

Quelles métadonnées techniques sont prises en compte par Google Images ?



Par Daniel Roch

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Certains types d'images contiennent ce que l'on appelle des métadonnées : il s'agit d'informations précises qui peuvent être lues par certains logiciels, applications, appareils photos et sites web. Mais ces données ont-elles un intérêt en référencement naturel, notamment sur Google Images ? Nous allons voir dans cet article les différents types de métadonnées qui existent et le poids qu'elles peuvent avoir dans le positionnement d'une image et d'un contenu...

Les images et leurs métadonnées

Qu'est-ce qu'une métadonnée ?

On peut inclure des métadonnées à l'intérieur d'un fichier image, sans recourir à du code HTML comme le célèbre attribut ALT de la balise IMG. Ces données sont relativement variées : modèle d'appareil photo, copyright, focale, titre, mots clés, description, vitesse d'obturation et bien d'autres...

Le format de métadonnées le plus connu est le format EXIF, très populaire pour les photographes et les amateurs d'art, mais il en existe d'autres : le format IPTC ou encore XMP.

Certaines données sont ajoutées automatiquement aux images, notamment grâce aux appareils photos numérique et certains téléphones portables qui vont par exemple ajouter des données EXIF pour chaque prise de vue. De l'autre côté, une autre partie des données (IPTC et XMP) doit être renseignée manuellement par un utilisateur avec un logiciel adapté, comme par exemple LightRoom ou encore IrfanView.

Il existe également d'autres types de métadonnées, cette fois-ci pour d'autres fichiers comme les PDF ou encore les fichiers MP3 (format de données ID3). Nous nous concentrerons dans cet article sur celles des images.

Quels sont les différents types de métadonnées ?

EXIF

Abréviation pour « *Exchangeable Image File Format* », il s'agit d'une spécification utilisée pour les images JPEG, TIFF et RIFF. Ces métadonnées sont la plupart du temps ajoutées automatiquement par les appareils photos aux images qu'un utilisateur peut prendre.

Parmi les données EXIF disponibles, voici celles qui sont ajoutées la plupart du temps par les différents modèles d'appareils :

- Constructeur ;
- Modèle ;
- Orientation ;
- Date et heure ;
- Compression ;
- Vitesse d'obturation ;

- Focale ;
- ISO ;
- Compensation d'exposition ;
- Utilisation du flash ;
- Etc.

Ce format de métadonnées n'est pas compatible avec le format JPEG 200 ni avec le format PNG.

IPTC et XMP

Le format IPTC est une abréviation de « *Information Interchange Model* ». Il s'agit ici d'un autre type de métadonnées que l'on doit ajouter manuellement aux images. Elles permettent, grâce à des logiciels adaptés, de rajouter des informations que l'on peut juger pertinentes pour les images :

- Titre ;
- Description ;
- Auteur ;
- Etc.

Le format XML est pour sa part une abréviation pour « *Extensible Metadata Platform* ». Il permet également, et peu ou prou, de renseigner les mêmes données, à un détail près : le XMP est un format XML, c'est-à-dire extensible par les différents outils qui vont les utiliser.

Les deux formats ont d'ailleurs permis un travail collaboratif entre Adobe et l'IPTC, aboutissant au final à l'« *IPTC Core Schema for XMP* » qui combine les deux approches, et qui permet d'étendre la portabilité de ces données également aux fichiers GIF, PNG, JPEG 2000, SVG et PDF. On l'appelle souvent « *IPTC XMP* », en remplacement de l'ancien format « *IPTC IIM* ».

Quand on parle d'IPTC à l'heure actuelle, on parle en réalité bien souvent de cette approche commune entre les deux formats.

Les problématiques de ces données

Avant de parler de l'impact éventuel en référencement naturel, il faut d'abord aborder toutes les problématiques globales que ces métadonnées vont provoquer.

1er défaut : chaque format ne fonctionne pas pour toutes les images. On peut retrouver par exemple des métadonnées Cross-IPTC dans un fichier GIF, mais on n'y trouvera pas de données EXIF.

2nd défaut : toutes les données IPTC sont manuelles. Sans intervention de l'utilisateur, aucune donnée ne sera ajoutée. C'est d'autant plus problématique qu'il s'agit là des informations les plus pertinentes au niveau sémantique : titre, description, mots clés, auteur...

3ème défaut : dès lors qu'un fichier image est modifié, sauvegardé, transféré ou mis en ligne, il est possible que les données soient supprimées. C'est souvent le cas lors d'un changement de format du fichier, ou encore quand vous compressez vos images pour gagner du poids. En d'autres termes, si vous tenez à ces informations, vérifiez qu'elles soient toujours présentes lors de chaque étape de vos manipulations.

Social Media site/system	Summary	Displays correctly?			Displays 4Cs?			Save As embedded?			Download embedded?		
		Exif	IPTC	IPTC	Exif	IPTC	XMP	Exif	IIM	XMP	Exif	IIM	XMP
500px - www.500px.com Tested in March 2013	Most relevant fields shown, copyright was overridden, in downloaded file we only found Exif metadata. Save As is disabled.	Grey	Yellow	Green	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
Dropbox - www.dropbox.com Tested in March 2013	No metadata shown, all embedded fields are preserved in the image files.	Grey	Grey	Red	Green	Green	Green	Green	Green	Green	Green	Green	Green
Facebook - www.facebook.com Tested in June 2013	Metadata not shown anymore, all embedded metadata stripped-off from image files.	Grey	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red
Flickr FREE account- www.flickr.com Tested in June 2013 (PRO account may show other results)	A few metadata fields shown, 'by' was overridden, for any downscaled rendition all embedded metadata are stripped-off from image files, only the Original rendition keeps all metadata.	Grey	Yellow	Yellow	Red	Red	Red	Red	Red	Red	Red	Red	Red
Google+ - plus.google.com Tested in March 2013	Primarily Exif metadata shown, all embedded fields are preserved.	Green	Green	Yellow	Green	Green	Green	Green	Green	Green	Green	Green	Green
Img.ly - www.img.ly Tested in October 2012	No metadata shown, all embedded fields are preserved in the Save As image files.	Grey	Grey	Red	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey
Instagram - instagram.com Tested in June 2013	Image taken by a smartphone, metadata edited with an app, then posted at Instagram: No metadata are shown, all metadata stripped-off from Save As files.	Grey	Grey	Red	Red	Red	Red	Grey	Grey	Grey	Grey	Grey	Grey
Joomeo - www.joomeo.com Tested in October 2012	Some metadata fields are shown and all embedded fields are preserved in the image files. Save As is disabled.	Green	Green	Yellow	Grey	Grey	Grey	Green	Green	Green	Green	Green	Green
Photobucket - www.photobucket.com Tested in October 2012	No metadata shown, embedded fields are preserved in downloaded image files but not in Save As files.	Grey	Grey	Red	Red	Red	Red	Green	Green	Green	Green	Green	Green
Pictify - www.pictify.com Tested in March 2013	No metadata shown, embedded fields are preserved in downloaded image files but not in Save As files.	Grey	Grey	Red	Red	Red	Red	Grey	Grey	Grey	Grey	Grey	Grey
Pinterest - www.pinterest.com Tested in March 2013	No metadata shown, all embedded fields are preserved in the Save As image files.	Grey	Grey	Red	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey
Tumblr - www.tumblr.com Tested in March 2013	No metadata shown, all embedded fields are preserved in the Save As image files.	Grey	Grey	Red	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey
Twitpic - twitpic.com Tested in March 2013	No metadata shown, only a few IPTC fields are preserved in Save As image files.	Grey	Grey	Red	Red	Yellow	Red	Grey	Grey	Grey	Grey	Grey	Grey
Twitter - www.twitter.com Tested in June 2013 (Link to the tweet)	No metadata shown, all embedded metadata stripped-off from image files.	Grey	Grey	Red	Red	Red	Red	Grey	Grey	Grey	Grey	Grey	Grey
Via.me - www.via.me Tested in October 2012	No metadata shown, all embedded fields are preserved in the Save As image files.	Grey	Grey	Red	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey
Yfrog - www.yfrog.com Tested in November 2012	No metadata shown, all embedded fields are preserved in the Save As image files.	Grey	Grey	Red	Green	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey

You want to know what the icons above stand for? In short: a green light indicates good results, a yellow one not so good ones and the red light that the handling of metadata should be improved, grey means: has not been tested. [Go to this page](http://www.embeddedmetadata.org/social-media-test-results.php) and read about our ranking in detail.

Fig.1. Prise en compte des métadonnées par les réseaux sociaux.
Source : <http://www.embeddedmetadata.org/social-media-test-results.php>

Pour bien comprendre cette dernière problématique, le site embeddedmeta-data.org avait fait un test pour voir si les métadonnées étaient ou non conservées lors de leur mise en ligne sur certains sites « sociaux ». Le constat est que cette conservation des informations est très variable d'un site à un autre. La figure 1, page précédente, illustre le résultat de ces tests :

- En gris : non testé ;
- En vert : pris en compte ;
- En jaune : une partie des données est perdue ;
- En rouge : presque toutes les données sont perdues.

Metadonnées des images et SEO : quel impact ?

Pour être parfaitement honnête, il est assez difficile de connaître l'impact SEO exact des informations contenus dans les images, et ce pour plusieurs raisons :

- Très peu d'articles traitent de ce sujet en profondeur, ou en parle souvent sans avoir testé réellement cet éventuel impact ;
- Bing n'en parle quasiment jamais ;
- Google a fait des annonces éparses et contradictoires sur le sujet.

Affichage

Le premier point que l'on peut regarder est de savoir si le moteur de recherche Google affiche ces métadonnées, surtout dans Google Image. En 2012, un article de Webmaster Central indiquait que les données EXIF pouvaient parfois être affichées dans le détail de l'image dans Google Image (<http://googlewebmastercentral.blogspot.fr/2012/04/1000-words-about-images.html>).

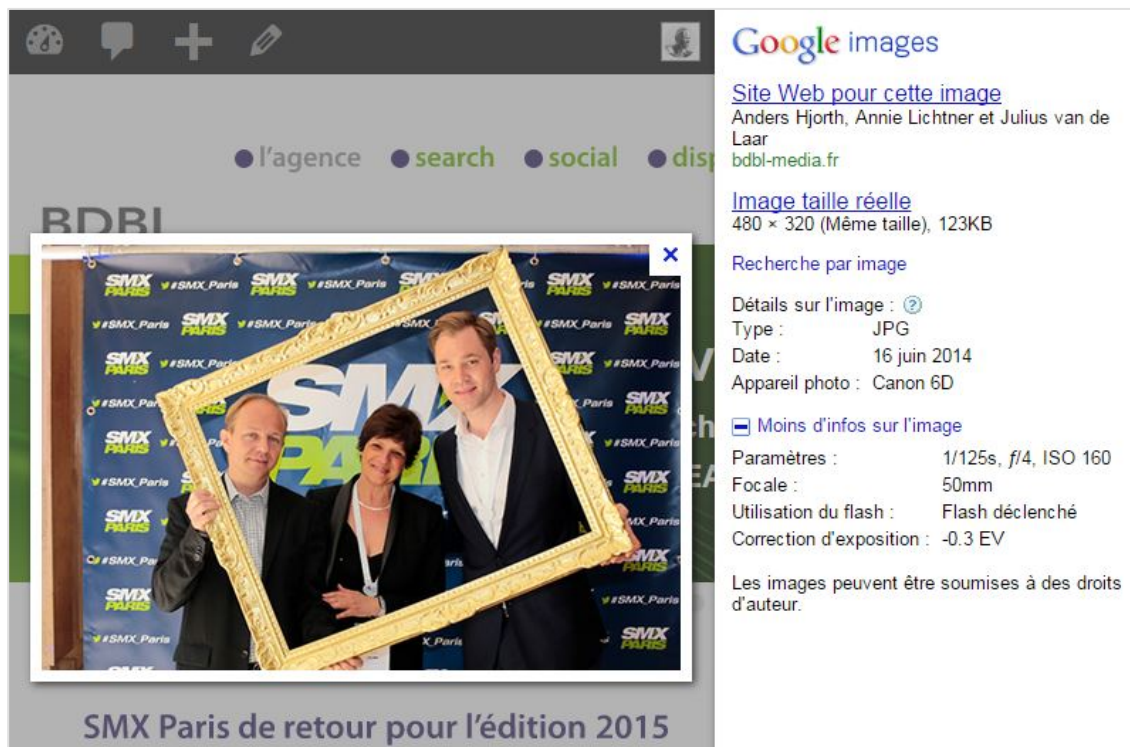


Fig.2. Affichage d'une image sur le moteur Google dédié à ce format. Source de l'image : <http://www.bdbl-media.fr/blog/referencement-dimages-un-point-sur-les-metadonnees.html>

En 2015, c'est toujours le cas. Dans l'exemple de la figure 2, on voit bien l'affichage à droite des informations EXIF sur le Flash, la correction d'exposition, la date de prise de vue ou encore l'appareil photo utilisé pour ce visuel.

D'ailleurs, l'utilisation de ces données pour de l'affichage n'est pas nouveau. Dans Picasa, un autre produit de Google, ces données sont également affichées pour l'utilisateur (voir <http://forum.webrankinfo.com/google-images-lit-les-iptc-dans-picasa-t148712.html>).

Indexation et positionnement

La plus importante question est la suivante : si Google lit et affiche ces métadonnées, les utilise-t-il pour son indexation et pour le positionnement des images ou des contenus qui les hébergent ?

A l'heure actuelle, la réponse est non. En 2012, l'annonce de Google indiquait que le moteur de recherche pouvait utiliser ces informations pour « aider les utilisateurs à trouver le contenu qu'ils recherchaient plus facilement ». Mais dans une vidéo datant de février 2014, Matt Cutts contredit, cet élément : « *It is something Google is able to parse out, and I think we do reserve the right to use it in rankings* » (<https://youtu.be/GMf6FmRus2M>).

En d'autres termes, Google est capable techniquement d'analyser ces informations (ce que nous savions déjà),

mais actuellement, ces données ne sont pas utilisées comme un facteur de positionnement. La subtilité est que le moteur de recherche se réserve toujours le droit de le faire dans les années à venir.

D'ailleurs, plusieurs tests de référencement naturel prouvent cette assertion : Google n'utilise pas les termes des métadonnées pour l'indexation ni pour le positionnement :

- Sur InfiniClick : <http://blog.infiniclick.fr/articles/referencement-images-etat-art.html> ;
- Sur SeoMix : <http://www.seomix.fr/donnees-exifs-images-seo/>.

Dans ces deux tests, les termes placés dans les différents champs IPTC ou EXIF ne permettent pas de faire remonter l'image ou la page sur le terme concerné.

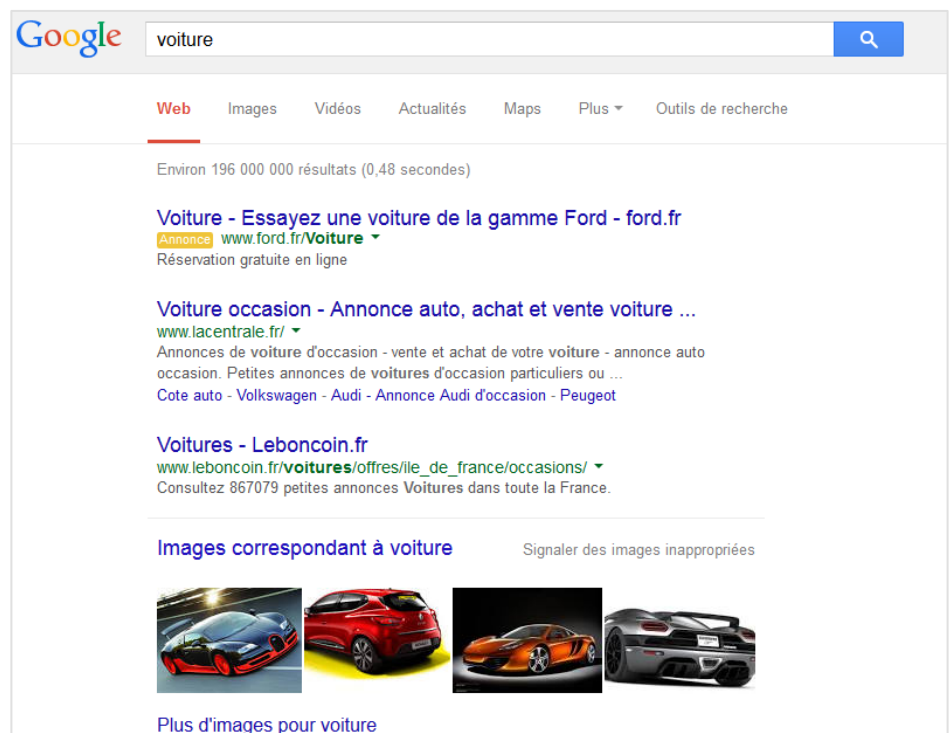


Fig.3. Les images sont très souvent proposées en « recherche universelle » directement dans les résultats web.

L'importance des images pour Google

A ce niveau, on pourrait donc penser que les métadonnées n'ont aucun intérêt en SEO, et c'est malheureusement le cas actuellement. Mais rien ne dit que Google ne changera pas d'avis...

D'ailleurs, le moteur a parfaitement compris l'importance des images pour l'utilisateur. Ce n'est pas pour rien s'il les met en avant de plus en plus souvent dans ses outils et interfaces. Il l'a tout d'abord fait avec Google Images. Auparavant clairement séparé de Google Web, les deux ont fusionné sur de nombreuses requêtes, affichant ainsi directement des résultats images dans une recherche web traditionnelle, comme le montre la figure 3 avec une recherche sur le terme « voiture ».

Mais Google va plus loin. Récemment, il a également redéveloppé la partie image de Google My Business (anciennement appelé Google Local Business Center ou Google Adresses). Lorsque vous modifiez une page présentant votre entreprise, l'interface de gestion des images a changé. Google ne vous demande non plus d'ajouter simplement vos images et photos, mais il vous demande surtout de les classer par type (figure 4) :

- Photo d'identité ;
- Photos d'intérieur ;
- Photos de l'extérieur ;
- Photo en action ;
- Photos d'équipe ;
- Photos supplémentaires (pour toutes vos autres photos et celles que vous aviez ajouté auparavant).

Attention cependant : le fait de classer vos images ne change pour l'instant rien dans l'affichage de vos pages Entreprise,

que ce soit sur Google+ ou dans Google Maps.

Garder ou non les métadonnées ?

Si l'impact de ces métadonnées est nul, faut-il les conserver dans les différents fichiers images ?

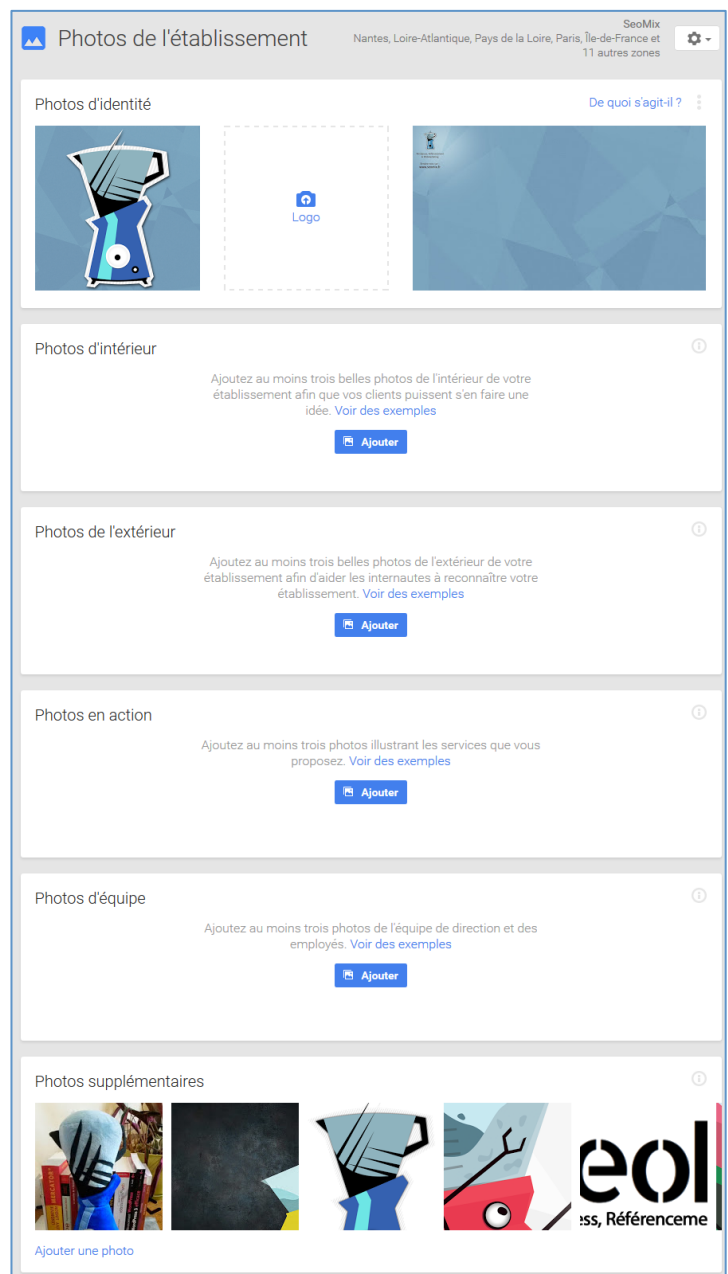


Fig.4. Classement des images dans Google My Business

Oui, il faut les garder !

Plusieurs raisons doivent vous pousser à conserver ces informations :

- Elles peuvent être lues et importées dans différents sites, logiciels et CMS, notamment 500px, Flickr, NextGen gallery pour WordPress... L'intérêt principal ici, c'est que ces sites ou outils peuvent alors copier/coller les informations contenues dans les images pour les ajouter directement dans le contenu HTML, générant ainsi un contenu unique et pertinent par rapport à l'image concernée ;
- Google affiche les données EXIF de certaines images dans son moteur d'images ;
- Le moteur de recherche se réserve le droit de les utiliser dans les années à venir ;
- Enfin, les métadonnées IPTC permettent d'ajouter un copyright. Si vous êtes l'auteur de ces images, cela vous permet de retracer bien plus facilement la source originale et les éventuels doublons ou vols de vos contenus.

Non, il ne faut pas les garder !

A l'inverse, deux raisons principales peuvent vous pousser à ne pas ajouter ces données, voir à les supprimer tout simplement :

- La raison évidente est qu'il est inutile de passer du temps à rajouter des métadonnées si elles n'ont aucun impact en référencement naturel : c'est du temps perdu que vous pourriez en effet passer sur d'autres actions bien plus pertinentes ;

- Deuxième raison : le poids des images. Le fait d'ajouter des données EXIF et IPTC va légèrement alourdir les images. Et on sait pertinemment qu'actuellement Google met en avant le fait que les sites web soient rapides à charger, et parfaitement adaptés pour les mobiles (donc légers).

Conclusion

Les métadonnées des images n'ont, comme vous l'aurez constaté, actuellement aucun impact direct sur le référencement des fichiers auxquelles elles sont attachés ou des contenus qui les hébergent. Vouloir ajouter ces données à vos fichiers est donc actuellement une perte de temps. Cependant, ne cherchez pas à les supprimer si elles sont déjà présentes : elles pourraient vous servir dans le futur.

Si vous souhaitez réellement référencer vos images, préférez de loin les points classiques à vérifier et à optimiser :

- L'attribut ALT (le texte alternatif) ;
- L'attribut Title (pour l'accessibilité, pas pour le SEO, et si l'image est cliquable) ;
- Le nom du fichier ;
- Le contexte autour de l'image, notamment sa légende ;
- Sa présence dans un fichier Sitemap.



Daniel Roch, *Consultant*

WordPress, Référencement et Webmarketing chez SeoMix

(<http://www.seomix.fr/>).