



Regards d'actualité au prisme des enjeux sociétaux sur les données historisées d'EGC

Sébastien Jean Valat, Cécile Favre

► To cite this version:

Sébastien Jean Valat, Cécile Favre. Regards d'actualité au prisme des enjeux sociétaux sur les données historisées d'EGC. 20ème édition de la conférence EGC (Extraction et Gestion des Connaissances), Jan 2020, Bruxelles, Belgique. pp.205-216. hal-04881918

HAL Id: hal-04881918

<https://hal.science/hal-04881918v1>

Submitted on 13 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Regards d'actualité au prisme des enjeux sociétaux sur les données historisées d'EGC

Sébastien Valat*, Cécile Favre**,***

* ATOS

sebastien.valat@atos.net

<https://atos.net/fr/>

** Université de Lyon, Lyon 2, ERIC EA 3083 (laboratoire d'informatique)

cecile.favre@univ-lyon2.fr

<https://eric.msh-lse.fr/>

*** Université de Lyon, Lyon 2, CMW UMR 5283 (laboratoire de sociologie)

<https://www.centre-max-weber.fr/>

Résumé. Au-delà du contenu scientifique produit dans le cadre d'EGC, EGC désigne aussi une communauté scientifique composée de personnes qui s'inscrivent dans un environnement traversé par des enjeux sociétaux qui interpellent en tant que citoyen-nes. Partant de cette réalité et de la demande du défi « 20 ans d'histoire pour quel avenir ? », il a paru pertinent de nous focaliser sur comment EGC pouvait se saisir de certains de ces enjeux. Nous axons notre analyse au regard de trois enjeux : le rapport au travail, l'égalité femmes-hommes et l'écologie, invitant ainsi à une réflexion plus large sur nos postures professionnelles.

1 Introduction

La conférence EGC a été créée en 2001 et 19 ans ont passé. Elle a su fédérer autour de son existence une communauté scientifique qui rassemble des personnes dont les intérêts de recherche la définissent. Cette communauté est constituée de personnes qui contribuent scientifiquement à la conférence et qui sont par ailleurs des citoyen-nes de la société, s'insérant dans un environnement social traversé par différents enjeux sociétaux pouvant nous interpeller.

Comme le précisait Pestre (2006) dans *l'Introduction aux Sciences studies* : « Malgré ce qu'énonce la philosophie spontanée des savants, il est difficile de laisser la science dans sa tour d'ivoire et de faire comme si c'était dans l'univers clos des laboratoires que se passait l'essentiel de ce qui la concerne. [...], la question du rapport des sciences [...] à l'économie et au politique (comme aux sociabilités et à la domination masculine) ne peut être évacuée ».

Partant de là, et de la demande du défi « 20 ans d'histoire pour quel avenir ? », nous avons souhaité adopter une perspective particulière pour nous focaliser sur comment EGC pouvait se saisir de certains de ces enjeux sociétaux. En effet, dans cet article, nous ne concentrons pas l'analyse sur le contenu scientifique produit (par exemple en termes d'analyse de l'évolution de thématiques, ou de la dimension collaborative), mais nous faisons le choix d'axer notre analyse au regard de trois enjeux sociétaux qui nous paraissent fondamentaux : le rapport au travail (articulation des temps de vie), l'égalité femmes-hommes et l'écologie.

Les travaux scientifiques en sciences humaines et sociales se multiplient autour de l'épuisement au travail, de l'articulation des temps de vie, sans que le métier d'enseignant-e-chercheur-e ne soit épargné par ces problématiques, et ce dans différents pays (Sousa, 2015; Hechiche-Salah et al., 2018). L'usage du mail apparaît être comme une des composantes pertinentes à analyser. C'est ce que font d'ailleurs Chaulet et Datchary (2014) dans un article intitulé *Moduler sa connexion : les enseignants-chercheurs aux prises avec leur courriel*, analysant sur la base d'entretiens l'usage du mail. Au vu des données fournies pour le défi EGC sur les mails de la liste de diffusion, il nous a paru pertinent de faire un focus sur l'analyse des heures d'envoi de ces mails, pour traiter de cette question du temps de travail.

Par ailleurs, l'enjeu de l'égalité femmes-hommes (déclarée grande cause du quinquennat du président de la république actuel), s'inscrit dans une société française qui a vu une médiatisation croissante sur le sujet. Collet (2019) livre dans son dernier ouvrage intitulé *Les oubliées du numérique* une analyse de la place des femmes dans le numérique, de façon historisée notamment. Si les efforts menés pour attirer davantage de femmes dans le domaine de l'informatique sont de plus en plus importants, l'enjeu de leur visibilité est également présent (en termes d'égalité professionnelle, en particulier pour l'évolution de carrière par rapport au plafond de verre). Nous nous intéressons alors à la place de celles-ci dans la communauté EGC. Ceci nous a amené-es à étiqueter manuellement par rapport au sexe non seulement les données fournies sur les auteur/trices, mais aussi des données complémentaires que nous avons récoltées sur les conférences invitées, les comités de programme et les différentes présidences.

Enfin, la dimension écologique constitue aujourd'hui un enjeu fondamental, notamment du point de vue du réchauffement climatique. Alors même que la réflexion globale en matière d'écologie numérique est croissante, il s'agit d'amorcer des réflexions vis-à-vis des émissions de carbone, et le transport constitue un des domaines où des actions sont et doivent être entreprises. L'ADEME (Agence De l'Environnement et de la Maîtrise de l'Energie), établissement public sous la tutelle conjointe du ministère de la Transition écologique et solidaire et du ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, précise qu'en France il est considéré que les transports sont responsables d'1/3 des émissions de gaz à effet de serre. Dans ce papier, nous nous intéressons alors à la dimension spatiale du point de vue du déplacement des auteur/trices aux différentes éditions de la conférence EGC.

Cet article est alors organisé de la façon suivante. Dans la section 2, nous précisons le travail de récupération et de préparation des données de manière générale. Puis, dans la section 3, nous abordons la dimension travail avec l'analyse des heures d'envoi de mail. Ensuite, dans la section 4, nous analysons la place des femmes dans la communauté EGC. Nous proposons ensuite, dans la section 5, d'évoquer la dimension écologique avec l'analyse en termes de transport des auteur/trices. Enfin nous apportons une conclusion de ce travail dans la section 6.

2 Récupération et préparation des données

Notre objectif était de pouvoir exploiter le maximum de données pour traiter les axes choisis, allant au-delà parfois des données fournies. La période considérée pour les trois thématiques n'est pas systématiquement la même en fonction des données accessibles.

Pour la dimension travail, les mails fournis vont de juillet 2006 à septembre 2018.

Sur la dimension de la place des femmes, nous avons cherché manuellement, pour compléter les données fournies, des données sur le site Web d'EGC et les sites Web des éditions. Nous

avons récolté ainsi les personnes ayant présenté les conférences invitées, la composition des comités de programme et les différentes présidences : les personnes invitées à la présidence d'honneur et celles qui ont présidé les comités de programme et d'organisation, nous permettant de faire une analyse de 2001 à 2019 inclus. Un important travail a été réalisé pour procéder à un étiquetage manuel de la variable sexe. L'objectif était d'avoir des données précises, là où les algorithmes d'étiquetage automatique dédiés ne produisent pas un résultat à l'unité près.

Concernant la dimension écologique, celle-ci a nécessité un travail important de préparation en développant un processus le plus automatisé possible. Il s'agissait de partir des documents PDF eux-mêmes pour accéder à l'affiliation des auteur/trices. Ainsi, compte-tenu de la disponibilité des données sur RNTI, cette analyse démarre à partir de 2006 et va jusqu'en 2019. La préparation de ces données nécessitant davantage de développements informatiques, des éléments seront précisés dans la partie dédiée à cette thématique en section 5.

3 Focus rapport au travail

Le droit à la déconnexion est une disposition issue de la loi El Khomri de 2016, dite « loi Travail ». Elle est entrée en vigueur le 1er janvier 2017. Un des objectifs de cette loi est d'adapter le droit du travail à l'ère du numérique. Concernant en particulier le droit à la déconnexion, l'objectif est de permettre aux salarié-es de concilier vie personnelle et vie professionnelle, tout en luttant contre les risques de burn out. Pour cela, ils/elles doivent avoir la possibilité de ne pas se connecter aux outils numériques et de ne pas être contacté-es par leur employeur en dehors de leur temps de travail (congrés payés, jours de RTT, week-end, soirées...).

Ce droit à la déconnexion concerne toutes les personnes salariées. Cependant, il est à noter que cette loi ne s'applique pas dans la fonction publique de façon réglementaire. Des initiatives peuvent néanmoins être prises dans les institutions. La question se pose alors sur la capacité à respecter ce temps de déconnexion, dans des environnements dont le temps de travail peut être une variable assez « élastique », et c'est le cas pour beaucoup d'enseignant-es-chercheur-es.

Le jeu de données de la liste de diffusion EGC comporte 7075 mails écrits entre le lundi 10 juillet 2006 et le dimanche 9 septembre 2018, provenant de 811 adresses mails différentes.

Pour travailler sur les horaires, nous avons utilisé le champ *date* des emails considérant qu'il reflète normalement l'heure locale ¹ pour la personne qui envoie le mail, contrairement au champ *received* donnant la date du côté serveur SMTP potentiellement distant.

Nous avons choisi de considérer l'envoi de mails dans le cadre professionnel de 8h à 20h comme « classique ». Nous avons pu dénombrer 1116 mails écrits en dehors de cette plage horaire, soit 16% des mails. Ceci est une trace du travail non négligeable effectué en dehors d'horaires classiques de travail, car si les mails aux listes de diffusion peuvent être considérés marginaux dans l'activité professionnelle, ces derniers sont la trace de cette activité.

Dans la figure 1, nous pouvons constater que l'augmentation de l'activité sur la liste de diffusion va de paire avec une augmentation du pourcentage de mails envoyés hors horaires classiques, qui oscille à présent autour de 15%.

Pour aller plus en détails, nous avons découpé la période non classique en 3 créneaux : de 20h à minuit, de minuit à 4h et de 4h à 8h. Les résultats sont fournis dans la figure 2. Nous pouvons y observer que la tranche 20h-minuit est celle qui est la plus utilisée, correspondant

1. en tenant compte du fuseau horaire local selon la norme RFC 4021, sous réserve que les mails envoyés ne l'aient pas été durant un déplacement à l'étranger avec décalage horaire

Regards d'actualité sur les données historisées d'EGC

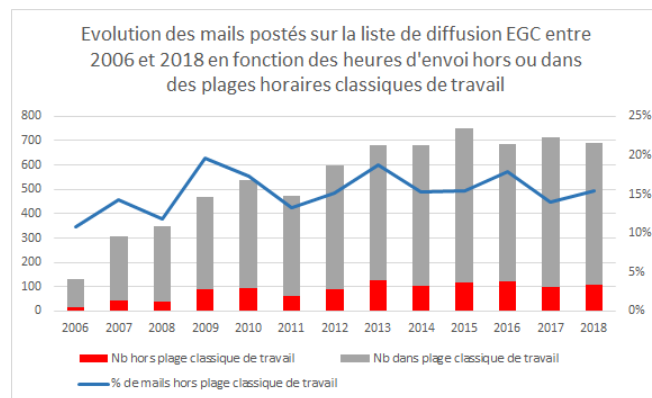


FIG. 1 – Évolution des mails postés sur la liste de diffusion EGC en fonction des plages horaires dédiées ou non classiquement au travail.

à une activité de travail en soirée. Parmi les 1116 mails hors créneaux classiques, 70% ont été écrits de 20h à minuit ; 15% de minuit à 4h et 15% de 4h à 8h. Ce travail en dehors d'horaires classiques doit nous interpeller sur le rapport au travail et notre organisation du temps de travail.

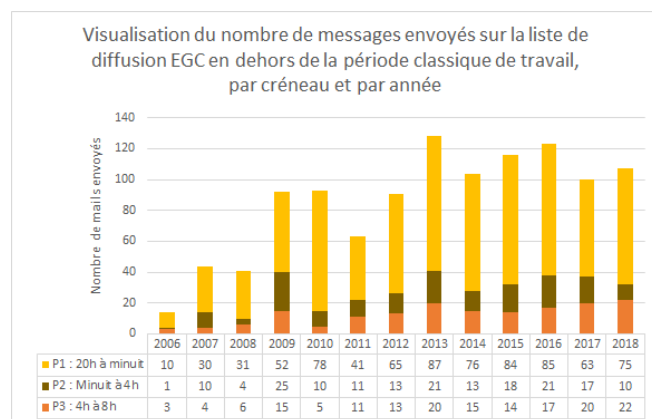


FIG. 2 – Evolution des mails postés sur la liste de diffusion EGC en fonction des plages horaires et des années pour les mails postés en dehors des plages classiques de travail.

4 Focus égalité femmes-hommes

La place des femmes dans l'enseignement supérieur et la recherche² reste une thématique d'actualité. Lors du défi 2016 sur les données d'EGC jusqu'à 2015, un premier travail avait

2. Rédaction en 2014 d'un Livre Blanc « Le genre dans l'enseignement supérieur et la recherche » par des chercheur-es de l'ANEF (Association nationale des études féministes)

été réalisé autour des femmes dans la communauté EGC (Cabanac et al., 2016). Il s'agit ici de reprendre des éléments d'analyse en prenant en compte les années depuis ce précédent défi et porter notre attention sur des aspects qui n'avaient pas été traités précédemment.

Tout d'abord, pour recontextualiser l'analyse, il s'agit de préciser que la communauté EGC concerne des personnes qui seraient plutôt catégorisées en sections 27 et 26 en terme de nomenclature CNU. D'après les dernières fiches démographiques disponibles, soit pour 2017-2018, fournies par le ministère³, la section 27 comprenait 24% de femmes et la section 26 comprenait 27% de femmes (PR et MCF confondus). Ceci peut servir de base de comparaison sur la composition actuelle en matière de représentation sexuée, positionnant ainsi les résultats en perspective du « vivier » d'enseignant-es-chercheur-es, même si les statuts de chercheur-es (CR et DR) ne sont pas considérés dans ces données du ministère.

4.1 Les auteurs et autrices d'EGC

Sur 19 éditions, parmi les co-auteur/trices des papiers (toutes catégories confondues qui sont présents dans les actes), nous trouvons 27% de femmes en moyenne.

Dans la figure 3, nous pouvons observer l'évolution de cette répartition sexuée en nombre et en pourcentage. Le pourcentage a oscillé entre 21 et 31%. Ainsi, nous pouvons considérer que la communauté EGC est dans la moyenne de ce qui s'observe en termes de répartition sexuée des sections CNU concernées.

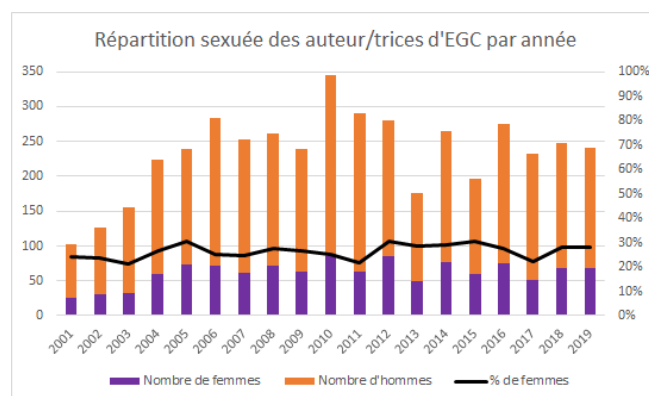


FIG. 3 – *Représentation sexuée des auteur/trices sur la période 2001-2019.*

Considérons à présent les personnes qui ont le plus contribué à la conférence sur les 19 ans, en nombre de papiers (quelle que soit la taille des papiers et moyennant les choix d'édition de certaines années⁴). La personne qui a le plus de papiers dans EGC sur la période 2001-2019 est un homme. Son premier papier remonte à la 2ème édition d'EGC, et il a en tout publié 37 papiers (hors papier extension pour la version étendue à l'international).

Dans la figure 4, nous pouvons observer la représentation sexuée des personnes ayant le plus contribué à EGC. Si l'on prend en compte les personnes ayant au moins 10 papiers sur les

3. <https://www.enseignementsup-recherche.gouv.fr/cid85019/fiches-demographiques-des-sections-du-cnu.html>

4. Par exemple, lors de certaines éditions, les actes n'ont pas comporté les papiers « démonstration de logiciel »

19 éditions, elles sont 48, dont 12 femmes (soit 25%). Nous restons donc sur les mêmes ordres de grandeur, ce qui montre que les femmes font bien partie aussi des personnes contribuant de façon importante et régulière à EGC.

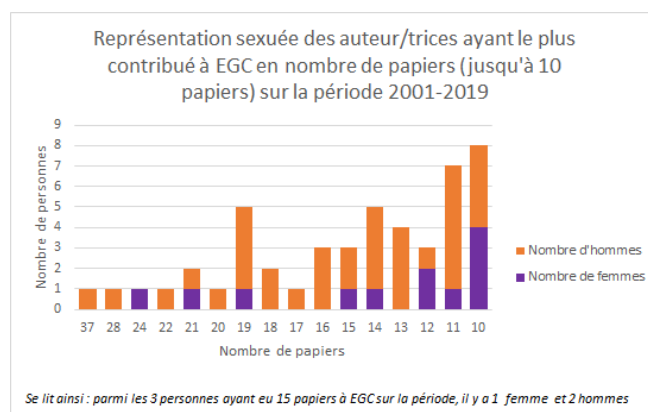


FIG. 4 – Représentation sexuée des auteur/trices ayant le plus contribué sur 2001-2019.

4.2 Présidences

Un autre aspect intéressant est le travail et la visibilité qui découlent du rôle dans une présidence. Nous avons recensé les présidences d'honneur, de comité de programme (CP) et d'organisation (CO). Dans la figure 5, nous pouvons observer le nombre de femmes et d'hommes impliqués. Ainsi, le pourcentage de femmes est de 17% pour la présidence d'honneur, 24% pour la présidence de comité de programme et 33% pour la présidence de comité d'organisation. Il est à noter que, parfois, des binômes de personnes ont pu être constitués, soit pour la gestion du comité de programme (surtout dans les premières éditions), soit pour la gestion de l'organisation. Si l'on considère le « prestige » pouvant être associé aux différents types de présidence, la dimension scientifique pouvant être globalement plus valorisée, il est intéressant de noter que l'on retrouve davantage de femmes sur l'aspect organisationnel des conférences, qui pourrait être mis en perspective d'analyses sociologiques sur la question des tâches et du genre.

4.3 Membres du comité de programme

Nous nous sommes également penché-es sur la présence des femmes dans le comité de programme. En étude sur le genre, un des indicateurs quantitatifs utilisés est le rapport de masculinité. Il est exprimé en nombre d'hommes pour 100 femmes (à la naissance, il est classiquement de 105 garçons pour 100 filles). Dans la figure 6, nous constatons que, depuis 2015, ce rapport de masculinité a diminué jusqu'à 200 hommes pour 100 femmes, tendant vers un meilleur équilibre sexué sur la place des femmes dans le comité de programme. En 2019, cela représente 34% de femmes dans le CP, c'est une manière de reconnaître leur compétence.

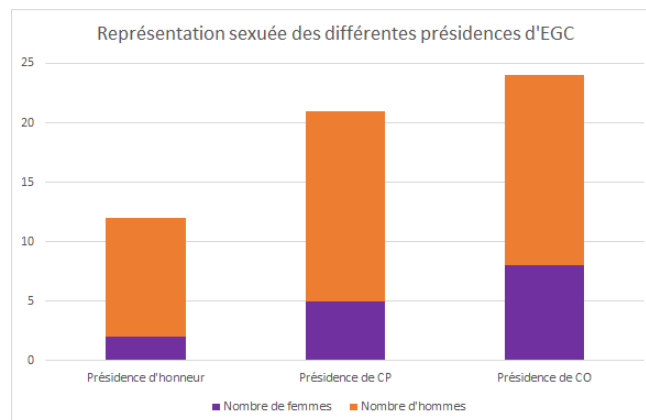


FIG. 5 – Représentation sexuée pour la présidence d'honneur, la présidence du Comité de Programme et la présidence du Comité d'Organisation.

4.4 Où sont les conférencières invitées ?

Dans les carrières académiques, la visibilité / le rayonnement scientifique sont des éléments importants, notamment du point de vue de l'avancement dans la carrière, ou l'accès à des primes telles que la PEDR (Prime d'Encadrement Doctoral et de Recherche).

Il apparaît que l'invitation pour présenter une conférence invitée dans le cadre d'un colloque peut contribuer à cet enjeu de visibilité. Ainsi, bien que cet aspect ne figure pas dans les données initiales fournies pour le défi, nous avons choisi de les collecter manuellement à la fois via le site Web EGC qui récapitule partiellement ces éléments dans le menu dédié⁵, et nous avons complété les données qui s'avéraient lacunaires par une recherche dans les publications, compte-tenu du fait que les conférences invitées font l'objet d'un résumé d'une page maximum dans les actes de la conférence. Pour l'année 2001 et 2002, nous n'avons pu être en mesure de déterminer l'éventuelle présence de personnes donnant une conférence invitée, supposant que sur les deux premières éditions, il était possible qu'il n'y en ait pas eu.

Ainsi, entre 2003 et 2019, 65 personnes se sont succédées en session plénière pour présenter une conférence invitée, selon la répartition par année présentée dans la figure 7. Parmi ces personnes, nous notons la présence de 16 femmes et 49 hommes, soit 25% de femmes. Compte-tenu du pourcentage de femmes dans la communauté, il y a une certaine forme de représentativité. Mais nous pourrions regretter que, dans certains cas, il y ait eu entre 3 et 5 personnes en conférence invitée, sans aucune femme (4 éditions concernées).

La présence en conférence invitée peut répondre à différentes logiques (y compris à celle du refus des personnes invitées) mais il apparaît nécessaire d'avoir une vigilance sur ce point, d'autant plus que la conférence EGC ne se limite pas à une seule invitation par édition.

5. <https://www.egc.asso.fr/category/publications/conferences-invitees>

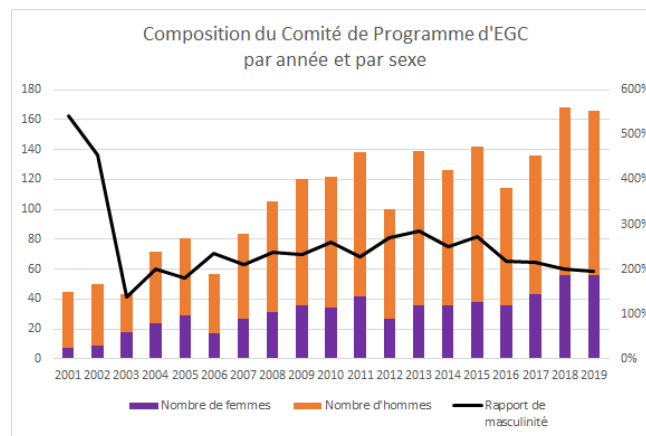


FIG. 6 – Composition sexuée du Comité de Programme par année et rapport de masculinité.

5 Focus environnement

Au fil des années qui passent, l'environnement prend une place de plus en plus importante dans les débats questionnant l'organisation de nos sociétés. Parmi les préoccupations actuelles, on dénote sur un plan politique les émissions de gaz à effets de serre, principalement le CO₂.

Dans le domaine de la recherche, une part des trajets à fortes émissions est en général attribuée aux conférences avec les déplacements nécessaires vers une zone géographique souvent éloignée de son lieu de travail commun. Différentes initiatives ou journées de réflexion⁶ émergent sur ce sujet et se structurent petit à petit pour questionner la manière d'exercer une profession mais également la dimension écologique de l'informatique plus largement.

EcoInfo⁷ est un groupement de services du CNRS qui a pour objectif général d'évaluer puis de réduire les impacts de l'informatique, notamment au sein de l'enseignement supérieur et de la recherche. Ce groupe a analysé ses propres émissions montrant la prépondérance des déplacements dans leurs émissions (EcoInfo, 2019).

Dans le cadre du défi EGC 2020, nous souhaitons pouvoir disposer d'une estimation en termes de distance et d'émission de carbone liées au déplacement pour la conférence. N'ayant pas accès aux participant-es de chaque édition, il s'agit pour nous d'aller vers le calcul d'un minimum en s'appuyant sur l'hypothèse que le déplacement a été effectué par la personne en première position sur le papier. Cette hypothèse de travail est nécessaire, même si elle peut être forte dans certaines situations où le papier est la production d'une collaboration entre des personnes d'organismes distants géographiquement. L'objectif visé n'est pas ici une estimation exacte, notamment car hormis les auteur/trices des papiers, d'autres personnes participent à la conférence, mais bien d'avoir une première idée des kilomètres parcourus et de ce que cela peut générer en termes d'émissions de carbone "au moins", en lien avec l'historique d'EGC.

En 2016, la publication de Kergosien et al. (2016) mentionnait une distribution des distances des auteur/trices au lieu de la conférence en ayant manuellement rempli leur base d'au-

6. <http://cedd-pes.com/nos-activites-2019/les-chercheurs-prennent-trop-lavion/>

7. <https://ecoinfo.cnrs.fr/>

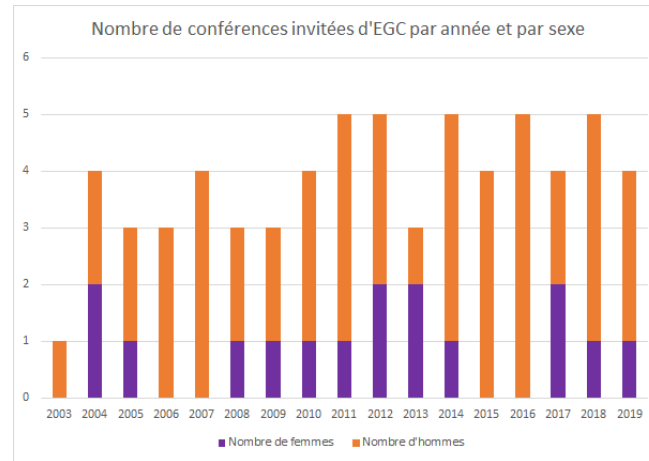


FIG. 7 – Historique des conférences invitées : représentation sexuée.

teur/trices et ville du laboratoire de rattachement sans, par ailleurs, mettre en perspective la dimension écologique. Ici, nous nous proposons non seulement d’automatiser ce processus afin d’obtenir les distances parcourues par les personnes en première position sur les papiers pour l’ensemble du corpus de publications EGC depuis 2006, et d’aller par ailleurs vers une estimation totale minimale d’émissions de carbone en se basant sur des estimations kilométriques par type de transport (exprimées dans l’unité gCO₂e/passager.km : grammes CO₂ équivalent par passager et par kilomètre).

5.1 Méthodologie d’extraction

Le jeu de données en entrée est constitué des informations DBLP sur les publications d’EGC et les PDF associés extraits de RNTI. La première étape consiste à extraire le texte de ces PDF pour obtenir les laboratoires de rattachement des auteur/trices. Afin de procéder à cette extraction, nous pouvons nous baser sur pdfminer. Cette démarche ne donne toutefois un résultat valide que pour 700 publications sur les 1390 analysées. Pour compenser ce manque, nous avons appliqué une méthode additionnelle sur les publications problématiques. La méthode la plus fiable que nous ayons trouvée consiste à générer une image de la première page de la publication avec la commande ‘convert’ fournie par ‘image magick’. Une fois cette image obtenue avec une résolution de 300dpi pour qu’elle soit suffisante, nous avons appliqué une reconnaissance de caractères avec l’outil ‘tesseract’ rendu open source par HP il y a quelques années. Sur nos tests, cet outil donne de bien meilleurs résultats que l’habituel GOCR fourni par la communauté GNU, avec beaucoup moins d’erreurs dans les fichiers textes générés.

Pour s’assurer d’une bonne découpe des auteur/trices, nous effectuons une correspondance avec les informations fournies par les fichiers JSON de la base DBLP, ce qui permet de détecter les erreurs liées aux publications où les informations d’auteur/trices sont mal formatées. Ces dernières ont été corrigées à la main (environ 300) pour pouvoir être repassées par les scripts automatiques. Les problèmes sont généralement liés aux virgules entre auteur/trices et caractères étoile (liant auteur/trice et laboratoire) mal reconnus.

5.2 Gestion des distances

Grâce à l'étape précédente, nous avons obtenu le laboratoire des auteur/trices et, pour ce qui nous intéresse, la personne en première position sur l'article. À partir de cette information, nous pouvons utiliser les services de Google pour géolocaliser le laboratoire de manière automatique. Ceci nous donne une extraction viable de 916 publications sur les 1390 qui constitue notre base de travail pour construire les estimations minimales des métriques qui suivent. Les erreurs proviennent soit d'une mauvaise extraction des laboratoires depuis les papiers, soit d'un échec d'extraction de localisation reporté par Google. Ceci représente une perte de 34% qui est relativement importante mais pourra être amélioré en ajoutant des fix automatiques sur les erreurs communes et récurrentes. Les taux de succès sont présentés dans le tableau 1.

Année	Nb traités	Nb publiés	% traitement	Année	Nb traités	Nb publiés	% traitement
2006	84	102	82%	2013	46	56	82%
2007	77	92	84%	2014	67	87	77%
2008	83	93	89%	2015	46	67	69%
2009	74	82	90%	2016	61	84	73%
2010	88	115	77%	2017	60	73	82%
2011	84	100	84%	2018	53	71	75%
2012	44	59	75%	2019	49	69	71%

TAB. 1 – *Traitement des papiers pour le calcul de distances sur la période 2006-2019*

Nous pouvons sommer les distances associées aux publications par année et les comparer à la circonférence terrestre qui est d'environ 40 075 km (Moureau et Brace, 2000). Les résultats sont reportés dans le graphique 8. Toutes les distances supérieures à 800 km ont été vérifiées, nous obtenons donc un minimum considérant qu'il y a des publications non traitées, ainsi que l'hypothèse selon laquelle la personne en première position sur le papier fait le voyage.

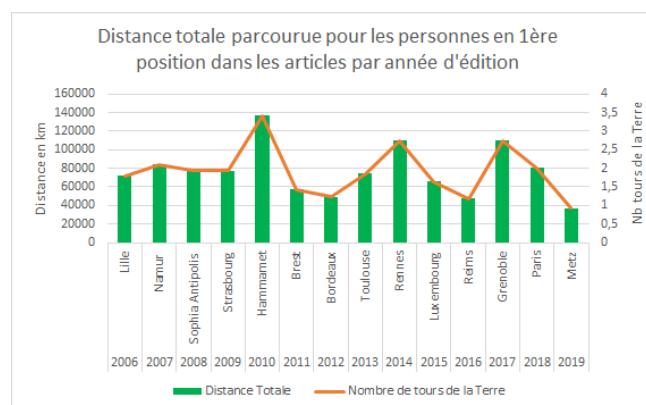


FIG. 8 – *Distance globale parcourue par édition, en km et en nombre de tours de la Terre.*

5.3 Empreinte carbone

Dans cette analyse, nous avons cherché une estimation minimale des émissions liées aux éditions EGC de 2006 à 2019 à des fins de sensibilisation. Pour parvenir à cette évaluation,

nous avons été amené-es à simplifier la manière dont sont traités les trajets d'un point de vue de leur estimation selon les informations accessibles. Ainsi, deux modes de transport sont pris en compte : le train et l'avion. Nous considérerons (certes de façon arbitraire) le train si une gare est trouvée à moins de 50 km du lieu de travail et du lieu de la conférence ; sinon, le moyen de transport considéré sera l'avion. Pour le train, le trajet considéré est celui entre les deux gares, le trajet pour aller à la gare étant ignoré. Pour le train, nous considérerons le type TGV (3,69 gCO₂e/passager.km, source eco-info/ADEME). Les TER consommant davantage (8,91 gCO₂e/passager.km) seront ignorés en cherchant une valeur minimale des émissions. Concernant les avions, nous prendrons les bases de l'aviation civile⁸ en considérant 2 types de trajets : ceux de plus de 1000km et ceux de moins de 1000km, les consommations à considérer étant différentes selon la catégorie. Nous prendrons donc pour ces deux catégories une moyenne de type Paris/Marseille (117 gCO₂e/passager.km) et Paris/New York (82 gCO₂e/passager.km) respectivement. Cette estimation prendra en compte la distance entre le laboratoire et le lieu de la conférence, approximant le trajet aux aéroports par ces mêmes moyennes.

Pour les trajets en train, une estimation plus juste peut être obtenue en utilisant les données ouvertes et API publiques de la SNCF, avec la liste des gares et les API d'obtention des horaires avec les émissions associées. Cette méthode n'a pu être appliquée que sur 500 des publications. Ainsi, nous nous limitons ici à l'estimation avec les moyennes par moyen de transport.

Les résultats donnent une estimation basse de 105 tonnes de CO₂ équivalent émises pour le total des trajets considérés. Le maximum à Tunis donnant 15 tonnes émises et les années les deux plus basses donnant respectivement 2 tonnes en 2004 (Clermont Ferrand) et 3 tonnes en 2019 à Metz. Nous notons qu'une prise en compte des trajets réels effectués en train pour les extractions faites avec succès ne change pas significativement ces résultats, les nombres étant dominés par les trajets en avion.

6 Conclusion

Dans ce papier, nous avons choisi de porter un regard sur les données EGC au travers d'enjeux sociétaux d'actualité. Ceci constitue l'originalité de ce papier, même si les méthodes mobilisées ne sont pas nécessairement complexes. Les analyses ainsi réalisées sur les questions de la déconnexion dans le rapport au travail, de l'égalité femmes-hommes et d'enjeux environnementaux au travers de l'empreinte carbone fournissent des éléments éclairants.

La récupération de données autres que celles fournies, de façon quasi automatique (dimension géographique) et manuelle (dimension égalité femmes-hommes) a nécessité un temps de travail important, restreignant les analyses menées qui pourraient donc être étendues.

Ainsi, en perspectives, des analyses pourraient être développées. Par exemple, sur le rapport au travail, une analyse centrée personne plutôt que sur le mail d'envoi serait pertinente, pour aller vers des éléments fins sur les profils de personnes (des adresses mails différentes peuvent correspondre à une seule et même personne : changement du nom de domaine de l'institution ou du nom en cas de mariage par exemple). En outre, nous avons centré notre analyse sur l'information des horaires d'envoi, une analyse au niveau des week-ends et des périodes de vacances serait intéressante par rapport à la définition du droit à la déconnexion. Concernant la dimension écologique, différents scénarios d'estimations pourraient être envisagés.

8. <https://eco-calculateur.dta.aviation-civile.gouv.fr/>

Pour finalement aller dans le sens d'une réponse au défi « EGC 2020 : 20 ans d'histoire pour quel avenir ? », les analyses fournies dans ce papier ne peuvent qu'inviter les chercheur-es de la communauté EGC et leurs responsables à se saisir des enjeux sociétaux qui traversent notre métier, notre posture professionnelle, à des niveaux à la fois individuel et collectif. Préserver son temps de sommeil, penser à (et agir pour) la place des femmes dans la communauté en termes de visibilité, faire le choix de l'écologie en matière de destination de conférences, autant de pistes qui nous amènent à une réflexivité sur nos manières d'exercer notre métier pour construire l'avenir d'EGC selon des choix conscients, et pourquoi pas engagés ?

Références

- Cabanac, G., G. Hubert, H. D. Tran, C. Favre, et C. Labbé (2016). Un regard lexicométrique sur le défi EGC 2016. In *EGC 2016*, Reims, France, pp. 419–424.
- Chalet, J. et C. Datchary (2014). Moduler sa connexion : les enseignants-chercheurs aux prises avec leur courriel. *Réseaux* 4/2014(186), 105–140.
- Collet, I. (2019). *Les oubliées du numérique. L'absence des femmes dans le monde digital n'est pas une fatalité*. Le Passeur.
- EcoInfo (2019). Quelle est notre propre empreinte carbone ?
<https://ecoinfo.cnrs.fr/2019/07/02/quelle-est-notre-propre-empreinte-carbone/>.
- Hechiche-Salah, L., E. G.-B. Ouerdian, T. Yahmadi, et S. B. Othman (2018). Quand le stress professionnel dégénère en souffrance au travail : cas des enseignants-chercheurs tunisiens. *@GRH* 2/2018(27), 57–82.
- Kergosien, E., M.-N. Bessagnet, C. Sallaberry, A. Le Parc-Lacayrelle, et A. Royer (2016). Analyse géographique de séries de publications : application aux conférences EGC. In *EGC 2016*, Reims, France, pp. 371–382.
- Moureau, M. et G. Brace (2000). *Dictionnaire des sciences de la terre (anglais-français, français-anglais)*. Editions Technip.
- Pestre, D. (2006). *Introduction aux Science Studies*. La Découverte, coll. « Repères ».
- Sousa, A. P. D. (2015). Productivisme et souffrance chez les enseignants-chercheurs au Brésil. *Pensée plurielle* 1/2015(38), 45–66.

Summary

Beyond the scientific content produced in the framework of the EGC conference, EGC is also a scientific community composed of people who are part of a social environment crossed by different societal issues, that challenge as citizens. Starting from this reality, as well as from the challenge of "20 years of history for which future?", it seemed appropriate to focus on how EGC could address some of these issues. We focus our analysis on three issues: the relationship to work, gender equality and ecology, inviting us to a broader reflection on our professional postures.