



Zeste de savoir

Une étude de l'attractivité de Zeste de
Savoir (épisode 2) : le cas des auteurs

9 mars 2021

Table des matières

1.	La méthode de travail	1
1.1.	Les données utilisées pour cette méthode	1
1.2.	Les limites de la méthode	2
2.	L'évolution des auteurs dans chaque type de contenus	2
3.	L'évolution des auteurs dans chaque domaine de savoir	3
4.	Les auteurs de Zeste de Savoir sont-ils fidèles?	3
5.	Qu'en est-il du rythme de publication dans tout ça?	4
6.	La productivité de la validation	4
6.1.	Les contenus publiés	4
	Contenu masqué	6

Chers agrumes, un [sujet à été ouvert sur le forum](#) de la part de @sgble dont l'objectif est de repenser le système de validation de Zeste de Savoir, parce que soit disant, ce dernier fonctionne mal et a donc pour effet de réduire l'attractivité du site pour les auteurs. Qu'en est-il vraiment?

Je me suis donc lancé (poussé la demande de @SpaceFox dans ledit sujet) dans une étude assez modeste pour essayer de comprendre à quel point Zeste de Savoir est attractif pour les auteurs. L'idée était surtout de mettre à jour les données de la [dernière étude](#).



Est-ce que le pari de l'ouverture (informatique, sciences, ...) qu'a fait la plateforme incite à écrire concrètement? La diversité des formats des contenus (tuto, article, billet) constituent-ils des catalyseurs d'écriture? Quelle fonctionnalité de la plateforme a eu une influence notable sur l'arrivée de nouveau contributeurs? Le système de beta/validation tel qu'on le connaît aujourd'hui porte-il ses fruits?

La suite du billet, commence par expliquer la méthode utilisée, puis présente les résultats des observations.

1. La méthode de travail

1.1. Les données utilisées pour cette méthode

1.1.1. Collecter l'information

J'avais besoin d'informations sur les auteurs de contenus, et n'ayant pas accès à la base de donnée du site (avec laquelle tout ceci aurait été plus pratique, mais moins *fun*) je me suis

2. L'évolution des auteurs dans chaque type de contenus

rabattu sur le *scraping* des pages du site. Je vois d'ici l'équipe technique qui me dévisage, mais ne vous inquiétez pas, je me suis contenté de bombarder le serveur de beta 🍊 .

J'ai repris mon script de la dernière fois en python3 qui matraque l'ensemble des pages de contenus pour extraire les métadonnées (date de publications, auteurs, catégories, ...) utiles à mon étude.

L'exécution de mon script me permet donc d'avoir des au format csv, qui sont l'ensemble des contenus publiés sur le site depuis sa création, avec leurs métadonnées.

1.1.2. Le traitement de l'information

C'est le moment le plus *fun*.

Avec les informations collectées, j'ai dû classer les catégories en domaines (SHS, Sciences de la nature, Informatique et Autres), et consolider le nombre d'auteurs cumulés toutes les semaines pour savoir comment ils ont évolués sur la durée.

1.2. Les limites de la méthode

- L'étude se base uniquement sur les projets de rédaction qui ont abouti a quelque chose: les membres dont les projets d'écriture n'ont jamais passé la validation ne seront donc pas considérés dans cette étude
- L'étude ne prend pas en compte l'historique de validation d'un contenu: c'est-à-dire qu'un contenu qui a subi une mise à jour majeure sera considéré comme ayant été publié a cette date pour la première fois. Notons que finalement ça a très peu d'impact sur les résultats obtenus.

2. L'évolution des auteurs dans chaque type de contenus

Zeste de Savoir permet aujourd'hui de rédiger dans trois types de contenus. Les tutoriels, les articles et les billets. Les articles et les tutoriels font partis des contenus qui sont validés par l'équipe et les billets sont des contenus qui ne sont pas passés par le filtre de la validation.

Voici une représentation de l'évolution du nombre d'auteurs dans le temps en fonction du type de contenus.

👁️ Contenu masqué n°1

On peut retenir que:

i

- Le nombre d'auteurs de billets (contenu non validé) à déjà dépassé le nombre d'auteurs d'articles et de tutoriels (depuis début 2020). L'attractivité des billets pour les auteurs n'est plus à démontrer.

3. L'évolution des auteurs dans chaque domaine de savoir

La particularité de Zeste de Savoir est de ne pas se limiter à une seule discipline. Sur ce site il y a trois gros domaines de savoir: Les sciences de la nature, les sciences humaines et sociales, et l'informatique. Tout le reste est classé dans "Les autres".

Il est donc intéressant de savoir si le pari de la pluridisciplinarité de Zeste de Savoir est toujours payant. Et si les auteurs sont réceptifs à cette ouverture.

© Contenu masqué n°2

Ces courbes nous apprennent certaines choses:

i

- Le domaine des sciences humaines et sociales n'attire plus les auteurs pratiquement depuis Juin 2020. Ce domaine de savoir semble fonctionner par période, car on observe des bonds de nouveaux auteurs pendant un temps, puis plus rien pendant plusieurs mois et ainsi de suite.
- L'informatique reste le truc de la communauté, les auteurs continuent d'arriver dans ce domaine.
- Les auteurs qui publie du contenus de type "Autres" qui prennent de plus en plus de place, preuve que la diversité semble être bénéfique.

4. Les auteurs de Zeste de Savoir sont-ils fidèles ?

Maintenant on va s'intéresser à la question de savoir si les auteurs du site rédigent juste une fois et quittent le navire, ou est-ce que leur première expérience les pousse à renouveler la chose.

Pour cela, j'ai construit un diagramme qui montre en abscisse le nombre de contenus publiés, et en ordonnée, le nombre d'auteurs qui ont publiés autant de contenus.

© Contenu masqué n°3

Pour ceux qui ont du mal à lire le graphique, la première barre signifie qu'il y a 119 auteurs qui ont publié un seul contenu sur le site, la seconde barre signifie qu'il y a 31 auteurs qui ont publié deux contenus sur le site, et ainsi de suite.

On peut faire les observations suivantes:

i

- La moitié de nos auteurs (119 sur 247) ont publié un seul contenu sur Zeste de Savoir et n'ont pas renouvelé l'expérience par la suite.
- La moitié des contenus publié sur ZdS est crée par 10 auteurs (il y'a 270 auteurs sur le site actuellement)

5. Qu'en est-il du rythme de publication dans tout ça?

J'en profite pour nommer les dix plus gros auteurs sur ce site (en nombre de contenus publiés, toute catégorie confondues).

Auteur	Nombre de contenus (tutoriel, article ou billet) publiés sur Zeste de Savoir
@SpaceFox	92
@Aabu	77
@informaticienzero	74
@artragis	67
@Auteur externe	56
@Gabbro	55
@Rockaround	52
@firm1	46
@Eskimon	42
@qwerty	40

Un grand merci à eux.

5. Qu'en est-il du rythme de publication dans tout ça ?

Pour remettre tout ceci en perspective avec les publications, un graphique qui montre l'évolution des contenus publiés par rapport à l'arrivée des auteurs.

👁️ Contenu masqué n°4

On remarque que globalement, malgré le nombre de contenus publiés qui augmentent, les nouveaux auteurs peinent à arriver. Serions nous moins attractif pour les auteurs?

6. La productivité de la validation

Maintenant, intéressons nous au nombre de contenus publiés.

6.1. Les contenus publiés

6. La productivité de la validation

Ici on observe que l'année 2017 a boosté sérieusement la production des contenus (c'est l'effet de l'arrivée des billets sur le site), même si cela semble régresser au fur et à mesure.

Il serait intéressant de suivre cette évolution par type de contenu.

6.1.1. La publication des tutoriels

👁 Contenu masqué n°5

i

Aucun tutoriel validé en 2021

Depuis 2018, la publication des tutoriels a été fortement ralentie pour atteindre le pire en 2021 (où l'on a encore aucun tutoriel publié)

6.1.2. La publication des articles

👁 Contenu masqué n°6

i

Un seul article publié en 2021

Depuis 2019, la publication des articles est elle aussi fortement ralentie et on est vraiment dans le creux de la courbe cette année (un seul article publié en 2021)

6.1.3. Les publications validées par Zds (tutoriels et articles)

👁 Contenu masqué n°7

i

Deux contenus validés par mois en moyenne

En 2019 et 2020, le rythme de la validation est en moyenne de deux contenus validés par mois. En 2021, quasi rien n'est sorti, problème de production ou problème de validation?

6.1.4. La publication des billets

👁️ Contenu masqué n°8

i

Dix billets publiés par mois en moyenne

Les billets ne jouent clairement pas dans la même catégorie que les contenus validés par le site. Depuis début 2019 on est sur une moyenne de 10 billets publiés par mois. On n'observe pas de grosse chute en 2021 sur les billets.

Comme on a pu le voir dans l'étude:

- Les billets ont toujours le vent en poupe et semblent attirer beaucoup d'auteurs ;
- Le plus gros des publications du site est porté par un petite poignée (une dizaine) d'auteurs ;
- Les contenus validés n'attirent plus depuis début 2019 ;
- On n'arrive toujours pas à fidéliser nos nouveaux auteurs (faire en sorte que les auteurs renouvelle leur expérience de rédaction).

Il devient urgent de trouver une solution pour remettre sur les rails la production des tutoriels et articles, parce qu'à ce rythme, on n'aura plus besoin de validateurs car il n'y aura plus de proposition d'articles/tutoriels.

Une question qui serait intéressante à se poser serait de savoir combien de contenus arrivent en beta pour pouvoir comparer ce qui est initié et ce qui est publié.

Merci encore aux auteurs qui font vivre le site.

Contenu masqué

Contenu masqué n°1

[Retourner au texte.](#)

Contenu masqué n°2

[Retourner au texte.](#)

Contenu masqué

Contenu masqué n°3

Retourner au texte.

Contenu masqué n°4

Retourner au texte.

Contenu masqué n°5

Retourner au texte.

Contenu masqué n°6

Retourner au texte.

Contenu masqué n°7

Retourner au texte.

Contenu masqué n°8

Retourner au texte.

Liste des abréviations

SHS Sciences Humaines et Sociales. 2