartenait. Une fois de plus, ce n'est possible que si on a conservé cette information. Mais assez souvent, si l'information a été effacée du catalogue, la catégorie parente peut se déduire de l'URL.

Une fois de plus, ce type de redirection est potentiellement déceptif pour un client potentiel, et faire apparaître un message expliquant pourquoi l'internaute se retrouve sur la page catégorie est là aussi, une bonne idée.

Pages 404/410 intelligentes ou redirections intelligentes : comment choisir ?

Notons que s'agissant de pages profondes et souvent peu « linkées », la « perte » de linkjuice due à la perte de backlinks des fiches produit peut se révéler anecdotique. Si on sait identifier (grâce aux logs par exemple) les fiches produits ayant des backlinks, on pourra donc décider de :

- Rediriger les fiches recevant des backlinks et proposer une « 404/410 intelligente » pour les autres ;
- Ou conclure que cela n'aura aucun impact et donc ne pas faire de redirection.

Attention aux redirections massives : le problème des soft 404

La redirection vers une catégorie parente est une méthode très répandue. Mais il faut que ces redirections restent peu nombreuses en proportion pour éviter une confusion avec une soft 404.

Les « soft 404 » sont un concept imaginé par Google pour tenir compte du fait qu'il arrive fréquemment qu'un webmaster ou un applicatif web ne sache pas renvoyer un code 404 en cas de ressource introuvable ou disparue, alors même que parfois la page affiche « je suis une page 404 » à l'internaute. Pour identifier ces cas, Google observe en particulier les cas où de nombreuses URL sont redirigées systématiquement vers une page unique (en particulier vers la home).

Les cas de soft 404 sont affichés dans un rapport ad hoc des Google Webmaster Tools.

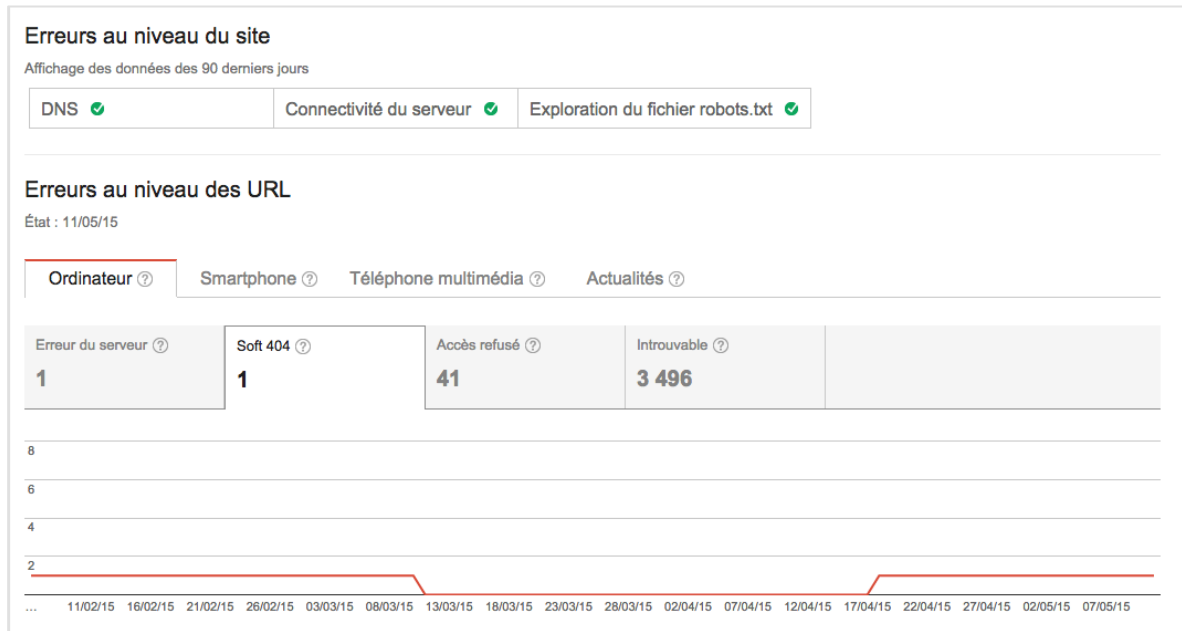


Fig.3. Le rapport sur les soft 404 dans les Webmaster Tools.

L'assimilation à une soft 404 des redirections fait perdre l'intérêt de ces redirections pour le SEO : autant renvoyer de vraies 404. Et cela permet d'éviter que la page de destination soit confondue avec une page 404 !

Conclusion : s'il faut, suite à un changement massif de catalogue, devoir rediriger un grand nombre d'URL de fiches produits vers la home ou des pages de rubrique, on évitera la solution des redirections (intelligentes ou non) pour recourir plutôt aux 404/410, voire à la meta « unavailable-after » (pour faire le ménage rapidement dans l'index).

Gérer les produits épuisés

Si le stock d'un produit se révèle épuisé, la situation est un peu différente. En cas de réapprovisionnement, le produit va redevenir disponible à la vente.

La bonne pratique consiste donc à essayer de gérer automatiquement les

situations d'indisponibilités liées aux stocks.

Si l'on sait que le stock de ce produit sera potentiellement réapprovisionné :

- Alors on garde la fiche produit, qui doit continuer de répondre en 200 (on ne veut surtout pas que l'URL disparaisse pour que la page réapparaisse plus tard avec le même contenu. Ce vibronnage nuira au bon référencement de la fiche produit. On ne veut pas non plus que la page réapparaisse avec une nouvelle URL : ce serait encore plus dommageable).
- On affiche un message clair du type « stock épuisé », en attente de réapprovisionnement ;
- On suggère des produits de substitution proches ;
- Et dans l'idéal, on propose de s'inscrire à une alerte par mail ou sms pour que le client soit averti du retour à la disponibilité du produit (ou on propose de le réserver, mais le problème est que l'empreinte carte bleue a une durée de vie limitée ce qui peut obliger le client à devoir repayer

plus tard avec un taux d'abandon élevé) ;

- Lorsque le produit redevient disponible : la page redevient « normale ».

Si le produit est définitivement épuisé et qu'on le sait : alors on gère la situation comme plus haut pour les sorties de catalogue « volontaires » à l'aide d'une 404 intelligente ou d'une redirection intelligente.

Si on ne sait pas toujours si le stock pourra être reconstitué : alors on gère une phase temporaire, où la page produit indique que le produit est en cours de réapprovisionnement comme expliqué ci-dessus. Après un certain délai, la page est désactivée (et non supprimée si possible, au cas où il faudrait recréer la page).

Très important : pensez à nettoyer le maillage interne

Si vous gérez la disparition des fiches produits avec des pages qui renvoient un code 404 ou 410, vous ne voulez pas en parallèle créer des liens brisés pointant vers des pages indisponibles qui seraient affichés sur d'autres pages de votre site.

Veillez bien à faire disparaître tous les liens qui pointent vers pages devenues indisponibles.

Si les URL vers des produits sortis du catalogue sont redirigées, c'est moins grave. Mais il est préférable de montrer directement les nouvelles URL plutôt que les anciennes.

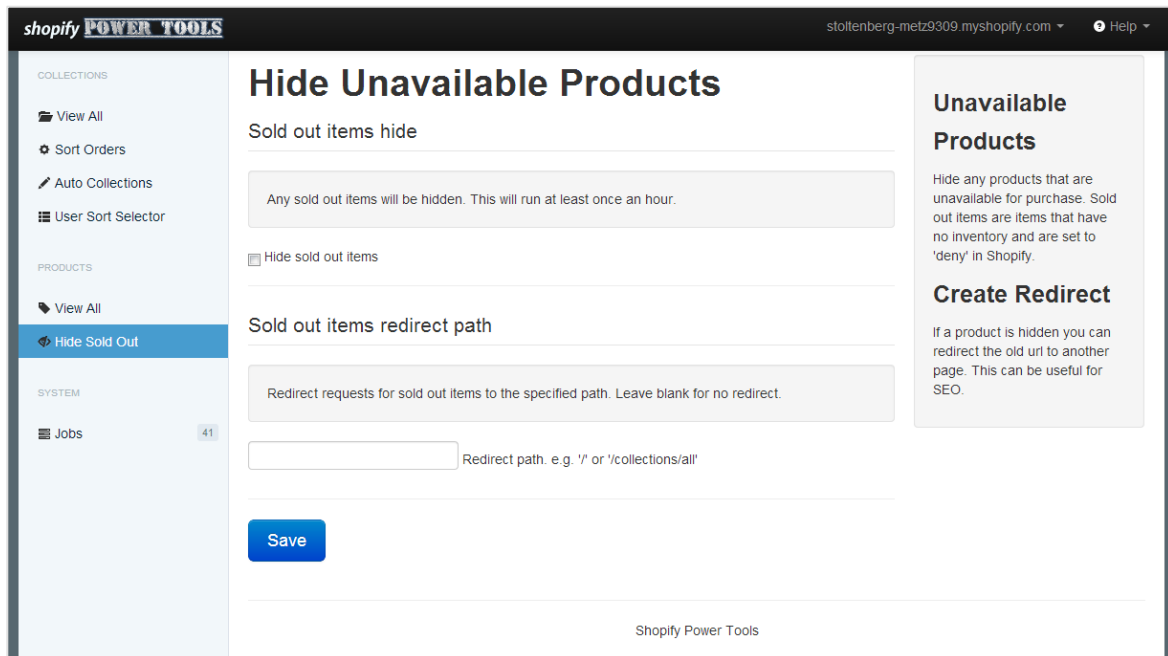


Fig.4. Un exemple de plugin pour Shopify pour gérer de façon basique les produits indisponibles ou épuisés : le système désactive automatiquement les produits dont le stock est épuisé, et propose de faire une redirection.

Quid des pages « listes de produit » ?

Au fur et à mesure que les produits listés sur une page de catégorie ou une page de catégorie filtrée disparaissent, et dans l'hypothèse où ces produits ne sont pas remplacés, il arrive un moment où la liste de produit devient vide. Comment gérer cela ?

- Si c'est une situation définitive : on transforme la page en « 404 intelligente », où on redirige vers une page listing équivalente (en mode « redirection intelligente »). Les liens vers cette page doivent disparaître du maillage.
- Si c'est temporaire ou si on ne sait prédire le devenir de la page : on garde la page dans l'arborescence quelque temps. Les liens internes pointant vers cette page sont conservés. La page continue de retourner un code 200. Et le contenu est « enrichi » à l'aide du moteur de recherche interne en présentant des résultats venant de la catégorie parente ou des suggestions. Par contre si toutes les pages de ce type se ressemblent, il sera peut-être obligatoire soit de faire varier les résultats d'une page à l'autre, soit de placer une meta robots avec la valeur d'attribut noindex pour éviter de « polluer » l'arborescence avec un grand nombre de quasi doublons.

Les plateformes e-commerce décident souvent de la solution finalement adoptée

Dans l'idéal, avec un petit peu d'effort et quelques développements, on peut concilier une bonne expérience utilisateur avec une bonne gestion de

l'impact SEO pour gérer les produits indisponibles ou épuisés.

Mais souvent, ce sont les limitations de la plateforme e-commerce qui décident du comportement réel de ces pages : redirection ou 404, page « intelligente » ou non etc.

Renvoyer une 404 quand le produit a disparu du catalogue est une solution qui, *a minima*, sera bien comprise par le moteur de recherche. Elle n'offre pas une expérience utilisateur idéale, sauf si on « enrichit » la page en suggérant un ou plusieurs produits de substitution. Si la plateforme ne permet pas d'autres fonctionnements, celui-là est acceptable et ne générera pas de vrais effets de bord pour le SEO.

Par contre, si le code renvoyé n'est pas « 404 », mais 200, il faudra corriger ce problème.

Ceci est vrai également si les pages « listes de produits » vides continuent à répondre 200 et à être maillées dans l'arborescence : ce comportement, assez fréquent, génère de nombreux pages doublons (il n'y a rien qui ressemble plus à une liste vide qu'une autre liste vide)... Ce problème doit être corrigé.

Si la plateforme génère plutôt une redirection automatique, on s'assurera que cette redirection ne retourne pas un code 302 mais bien 301, et qu'il n'y a pas de redirections multiples, voire de boucles infinies.

Le mieux est l'ennemi du bien...

Il n'existe donc pas une seule façon de gérer les produits indisponibles sur un site marchand. Différentes approches sont possibles, certaines simples à implémenter, d'autres non. Aucune n'est parfaitement idéale : renvoyer une 404 peut vous faire perdre un peu de linkjuice, un trop grand nombre de redirections peut générer des soft 404... Dans tous les cas, c'est l'amélioration de l'expérience utilisateur qui se révèle parfois complexe à obtenir, car les objectifs SEO restent simples :

- Stabiliser les pages qui risquent de disparaître pour réapparaître un peu plus tard et préserver leur référencement ;
- Nettoyer assez rapidement l'index du moteur de recherche des pages qui ne figurent plus sur le site ;
- Et si possible, conserver le linkjuice des pages qui ont disparu (dans la mesure où cela fait sens).

Il est regrettable de constater que sur nombre de plateforme e-commerce, la gestion de ces cas est assez « contrainte » et qu'il n'est pas possible d'implémenter toutes les approches. Le résultat est que l'expérience client est souvent décevante sur les sites e-commerce, notamment quand un internaute a découvert un lien dans une page de résultats de moteur de recherche.

Mais il ne tient qu'à vous d'améliorer cela, en suivant ces bonnes pratiques...

Bibliographie et liens

Un article de Moz du 11 avril 2012 sur les contenus expirés :

<https://moz.com/blog/how-should-you-handle-expired-content>

Une des dernières vidéos tournée par Matt Cutts, qui parle du sujet de cet article :

https://www.youtube.com/watch?v=9tz7Eexwp_A

Vidéo de Matt Cutts à propos de l'usage des 410 vs 404 :

<https://www.youtube.com/watch?v=xp5Nf8ANfOw>

La présentation de la balise meta « unavailable after » sur le blog pour les webmasters de Google :

<http://googleblog.blogspot.fr/2007/07/robots-exclusion-protocol-now-with-even.html>

La page d'aide de Google sur les « soft 404 » :

<https://support.google.com/webmasters/answer/181708?hl=fr>



Philippe YONNET, *Directeur Général de l'agence Search-Foresight, groupe My Media*
(<http://www.search-foresight.com>).