

Modélisation thématique

Guillaume Poirier

Le sujet : Office du tourisme d'Ille-et-Vilaine

Consultation citoyenne “comment agir pour un tourisme plus durable en France”
(1493 réponses)

Quelles actions simples pourrait-on mener en priorité pour tendre vers un tourisme responsable ?



Le contexte en quelques mots : le déplacement ça coûte..



Vélo ou marche

0 kg CO₂e



RER ou Transilien

4.89 kg CO₂e



Autocar

14.7 kg CO₂e



TGV

1.47 kg CO₂e



Vélo (ou trottinette) à assistance

5.48 kg CO₂e



Voiture (moteur thermique)

- Avec 3 covoitureurs +

27.2 kg CO₂e



Tramway

2.14 kg CO₂e



Bus (moteur électrique)

10.9 kg CO₂e



Scooter ou moto légère

38.2 kg CO₂e



Métro

2.22 kg CO₂e



Voiture (moteur électrique)

12.9 kg CO₂e



Moto

95.7 kg CO₂e



Intercités

4.49 kg CO₂e



TER

13.8 kg CO₂e



Avion (court courrier)

129 kg CO₂e

■ construction ■ usage

Le contexte en quelques mots : quelques chiffres



Année	Changement d'affectation du sol	Energies fossiles et production de ciment	Total (CO2)	Notes
1960	9.3	6.8	16.1	
1970	14.8	5.1	19.9	
1980	19.3	4.4	23.7	
1990	22.5	5.2	27.7	Année de référence Protocole de Kyoto
2000	25.2	5.4	30.6	
2010	32.8	5.2	38	
2015	34.8	5.6	40.4	Accord de Paris
2020	34.2	4.3	38.5	Pandémie de COVID-19
2021	36	4.3	40.3	
2022	36.4	4.3	40.7	
2023	36.8	4.1	40.9	Projection

Emissions de CO2 en milliards de tonnes (Gt) de CO2
Source : Global Carbon Project 2023 - [Friedlingstein et al](#) (2023).



Puits carbone : ~ 50 % du total des émissions en 2020 (Rapport du GIEC)

Part du tourisme : 11% des émissions GES en 2022 (ADEME)

Les données : description générale

- Recolte des données depuis Make.org
- 1493 propositions (phrases) d'environ une dizaine de mots
- Chaque propositions est voté :
- d'accord :
 - Coup de coeur
 - Banalité
 - Réaliste
- Sans opinion :
 - JPas d'avis
 - Pas compris
 - Indifférent
- Pas d'accord :
 - Infaisable
 - Surtout pas
 - Banalité

Process de traitement

1. Pre-processing (Tokenisation, Normalisation, Lemmatisation, Stops words)
2. Présentation en nuage de mots
3. Vectorisation des mots : représentation des mots dans un espace numérique
4. Représentation en 2D des mots
5. Représentation en 2D des phrases
6. Clustering : Regroupement des proposition en différents thèmes
7. Représentation en 2D des clusters
8. Topic modeling avec LDA : une autre méthode statistique

Première représentation : nuage de mots



Deuxième représentation : mots en 2D

Les 10 mots les plus proches de déchet / voir rapport html

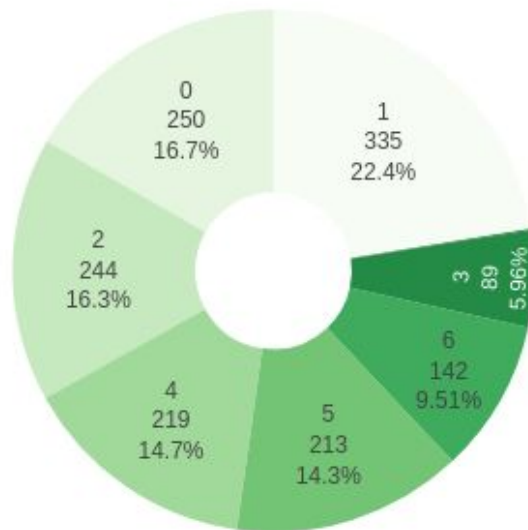
Terme	Score
Ramasser	0.979
Montrer	0.978
Déchets	0.976
Besoin	0.976
Scolaire	0.975
Hôtellerie	0.975
Non	0.975
Établir	0.974
Ramassage	0.974
Amende	0.974

Troisième représentation : phrases en 2D

L'idée est simple : pour représenter une phrase, on fait la moyenne des vecteurs des mots / voir rapport html

Troisième représentation : clusters en 2D

Poids des clusters



Troisième représentation : clusters en 2D

Les mots les plus présents par clusters

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
responsable (61), durable (47), local (29), acteur (28), creer (22), site (21), eco (20), mettre (18), chaque (17), vacance (15), tourism (15), territoire (14), label (14), mass (14), developper (14)	voyage (28), mettre (27), impact (26), developper (25), site (23), place (23), vacance (22), permettre (22), proposer (19), hebergement (18), activite (16), lieux (16), pouvoir (16), certain (14), velo (14)	camping (35), limiter (29), arreter (28), interdire (27), car (23), pouvoir (23), eau (22), developper (21), grand (20), autre (19), plastique (18), hotel (17), petit (16), aire (16), nombre (16)	cyclable (35), developper (28), piste (27), camping (20), velo (18), voie (15), car (14), ville (12), itineraire (10), petit (9), territoire (8), local (8), rural (8), vert (8), entrer (8)	local (80), environnement (33), sensibiliser (30), dechet (24), produire (22), respecter (18), inciter (16), action (16), favoriser (16), nature (15), bon (15), circuit (14), proposer (14), region (13), etre (13)	transport (93), train (85), velo (48), developper (31), commun (26), avion (24), acce (21), interdire (21), vacance (21), vol (19), voiture (18), sejour (18), faciliter (17), rendre (16), taxe (16)	naturel (42), lieux (35), limiter (24), mettre (24), poubelle (24), tri (24), site (22), parc (17), dechet (14), toilette (14), espace (12), frequentation (11), mieux (10), afin (10), eviter (9)	assouplir (1), carcan (1), administratif (1)

Troisième représentation : clusters en 2D

Les votes par clusters / voir rapport html

	content	agree_count	neutral_count	disagree_count	agree_proportion_percent	difference_from_average_percent
cluster_label						
0	250	27353	7519	6090	66.776525	-4.001598
1	335	32702	9261	9101	64.041203	-6.736919
2	244	32791	7498	7827	68.149888	-2.628235
3	89	17415	3043	2495	75.872435	5.094312
4	219	49475	7651	5702	78.746737	7.968614
5	213	30801	6922	9032	65.877446	-4.900676
6	142	30844	4943	4309	76.925379	6.147256
7	1	25	13	5	58.139535	-12.638588

Quatrième représentation : une autre méthode

voir rapport html