

“Estimating the carbon footprint of computer science conferences”

Loukil Rahma - Houlné Vincent - Gonzalez Paul

Projet :

Notre projet, piloté par *Antoine Amarilli* maître de conférences en informatique à Télécom Paris, vise à évaluer les émissions de carbone engendrées par les activités de recherche en informatique ("Computer Science"). Contrairement à d'autres domaines de recherche, les chercheurs en informatique ont l'habitude de se rendre à des conférences pour présenter et valider leurs travaux avant de les publier. Cette pratique implique une participation active aux événements organisés à travers le monde, ce qui entraîne des émissions de carbone considérables. Pour quantifier ces émissions, nous exploitons la base de données DBLP qui répertorie toutes les conférences en "Computer Science". Pour chaque article présenté lors de ces conférences, nous recueillons des informations sur l'affiliation des chercheurs (pays, ville, institution) afin de calculer l'empreinte carbone de leur déplacement jusqu'au lieu de la conférence. Ce processus est itéré à l'aide d'un programme pour tous les articles présentés lors de toutes les conférences répertoriées dans la base de données, ce qui représente des milliers de données à traiter.

Nous estimons que notre projet dépasse la simple identification de l'empreinte carbone des différentes conférences. Il met en lumière l'ampleur de leur impact sur l'environnement. Cela permet de sensibiliser les différents acteurs du domaine sur l'impact écologique significatif qu'ont ces déplacements. L'absurdité est qu'il soit nécessaire d'animer des conférences pour valider ses travaux au lieu, comme les autres domaines de recherche, de les publier dans des revues scientifiques. Nous souhaitons, au-delà de calculer l'empreinte carbone, délivrer un site web et des bases de données (Conférence, lieu, article, chercheur, affiliation, empreinte carbone) traitées et ordonnées qui pourront être consultées. Nous croyons sincèrement que la communication des résultats encouragera les différents acteurs à adopter des pratiques plus responsables.

Social :

Nous avons remarqué que la recherche en informatique est très codifiée. Le fait qu'il soit admis et normalisé d'animer des conférences alors même que les articles ne sont pas publiés pose aussi des problèmes et injustices sociales. En effet, les chercheurs ne bénéficiant pas des moyens

nécessaires ou ne souhaitant pas assister à ces conférences sont pénalisés. Les conférences assurent une visibilité de leur thèse et la valorisation de la part de la communauté scientifique. De plus, ils privilégient l'accès en présentiel pour ces conférences afin de bénéficier de toute la partie échanges informels.

Recul sur notre projet et son empreinte carbone :

Pour traiter la base de données de DBLP** qui est en libre accès, nous l'avons parsé* à l'aide de programmes. Ce ne fut pas la tâche la plus longue, mais comme 95% des chercheurs n'ont pas d'affiliation, il faut passer par la base de données d'ORCID*** pour trouver les localisations manquantes. Or, la base de données zippée d'ORCID*** fait plus de 150 Go... ce qui représente dézippée une taille astronomique. La cerise sur le gâteau étant que nous sommes obligés de parcourir l'entièreté des fichiers, ce qui implique de laisser tourner un ordinateur des dizaines (si ce ne sont pas des centaines) d'heures, le temps que l'opération de dézippage finisse.

On sait qu'un ordinateur de bureau en utilisation "normale" consomme en moyenne 0.538 kWh/jour, et 0.20 kWh/jour pour un ordinateur portable. Or le programme était lancé depuis un ordinateur portable, mais il tournait sur une machine de l'école (car impossible de dézipper des fichiers d'une telle taille sur nos machines) et ce pendant des heures. In fine, la consommation d'énergie de notre projet, même si elle reste dérisoire face à l'empreinte carbone que l'on calcule, est néanmoins loin d'être nulle.

C'est le serpent qui se mord la queue. Tout comme les chercheurs qui voyagent à l'autre bout du monde pour faire des conférences sur l'impact du numérique et sa participation au dérèglement climatique, nous sommes obligés dans nos recherches de consommer de l'énergie. Malheureusement, il semble difficile de s'amputer de telles ressources.

Conclusion :

Ce projet met en évidence le paradoxe environnemental du numérique : comment s'inscrit le numérique dans la transition écologique ? Comment ne pas devenir, à l'instar de certain chercheur en informatique, ce que l'on critique ? Les états d'esprit doivent changer et les consciences s'éveiller.

***Parser** : En informatique, réaliser l'analyse de la syntaxe d'une page et en récupérer les données.

****DBLP** : est un alphanumérique archivant les auteurs et les contributeurs de la communication scientifique dans le domaine de la recherche informatique.

*****ORCID** : est un alphanumérique archivant les auteurs et les contributeurs de la communication scientifique.