

Une typologie des lecteurs de Wikipedia

Laurent Mell & Nicolas Jullien

Cette enquête a bénéficié du soutien de la Région Bretagne et du PEPR eNSEMBLE, qui travaille sur le futur de la collaboration numérique (projet ciblé : CONGRATS).

Avec plus de 700 millions de visiteurs quotidiens dans le monde et près de 4 millions de lecteurs uniques par jour en France, Wikipédia occupe une place non négligeable dans le processus de construction des savoirs (Jemielniak et Aibar, 2016 ; Sahut, 2016). Interroger la place et le rôle de Wikipédia dans les pratiques de lecture devient éminemment intéressant puisque la valeur des informations proposées est régulièrement remise en question et sa lecture reléguée au regard des pratiques culturelles légitimes. Qu'est-ce qu'être un lecteur de Wikipédia ? Quelle valeur est attribuée au contenu proposé sur cette plateforme ? Son usage s'inscrit-il dans une dynamique de pratiques de lecture « ordinaires » (Barth-Robot, 2023 ; Thumala Olave, 2018) et plus largement de pratiques culturelles (Bennett et al., 2009 ; Coulangeon, 2010), ou vient-il renégocier les modes d'accès à l'information ainsi que la construction des connaissances ?

C'est pour répondre à ces questions qu'une enquête, coordonnée par le centre de recherche public Marsouin¹, a été menée auprès des usagers de Wikipédia durant l'été 2023. Les données ont été collectées à l'aide d'un questionnaire en ligne entre les mois de juin et de juillet. Le lien vers le questionnaire a été distribué via une bannière publiée en 8 langues sur la page Wikipédia. L'enquête comprenait 200 questions sur : ce que les gens faisaient sur Wikipédia avant de cliquer sur le lien vers le questionnaire ; la manière dont ils utilisent Wikipédia en tant que lecteurs (utilisations professionnelles et personnelles) ; leur opinion sur la qualité, la couverture thématique et l'importance de l'encyclopédie ; la création de Wikipédia (comment ils pensent qu'elle est créée, s'ils y ont déjà contribué et comment) ; leurs activités sociales, sportives, artistiques et culturelles, en ligne et hors ligne ; leurs caractéristiques socio-économiques, y compris leurs convictions politiques, et leur propension à faire confiance à l'encyclopédie. Plus de 200 000 personnes ont ouvert le questionnaire, 100 332 ont commencé à y répondre, constituant ainsi notre ensemble de données, et 10 576 l'ont terminé. Ces 10576 réponses constituent notre échantillon de répondants².

¹ Plus de détails sur Marsouin à cette adresse : <https://www.marsouin.org>

² Présentation de l'enquête :

https://meta.wikimedia.org/wiki/Research:Surveying_readers_and_contributors_to_Wikipedia

Méthode de traitement des données.

Il est important de préciser d'abord que notre échantillon est ce qu'on appelle « un échantillon de convenance ». Il ne nous dit rien sur la répartition effective dans la population qui utilise Wikipédia dans les huit langues de l'enquête. Par contre, nous faisons l'hypothèse que les répondants représentent bien l'ensemble des profils et des usages de l'encyclopédie. Nous allons ainsi analyser les relations statistiques entre des usages (ici contribuer) et des caractéristiques socio-démographiques, culturelles, professionnelles, etc. des personnes ayant ces usages.

Afin de repérer les facteurs de différenciation des pratiques de lecture de Wikipédia, nous avons choisi de réaliser une typologie³ à partir d'une sélection de variables décrivant ces pratiques : l'accès à Wikipédia ; le mode d'accès au contenu ; l'évaluation de l'information ; la qualité de l'article ; les besoins professionnels ou d'études nécessitant d'accéder à de l'information ou des connaissances écrites ; l'usage de Wikipédia pour répondre à ces besoins professionnels ou d'études ; les outils numériques (ordinateurs, téléphones) pour accéder à Wikipédia pour ces activités professionnels ou d'étude ; les mêmes variables pour les besoins personnels ; le rapport aux sources Wikipédia. Les profils issus de cette classification sont ensuite interprétés au regard d'un ensemble de variables illustratives portant sur : les caractéristiques sociodémographiques ; les compétences numériques ; les usages numériques ; les pratiques de lecture ; les pratiques culturelles ; les discussions concernant Wikipédia ; les perceptions externes de Wikipédia ; les modes d'accès à l'information médiatique ; la fiabilité perçue de ces sources d'information ; et les pratiques de contribution.

Nous avons ainsi identifié cinq classes de lecteurs de Wikipédia (Tableau 1), construites à partir de leurs usages. Nous n'avons pas la possibilité de savoir la proportion de chacune de ces classes dans la population des lecteurs de Wikipédia. C'est pour cette raison que nous n'indiquons pas la proportion de chaque classe dans notre échantillon : elle ne donne aucune information sur la répartition réelle. Néanmoins, nous estimons, comme indiqué ci-dessus, qu'elles représentent assez bien la diversité des types de lecteurs rencontrés sur Wikipédia.

La présentation de chacune des classes ci-dessous se décompose en trois parties. La première, en italique, présente leurs usages de Wikipédia (autrement dit les variables actives ayant participé à la construction des différentes classes). La deuxième détaille leur profil socio-économique, leurs compétences numériques, leurs pratiques de lecture ainsi que leur mode d'accès aux médias et à l'information (ces variables n'entrent pas dans la construction des classes, et il s'agit de voir si certaines modalités de ces variables sont sur-représentées dans

³ Nous avons élaboré notre typologie suivant une méthode mixte combinant analyse factorielle et classification (Lombardo et Daly, 2021). D'abord, nous avons effectué une analyse factorielle multiple (AFM) à partir de 53 variables actives structurées en 11 groupes et de 79 variables illustratives rassemblées en 10 groupes. À partir des résultats de l'AFM, nous avons réalisé une classification ascendante hiérarchique (CAH) selon la méthode de Ward. Les classes issues de la CAH ont été ensuite consolidées avec la méthode de partitionnement des K-means et nous a permis de dégager cinq classes.

les classes construites). La troisième partie est consacrée à la synthèse des résultats présentés en amont mais surtout notre interprétation de la classe dans une perspective d'ensemble.

Tableau 1 : Typologie des lecteurs de Wikipédia

	Nom
Classe 01	<i>Les Petits lecteurs</i>
Classe 02	<i>Les Conformistes</i>
Classe 03	<i>Les Ambivalents</i>
Classe 04	<i>Les Sur-informés</i>
Classe 05	<i>Les Confiants</i>

Classe 01 (les Petits lecteurs)

Que ce soit dans un contexte personnel, professionnel ou d'études, les Petits lecteurs font part d'un besoin modéré à se former, se cultiver et se renseigner. Quand ce besoin existe, Wikipédia est peu ou pas mobilisé. Cette mise à contribution parcellaire de Wikipédia est également visible lorsqu'il est question d'employer des textes issus de Wikipédia pour des documents à écrire dans le cadre du travail ou des études.

Les *Petits lecteurs* se caractérisent par une part importante de personnes âgées, près de la moitié des membres de la classe a plus de 55 ans. Il s'agit de la classe la moins diplômée et également celle avec la part la plus importante de catégories socio-professionnelles les moins favorisées. Ils expriment le sentiment d'aisance financière le plus faible des cinq classes. Les *Petits lecteurs* présentent des compétences numériques limitées, se traduisant par un niveau d'aisance numérique bien inférieur aux quatre autres classes mais néanmoins en adéquation avec un usage limité de Wikipédia. Les *Petits lecteurs* renvoient l'image de la classe la « moins informée ». Ils se présentent comme des faibles consommateurs d'information, où la lecture papier et numérique est la moins développée, avec une majorité de non-lecteurs quel que soit le contenu. Par ailleurs, ils s'informent particulièrement par le biais de la télévision et des moteurs de recherche, tout en développant un sentiment de fiabilité générale de l'information relativement réduit. Leurs pratiques culturelles, sportives, artistiques et sociales sont relativement faibles.

Interprétation : Globalement, les *Petits lecteurs* sont des lecteurs qui identifient Wikipédia comme un site internet parmi d'autres et le mobilisent de manière sporadique. Leur faible consommation d'information – sur Wikipédia et au-delà – est liée à de faibles besoins (inférieurs aux autres classes) ou des compétences restreintes en termes de lecture, de recherche d'information, mais également de maîtrise des dispositifs numériques. Ils correspondent aux « faibles lecteurs » identifiés dans les enquêtes traditionnelles (Poissenot, 2019) et sont plutôt issus des classes populaires. Inscrits dans un rapport ordinaire des classes populaires au savoir et à la forme écrite (Lahire, 1993), les *Petits lecteurs* sont plus enclins à remettre en question la fiabilité de l'information sur Wikipédia et, plus largement, sa légitimité en tant qu'institution de construction d'un savoir encyclopédique.

Classe 02 (les Conformistes)

De la même manière que les Petits lecteurs, les Conformistes expriment le besoin d'effectuer des recherches par curiosité comme de se former ou de se cultiver pour des besoins professionnels ou d'études. Mais à la différence des premiers, Wikipédia y répond partiellement. Les besoins personnels sont plus importants et diversifiés : se former ou se cultiver, se renseigner, s'informer sur la politique et, plus modérément, faire un choix ainsi que trancher un désaccord. Mais, là encore, Wikipédia est mobilisé de manière sélective. Par ailleurs, les Conformistes font partie de la classe la moins encline à citer les sources auxquelles elle se réfère sur Wikipédia.

Les *Conformistes* se distinguent par une proportion notablement élevée de femmes en comparaison des autres classes. Ils constituent également la classe la plus âgée, avec une proportion significative de retraités. Les *Conformistes* affichent un sentiment d'aisance financière particulièrement marqué, et une part importante des membres de cette classe a un peu ou beaucoup de temps libre. Ils se distinguent des *Petits lecteurs* par une relative aisance numérique, même s'ils conservent des pratiques numériques assez « traditionnelles » avec une consommation plus marquée de la télévision et de la radio. Ils montrent un engouement sélectif pour le format papier, en se concentrant particulièrement sur les romans ou concernant l'actualité. Ils s'informent régulièrement mais adoptent une approche ciblée des médias en vue d'accéder à l'information. Ils correspondent à une des deux classes avec le niveau de confiance le plus élevé dans les médias en général mais celle avec la confiance la plus marquée dans les médias « traditionnels » (tv, radio, presse). Sur le plan des pratiques culturelles et artistiques, les *Conformistes* manifestent peu d'intérêts mais développent d'autres formes d'engagement au niveau sportif et social.

Interprétation : Les *Conformistes* s'identifient à un groupe de retraités actifs qui reconnaissent Wikipédia comme une institution de la connaissance encyclopédique. Wikipédia est mobilisé de manière sélective pour répondre aux besoins personnels ou professionnels, en complément d'autres sources. Les *Conformistes* s'inscrivent dans une dynamique de « butinage informationnel », facilitée par la part importante de temps libre à disposition. Wikipédia n'est qu'une source d'informations parmi d'autres, tout en étant reconnue comme une institution légitime du savoir. À ce titre, cette classe adoptent un comportement conforme de consommation de la connaissance dans une institution qui, *a priori*, semble plus accessible que les établissements du savoir et de la culture « traditionnelles ».

Classe 03 (les Ambivalents)

Les Ambivalents se distingue par une diversification notable de leurs besoins professionnels ou d'études, dépassant celle des Petits lecteurs et des Conformistes. Leurs usages incluent la formation ou l'enrichissement personnel, la recherche d'informations variées, mais aussi la vérification précise de faits ou de données. En comparaison des deux premières classes,

Wikipédia y répond davantage sans toutefois les couvrir intégralement. Dans ce contexte, les contenus de Wikipédia sont parfois mobilisés et les sources sont ponctuellement citées. Les besoins personnels sont également variés, à l'image des Conformistes. Ici encore, Wikipédia est mobilisé de façon sélective, son utilisation étant légèrement moins marquée pour effectuer des choix ou pour suivre l'actualité politique.

Les *Ambivalents* se caractérisent par leur profil très jeune ainsi qu'une proportion significative de membres de cette classe en emploi. Cette classe enregistre également la part la plus élevée de personnes en études, illustrant un engagement fort dans la construction d'un avenir professionnel ou le développement d'un parcours d'études. Cette double implication — emploi ou études — se traduit par un temps libre restreint. Cette classe se distingue par un bon niveau d'aisance numérique mais développant des usages numériques modérés pour le travail ou les études, comme pour le divertissement. En matière de lecture, les *Ambivalents* adoptent une approche sélective, au même titre que les *Conformistes*. Si l'intérêt pour la lecture papier est modéré, celui pour la lecture numérique est plus développé. Il en va de même pour les loisirs. Si les activités culturelles et artistiques restent relativement modérées, les activités sportives et sociales sont privilégiées. À l'instar des *Petits lecteurs*, ils affichent l'un des niveaux de confiance dans Wikipédia le plus bas mais, contrairement à ces derniers, Wikipédia répond partiellement à leurs besoins (professionnel comme personnel).

Interprétation : Les *Ambivalents* se caractérisent par un besoin réel de se cultiver et de s'informer mais adoptent une posture critique vis-à-vis de l'information trouvée sur Wikipédia. Ils semblent ponctuellement en difficulté avec le format encyclopédique proposé. L'approche de cette classe vis-à-vis de Wikipédia reflète une posture ambivalente : un usage pragmatique de consommation d'information en vue de réaliser une tâche (professionnelle, académique, personnelle) mais des difficultés à s'appropriier le contenu et le transformer en savoir en savoir utile, d'autant plus pour une classe qui souffre d'un temps libre restreint.

Classe 04 (les Sur-informés)

Les Sur-informés se caractérisent par des besoins particulièrement diversifiés et importants, tant dans les sphères professionnelles que d'études (se former, se cultiver, se renseigner mais également vérifier des informations ou trancher un désaccord). Dans ces contextes, Wikipédia parvient à satisfaire une grande partie de ces attentes. Sur le plan personnel, les besoins se révèlent également nombreux et variés, dépassant ceux des Conformistes et des Ambivalents. Les membres de cette classe utilisent Wikipédia de manière sélective, l'outil étant perçu comme un complément pertinent mais non exclusif. Cette classe s'illustre par sa forte tendance à citer les sources issues de Wikipédia dans le cadre professionnel ou académique.

Les lecteurs appartenant à cette classe se distinguent par des caractéristiques sociodémographiques marquantes, qui témoignent d'un positionnement social privilégié. Cette classe est la plus diplômée et rassemble la part la plus importante d'individus ayant un master ou un doctorat. Cette distinction s'accompagne d'une forte représentation des

catégories socio-professionnelles supérieures (directeurs, cadres de direction, gérants, professions intellectuelles ou scientifiques). Le sentiment d'aisance financière y est également particulièrement marqué. Par ailleurs, elle présente la proportion de femmes la plus basse parmi l'ensemble des classes étudiées. Cependant, les *Sur-informés* font face à une contrainte notable : le temps libre. Ils figurent parmi ceux disposant du moins de temps pour des activités de loisirs. À l'image des *Ambivalents*, ils présentent un bon niveau d'aisance numérique mais mobilisent Wikipédia de manière sélective pour répondre aux besoins. Dans cette classe, la pratique de la lecture papier est sélective. Si l'intérêt pour ce support demeure modéré, la lecture numérique y est nettement plus développée, en adéquation avec des usages numériques très diversifiés. Cette classe est celle qui s'informe le plus, tous supports confondus. Elle se distingue particulièrement par son recours fréquent aux publications scientifiques.

Interprétation : Cette classe incarne un profil socio-économique privilégié, avec un haut niveau de qualification, une réussite professionnelle et une aisance financière significative, mais elle doit composer avec des contraintes temporelles. Les *Sur-informés* se caractérisent par une quête d'information constante, appuyée sur des pratiques numériques variées et une confiance affirmée dans les médias, particulièrement les médias « traditionnels ». Cette consommation importante d'information s'appuie sur la reconnaissance de Wikipédia en tant qu'institution de référence mais également sur une confiance dans la capacité à savoir sélectionner l'information et l'interpréter, à la différence des *Ambivalents*. Cette attitude marquée à citer ses sources pourrait s'expliquer par la concentration importante de contributeurs parmi les *Sur-informés*, conscients du temps nécessaire à la structuration de la connaissance sur la plateforme.

Classe 05 (les Confiants)

Les Confiants apparaissent comme la classe où les besoins personnels, professionnels et d'études sont les plus significatifs et variés. Ces besoins couvrent une large gamme d'objectifs, tels que trancher un désaccord, vérifier une information, se cultiver, se renseigner, effectuer un choix ou s'informer sur des sujets politiques. Dans ce contexte, Wikipédia occupe une place centrale. Cette classe est celle qui mobilise le plus Wikipédia et où Wikipédia y répond le mieux. À l'instar des Sur-informés, les Confiants sont prompts à citer les sources issues de Wikipédia.

Cette classe est l'une de celles où la proportion de femmes est la plus élevée. Également l'une des classes les plus jeunes, elle incarne une population en pleine activité ou dans la poursuite d'études. Avec un niveau de diplôme relativement élevé, une proportion non négligeable de catégories socio-professionnelles supérieures et un sentiment d'aisance financière marqué, les *Confiants* reflètent une image jeune, qualifiée et, pour une partie, économiquement stable, mais confrontée aux exigences temporelles d'un rythme de vie qui limite leur temps libre disponible. C'est la classe la plus à l'aise avec le numérique et celle avec les usages les plus diversifiés. Cette classe se distingue par son rapport singulier à la lecture et à l'information. La pratique de la lecture papier y est sélective, témoignant d'un intérêt modéré

pour ce support traditionnel. En revanche, elle excelle dans la lecture numérique, se plaçant comme la classe où cette pratique est la plus développée. Elle s'informe de manière régulière, tout en exploitant de multiples canaux. Les moteurs de recherche jouent un rôle central dans cette démarche, cette classe étant celle qui les mobilise le plus pour accéder à de l'information. Par ailleurs, elle s'illustre par son niveau de confiance élevé envers les médias en général mais surtout la confiance plus élevée dans les médias "numériques" (particulièrement concernant les publications scientifiques, les moteurs de recherche ainsi que les réseaux sociaux).

Interprétation : Les *Confiants* incarnent un profil connecté et curieux, misant sur les ressources numériques et les médias en ligne dans sa quête d'information. Cette classe incarne un usage intensif et centré sur Wikipédia, en exploitant intensément ses ressources pour répondre à des besoins multiples et variés.

Conclusion

Cette typologie des lecteurs de Wikipédia met en lumière des disparités et révèle des dynamiques d'usages profondément structurées suivant les caractéristiques sociodémographiques, les pratiques culturelles et les modes de consommation informationnels des individus. Les cinq classes identifiées illustrent à la fois la diversité des usages de l'encyclopédie en ligne comme les rapports différenciés au savoir et à l'information.

Les *Petits lecteurs* représentent une classe peu consommatrice d'information, limitée dans ses pratiques numériques comme de lecture, et éloignée d'un usage régulier ou approfondi de Wikipédia. Inscrits dans un rapport ordinaire des classes populaires à la connaissance légitime et à la forme écrite, ils peinent à s'appropriier le savoir encyclopédique. À l'opposé, les *Confiants* incarnent une classe jeune et connectée, plaçant Wikipédia au cœur de ses usages informationnels, tant personnels que professionnels. Ce groupe, qui mobilise largement l'encyclopédie pour répondre à ses besoins, reflète un rapport affirmé et confiant au numérique, à la lecture numérique et aux médias en ligne. Entre ces deux profils, les *Conformistes*, *Ambivalents*, et *Sur-informés* adoptent des usages hétérogènes de Wikipédia, témoins de leurs positionnements spécifiques. Les *Conformistes* exploitent Wikipédia de manière sélective, en complément d'autres médias traditionnels qu'ils valorisent davantage. Les *Ambivalents*, bien qu'ils fassent preuve de méfiance envers Wikipédia et peinent parfois à s'appropriier le contenu, l'intègrent de manière pragmatique dans leurs démarches informationnelles. Enfin, les *Sur-informés*, avec leurs pratiques numériques diversifiées et leur appétit constant pour l'information, reconnaissent Wikipédia comme une institution de savoir incontournable.

Plus globalement, cette typologie interroge quant à la manière dont Wikipédia structure son contenu et présente l'information, compte tenu de la diversité des modes d'appropriation des lecteurs de la plateforme. Ces premiers résultats montrent que le processus de différenciation social des pratiques de lecture opère également sur la plateforme et que

l'accès à l'information sur Wikipédia renvoie à des pratiques "ordinaires" de lecture. Néanmoins, cette typologie de lecteurs de Wikipédia met également en lumière des modes d'appropriation spécifiques de l'information qui s'écartent des représentations sociales de la lecture. Ces résultats préliminaires ouvrent la voie à des réflexions sur la manière de saisir les liens entre besoins informationnels, compétences numériques et rapports différenciés au savoir. Longtemps critiqué et relégué par les tenants du savoir légitime (institution scolaire, milieux académiques, professionnels de l'information, acteurs politiques, etc.), Wikipédia tend progressivement à en adopter les valeurs. Mais s'inscrire dans cette continuité ne restreint-il pas l'accès au savoir aux publics les plus en difficulté avec la forme écrite et la culture lettrée ?

Bibliographie

Barth-Rabot, C. (2023). *La lecture. Valeur et déterminants d'une pratique*, Paris, Armand Colin.

Bennett, T., Savage, M., Silva, E., Warde, A., Gayo-Cal, M., & Wright, D. (2009). *Culture, Class, Distinction*. Routledge.

Coulangeon, P. (2010). *Sociologie des pratiques culturelles*. La Découverte.

Jemielniak, D., & Aibar, E. (2016). « Bridging the gap between Wikipedia and academia ». *Journal of the Association for Information Science and Technology*, 67(7), 1773-1776.

Lahire, B. (1993), « Lectures populaires : les modes d'appropriation des textes », *Revue française de pédagogie*, 104, 17-26.

Lombardo, F., Daly, M. (2021). « Analyzing suicide life stories on Wikipedia with Highway_star and other textual visualization tools ». *SN Soc Sci* 1, 268.

Poissenot, C. (2019). *Sociologie de la lecture*. Paris, Armand Colin.

Sahut, G. (2016). « Construire une encyclopédie avec un wiki ? Regards rétrospectifs sur la politique éditoriale de Wikipédia ». *I2D - Information, données & documents*, Volume 53(4).

Thumala Olave, M. A. (2018). « Reading Matters: Towards a Cultural Sociology of Reading ». *American Journal of Cultural Sociology*, 6(3), 417-454.