

ANALYSE DES DONNÉES POUR LE RECENSEMENT DES PLATEFORMES TECHNOLOGIQUES DE BIOPHYSIQUE

SOUS LA DIRECTION DE M. MARC-ANDRÉ DELSUC

UNE ÉTUDE RÉALISÉE PAR JOFFREY BÉCART & CHLOE GERARD

INTRODUCTION

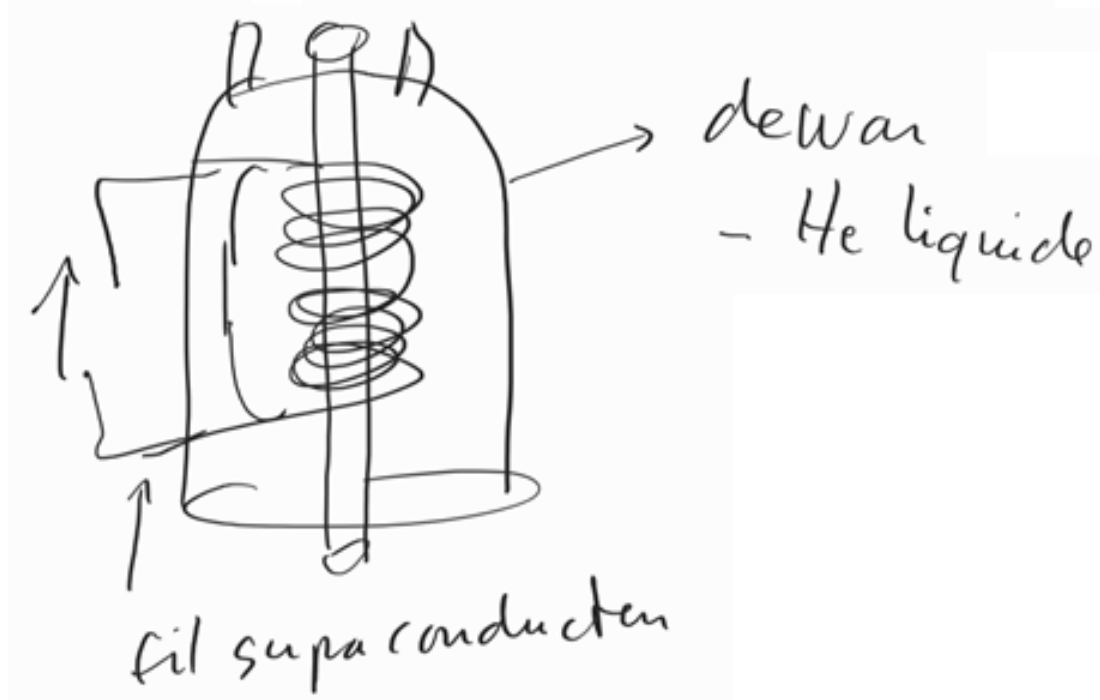


Schéma de principe d'un aimant supraconducteur utilisé en RMN ou en FT-ICR.

Ce projet repose sur une analyse de données pour le recensement des plateformes technologiques de biophysique afin de répondre à la problématique suivante : existe-t-il une corrélation entre le coût de l'équipement et l'influence d'un groupe ?

Cette étude a été faite sous la direction de M. Delsuc, directeur de recherche au CNRS, scientifique sénior à l'IGBMC de Strasbourg et cofondateur de Casc4de.

Pour répondre à cette problématique, nous nous basons sur le coût de l'équipement : les aimants utilisés en FT-ICR. Le FT-ICR (Fourier Transform – Ion Cyclotron Resonance Mass Spectrometry = spectrométrie de masse à résonance cyclotronique ionique) est un instrument produisant un champ magnétique, exprime en Tesla, permettant de déterminer très précisément la masse et la composition des molécules présentes dans un mélange.



DÉMARCHE

Pour estimer le coût de l'équipement, nous allons regarder la puissance de l'aimant : plus l'aimant est puissant, plus il est cher. Nous allons ainsi mesurer cette donnée en Tesla.

Remarque : il est important de préciser que le prix des aimants n'est pas proportionnel à leur puissance mais est exponentiel. Autre précision d'importance : des champs magnétique de puissances différentes ne permettent pas de faire les mêmes choses.

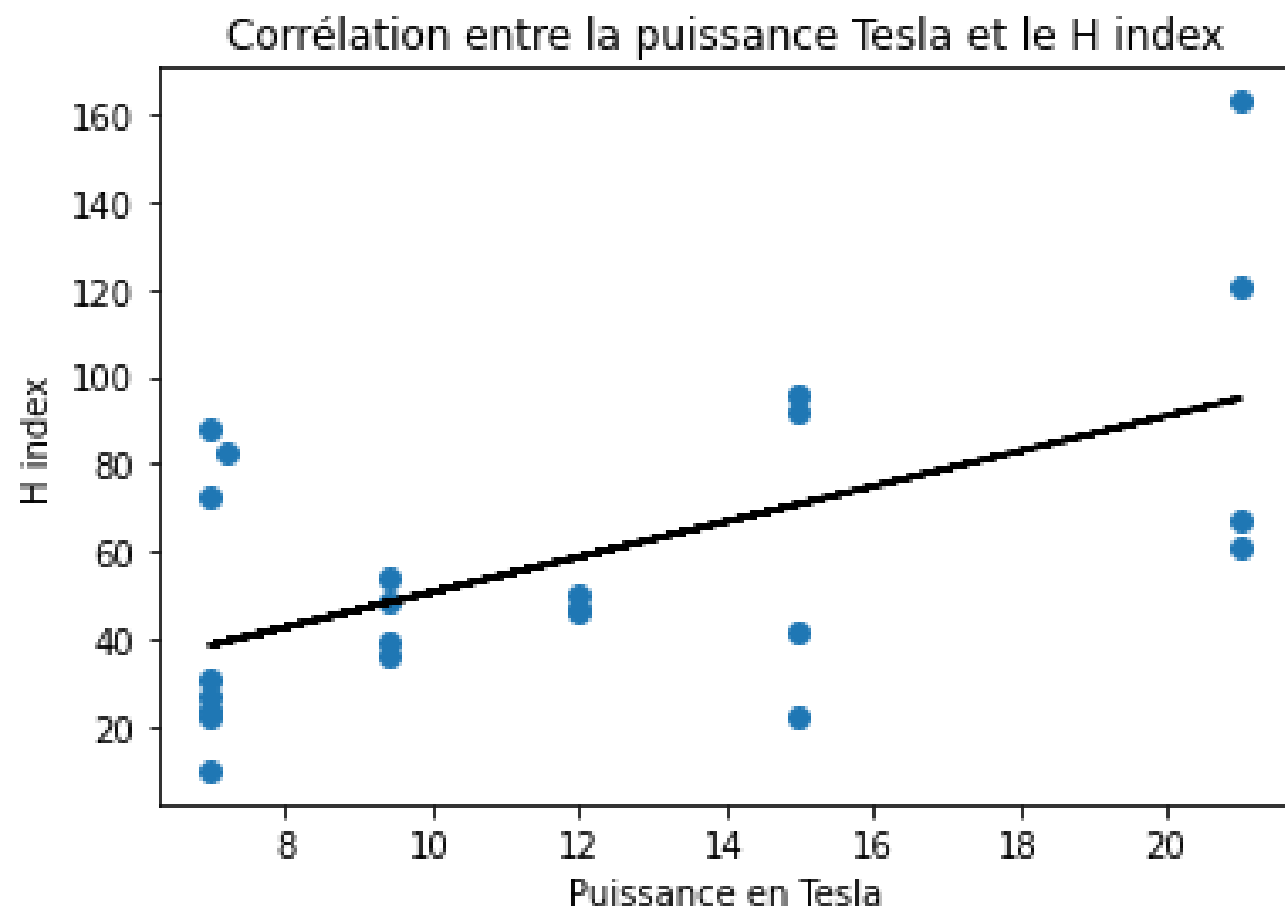
C'est ainsi que nous avons démarré nos recherches. A l'aide de Google Scholar, nous avons fait notre propre référencement d'articles : un tri a priori. C'est de cette façon que nous avons pris en compte : les auteurs, le nombre de fois où l'auteur a été cité, les années d'activité, la puissance en Tesla des aimants.

Remarque : le travail de collecte de données s'est avéré très compliqué. En effet, de nombreux articles ne sont pas accessibles. Or, c'est dans ces publications qu'il nous était possible de trouver notre information cruciale : la puissance des aimants en Tesla.

Nous avons basé certaines de nos analyses sur l'indice H (auss appelé H index) qui nous est utile pour mesurer l'influence. Il représente le nombre d'articles H cités au moins H fois chacun. Bien que controversé, l'utilisation de l'index H reste pertinent au sein d'une communauté scientifique homogène comme celle-ci.

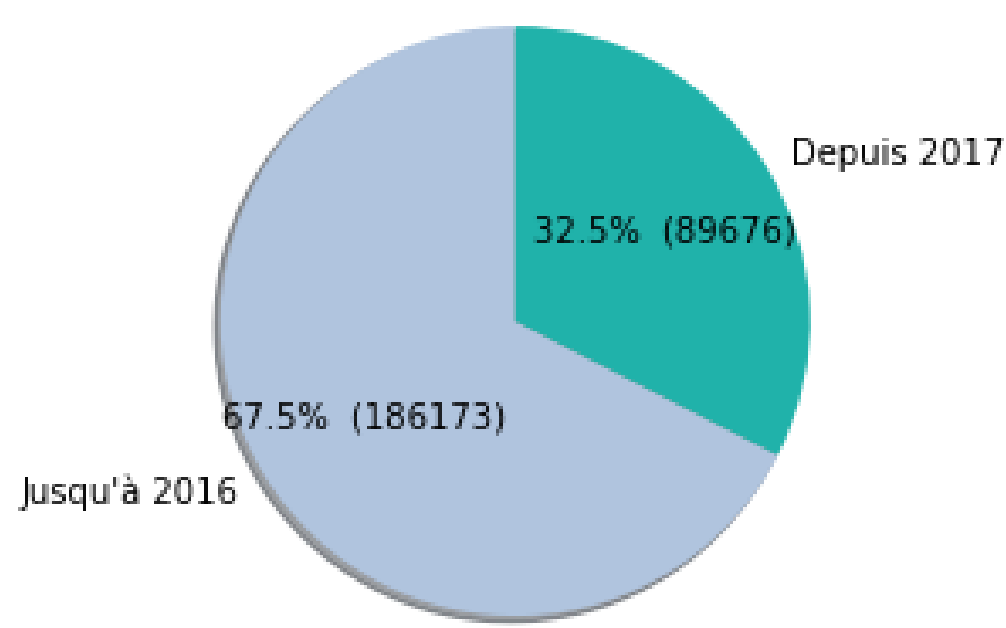
APPLICATION

Une fois la récupération de nos données, nous avons pu commencer à les analyser.



Nous avons pu observer une corrélation entre un H index élevé et la puissance de l'aimant du laboratoire (en Tesla) en obtenant un coefficient de Pearson de 0,58.

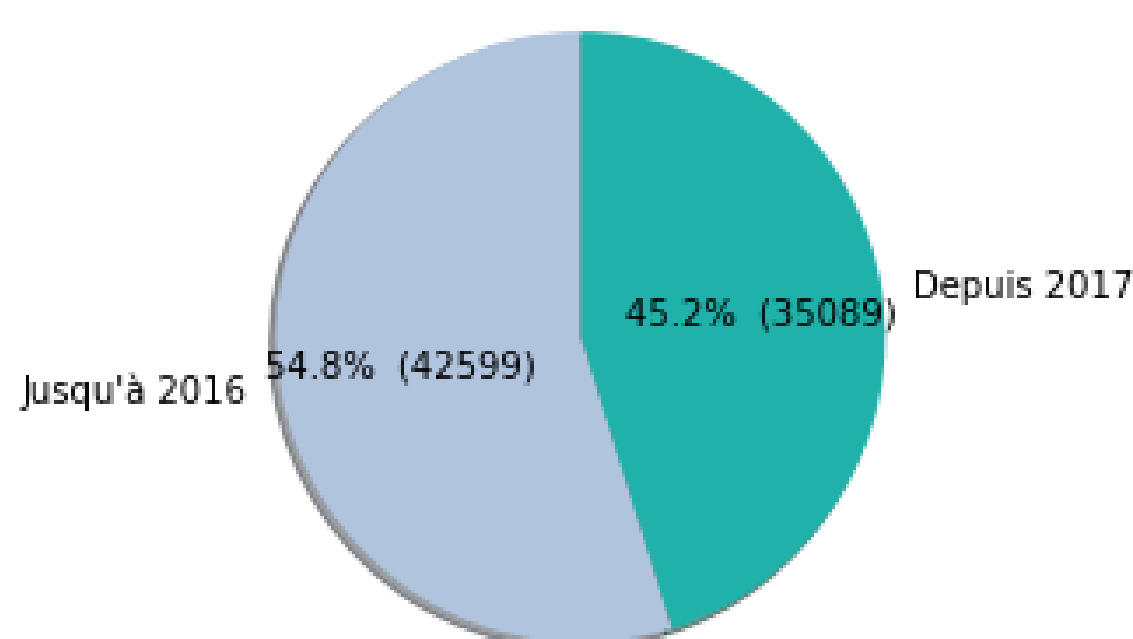
Nombre de citations jusqu'à 2016 et depuis 2017 dans le groupe ayant beaucoup de moyens



Dans les groupes ayant plus de moyens (des aimants plus puissants), nous observons qu'ils ont été cités 89 676 fois ces cinq dernières années, contre 186 173 fois auparavant. Cependant, avant 2017, les données remontent jusqu'à 1985, c'est pourquoi le nombre de citations est beaucoup plus élevé.

Il faut cependant comparer ces données à celles recueillies pour les groupes ayant moins de moyens.

Nombre de citations jusqu'à 2016 et depuis 2017 dans le groupe ayant moins de moyens

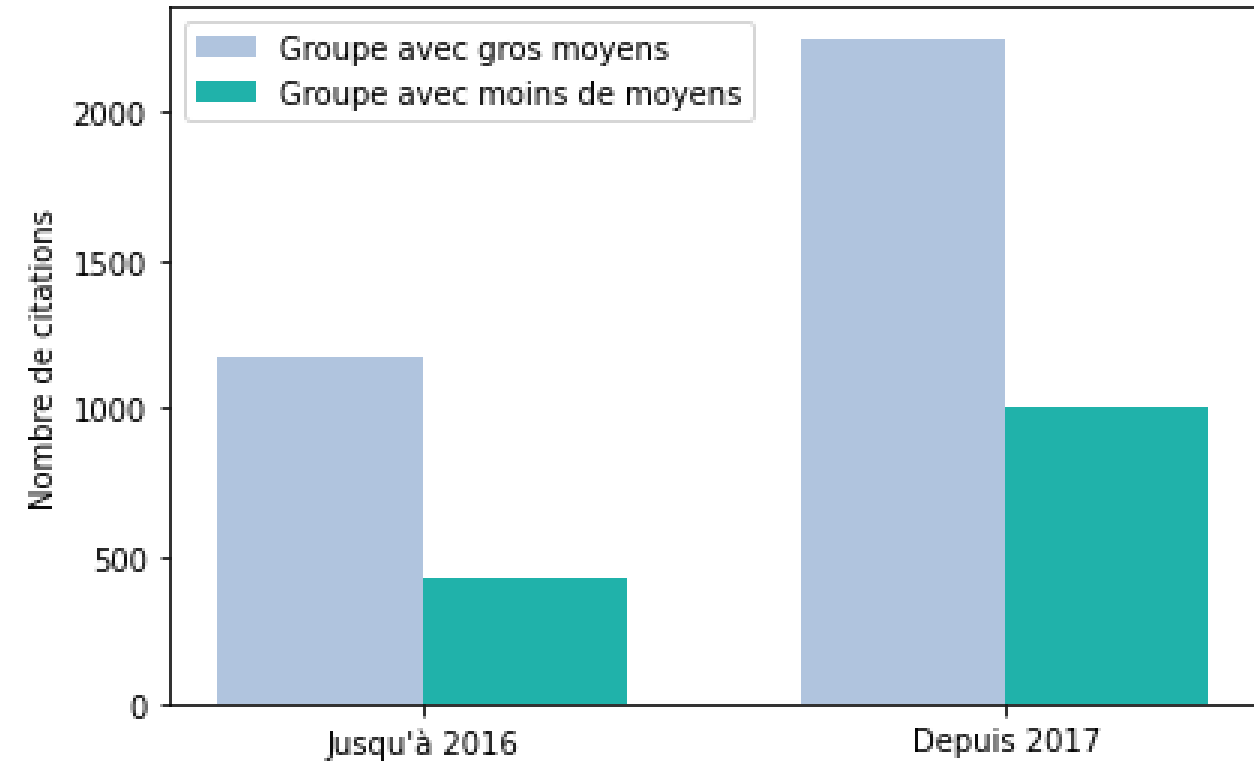


Dans les groupes ayant moins de moyens, nous observons que, bien que toujours moins cités depuis 2017, l'écart est réduit. (45.3% des citations en cinq ans !)

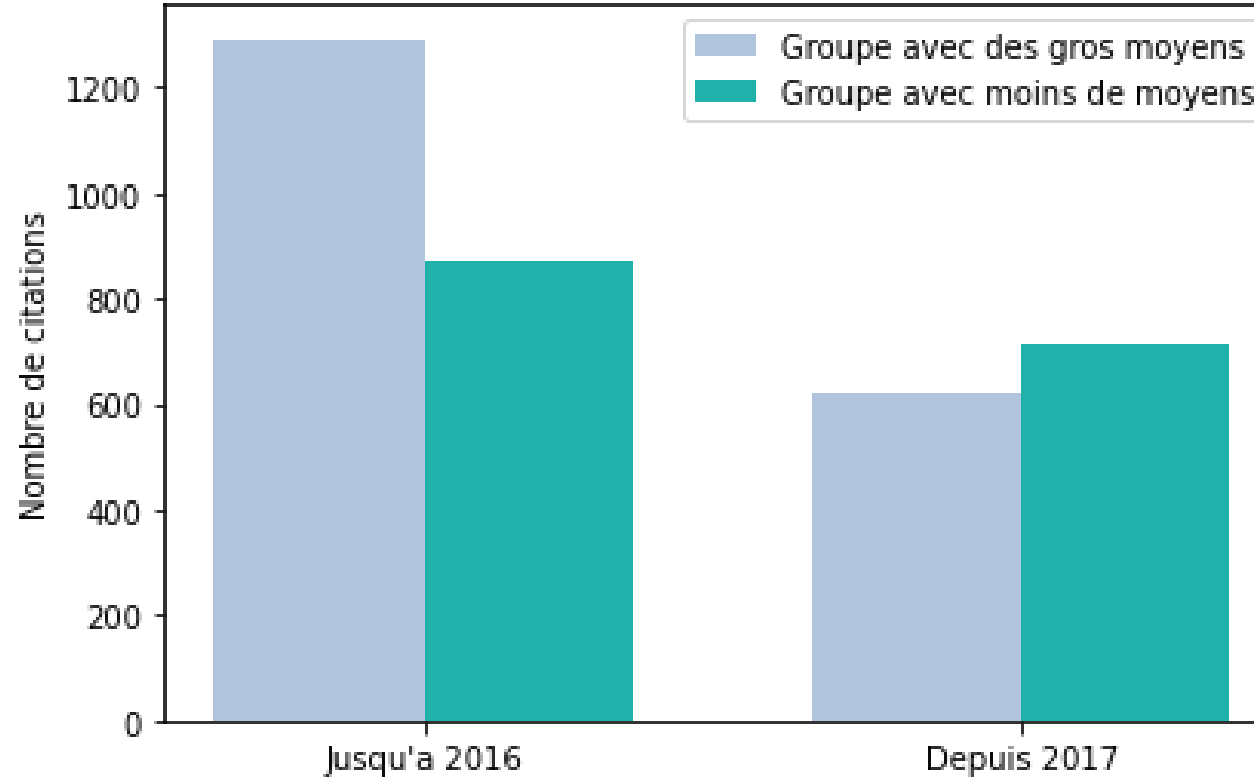
Il est possible d'expliquer cette différence de rapport de nombre de citations entre les deux groupes du fait que les carrières des chercheurs dans les groupes ayant plus de moyens ont, de manière générale, commencée plus tôt que celles des chercheurs des groupes ayant moins de moyens.

Ainsi les chercheurs appartenant aux grands groupes ont été cités de nombreuses fois toute au long de leur carrière tandis que ceux appartenant à des groupes plus modestes connaissent une augmentation bien qu'ils restent moins cités.

Nombre de citations par an et par chercheur dans chacun des groupes, avant et depuis 2017

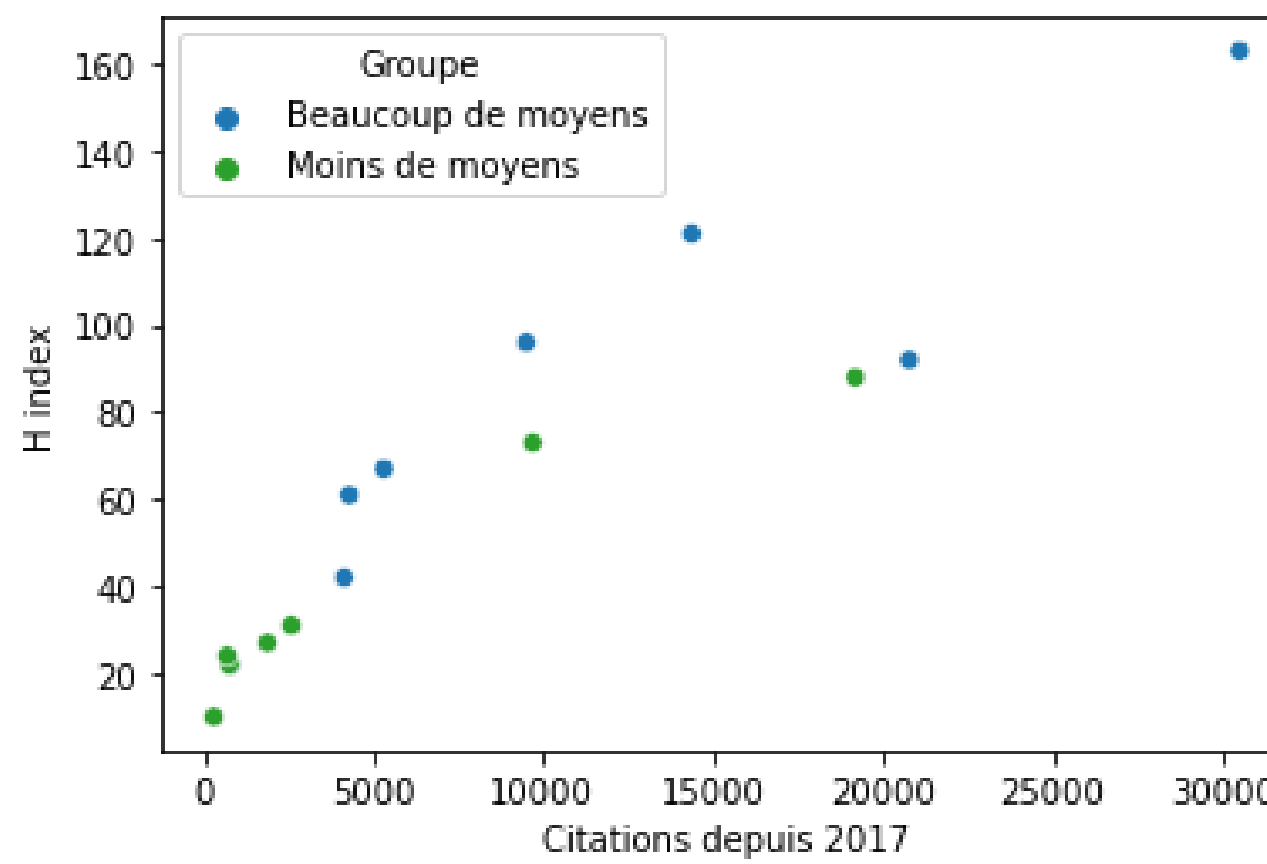
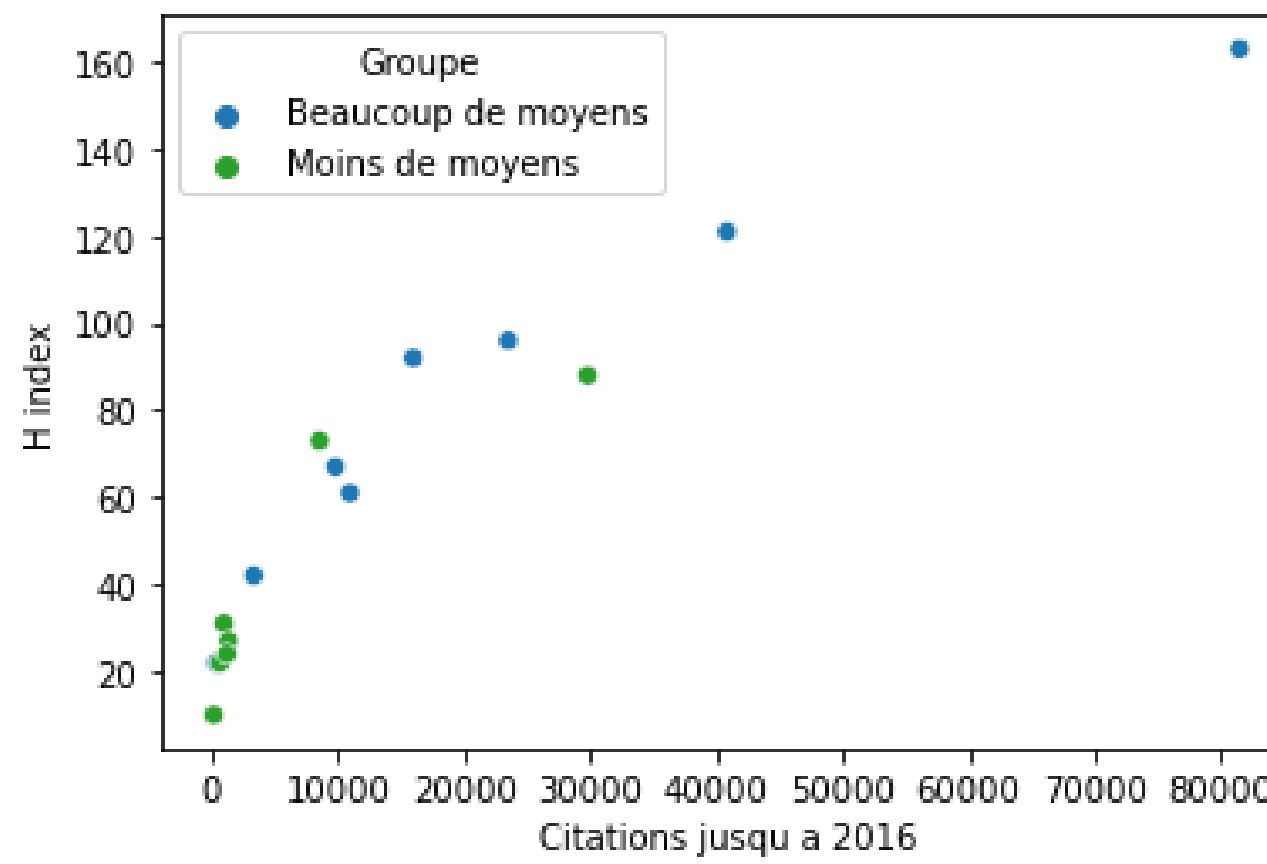


Nombre de citations par Tesla et par chercheur selon leur groupe, avant et depuis 2017



Ces graphiques ci-dessus mettent en avant le fait que les chercheurs avec moins de moyens sont capables d'obtenir d'excellents résultats lorsque nous calculons le nombre de citations qu'ils obtiennent par Tesla, rivalisant même avec le groupe des « grosses pointures » du domaine.

S'il est vrai qu'il est assez compliqué de comparer pour les citations avant 2017, puisque les technologies utilisées ont évolué à travers le temps, nous pouvons remarquer que, sur les cinq dernières années, les chercheurs avec moins de moyens ont un nombre de citations par Tesla qui est comparable, voire légèrement supérieur, à celui des chercheurs possédant les meilleurs aimants.



La courbe que forment le H index et le nombre de citations (jusqu'à 2016) est facilement visible ici.

Les chercheurs qui ont de nombreuses citations publient souvent et ont, par conséquent, un H index élevé. Ceci démontre leur influence (comme Richard D. Smith, bien connu dans le milieu tout en haut à droite du graphique).

Au final, on observe très peu de variation entre la période avant 2016 et depuis 2017.



Ryan P. Rodgers, Brice Bouyssiere, Carlos Afonso, Sebastien Tessier, and Pierre Giusti
Energy & Fuels202135 (22), 17973-17978

Alan G. Marshall ? Co-inventeur et meneur du développement de la FT-ICR. Il est considéré comme le "grand-père et le fondateur" du domaine, avec un H index de 120.

CONCLUSION

Finalement, au premier regard, il est facile de conclure que ce sont les groupes avec plus de moyens qui sont plus cités et donc ont de meilleurs résultats.

Cependant, en regardant au nombre de citation par Tesla, les chercheurs les moins bien dotés sont tout autant cités, si ce n'est plus, que ceux possédant un plus gros budget.

Comment l'expliquer ? La productivité récente est meilleure sur des petites machines car elle correspond à l'utilisation de nouvelles méthodologies. Aujourd'hui, il y a un certain abandon du gigantisme. C'est pourquoi, avec l'évolution du matériel informatique, il est possible de compenser. Les groupes avec moins de moyens ont recours à des solutions créatives.

REMARQUES

Il est intéressant de remarquer, au-delà des facteurs d'influence relatifs aux citations, que d'autres observations ont été faites grâce à cette recherche de données.

D'une part, grâce à l'évolution des technologies, une nouvelle générations de chercheurs a émergé dans les années 2000, ce qui influe fortement sur le nombre d'articles écrit - et donc du nombre de citations.

D'autre part, que le nombre de chercheuses dans les chercheurs les plus cités est bien moindre par rapport à leur collègues masculins, cependant, ce nombre a tendance à augmenter dans les dernières années (bien qu'il reste encore faible, le gender ratio étant de 5:1 en faveur des hommes).

Vous pouvez flasher ce QR code pour connaître tout notre travail en détail !

