



**HAL**  
open science

# Les minorités linguistiques sur le Web, un cas d'étude pour les humanités numériques

Cécile-Marie Martin

► **To cite this version:**

Cécile-Marie Martin. Les minorités linguistiques sur le Web, un cas d'étude pour les humanités numériques. Humanistica 2023, Association francophone des humanités numériques, Jun 2023, Genève, Suisse. hal-04649301

**HAL Id: hal-04649301**

**<https://hal.science/hal-04649301v1>**

Submitted on 16 Jul 2024

**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 4.0 International License

# Les minorités linguistiques sur le Web, un cas d'étude pour les humanités numériques

Anne-Cécile Martin

Université des Antilles

cecile.martin@univ-antilles.fr.

## Résumé

Nous souhaitons aborder les questions de mondialisation, de transferts culturels, d'histoire transnationale, de tradition et d'innovation. Pour cela, nous évoquons des projets Web réalisés par et pour des minorités linguistiques, qui permettent de faire dialoguer les humanités numériques avec la création contemporaine, dans une optique de recherche en sciences humaines, notamment de pratiques computationnelles de valorisation de la culture et de la recherche.

## 1 Introduction

Si pour certains le Web consacre l'idée d'un village global (MacLuhan, 2004), ou encore d'un dispositif de mondialisation de la culture (Warnier, 2017), au sens d'un espace d'expression affranchi des frontières nationales, dans les faits le caractère disruptif de ce territoire peut être relativisé. De nombreux éléments sont en effet intrinsèquement liés à la structure étatique, des noms de domaines aux adresses mails en passant par les modalités d'accès au réseau. Ainsi, la mise en œuvre d'un modèle du Web pensé comme un espace universel (Berners-Lee et Cailliau, 1990; Berners-Lee et al., 1992), nécessite de prendre en compte la complexité de ce dispositif qui comporte une multiplicité de niveaux (Brondizio et al., 2013).

	2013 Jan	2015 Jan	2017 Jan	2019 Jan	2021 Jan	2023 Jan
English	54.9%	55.6%	53.2%	54%	60.4%	57.7%
Spanish	4.6%	4.7%	5%	4.9%	4%	4.5%
Russian	5.5%	5.8%	6.4%	6%	8.6%	5.3%
German	6.5%	6%	5.4%	6%	2.4%	3.9%
French	4.6%	4%	4%	4%	2.6%	3.9%
Japanese	4.5%	5%	5.7%	3.4%	2.1%	3.2%
Portug.	2.3%	2.4%	2.6%	2.9%	1.3%	1.8%
Turkish	1.3%	1.4%	1.6%	1.2%	3.6%	2.6%
Italian	1.6%	1.9%	2.3%	2.3%	0.8%	1.7%
Persian	0.8%	0.8%	1.5%	2%	3%	2.2%

TABEAU 1 – Tendances annuelles historiques des statistiques d'utilisation des langues de contenu pour les sites web. Source : W3 Techs.

Parmi les freins identifiés pour démocratiser l'ac-

cess à cet espace, force est de constater que le nombre de langues mises en avant pour les activités d'écriture sur le Web, la création de contenus, la recherche et la navigation se trouvent au premier plan. De là, il paraît judicieux de se demander dans quelle mesure le Web peut-il être un espace d'expression pour les minorités linguistiques alors que la gestion de ce dispositif est assurée par des acteurs utilisant majoritairement l'anglais pour publier (57,7% des contenus en janvier 2023, cf. tab 1) ou programmer (les 50 langages les plus utilisés en janvier 2023 sont à base syntaxique anglaise, cf. tab 2).

2003	2013	2023
Java	C	Python
C	Java	C
C++	C++	C++
PHP	C#	Java
SQL	PHP	C#
Javascript	Python	Visual Basic
C#	Javascript	Javascript
Python	Visual Basic	PHP

TABEAU 2 – Évolution des langages de programmation les plus populaires entre 2003 et 2023. Source : TIOBE.

Pour cette présentation, nous avons sélectionné différents exemples d'initiatives développées par des minorités linguistiques (Lespinay, 2016), telles que la communauté warumungu en Australie, les communautés wayana et apalaï en Guyane française, la communauté māorie en Nouvelle Zélande.

Le choix de ces projets s'explique par la démarche commune des populations qui en sont à l'initiative, à savoir de s'engager pour être reconnues en tant que communautés fédérées autour d'un patrimoine et d'une identité au sein d'un ensemble plus vaste, la dimension nationale, dans laquelle elles ont été généralement intégrées contre leur gré.

## 2 Mukurtu, un système de gestion de contenu conçu avec et pour les populations autochtones (Australie)

Le système de gestion de contenu (SGC) ou *content management system* (CMS) Mukurtu<sup>1</sup> est le fruit d'une collaboration entre la communauté aborigène warumungu, le développeur Craig Dietrich et la chercheuse Kimberley Christen (2005). Après avoir observé les prémices de la mise en ligne du patrimoine warumungu par les institutions nationales, l'objectif de l'équipe était de développer une solution pour administrer le patrimoine numérisé de la communauté, une solution qui devait être en phase avec la culture warumungu. En effet, dans l'approche classique de la conception d'un site web, la navigation est conçue en cherchant à anticiper le comportement de l'utilisateur et il est courant de considérer que la rapidité avec laquelle ce dernier peut trouver une ressource est un gage de qualité. Mais dans la culture warumungu, pour accéder aux contenus, l'utilisateur doit passer par des étapes d'identification qui amènent à reconsidérer le processus de navigation. L'utilisateur peut être identifié par son genre, son âge et son rapport à la communauté : est-il membre de la communauté ou non ? quelle place occupe-t-il au sein de celle-ci ? L'accès est conditionné par un certain nombre de paramètres, certains liés à l'identité du visiteur, d'autres à des facteurs externes tels que la période de l'année ou la localisation géographique.

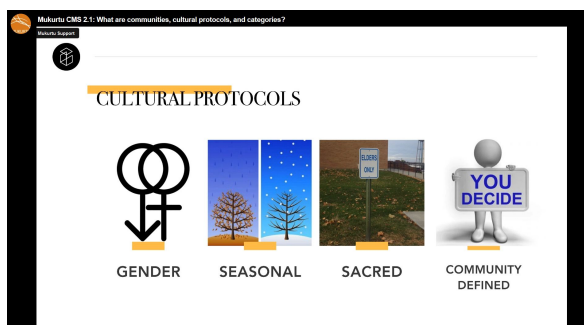


FIGURE 1 – Les protocoles culturels proposés dans Mukurtu couvrent 4 catégories : le genre, la saison, le caractère sacré, la communauté.

Le CMS Mukurtu a donc été conçu pour répondre aux besoins des membres des communautés, aussi bien dans le rôle de créateur et d'administrateur de contenu que de visiteur. La navigation a été « déconstruite », repensée, pour anticiper une démarche spécifique. L'équipe a réussi l'exploit de

1. <https://mukurtu.org>.

coordonner la réponse à ces besoins tout en assurant l'intégration de cette solution dans le système Web, et en facilitant sa réappropriation par d'autres communautés. Le travail de conception du site a été analysé d'un point de vue technique et culturel, amenant à la conception de protocoles culturels (cf. fig. 1). L'architecture du CMS a été conçue pour une articulation optimale de ces trois niveaux : base de données, interface administrateur, interface visiteur.

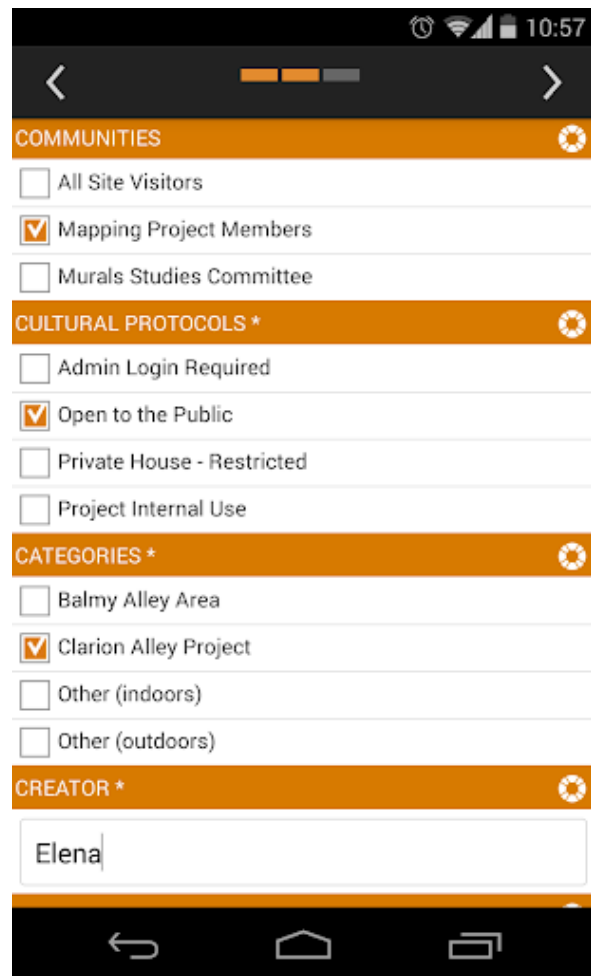


FIGURE 2 – Interface administrateur de la version mobile de Mukurtu.

Une fonction permet ainsi aux administrateurs de gérer les droits d'accès en appliquant des filtres par contenu ou pour une sélection de contenus, pour une période donnée et de manière réversible. Il est également possible d'utiliser la même base pour développer différents sites web, des pages privées ou publiques. Une attention a été portée à l'intégrité des contenus, en cas d'exploitation de ces contenus par des tiers, avec l'ajout simplifié de métadonnées. Un soin particulier a été apporté à la question de l'appropriation de l'outil, mais aussi

à la question de l'autonomie et de la littératie numérique (Drot-Delange, 2014). La partie logicielle est libre de droits et des formations sont proposées pour apprendre à l'utiliser. Le développement de cet outil de création de site web a été l'occasion de porter un regard neuf sur certaines idéologies à l'œuvre dans les principes de conception des sites web.

Ceux-ci sont en effet développés pour servir plusieurs objectifs, tels que le partage mais aussi la protection de ressources qui souvent, dès lors qu'elles sont publiées sur le Web, peuvent devenir des biens communs. Cependant, cette idéologie du libre accès, défendue par les pionniers du Web pour s'opposer à sa marchandisation et à sa privatisation, peut dans certains cas s'apparenter, comme la situation de départ de ce projet le montre, à une forme de domination sur le patrimoine culturel des minorités autochtones, puisque celles-ci dans leur grande majorité n'ont pas été consultées pour savoir si elles entendaient y souscrire, et de quelle manière (Martin, 2018).

### 3 WATAU, un portail fait avec et pour les Wayana et Apalaï (Guyane française)

Dans le cadre du projet SAWA (Savoirs Autochtones Wayana-Apalaï)<sup>2</sup>, les Wayana et Apalaï ont participé à la création de WATAU<sup>3</sup>, un portail destiné à leurs communautés et notamment aux plus jeunes. Ancré dans un processus d'autodétermination et d'affirmation identitaire, SAWA a été un projet collaboratif réunissant des représentants des communautés, des chercheurs et des conservateurs. Il s'est déroulé entre 2016 et 2020 et avait pour objectif principal de valoriser et de faciliter l'accès pour les populations Wayana et Apalaï à un ensemble d'enregistrements sonores, de collections audiovisuelles et photographiques, ainsi qu'à des collections d'objets représentatifs de leur culture.

Le projet s'est accompagné d'une réflexion sur les pratiques de restitution et leur impact sur la transmission des savoirs traditionnels, ainsi que sur les modalités d'appropriation des nouveaux outils et technologies de médiation dans les communautés autochtones. Son originalité est de donner un rôle central aux communautés autochtones en créant les conditions d'une participation active d'une équipe wayana-apalaï. Non seulement cette équipe a été

2. <https://watau.fr/s/watau-fra/page/projet-SAWA>.

3. <https://watau.fr>.

impliquée dans le choix des contenus, mais surtout elle a participé aux choix graphiques et à la mise en page du site. Ainsi, WATAU a été conçu comme un portail multilingue dont les formes et les conditions d'accès ont été définies en premier lieu par les participants amérindiens dans leurs propres langues. Enfin, la démarche du projet est exemplaire en matière de littératie numérique, puisque des ateliers encadrés par un ethnolinguiste ont permis de traduire le lexique numérique en langues wayana et apalaï (Camargo et al., 2020). Ceci a permis d'éviter de recourir aux emprunts lexicaux, que ce soit à l'anglais, au français ou au portugais. Ces ateliers ont ainsi favorisé la démarche pour ces communautés de s'approprier ces savoir-faire dans leur langue maternelle.



FIGURE 3 – Page d'accueil du site WATAU, interface visiteur en langue wayana

### 4 Emotikis, une sélection d'émojis réalisés par une institution maorie pour la communauté maorie (Nouvelle-Zélande)

Jouant un rôle important dans la communication numérique, dans les courriels ou la messagerie instantanée, les émojis sont un outil puissant pour l'inclusion technologique. Apparus dans les années 1990, leur usage s'est démocratisé à partir des années 2000. Au fil des ans, des voix se sont élevées pour dénoncer le manque de diversité, notamment en matière de couleur de peau ou d'orientation sexuelle. Ces critiques ont conduit à l'élaboration régulière de nouveaux pictogrammes. Mais pour pallier les limites d'un système occidental auto-centré, certaines communautés d'utilisateurs ont également mis en place d'autres moyens d'exprimer leurs spécificités culturelles. Des stratégies d'appropriation et de détournement de jeux de caractères imposés ont ainsi été mises en évidence dans certaines cultures, comme c'est le cas pour le monde arabophone ou pour la Chine depuis les

années 1980 (Yu, 2021).

Plus récemment, en 2016, une équipe du centre culturel Te Puia Māori en Nouvelle-Zélande a entrepris une autre démarche. Ils se sont lancés dans un processus créatif. En effet, l'équipe de Te Puia a conçu plus de 150 émojis qu'ils ont appelés Emotikis. Il s'agit de la représentation d'objets traditionnels de la culture māorie, tels que des canoës, des armes traditionnelles et un ensemble de tikis, la figure mi-humaine mi-dieu, de la couleur des rochers Pounamu, à l'origine de l'humanité dans le mythe māori de la création du monde. Ces Emotikis forment un large éventail d'artefacts et d'expressions faciales. L'ensemble comprend même un certain nombre de gifs animés.

Si ces émojis māori peuvent sembler avoir été conçus pour s'amuser, la directrice générale du centre culturel, Kiri Atkinson-Crean, rappelle qu'« ils donnent aux jeunes Māori un moyen de s'engager les uns avec les autres en ligne en utilisant des signifiants de leur propre culture » (Lewis, 2016).



FIGURE 4 – Emotikis utilisables dans les applications de messagerie, courriel et traitement de texte

## 5 Conclusion

Les exemples d'initiatives permettant aux minorités autochtones d'établir leur représentativité sur le Web ont servi à illustrer notre propos et à identifier les possibilités d'une présence numérique (Merzeau et al., 2010) pour ces communautés. Tout d'abord nous avons vu que les éléments présentés sont liés aux communautés qui les ont développés, mais que les démarches et les livrables sont pensés pour être transférables dans d'autres cultures. Par ailleurs, ces projets tendent à prouver qu'il n'est pas nécessaire de s'engager dans la voie de l'acculturation pour assurer la présence des minorités linguistiques sur le Web, et démontrent plutôt l'intérêt de mettre en œuvre une forme alternative de modernité (Habermas, 1998). Ces exemples ont permis

de donner un aperçu de l'intérêt du développement des compétences dans le domaine de la médiation, de la communication et de la formation au numérique. Entendues comme un ensemble d'outils et de pratiques favorisant la littératie numérique, nous faisons l'hypothèse que les humanités numériques peuvent soutenir le développement de compétences pour comprendre et utiliser les technologies numériques, favorisant ainsi la réduction de la fracture numérique dans sa dimension multiculturelle.

En mobilisant des savoirs issus des sciences politiques, des sciences de l'information et de la communication, des sciences de l'éducation et des sciences ethnolinguistiques pour traiter de l'actualisation de dispositifs numériques de création et de partage de connaissances (Paquelin, 2004), les activités d'analyse de processus de démocratisation du Web sur le plan multiculturel sont au cœur du champ de réflexion des humanités numériques. L'inscription dans une telle démarche, en contexte didactique, semble pouvoir offrir aux représentants de communautés minoritaires linguistiques qui expriment le souhait un soutien au développement de « méta-compétences » en situation de transmission d'information dans un contexte multilingue, en facilitant notamment le dialogue entre représentants de différentes communautés mais aussi de différents statuts et corps de métiers.

De plus, en participant à l'encadrement d'un travail de réflexion pour l'élaboration des modalités d'information sur les aspects techniques (infrastructure réseau, fonctionnement des équipements, des protocoles) et de gouvernance du Web en contexte multilingue, notre hypothèse est que les humanités numériques peuvent œuvrer pour réaliser un maillage soutenant la construction d'un espace d'expression commun. Si la dimension universelle du Web doit être relativisée (Liquète, 2014), il s'agirait alors de tenter de maintenir une cohérence et un équilibre entre des zones autonomes chaînées entre elles, l'accès étant conditionné par l'acquisition de compétences « véhiculaires » techniques et lexicales, garantes d'une forme de chaînage interculturel.

Ces projets montrent qu'encourager la publication mais aussi la traduction d'informations sur le fonctionnement du Web pourrait s'avérer judicieux, car c'est en ayant une vision claire du fonctionnement de cet ensemble et du rôle qu'y jouent acteurs privés et publics que certaines barrières à l'entrée pour les minorités linguistiques pourraient être le-

vées. L'inscription dans une telle démarche offre des opportunités aux représentants des communautés linguistiques minoritaires pour développer une utilisation autonome du Web, d'une manière qui ne dilue pas ou ne compromet pas leur identité, mais qui au contraire la renforce.

## Bibliographie

- Tim Berners-Lee et Robert Cailliau. 1990. *WorldWideWeb. Proposal for a HyperText Project*. CERN, Genève.
- Tim Berners-Lee, Robert Cailliau, Jean-François Groff, et Bernd Pollermann. 1992. *World wide web : The information universe*. *Internet Research*, 2(1) :52–58.
- Eduardo S. Brondizio, Elinor Ostrom, et Oran R. Young. 2013. *Connectivité et gouvernance des systèmes socio-écologiques multiniveaux : le rôle du capital social*. *Management & Avenir*, 7(7) :108–140.
- Éliane Camargo, Veronice Holguin Lew, Sara Tandar, et Équipe Wayana et Apalaï. 2020. *L'amazonie amérindienne dans l'ère du numérique : le portail multilingue watau*. *Patrimoines du Sud*, 12.
- Kimberley Christen. 2005. *Gone digital. aboriginal remix and the cultural commons*. *International Journal of Cultural property*, 12(3) :315–345.
- Béatrice Drot-Delange. 2014. *Littératie informatique : quels ancrages théoriques pour quels apprentissages? Spirale*. *Revue de recherches en éducation*, 53 : Littéracies en contexte d'enseignement et d'apprentissages :121–132.
- Jürgen Habermas. 1998. *L'intégration républicaine. Essais de théorie politique*. Fayard, Paris.
- Laurence Ifrah. 2010. *L'information et le renseignement par Internet*, chapitre Histoire des moteurs de recherche. Presses Universitaires de France, Paris.
- Charles de Lespinay. 2016. *Les concepts d'autochtone (indigenous) et de minorité (minority)*. *Droit et cultures*, 72 :19–42.
- Danny Lewis. 2016. *Emotikis and new keyboards bring indigenous cultures to text messaging. from maori emojis to first nations languages*. *Smithsonian Magazine*.
- Vincent Liguète. 2014. *Les Oubliés de l'internet. Culture et langues sur l'Internet oubli ou déni?*, chapitre Les composantes communicationnelles et cognitives des cultures de l'information face à la massification de la culture. Les Etudes hospitalières, Bordeaux.
- Marshall MacLuhan. 2004. *Pour comprendre les médias*. Editions du Seuil, Paris.
- Cécile-Marie Martin. 2018. *Le système mukurtu : une ouverture sur l'interculturel*. *Hermès*, 3(3) :238–243.
- Louise Merzeau, Hervé Le Crosnier, Vincent Denis, Jean-Marc Picard, Brigitte Juanals, et Sylvère Mercier. 2010. *Une nouvelle dimension de l'information*. *Documentaliste-Sciences de l'Information*, 1(1) :32–41.
- Didier Paquelin. 2004. *Le tutorat : accompagnement de l'actualisation du dispositif*. *Distances et savoirs*, 2 :157–182.
- Alain Rallet et Fabrice Rochelandet. 2004. *La fracture numérique : une faille sans fondement?* *Réseaux*, 5(6) :19–54.
- Isabelle Rigoni. 2010. *Éditorial. les médias des minorités ethniques. représenter l'identité collective sur la scène publique*. *Revue Européenne des Migrations Internationales*, 26(1) :7–16.
- Alexandre Serres. 2000. *Aux sources d'Internet : l'émergence d'ARPANET. Exploration du processus d'émergence d'une infrastructure informationnelle. Description des trajectoires des acteurs et actants, des filières et des réseaux constitutifs de la naissance d'ARPANET. Problèmes critiques et épistémologiques posés par l'histoire des innovations*. Thèse de doctorat, Université Rennes 2, Rennes.
- Tim Smith et François Flückiger. 2021. *Les versions publiques du logiciel du web*. CERN, Genève.
- Ferdinand Tönnies. 2010. *Communauté et société*. PUF, Paris.
- Geneviève Vinsonneau. 2012. *Mondialisation et identité culturelle*. De Boeck Supérieur, Louvain-la-Neuve.
- Jean-Pierre Warnier. 2017. *La mondialisation de la culture*. La Découverte, Paris.
- Max Weber. 1995. *Economie et société*. Pocket, Paris.
- Mengyang Yu. 2021. *Les spécificités des émoticônes chinoises*. *La linguistique*, 57(3) :165–187.