



HAL
open science

Analyse bibliométrique pour saisir la place de la fréquentation des espaces naturels dans le champ de recherche des Informations Géographiques Volontaires

Robin Lesné, Marc Langenbach, Pascal Mao, Hugues François, Nicolas Robinet

► To cite this version:

Robin Lesné, Marc Langenbach, Pascal Mao, Hugues François, Nicolas Robinet. Analyse bibliométrique pour saisir la place de la fréquentation des espaces naturels dans le champ de recherche des Informations Géographiques Volontaires. *Revue de Géographie Alpine / Journal of Alpine Research*, 2024, 111 (3), pp.1-19. 10.4000/rga.12517 . hal-04460729

HAL Id: hal-04460729

<https://hal.science/hal-04460729v1>

Submitted on 24 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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Analyse bibliométrique pour saisir la place de la fréquentation des espaces naturels dans le champ de recherche des Informations Géographiques Volontaires

Robin Lesné, Marc Langenbach, Pascal Mao, Hugues François et Nicolas Robinet



Édition électronique

URL : <https://journals.openedition.org/rga/12517>

DOI : 10.4000/rga.12517

ISSN : 1760-7426

Traduction(s) :

Bibliometric Analysis to Understand the Place of the Visits of Natural Areas in the Field of Volunteered Geographic Information Research - URL : <https://journals.openedition.org/rga/12550> [en]

Éditeur :

UGA Éditions/Université Grenoble Alpes, Association pour la diffusion de la recherche alpine

Ce document vous est fourni par INRAE Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement



Référence électronique

Robin Lesné, Marc Langenbach, Pascal Mao, Hugues François et Nicolas Robinet, « Analyse bibliométrique pour saisir la place de la fréquentation des espaces naturels dans le champ de recherche des Informations Géographiques Volontaires », *Journal of Alpine Research | Revue de géographie alpine* [En ligne], 111-3 | 2023, mis en ligne le 01 février 2024, consulté le 24 janvier 2025. URL : <http://journals.openedition.org/rga/12517> ; DOI : <https://doi.org/10.4000/rga.12517>

Ce document a été généré automatiquement le 15 avril 2024.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

Analyse bibliométrique pour saisir la place de la fréquentation des espaces naturels dans le champ de recherche des Informations Géographiques Volontaires

Robin Lesné, Marc Langenbach, Pascal Mao, Hugues François et Nicolas Robinet

NOTE DE L'AUTEUR

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme « France 2030 » portant la référence ANR-15-IDEX-0002 via le LabEx ITTEM.

Les auteurs tiennent à adresser leurs remerciements aux stagiaires (Eloïse Flachet, Benjamin Gaudey, Thomas Hurez, Yaelle Michard et Théo Pastrello) investis au sein du projet de recherche IGVLab.

Introduction

De plus en plus, les personnes fréquentant les espaces naturels, les sportifs en premier titre (Mao & Obin, 2018), s'équipent, mobilisent des dispositifs connectés et interagissent et se structurent dans des collectifs virtuels. Ils produisent ainsi des traces numériques, c'est-à-dire des empreintes laissées derrière soi en naviguant sur Internet ou en ayant recours à un service à partir d'un terminal numérique. Ces traces disposent de contenu informationnel, expérientiel et/ou spatial, devenant alors des traces géonumériques (Mericskay & Roche, 2011; Mericskay *et al.*, 2018), ou

Informations Géographiques Volontaires (IGV). L'exploitation de ces données dessine de nouvelles perspectives pour la recherche, par exemple en matière de mobilité urbaine (Le Breton *et al.*, 2022), y compris sportive (Musakwa & Selala, 2016 ; Dadashova *et al.*, 2020 ; Hong *et al.*, 2020 ; Kyuhyun & Ipek Nese, 2021 ; Nelson *et al.*, 2021 ; Venter *et al.*, 2021), ou de comportements spatiaux des touristes en milieu urbain (Mondo *et al.*, 2020). Ce web collaboratif donne accès à de grandes quantités de données documentant des phénomènes par ailleurs difficilement observables (Houllier & Merilhou-Goudard, 2016), notamment en montagne du fait de la dispersion et de la faible densité des usages. Or la connaissance de cette fréquentation est indispensable pour la gestion de cet espace.

- 1 Parallèlement, les méthodes d'évaluation de la fréquentation des espaces (*in situ*, par des dispositifs automatiques ou par enquêtes, ou à distance, par photos aériennes ou à l'aide de traceurs GPS) montrent aujourd'hui toutes leurs limites (Meur-Férec *et al.*, 2001), que ce soit dans leur mise en place comme dans leur efficacité. Elles sont notamment largement inopérantes sur des espaces ouverts où la divagation et l'évolution dans le temps du profil topographique des itinéraires sont fréquentes (Audouit *et al.*, 2016), ce qui a conduit à un déficit de connaissance objectivée de leur fréquentation. Dans ce contexte, les questionnaires se tournent vers de nouvelles solutions : les données téléphoniques, porteuses de potentiel (Nettles *et al.*, 2022), s'avèrent peu pertinentes dans les espaces de montagne peu aménagés du fait de la faiblesse de la couverture téléphonique ; et les données des réseaux sociaux, qui se révèlent plus efficaces et assurent continuité, analyse en temps réel et transférabilité (Hausmann *et al.*, 2018), soulèvent des enjeux méthodologiques du fait d'inégalités dans leur production (Tenkanen *et al.*, 2017). Reste le recours aux IGV, grands absents des outils d'observation de la fréquentation en montagne hormis dans deux publications (Rupf & Stäubli, 2018 ; Davoine & Garat, 2022).
- 2 Il existe donc un réel intérêt à documenter plus largement le phénomène des IGV au sein de la littérature académique dans le but de poser les bases de futures recherches permettant de combler les lacunes dans le domaine de la connaissance de la fréquentation des espaces naturels à partir des traces géonumériques. Ainsi, notre étude vise à analyser de la manière la plus exhaustive possible la production scientifique sur les IGV afin de proposer une compréhension globale du savoir et d'identifier des orientations pertinentes (Arksey & O'Malley, 2005). Autrement dit, l'objectif de notre article n'est pas de réaliser une revue de la littérature traitant des IGV et de la fréquentation des espaces naturels, mais de saisir la place du sujet de la fréquentation des espaces naturels au sein du vaste champ de recherche des IGV.

Matériel et méthodes

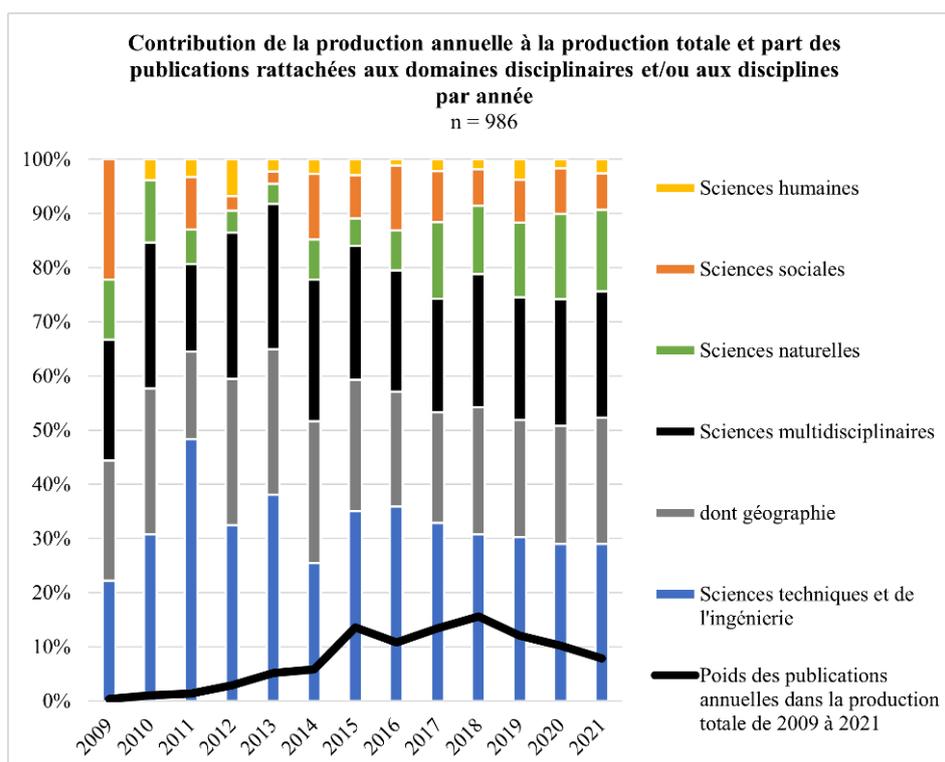
- 3 L'analyse bibliométrique étudie quantitativement la production scientifique sur un sujet à partir d'un corpus de publications. Elle présente un intérêt dans l'objectif de défricher un sujet précis. Elle est donc adaptée aux IGV dont la littérature est récente et relativement peu importante. Pour la réaliser, nous retenons la méthode de la « *scoping study* » (Arksey & O'Malley, 2005), assimilable à la cartographie d'un champ de recherche cherchant à en comprendre la construction (Munn *et al.*, 2018). En cela, elle constitue la source d'un apport nouveau à la science. Notre étude procède en cinq étapes (Badger *et al.*, 2000 ; Arksey & O'Malley, 2005).

- 4 La première est de définir la question de recherche. Afin de proposer un tour d'horizon le plus large possible de la littérature scientifique sur les IGV, la question de recherche est volontairement large. La deuxième étape est l'identification systématisée (Pahlevan-Sharif *et al.*, 2019) et transparente des ressources. Elle a été réalisée en avril 2022 par une recherche des expressions « information géographique volontaire », « *volunteered geographic information* », « *volunteered geographical information* » et « *VGI* » dans les titres, mots-clés et/ou résumés des publications de tous types du *Web of Science* (WoS). Nous sommes partis du principe que le mot-clé « IGV » constitue le pivot de notre analyse bibliométrique et que le WoS est une des bases de la littérature scientifique anglophone de référence pour réaliser ce type d'analyses. Le troisième temps consiste à sélectionner les ressources. Aucune borne temporelle et aucune contrainte linguistique n'ont été retenues, ni aucune limite thématique ou disciplinaire. 1 001 publications composent alors notre corpus. En quatrième, nous avons récolté les mêmes données pour toutes les publications afin d'assurer leur comparabilité (Arksey & O'Malley, 2005). Nous analysons neuf métadonnées disponibles pour chaque publication : l'année de publication ; les catégories disciplinaires définies par le *Web of Science*, *a posteriori* réorganisées en 31 catégories ajustées et réunies au sein de cinq domaines disciplinaires ; le type de document ; la langue de la publication ; la revue ou l'ouvrage de publication ou la conférence de présentation ; la liste des auteurs ; les pays des institutions de rattachement des auteurs ; les contenus des sections « *Fundings* » ; et les mots-clés renseignés par les auteurs, *a posteriori* réorganisées en mots-clés ajustés. La dernière étape correspond à l'exposé analytique des résultats en trois points. Il débute par un tour d'horizon temporel et disciplinaire de la littérature scientifique sur les IGV. Vient ensuite l'analyse du contexte de production des publications, entre espaces de diffusion des publications, autorat et financement de recherche. Finalement, nous proposons une lecture thématique de la littérature académique sur les IGV à partir de l'étude des mots-clés du corpus.
- 5 Le WoS est une plateforme numérique gérée par l'entreprise Clarivate Analytics et qui permet d'accéder à une base de données bibliographiques. Il indexe notamment des articles de revues, des ouvrages et des actes de conférences scientifiques. Précisons que le corpus connaît un défaut d'exhaustivité lié à nos choix. D'une part, l'identification des publications à partir du WoS est largement lacunaire quant à la littérature non anglophone. Par ailleurs, *Scopus* est une base plus riche en sciences humaines et sociales, en revues « régionales » (Frenken & Hoekman, 2014), en publications à plus faible notoriété et en littérature non anglophone, mais nous n'avons pas pu la retenir car nos établissements ne nous y donnent pas accès. Nous savons tout au plus que *Scopus* recense 2 009 publications respectant les mêmes critères de recherche et datant de 2021 ou avant, contre 986 sur le WoS. Il serait donc pertinent de réaliser l'étude qui suit en considérant également les publications référencées sur *Scopus*. D'autre part, la réalisation de la récolte des données en avril 2022 ne permet pas de considérer l'année 2022 ($n=15$) dans l'analyse temporelle et pourrait même écarter quelques publications de 2021 mises en ligne tardivement, compte tenu du délai d'environ deux ans entre la sortie des publications scientifiques et leur enregistrement sur le WoS.

Évolution générale de la production scientifique

- 6 L'intérêt scientifique pour les IGV est récent puisque les premiers articles recensés sur le WoS sont, malgré un article paru en 2008 dans *GeoJournal* (Tulloch, 2008), publié en 2009. La démocratisation d'internet et des outils mobiles et/ou connectés au début du XXI^e siècle explique certainement l'augmentation très forte du nombre de publications jusqu'en 2015. Par la suite, cette croissance ralentit et nous observons depuis 2018 une diminution de la production scientifique.
- 7 Après regroupement des 86 catégories du WoS (une publication peut être associée à plusieurs catégories) en 31 catégories ajustées puis en 5 domaines disciplinaires, deux domaines dominent l'étude des IGV aujourd'hui (Figure 1). Le premier est celui des sciences techniques et de l'ingénierie, dont 80 % des publications relèvent. Dans le détail, environ un tiers des publications est associé aux *computer sciences* et/ou aux sciences de l'imagerie et ce taux est de 10 % pour les sciences de l'ingénieur. Le second domaine scientifique est multidisciplinaire, domaine regroupant les catégories de diverses sciences dont la géographie. Précisons que cette dernière représente près de 97 % des publications de ce domaine multidisciplinaire. Derrière viennent les sciences de la nature et les sciences sociales. Leur intérêt pour les IGV est relevé toutes les années depuis 2011 et le WoS rattache respectivement environ 27 % et 20 % des publications à ces domaines scientifiques. Il est toutefois important de préciser que la géographie est exclue des sciences sociales dans cette analyse, ce qui invite à relativiser le rattachement de seulement 20 % des publications du corpus aux sciences sociales. Quant aux sciences humaines, elles sont très faiblement représentées dans la production académique sur les IGV, mais également dans le WoS en général ; aussi faut-il également relativiser ce résultat.
- 8 La géographie est la catégorie ajustée la mieux représentée dans le corpus, ce qui semble logique compte tenu de la nature géographique des IGV. Le WoS y rattache plus de la moitié des publications du corpus et en fait la discipline la plus représentée dans les publications de toutes les années hormis 2011. Pour autant, elle n'a ni le monopole des publications ni le statut de précurseur : plus de 40 % des publications du corpus ne sont pas rattachées à la géographie, et cela dès les premières années de publication sur les IGV.

Figure 1. Évolution de la production scientifique sur les IGV par année et par domaine disciplinaire



Contextualisation des publications académiques

Domination du champ des Systèmes d'Information Géographique (SIG) dans l'autorat et les espaces de diffusion

- 9 718 publications du corpus sont des articles, issus de 219 revues scientifiques. Les 13 revues les plus représentées regroupent 9 à 104 articles chacune. Si les sciences de l'espace (SIG, géographie, sciences de l'information géographique, aménagement et urbanisme) sont omniprésentes dans leurs thématiques, ce sont plus précisément les SIG qui dominent largement, puisque représentés par 5 revues, dont *ISPRS International Journal of Geo-Information* (n=104), *International Journal of Geographical Information Science* (n=45) et *Transaction in GIS* (n=40). Le constat est similaire du côté des 220 publications issues d'actes de colloques scientifiques. Plus précisément, quelques événements focalisés sur les IGV jouent un rôle important dans la production académique : *Volunteered Geographic Information and The Future of Geospatial Data* (n=12) ; *Openstreetmap in GIScience: Experiences, Research, and Applications* (n=7) ; *Fundamentals of Human Factors Design for Volunteered Geographic Information* et *Mobile Information Systems Leveraging Volunteered Geographic Information for Earth Observation* (n=5).
- 10 Ces publications, pour 99 % en anglais, sont écrites par 2 378 auteurs, dont 22 % sont impliqués dans au moins deux publications différentes. Les 10 auteurs signant au minimum 10 publications chacun forment trois groupes. Avec respectivement 34 et 30 publications, Zipf et See sont les deux auteurs les plus importants sur les IGV aujourd'hui et sont tous deux spécialisés en SIG. Viennent ensuite Mooney (n=19) et Fritz (n=17). Si le second relève du champ des SIG, le premier est rattaché à un

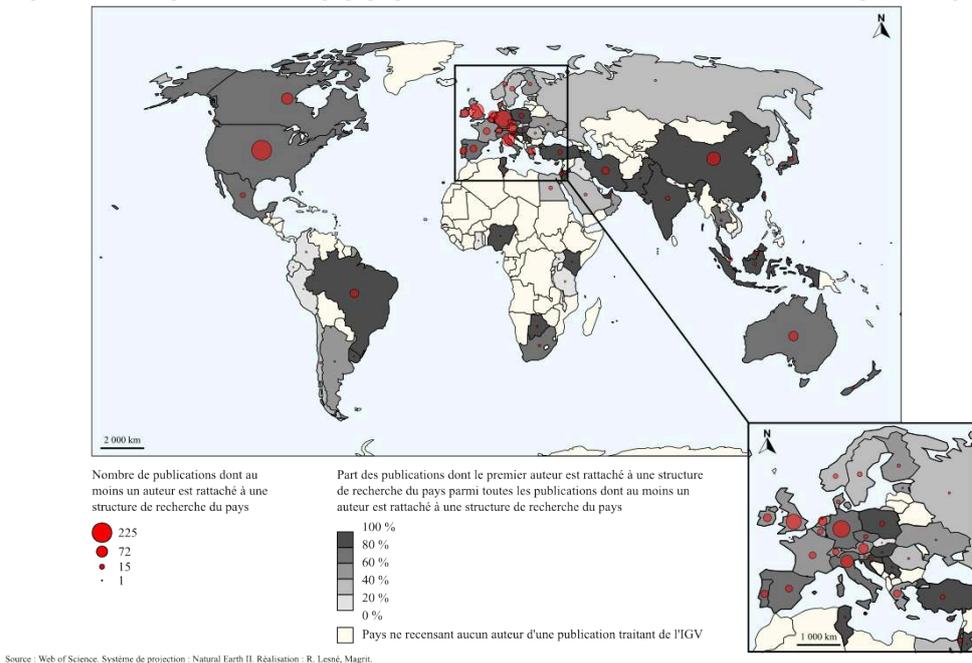
département de sciences informatiques. Enfin, cinq auteurs signent 14 à 10 articles chacun et travaillent dans les sciences de l'espace (Antoniou), majoritairement dans les SIG (Arsanjani, Harklay, Minghini et Hochmair). Finalement, parmi les 10 auteurs les plus productifs, seul Albuquerque (de) n'est pas spécialiste de SIG et/ou d'informatique. Les disciplines, thématiques et objets de spécialité de cet autorat très actif dans la production scientifique corroborent la domination du champ disciplinaire de la géographie et du monde des SIG sur le corpus.

Géographie de l'autorat

- 11 L'analyse géographique de l'autorat (Figure 2) distingue trois pays parmi les adresses des institutions de rattachement des auteurs du corpus déclarées : les États-Unis, l'Allemagne et le Royaume-Uni. Ils sont représentés dans 13 % à 22 % des publications et 6 des 10 auteurs les plus productifs y sont rattachés. Trois autres pays sont particulièrement bien représentés puisque présents dans l'autorat de 7 % à 11 % des publications : la Chine, l'Italie et le Canada. La France arrive seulement en 15^e position. D'une manière générale, si l'Europe, surtout occidentale, et l'Amérique du Nord sont deux continents très impliqués dans la production scientifique sur les IGV, d'autres régions du monde y participent également de manière relativement importante. Il s'agit du Moyen-Orient, avec l'Iran (n=30) ainsi qu'Israël et la Turquie (n=12) en tête, mais également le reste du continent asiatique puisque le Japon, l'Inde et Singapour sont chacun représentés dans au moins 1% des publications, tout comme le Brésil. Ces résultats ressemblent dans les grandes lignes à la cartographie mondiale des publications scientifiques (Eckert *et al.*, 2014), bien que quelques sous-représentations ou absences notables se font remarquer, en premier lieu l'Inde et plus largement l'Asie centrale et une grande partie des pays africains, mais aussi l'Amérique du Sud où le Brésil est impliqué dans 80 % de l'activité scientifique sur les IGV du sous-continent.

Figure 2. Cartographie de l'autorat du corpus

La production scientifique sur l'information géographique volontaire dans le monde en 2022 : des contributions et des coopérations inégales



- 12 L'étude du pays de rattachement du premier auteur de chaque publication permet d'identifier deux dynamiques. D'un côté, parmi les pays les plus représentés dans le corpus, certains sont des pays où les chercheurs sont très rarement positionnés ailleurs qu'en première place de l'autorat d'une publication, ce qui peut correspondre à une position de leader dans la conduite de recherches traitant des IGV et/ou à une faible dynamique collaborative internationale. En effet, la recherche académique est entrée dans une quatrième ère marquée par un recul du poids des publications exclusivement nationales dans la production scientifique des pays anciennement développés tandis que les pays « émergeant » de la recherche scientifique ont une production croissante depuis les années 1980 en grande partie alimentée par des publications exclusivement nationales (Adams, 2013). Avec des scores de 84 % à 97 %, la Chine, le Brésil et l'Iran correspondent à ces pays qui jouent un rôle important dans la production scientifique mondiale sur les IGV presque exclusivement en pilotant des travaux de recherche. D'un autre côté, hormis la Chine, les pays aux niveaux de contribution les plus importants sont également moteurs dans la participation et/ou le pilotage de travaux sur les IGV, mais dans des proportions moindres. Ainsi, les auteurs d'Allemagne, du Royaume-Uni et du Canada sont premiers auteurs de 50 % à 80 % des publications auxquelles ils contribuent. Pour aller plus loin, une géographie de ces collaborations internationales serait réalisable : à quel point les auteurs d'un pays participent-ils à des co-publications avec d'autres chercheurs internationaux ? Quels pays, à travers leurs auteurs, collaborent ? Et dans quelles proportions absolues et relatives à leurs contributions nationales à la littérature scientifique sur les IGV ?

Le soutien des organismes finançant la recherche aux travaux sur les IGV

- 13 Près de la moitié des publications a une section « *Fundings* » complétée. Pour autant, il s'agit d'une source d'information à la fiabilité limitée dans la mesure où, d'une part, les règles et pratiques quant à sa complétion sont hétérogènes et elle n'est pas systématiquement complétée par les auteurs ayant bénéficié de financements directs ou indirects dédiés à une recherche par/sur les IGV et, d'autre part, l'interprétation binaire réduisant sa complétion à l'existence d'un financement et son absence de complétion à l'absence de financement entraînerait des biais méthodologiques trop importants. Tout au plus, une lecture non exhaustive des sections complétées nous permet d'affirmer que les IGV sont bien un sujet qui suscite l'intérêt de divers organismes finançant la recherche sans que nous soyons en mesure de quantifier ou de qualifier outre mesure ce phénomène. Parmi ces organismes, nous retrouvons des programmes de recherche publics régionaux, nationaux et internationaux ainsi que des institutions académiques, mais aussi des entités territoriales de diverses échelles ou encore des organismes parapublics et privés, notamment à vocation de protection de l'environnement.

Analyse thématique de la littérature universitaire

Les mots-clés pour saisir les thématiques

- 14 Nous recensons 2 326 mots-clés renseignés par les auteurs et singuliers dans leur composition alphanumérique, bien que parfois identiques dans leur signification. Nous avons donc harmonisé les différentes orthographes et traductions d'un même mot-clé, par exemple « *crowdsourcing* », « *crowd sourcing* », « *crowd-sourcing* », « *crowdthourcing* », etc. Ensuite, nous avons rassemblé les mots-clés différents, mais très proches dans leur signification sous un même mot-clé ajusté. Par exemple, le mot-clé ajusté « *crowdsourcing* » remplace des mots-clés comme « *crowd and community sourcing* », « *crowdsorce* », « *passive crowdsourcing* » ou encore « *crowdsourced* ». Enfin, nous avons regroupé les déclinaisons spécifiques d'un même mot-clé lorsque celles-ci présentaient de très faibles récurrences. Ainsi, « *crowdsourced earth observation product* », « *crowdsourced platforms* », « *crowdsourced road view photo* » et « *Crowdsourcing air temperature* » ont été simplifiées au mot-clé ajusté « *crowdsourcing* » tandis que nous avons conservé de manière singulière les idées de données et de cartographies issues du crowdsourcing dans des mots-clés ajustés distincts. Par ailleurs, les IGV étant au centre de la recherche bibliographique, nous avons écarté les mots-clés mobilisant la notion "IGV". Le résultat final est une liste de 1 592 mots-clés ajustés.
- 15 D'une manière générale, six mots-clés ajustés se démarquent nettement. Ils concentrent chacun plus de 1 % de l'usage total des mots-clés dans le corpus et sont mobilisés dans plus de 5 % des textes : *open street map* (n=169) ; *crowdsourcing* (n=137) ; *geographic information system* (n=79) ; *data quality* (n=69) ; *social media/networks* (n=61) ; et *citizen science* (n=56). Ils donnent un premier aperçu des entrées thématiques majeures. Les IGV sont traitées au regard des principaux outils avec lesquels elles sont étroitement liées, à savoir la plateforme cartographique collaborative OpenStreetMap et les réseaux sociaux numériques. De plus, leur spécificité étant leur caractère

« volontaire », les IGV s'inscrivent dans les champs du *crowdsourcing* et des sciences citoyennes. Par ailleurs, il s'agit d'un sujet soulevant des enjeux méthodologiques importants, que ce soit en matière de traitement – d'où la prégnance des SIG – ou en matière de réflexion sur la donnée en elle-même – d'où la prégnance de la question de sa qualité.

Cartographie thématique de la littérature

- 16 Pour analyser les thématiques du corpus, nous proposons une « cartographie » des mots-clés ajustés réalisée avec le logiciel Gephi. Chacun constitue un nœud du réseau, représenté par un cercle dont la taille est proportionnelle à la fréquence de sa mobilisation dans le corpus. Le fait que deux mots-clés ajustés soient mobilisés dans une même publication constitue un lien entre deux nœuds du réseau. Ce lien peut être plus important si l'utilisation combinée de ces deux mots-clés ajustés est plus fréquente, ce qui s'illustre par une largeur plus importante du trait reliant les cercles. Le réseau a d'abord été réorganisé spatialement via l'algorithme de spatialisation Force Atlas car il accentue les complémentarités entre les mots-clés ajustés uniquement sur base des forces des liens et il est adapté aux réseaux de taille inférieure à 10 000 nœuds (tandis que l'algorithme de Yifan Hu est recommandé pour des graphes de 100 à 100 000 nœuds et qu'il cherche à simplifier le graphe en plus de se baser sur les forces des liens)¹. Puis nous avons partitionné le réseau en plusieurs communautés à l'aide de l'outil « modularité » du logiciel Gephi, celui-ci appliquant l'algorithme de Louvain pour le partitionnement. La méthode a été utilisée en variant les résolutions afin d'identifier la résolution idéale entre l'obtention d'une quantité de communautés suffisamment grande pour refléter la diversité des publications du corpus, mais aussi suffisamment petite pour résumer cette diversité. Avec une résolution de 1, le partitionnement propose un découpage du réseau en 15 communautés, mais nous écartons les microréseaux isolés de l'analyse, ce qui mène à la classification finale de 1 592 mots-clés ajustés en 12 communautés. Afin d'en faciliter la lecture, nous avons enfin appliqué un filtre par plage de degrés : seuls les mots-clés ajustés dont le nombre de liens est supérieur ou égal à vingt sont affichés (Figure 3). Bien évidemment, approcher les thématiques d'un corpus de 1 001 publications à partir des mots-clés des auteurs présente des biais, notamment leur caractère synthétique et polysémique : les mots-clés ne permettent pas d'accéder finement aux idées d'une publication ; et un même mot-clé peut avoir des significations différentes d'une publication à l'autre. Nous retenons donc cette approche comme un premier tour d'horizon et en surface des thématiques du corpus, avec les limites qu'elle comporte.

communauté pointent vers des recherches portant notamment sur l'aménagement, les infrastructures et les villes. Cette communauté est regroupée dans une partie du réseau, mais elle est néanmoins présente, par quelques éléments seulement, sur toute la surface du réseau. Cela semble traduire une faible interrelation, uniquement ponctuelle, avec les autres communautés. Plus précisément, il semble davantage y avoir un pan de la recherche par télédétection qui mobilise les IGV qu'un pan de la recherche par ou sur les IGV mobilisant également la télédétection. En effet, cette communauté apparaît très faiblement connectée aux autres communautés mobilisant les IGV via des travaux appliqués à la gestion de crise, la participation citoyenne ou encore aux loisirs.

- 20 La communauté « Mobilisation des données massives/ouvertes » est identifiable en premier lieu par sa présence transversale et en filigrane des autres communautés puis en second lieu du fait du poids de quatre de ses mots-clés ajustés dans le réseau. Ceux-ci s'articulent autour des notions, d'une part, de données massives et de données ouvertes et, d'autre part, d'aménagement urbain et de participation. Ce faisant, cette communauté met en exergue tout un champ des travaux mobilisant les IGV qui traitent des questions d'aménagement du territoire en milieu urbain à l'aide de techniques de participation citoyenne et via des données massivement produites par les citoyens. Elle couvre une surface assez étendue du réseau et à proximité d'autres communautés. Ainsi, plusieurs mots-clés ajustés qui la composent (les dispositifs portatifs, notamment) sont proches de mots-clés ajustés caractérisant d'autres communautés (la participation), ce qui indique une forte interrelation avec les autres communautés.
- 21 La communauté « Ouverture des SIG » est importante en grande majorité suivant le poids élevé du mot-clé ajusté *GIS* (SIG). Ce résultat met en évidence l'aspect central des SIG dans les travaux scientifiques sur les IGV aujourd'hui, ce qui semble tout à fait logique d'un point de vue disciplinaire en géographie. Combiné avec les notions de SIG web et/ou participatif, voire encore d'*open source*, cette communauté illustre le lien étroit des travaux scientifiques traitant des IGV avec la question de l'ouverture des SIG à des données produites par les citoyens, mais aussi à leur manipulation par les citoyens eux-mêmes. Le positionnement de cette communauté au sein du réseau traduit cette réalité puisque, *GIS* en position centrale mis à part, ses mots-clés sont spatialement proches d'autres communautés reflétant les champs d'application des recherches mobilisant les IGV tels que l'aménagement participatif, la gestion des risques ou encore les loisirs.
- 22 La communauté « Gestion et stockage des données » n'est pas d'une grande importance, en termes d'intensité et donc de nombres de mots-clés ajustés, mais elle se compose néanmoins de thèmes importants d'un point de vue méthodologique et disciplinaire en géographie. Ainsi, ses principaux thèmes sont l'information géographique (*geographic information*), les infrastructures de données (*data infrastructures*) ou encore les métadonnées (*metadata*). Ces catégories sont toutes fondamentales pour bâtir un raisonnement géographique objectif au travers des IGV. À l'échelle de l'organisation globale du réseau, cette communauté se localise d'une manière similaire à celle d'OSM et semble donc indiquer une proximité thématique, mais aussi une complémentarité avec des thèmes tels que cette plateforme cartographique collaborative ou d'autres modes de valorisation des données.
- 23 La communauté « Management des crises, risques et catastrophes » bénéficie d'un poids modeste principalement casé sur des mots-clés ajustés tels que *social media/networks*, *Twitter* ou encore *emergency*. Cette communauté gravite donc autour des

questions de gestions des catastrophes, des urgences et des crises à l'aide de données récupérées sur les plateformes que sont les réseaux sociaux, Twitter en tête. Cette organisation illustre la structuration d'une communauté de recherche où les IGV sont des outils de la gestion des crises, notamment des catastrophes naturelles. Suivant l'analyse de la localisation des mots-clés dans le réseau et parmi ses communautés importantes, celle-ci est singulière par sa localisation au centre de la communauté « Mobilisation des données massives/ouvertes ». L'orientation de mobilisation scientifique des IGV qu'elle dessine semble donc constituer un sujet spécifique étroitement connecté au champ des travaux traitant des questions d'aménagement du territoire par les données citoyennes massives ou de la participation citoyenne, au point de sembler en constituer un sous-champ.

- 24 La communauté « Sciences et données citoyennes » est regroupée autour des mots-clés ajustés tels que *citizen science* et *user content*, mais aussi *accuracy*. Leur co-présence dans cette communauté l'inscrit dans le champ des enjeux méthodologiques liés aux données crowdsourcées principalement représentés par la communauté « Source et traitement des données crowdsourcées ». Les enjeux méthodologiques identifiés ici semblent donc être le corolaire des travaux thématiques basés sur les données crowdsourcées et la proximité des deux communautés indique bien la dynamique de fixation des enjeux méthodologiques notamment via l'application à différents domaines de la géographie.
- 25 Structurée autour de trois mots-clés ajustés localisés sur le réseau, la communauté « Mobilités cyclables » apparaît assez étendue à l'échelle du réseau. Elle regroupe principalement les mots-clés *bicycling* et *machine learning*. Ce faisant, cette communauté reflète l'existence d'un champ de recherche limité en taille et mobilisant les IGV générées lors de mobilités cyclistes, notamment à partir de la plateforme *Strava* (mot-clé ajusté non affiché du fait du filtre), données ensuite valorisées par des dispositifs de gestion de larges jeux de données et d'apprentissage machine. Ce champ thématique correspond à celui de l'aménagement urbain. Sa position à proximité de la zone occupée par la communauté « Source et traitement des données crowdsourcées » semble indiquer qu'elle aussi en constitue un sous-champ, ici récoltées lors d'activités de loisir ou de mobilité puis valorisées dans un autre domaine. Toutefois, sa localisation au sein du réseau traduit également, comme pour la communauté précédente, son interrelation avec la communauté « Sciences et données citoyennes » qui illustre le besoin de traiter de manière concomitante les enjeux méthodologiques liés aux données citoyennes.
- 26 Structurée autour de quatre mots-clés ajustés affichés sur le réseau, la communauté « Cartographie, occupation du sol et inondation » fait apparaître une organisation principalement autour des mots-clés *land cover* et *participatory mapping*. Cette communauté reflète l'existence d'un champ de recherche où les IGV sont mobilisées dans une perspective de cartographie participative et au service de l'identification de l'occupation des sols et de la prévention des risques naturels (inondations, glissements de terrain, etc.). Sa position au sein du réseau en fait logiquement une communauté connectée à plusieurs autres communautés, principalement celles portant sur les enjeux méthodologiques des données crowdsourcées et sur le volet participatif des travaux utilisant des SIG dans une perspective aménagiste.
- 27 Structurée autour de quelques mots-clés seulement, la communauté « Loisirs, tourisme et photographies » apparaît comme mineure. Sa logique se dessine cependant de

manière très distincte puisque cette communauté traduit l'existence d'un champ de recherche, certes marginal par le nombre d'occurrences des mots-clés, traitant d'une catégorie spécifique des IGV que sont les photos géolocalisées. Ce faisant, le regroupement de ces mots-clés indique une préoccupation pour les données produites dans un contexte de loisir, mais aussi leur intégration aux études sur la fréquentation des espaces touristiques. Toutefois, la taille modeste de cette communauté illustre la place mineure qu'elle occupe aujourd'hui au sein de la recherche scientifique associée aux IGV. Sa localisation au sein du réseau montre sa connexion avec d'autres communautés, notamment celles portant sur les enjeux méthodologiques des données crowdsourcée et citoyennes, ou encore celles sur l'aménagement du territoire et sur les mobilités cyclistes.

- 28 La communauté « Les citoyens comme capteurs et la santé » apparaît comme mineure et à la cohérence thématique difficile à définir. Avec ses mots-clés ajustés *citizen sensors* et *health*, elle semble toutefois refléter l'existence d'un champ de recherches marginal traitant des IGV et portant sur la santé et la mobilisation de citoyens comme capteurs de santé publique. Sa position à l'échelle du réseau et à proximité de la communauté « Mobilisation des données massives/ouvertes » semble indiquer qu'elle en constitue un sous-champ du champ de l'aménagement du territoire via des données participatives. Sa place au sein du réseau traduit également une interrelation avec les communautés portant sur la gestion et le stockage des données et portant sur la gestion des crises grâce aux données des réseaux sociaux.

Conclusion

- 29 La recherche sur les IGV, bien que récente, est désormais nettement implantée dans le paysage scientifique international. Les avancées, tant thématiques que méthodologiques, permettent donc aujourd'hui de disposer d'un socle de connaissances favorisant le montage de projets scientifiques de plus grandes envergures. Par ailleurs, l'identité ancrée en géographie des travaux sur les IGV et leurs dimensions participatives et impliquant des travaux en informatique fondent le caractère central de la multidisciplinarité dans la recherche sur cet objet aujourd'hui. Si les géographes s'affirment comme des chercheurs incontournables sur cet objet, notamment en SIG, les spécialistes des sciences citoyennes ou encore de l'informatique jouent également un rôle particulièrement important. Parmi ces contributeurs à la recherche sur les IGV, nous notons à ce jour un « retard » français, notamment au regard de l'investissement des chercheurs des pays voisins. La domination de la langue anglaise au sein du corpus étudié nous invite cependant à relativiser ce point, ce manque pouvant n'être qu'apparent dans la mesure où plusieurs revues francophones hébergent des articles traitant des IGV. Enfin, nous retenons de l'analyse des mots-clés que 4 grandes thématiques dominent la mobilisation des IGV dans le cadre de la recherche scientifique, à savoir OpenStreetMap, la mobilisation des données crowdsourcées, la participation citoyenne et l'étude du globe terrestre. Face à ces thématiques, celles des loisirs et/ou de la fréquentation des espaces naturels apparaissent donc comme largement minoritaires, y compris au sein de la question de l'aménagement et de la gestion de l'espace. Ainsi, à l'issue de notre analyse bibliométrique, nous ne retenons que 45 publications parmi les 1 001 du corpus comme directement en prise avec la question de la fréquentation des espaces naturels. Cette sélection résulte du processus

suivant : extraction des 144 mots-clés ajustés rattachés aux communautés « Loisirs, tourisme et photographies » ainsi que « Mobilités cyclables » ; identification des 43 mots-clés de cette liste correspondant à la thématique de l'observation de la fréquentation des espaces naturels ; extraction des 82 publications du corpus comportant au moins un de ces mots-clés ; et lecture des titres et résumés afin d'identifier les 45 publications. L'ensemble de ces résultats sont à relativiser du fait des biais liés au fait que le matériau exploité a été récolté sur le WoS. Cette principale limite nous incite à envisager de réaliser la même étude sur le corpus des plus de 2 000 publications recensées dans *Scopus* sur le même sujet et d'en croiser les résultats avec ceux exposés dans le présent article.

- 30 Enfin, si les travaux sur les IGV constituent collectivement un champ désormais constitué, il reste à y inscrire les travaux sur l'observation de la fréquentation des espaces naturels en profondeur. En effet, plusieurs initiatives engagées par des institutions en charge du sport montrent que les IGV, ici sous la forme de traces géonumériques, deviennent des outils légitimes de l'action publique. Mentionnons par exemple, en France, le dispositif Outdoorvision porté par le ministère des Sports et des Jeux Olympiques et Paralympiques ou encore, dans plusieurs pays d'Europe Centrale, la place importante occupée par l'application OutdoorActive dans la promotion, mais aussi la gestion du tourisme sportif. Dans ce cadre, si leur portée heuristique reste encore à définir avec précision, ils ont leur place pour observer la fréquentation de territoires difficiles à appareiller de manière traditionnelle (Norman et Pickering, 2017, 2019 ; Norman *et al.*, 2019) suivant l'évolution continue des itinéraires ou encore la forte diffusion des pratiquants dans l'espace. Il s'agit ici d'un sous-champ des travaux sur les IGV en cours de constitution, comme l'illustrent les travaux exploratoires et souvent concomitants développés en France (publication d'un numéro spécial de la *Revue de Géographie Alpine*, organisation de séminaires) et en Europe (session dédiée dans les colloques Monitoring and Management of Visitors – MMV).

BIBLIOGRAPHIE

- Adams J., 2013.- « Collaborations: The fourth age of research », *Nature*, n° 497, p. 557-560. DOI : <https://doi.org/10.1038/497557a>.
- Arksey, H., O'Malley, L., 2005.- « Scoping studies: towards a methodological framework », *International Journal of Social Research Methodology*, n° 8(1), p. 19-32.
- Audouit, C., Rufin-Soler, C., Le Falher, G., Flanquart, H., Deboudt, P., 2016.- « Perception et gestion des espaces littoraux préservés : l'apport des études de fréquentation (Nord et Languedoc Roussillon, France) », *Vertigo*, vol. 16, n° 2. DOI : <https://doi.org/10.4000/vertigo.17721>.
- Badger, D., Nursten, J., Williams, P., Woodward, M., 2000.- « Should all literature reviews be systematic? », *Evaluation & Research in Education*, vol. 14, n°s 3-4, p. 220-230.

- Dadashova, B., Griffin, G. P., Das, S., Turner, S., Sherman, B., 2020.- « Estimation of Average Annual Daily Bicycle Counts using Crowdsourced Strava Data », *Transportation Research Record*, n° 2674(11), p. 390-402.
- Davoine, P.-A., Garat, P., 2022.- « Analyser les pratiques de la randonnée à ski par les données géographiques issues des médias sociaux : l'exemple du forum Skitour à l'échelle du massif des Écrins », dans Kreziak D., Reynier V., Bourdeau P. and Joye J.-F. (dir.), *Le ski de randonnée brouille les pistes*, Éditions du Fournel, p. 76-94.
- Eckert, D., Grossetti, M., Jégou, L., Maisonobe, M., 2014.- « Les villes de la science dans le monde », *M@ppemonde*, n°116(4). En ligne : <http://mappemonde-archive.mgm.fr/num44/articles/art14401.html>, consulté le 16 janvier 2024.
- Frenken, K., Hoekman, J.- « Spatial Scientometrics and Scholarly Impact: A Review of Recent Studies, Tools, and Methods », dans Ding Y., Rousseau R. and Wolfram D. (dir.), *Measuring Scholarly Impact*, Springer International Publishing, p. 127-146.
- Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., Di Minin, E., 2018.- « Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas », *Conversation Letters*, vol. 11, n° 1. DOI : <https://doi.org/10.1111/conl.12343>.
- Hong, J., McArthur, D. P., Livingston, M., 2020.- « The evaluation of large cycling infrastructure investments in Glasgow using crowdsourced cycle data », *Transportation*, n° 47, p. 2859-2872.
- Houllier, F., Merilhou-Goudard, J.-B., 2016.- *Les sciences participatives en France : États des lieux, bonnes pratiques et recommandations*, Inra. En ligne : <https://hal.archives-ouvertes.fr/hal-02801940/>, consulté le 16 janvier 2024.
- Kyuhyun, L., Ipek Nese, S., 2021.- « Strava Metro data for bicycle monitoring: a literature review », *Transport Reviews*, n° 41(1), p. 27-47.
- Le Breton, M.-A., Bailleul, H., Le Corf, J.G., Mericskay, B., 2022.- « La gouvernance des données urbaines entre territoire de projets et projet de territoire. L'exemple de Rennes Métropole », *Flux*, n° 127, p. 65-84.
- Mao, P., Obin, O., 2018.- « La transition numérique des sports de nature, vers des sportsnature 3.1 », *Nature et Récréation*, n° 6, p. 13-25.
- Mericskay, B., Noucher, M., Roche, S., 2018.- « Usages des traces numériques en géographie : potentiels heuristiques et enjeux de recherche », *L'Information géographique*, n° 82, p. 39-61. DOI : <https://doi.org/10.3917/lig.822.0039>.
- Mericskay, B., Roche, S., 2011.- « Cartographie 2.0 : le grand public, producteur de contenus et de savoirs géographiques avec le web 2.0 », *Cybergeo*. DOI : <https://doi.org/10.4000/cybergeo.24710>.
- Meur-Férec, C., Hoedt, M., Danes, L., 2001.- « Fréquentation des sites naturels littoraux. Vers une meilleure connaissance du public : l'exemple de deux sites dunaires à l'Est du Dunkerque », *Hommes et Terres du Nord*, n° 2001(2), p. 81-89.
- Mondo, M., Noucher, M., Le Campion, G., Vacher, L., Vye, D. D., 2020.- « Critical approach of digital footprints to study spatial practices of urban tourist areas: a case study of Instagram data in Biarritz (France) », dans Passerini G. and Ricci S. (dir.), *The Sustainable City XIV*, WIT Press, p. 319-333. DOI : <https://doi.org/10.2495/SC200271>.
- Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E., 2018.- « Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping

- review approach », *BMC Medical Research Methodology*, n° 18, art. 143. DOI : <https://doi.org/10.1186/s12874-018-0611-x>.
- Musakwa, W., Selala, K. M., 2016.- « Mapping cycling patterns and trends using Strava Metro data in the city of Johannesburg, South Africa », *Data in Brief*, n° 9, p. 898-905.
- Nelson, T. A., Roy, A., Ferster, C. J., Fischer, J., Brum-Bastos, V., Laberee, K., Yu, H., Winters, M., 2021.- « Generalized model for mapping bicycle ridership with crowdsourced data », *Transportation Research Part C: Emerging Technologies*, n° 125. DOI : <https://doi.org/10.1016/j.trc.2021.102981>.
- Nettles, J. M., Brownlee, M. T. J., Sharp, R. L., Verbos, R. I., 2022.- « The utilization distribution: wildlife research methods as a tool for understanding visitor use in remote parks and protected areas », *Human Dimensions of Wildlife*, vol. 27, n° 2, p. 151-163.
- Norman, P., Pickering, C. M., 2017.- « Using volunteered geographic information to assess park visitation: Comparing three on-line platforms », *Applied Geography*, n° 89, p. 163-172. DOI : <https://doi.org/10.1016/j.apgeog.2017.11.001>.
- Norman, P., Pickering, C. M., 2019.- « Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data », *Journal of Environmental Management*, n° 249. DOI : <https://doi.org/10.1016/j.jenvman.2019.109413>.
- Norman, P., Pickering, C. M., Castley, G., 2019.- « What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? », *Landscape and Urban Planning*, n° 185, p. 180-190. DOI : <https://doi.org/10.1016/j.landurbplan.2019.02.015>.
- Pahlevan-Sharif, S., Murab, P., Wijesinghec, S. N. R., 2019.- « A systematic review of systematic reviews in tourism », *Journal of Hospitality and Tourism Management*, n° 39, p. 158-165.
- Rupf, R., Stauble, A., 2018.- « Monitoring methods of winter backcountry recreation in a wildlife sanctuary », *9^e conférence internationale MMV (Managing and Monitoring of Visitors in Recreational and Protected Areas)*, Bordeaux.
- Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., Toivonen, T., 2017.- « Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas », *Scientific Reports*. DOI : <https://doi.org/10.1038/s41598-017-18007-4>.
- Tulloch, D., 2008.- « Is VGI participation? From vernal pools to Video Games », *GeoJournal*, vol. 72, n° 3, p. 161-171.
- Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H., Nowell, M. S., 2021.- « Back to nature: Norwegians sustain increased recreational use of urban green space months after the COVID-19 outbreak », *Landscape and Urban Planning*, n° 214. DOI : <https://doi.org/10.1016/j.landurbplan.2021.104175>.

NOTES

1. <https://github.com/gephi/gephi/wiki/Spatialisations-%28FR%29>

RÉSUMÉS

La fréquentation des espaces naturels est marquée par une numérisation croissante et massive de la société. Elle entraîne la production et le partage de traces numériques qui, dès lors qu'elles contiennent des données permettant leur spatialisation, deviennent des Informations Géographiques Volontaires (IGV). Permettant d'appréhender des phénomènes difficilement observables, leur exploitation dessine des perspectives de recherche, d'autant plus pour les espaces de montagne, marqués par un enjeu de connaissances de cette fréquentation et par le manque d'efficacité d'autres dispositifs. Documenter le phénomène des IGV au sein de la production scientifique peut donc permettre de poser les bases de ces futures recherches. Notre objectif est de proposer une compréhension globale de la littérature académique sur les IGV. Nous réalisons une analyse bibliométrique à partir d'une *scoping study* des 1 001 publications recensées sur le *Web of Science* traitant des IGV. Premièrement, nous identifions l'évolution de leur production depuis 2009 et leur rattachement très important à la géographie et aux sciences techniques et de l'ingénierie. Deuxièmement, nous mettons en évidence quelques tendances quant à leur contexte de production, notamment l'influence des sciences de l'information géographique dans l'autorat et les espaces de diffusion des travaux ainsi que la domination de quelques pays dans la production de cette littérature. Troisièmement, nous repérons des thématiques prégnantes au sein du corpus à travers l'analyse des mots-clés et le caractère lacunaire de la question de la fréquentation récréative des espaces naturels. Ainsi, notre article met en évidence la maturité du champ d'études des IGV, mais aussi l'opportunité qui se dessine en matière de mobilisation de cet outil dans des travaux scientifiques traitant de la fréquentation et de la gestion des espaces naturels.

INDEX

Mots-clés : Information géographique volontaire, fréquentation des espaces naturels, scoping study, analyse bibliométrique

AUTEURS

ROBIN LESNÉ

Univ. Littoral Côte d'Opale, Univ. Lille, ULR 4477 - TVES – Territoires Villes Environnement & Société, F-59140 Dunkerque, France

MARC LANGENBACH

Univ. Grenoble Alpes, CNRS, Sciences Po Grenoble*, Pacte, 38000 Grenoble, France

* School of Political Studies Univ. Grenoble Alpes

PASCAL MAO

Univ. Grenoble Alpes, CNRS, Sciences Po Grenoble*, Pacte, 38000 Grenoble, France

* School of Political Studies Univ. Grenoble Alpes

HUGUES FRANÇOIS

Univ. Grenoble Alpes - INRAE, Lessem

NICOLAS ROBINET

Univ. Grenoble Alpes, CNRS, Sciences Po Grenoble*, Pacte, 38000 Grenoble, France

* School of Political Studies Univ. Grenoble Alpes