

Utilisation des outils d'extraction et filtrage de données HERE

Version 3

Ces outils permettent, en mode totalement autonome, l'extraction et le filtrage de données brutes de vitesses HERE.

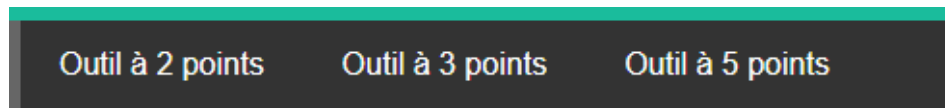
Les étapes à suivre afin d'obtenir les données de vitesses sont :

- 1- Identifier les liens HERE
- 2- Regrouper les liens HERE dans un fichier
- 3- Extraire les vitesses associées aux liens sélectionnés et filtrage fin des données

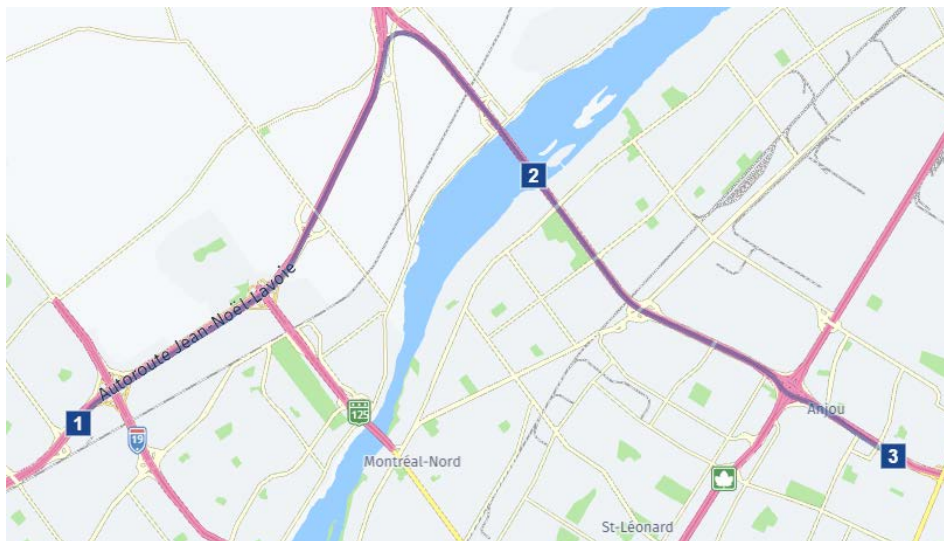
1 Outil d'identification de liens HERE

L'outil est basé sur une page Web <https://visualgis.w3spaces.com/index.html> qui permet de construire en mode interactif des trajets constitués de liens HERE. **Les codes extraits de cette page ne doivent être utilisés que pour extraire des données de vitesse compatibles avec la dernière version du réseau HERE.**

En fonction de la complexité du trajet à définir, sur la page Web, il y a la possibilité de construire des trajets à l'aide de 2, 3 ou 5 points obligatoires de passage identifiés par des marqueurs colorés en bleu.

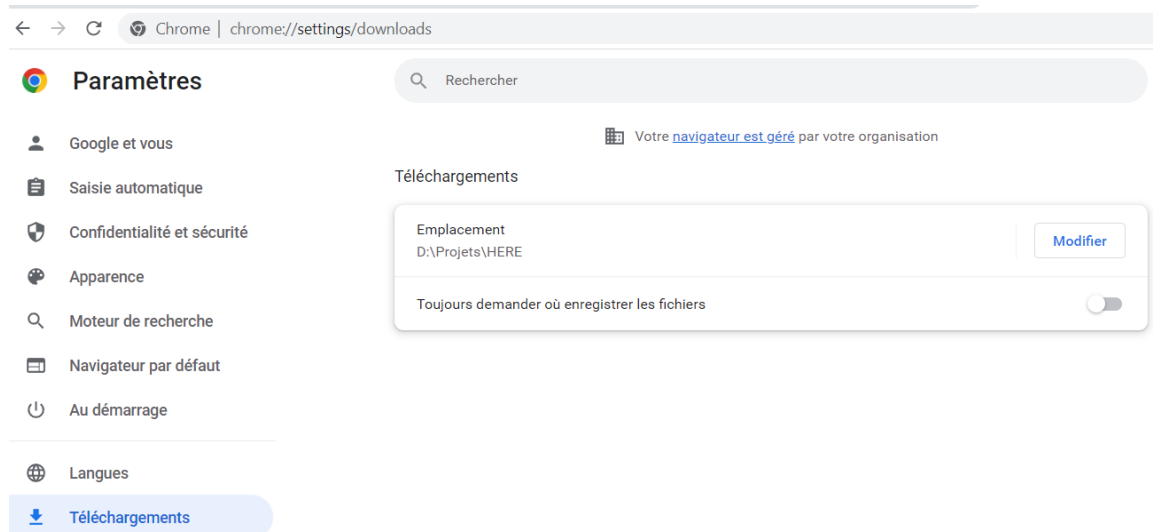


Le sens du trajet défini est toujours du marker no. 1 vers les suivants :



1.1 Configurer le répertoire de téléchargement

Pour utiliser facilement l'outil d'extraction et l'outil de filtrage, il est recommandé de définir dans le navigateur un répertoire spécifique pour le téléchargement des fichiers contenant des liens HERE.



Dans le répertoire désigné, tous les fichiers contenant les liens de chaque trajet seront sauvegardés un par un.

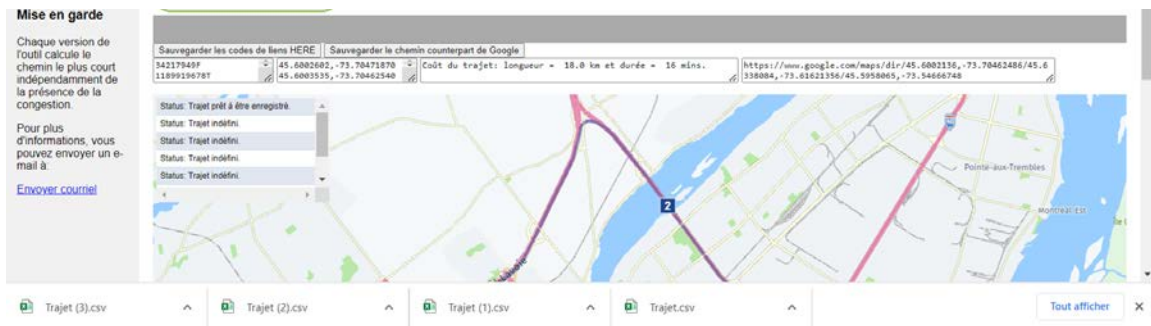
1.2 Définir et sauvegarder les trajets

Les trajets sont définis en déplaçant les marqueurs sur la carte puis en cliquant sur la carte avec la souris. En fonction de la complexité du trajet, il est possible de choisir trois types de marqueurs : à 2, à 3 et à 5 points obligatoires de passage.

Dès que le trajet montré sur la carte reproduit le trajet voulu, il est possible de sauvegarder les liens HERE associés au trajet en cliquant sur le bouton <Sauvegarder les codes de liens HERE>.

Sauvegarder les codes de liens HERE		Sauvegarder le chemin counterpart de Google	
34217949F	45.6002602,-73.70471870	45.6003535,-73.70462540	Coût du trajet: longueur = 18.0 km et durée = 16 mins.
1189919678T			

Les noms de tous les fichiers sauvegardés dans le répertoire cible sont une série comme celle-ci : Trajet.csv, Trajet (1).csv, trajet (2).csv, ...

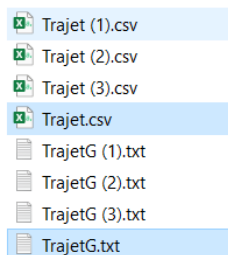


Avant de commencer une séance de codification, il est fortement recommandé de vider le répertoire cible de tout fichier qui commence par le mot Trajet et le type est csv. Pour garder une trace entre le nom de chaque fichier de trajet et le nom du trajet, il est recommandé de créer un fichier texte avec deux colonnes identifiant les trajets comme ceci :

Nom_fichier	Cod_trajet
Trajet.csv	Trajet_A
Trajet (1).csv	Trajet_B
*****	*****
Trajet (k).csv	Trajet_X

Le fichier de correspondance permettra plus tard de bien associer chaque donnée de vitesse au bon trajet.

Il y a aussi la possibilité d'enregistrer la contrepartie Google pour chaque itinéraire défini. Si c'est le cas il faut le faire systématiquement pour garder la bonne correspondance.

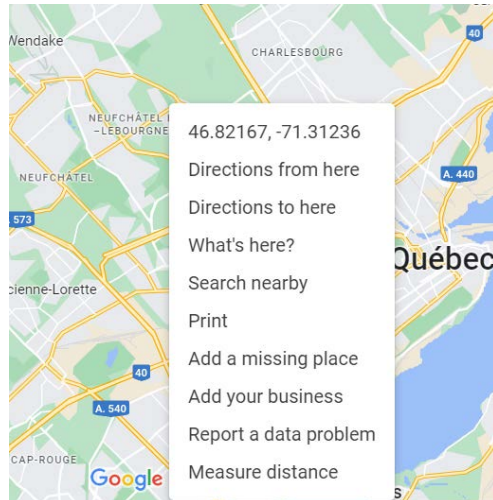


Comme il est mentionné dans la page Web, le chemin trouvé est calculé sans prise en compte de la présence du trafic, et donc est basé sur la vitesse à écoulement libre.

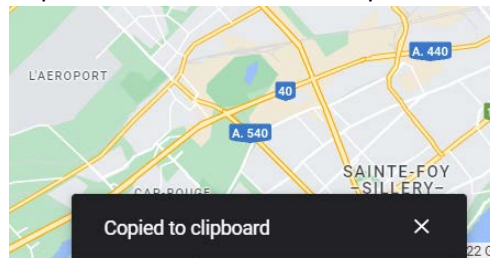
1.3 Changer la zone de codification

Par défaut, la zone de codification est placée sur l'île de Montréal. Pour changer la zone de codification dans une autre région du Québec ou de Canada, il faut suivre la procédure suivante.

- Aller sur Google Maps et choisir un point placé dans la nouvelle région de codification



- Copier les coordonnées en cliquant sur les valeurs affichées



- Aller dans la page Web, cocher <Oui> pour confirmer le déplacement et remplacer les coordonnées en rouge par les nouvelles coordonnées

Outil à 3 points

Identification des liens en utilisant 3 points de passage obligatoire

Déplacer les repères vers de nouvelles coordonnées ?

☒ Oui

☐ Non

46.821670, -71.312360

- Cliquer sur la carte avec la souris et cocher <Non> pour annuler le déplacement à nouveau

Déplacer les repères vers de nouvelles coordonnées ?

☐ Oui

☒ Non

46.821670, -71.312360

//

Sauvegarder les codes de liens HERE		Sauvegarder le chemin counterpart de Google	
1241151289F	46.8216382, -71.31233930	46.8216382, -71.31233930	Coût du trajet: longueur = 0 m et durée =

Status: Trajet indéfini.

Status: Trajet indéfini.

Status: Trajet indéfini.

Status: Trajet indéfini.

Status: Les coordonnées ont été replanées.



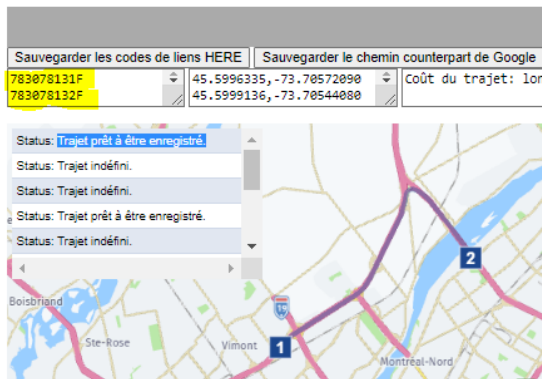
- Zoomer et déplacer les marqueurs qui sont superposés dans les mêmes coordonnées



1.4 Confirmation que le trajet est prêt pour être enregistré

Le processus de codification nécessite une séance des zooms avant et arrière et de positionnement de marqueurs afin d'obtenir le trajet désiré. Après chaque déplacement d'un repère, il faut cliquer avec la souris sur la carte pour avoir le nouveau chemin. La sauvegarde d'un chemin peut se faire quand les deux conditions suivantes sont établies :

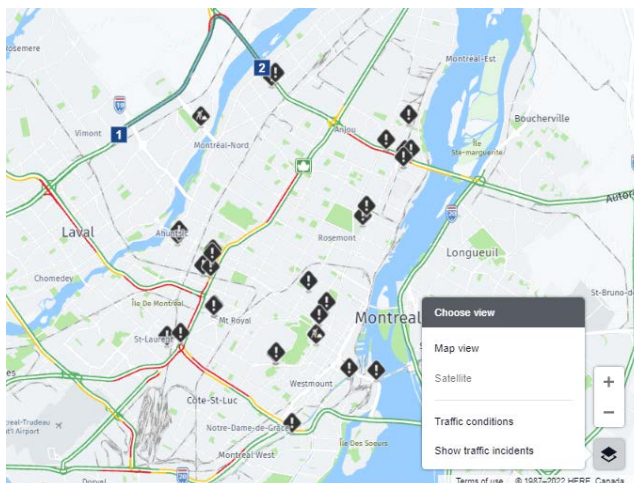
- Le message <Trajet prêt à être enregistré.> est affiché
- Les codes HERE (en jaune) sont affichés dans le premier onglet à gauche



1.5 Divers

Pour des besoins particuliers, les coordonnées LAT/LONG de l'itinéraire peuvent être extraites de l'onglet situé à droite de l'onglet qui conserve les codes HERE

La page Web offre aussi la possibilité d'avoir un aperçu sur la condition de circulation au moment de la codification ainsi que les incidents.



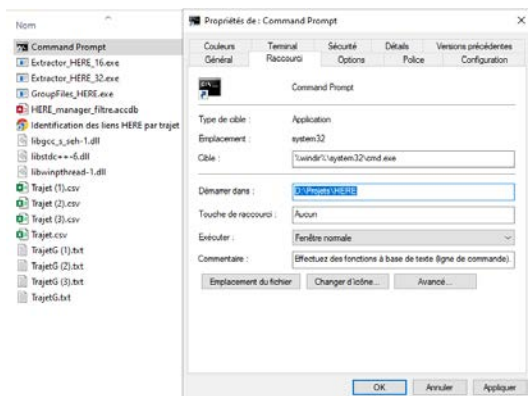
2 Outil pour regrouper les codes de liens HERE

Avant de commencer l'étape de regroupement de codes de liens HERE il est recommandé de suivre la procédure suivante :

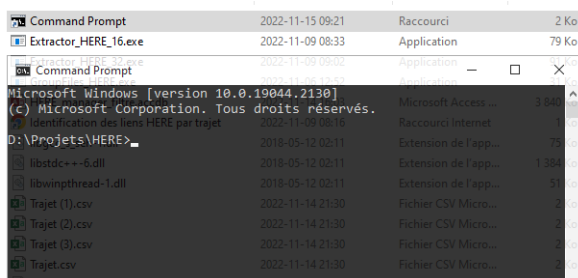
- Déplacer tous les fichiers à partir du répertoire \DONNÉES_MASSIVES\HERE\Outils dans le répertoire de travail

DMST > DONNEES_MASSIVES > HERE > Outils				
Rechercher dans : Outils				
Nom	Modifié le	Type	Taille	
Command Prompt	2022-11-09 18:36	Raccourci	2 Ko	
Extractor_HERE_16.exe	2022-11-09 08:33	Application	79 Ko	
Extractor_HERE_32.exe	2022-11-09 09:02	Application	91 Ko	
GroupFiles_HERE.exe	2022-11-06 12:52	Application	31 Ko	
HERE_manager_filtre.accdB	2022-11-14 16:03	Microsoft Access ...	3 840 Ko	
Identification des liens HERE par trajet	2022-11-09 08:16	Raccourci Internet	1 Ko	
libgcc_s_seh-1.dll	2018-05-12 02:11	Extension de l'app...	75 Ko	
libstdc++-6.dll	2018-05-12 02:11	Extension de l'app...	1 384 Ko	
libwinpthread-1.dll	2018-05-12 02:11	Extension de l'app...	51 Ko	

- Changer le répertoire de démarrage du fichier des commandes DOS pour le répertoire de travail



- Faire double clic sur <Command prompt>

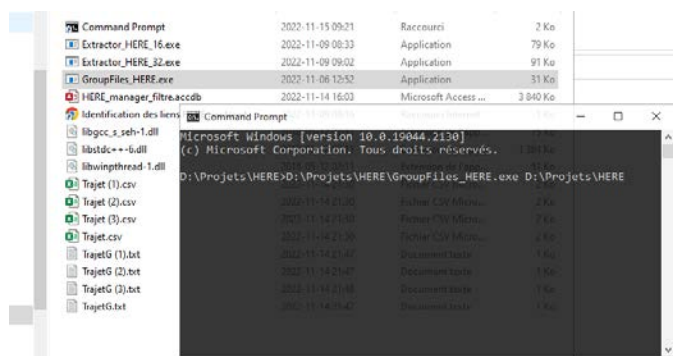


2.1 Faire le regroupement sans détruire le fichier index.csv existant

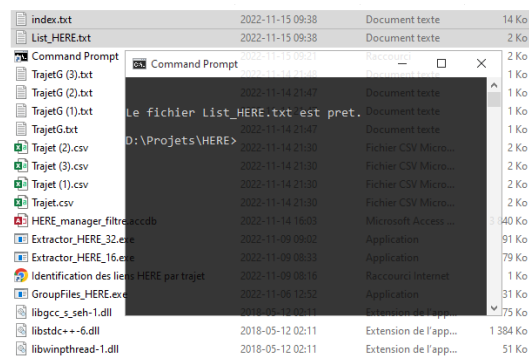
L'outil GroupFiles_HERE.exe permet la création d'un fichier de correspondance entre le code de chaque lien HERE et le nom du fichier père et d'un fichier contenant les codes HERE de tous les fichiers trouvés dans le répertoire de travail qui ont ce pattern : Trajet<>.csv

Les paramètres fournis dans la ligne de commande de cet outil sont :

- Le répertoire cible (obligatoire)
- La lettre **d** (delete) qui n'est pas obligatoire. Si le 2^e paramètre est manquant, l'outil va créer un fichier de correspondance nommée index.txt. Si le 2^e paramètre est la lettre **d** l'outil va écraser un éventuel fichier index.csv et va créer un nouveau fichier index.csv



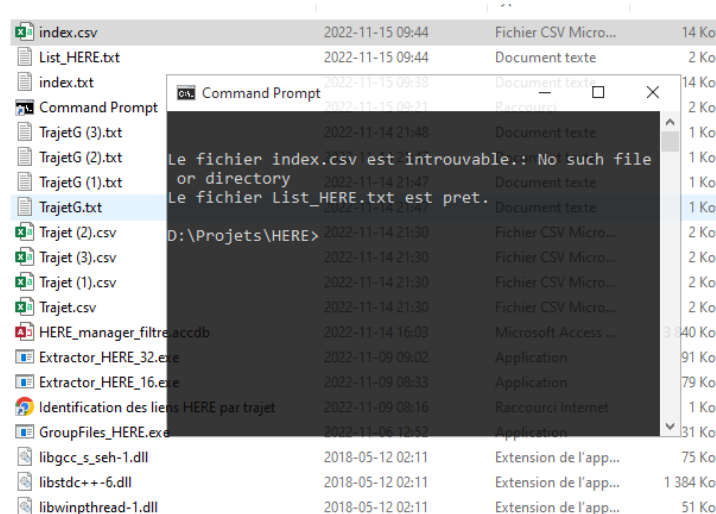
Quand le 2^e paramètre n'est pas fourni, l'outil va produire deux fichiers index.txt et List_HERE.txt sans détruire un éventuel fichier index.csv.



Cette possibilité de ne pas effacer un ancien fichier index.csv est de ne pas perdre un éventuel travail d'arrimage entre les codes de liens HERE et les codes de trajets.

2.2 Faire le regroupement avec la destruction d'un fichier index.csv existant

Si le 2^e paramètre est fourni, la lettre d, un nouveau fichier index.csv est créé ainsi qu'un nouveau fichier List_HERE.txt en faisant le balayage de tous les fichiers Trajet<>.csv situés dans le répertoire cible (dans notre cas le répertoire de travail).



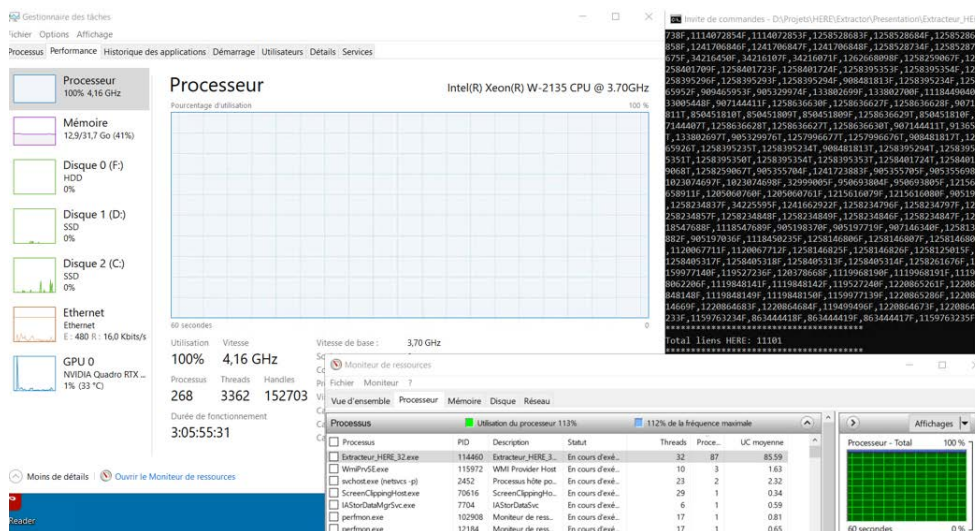
Le fichier index.csv est déjà prêt pour saisir dans la colonne **tag_route** le nom du trajet qui correspondait au fichier trajet<>.csv pendant le processus de codification.

	A	B	C	D
1	link_HERE	file_name	tag_route	
2	34217949F	trajet (1).csv		
3	1189919678T	trajet (1).csv		
4	783078129F	trajet (1).csv		
5	783078130F	trajet (1).csv		
6	34217366F	trajet (1).csv		
7	34262976F	trajet (1).csv		
8	34262977F	trajet (1).csv		
9	34216678F	trajet (1).csv		
10	912885636F	trajet (1).csv		
11	912885637F	trajet (1).csv		
12	34215862F	trajet (1).csv		
13	1123024822F	trajet (1).csv		
14	1310295909F	trajet (1).csv		
15	1310295910F	trajet (1).csv		
16	1258872261F	trajet (1).csv		
17	783385119F	trajet (1).csv		
18	34214881F	trajet (1).csv		
19	1320553522F	trajet (1).csv		
20	1320553523F	trajet (1).csv		
21	905973904F	trajet (1).csv		
22	1314372905F	trajet (1).csv		

Le fichier List_HERE.txt sera utilisé comme le 2e paramètre d'input pour l'outil qui va extraire les données de vitesses HERE à partir des fichiers envoyés par Transport Canada pour tous les codes de liens HERE existants dans ce fichier d'input.

3 Outil d'extraction de données de vitesses HERE

Cet outil permet l'extraction de données de vitesses HERE à partir des gros fichiers CSV qui peuvent avoir des dimensions entre 2 Go (mensuel) jusqu'à 20 Go (annuel) pour une liste de codes de liens HERE qui peut avoir maximum 15 000 liens dans un temps raisonnable d'ordre de dizaines de minutes. Il est conçu avec une architecture « [parallel processing](#) » en C++. Pour de raisons de performance, il y a deux outils disponibles, avec 16 et 32 threads.



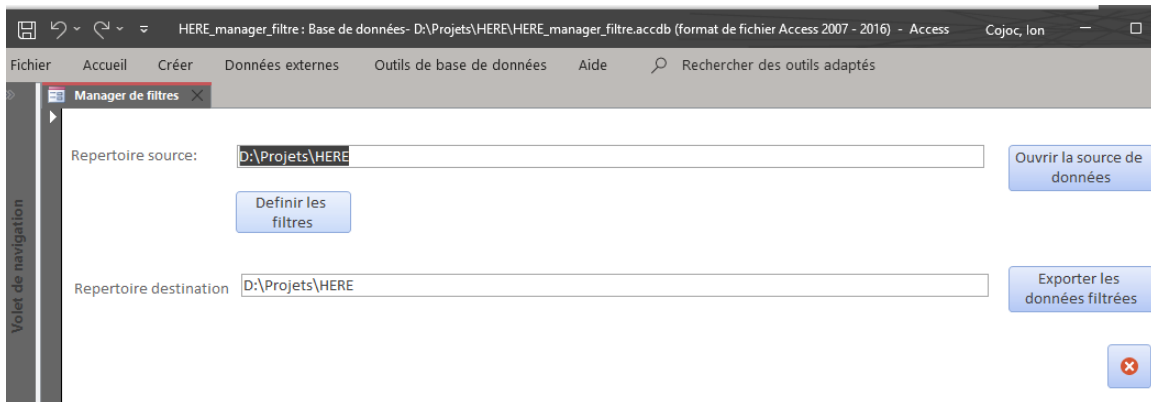
L'architecture multithreads des processeurs actuels permet d'exécution en parallèle de processus pour maximiser le pourcentage d'utilisation du processeur et réduire dramatiquement le temps d'exécution.

Si le fichier de codes de liens HERE à extraire est en bas de 100 liens, il est recommandé la version 16 threads et au-delà de 100 liens il est recommandé la version 32 threads.

L'outil d'extraction va produire toujours le fichier d'extraction, nommée Extract_HERE.csv , dans le répertoire hôte du fichier de liste de code, dans notre cas le répertoire de travail.

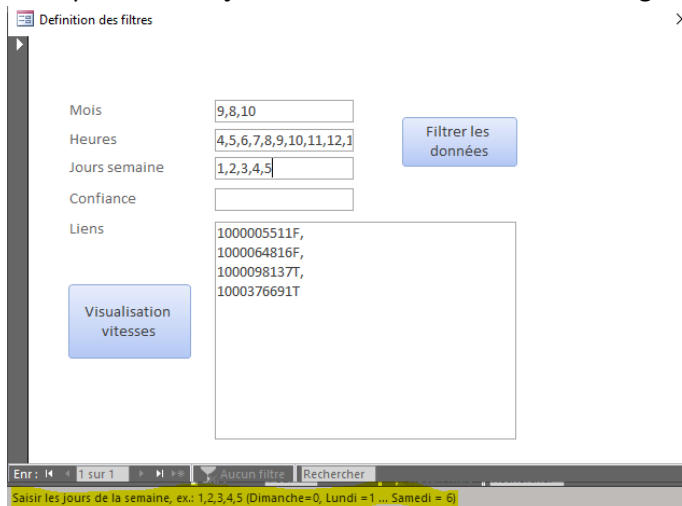
4 Outil supplémentaire de filtrage des fichiers de données HERE

Dans des cas où on a besoin de filtrer davantage le fichier produit par l'outil d'extraction, un petit utilitaire développé en MS Access peut nous aider pour cette tâche.



L'utilisation d'outil est simple et nécessite quatre étapes.

1. Définir le répertoire source qui contient un fichier de type Extract_Here.csv
2. Cliquer sur le bouton <Ouvrir la source de données>
3. Cliquer sur le bouton <Définir les filtres>. Le formulaire permet d'appliquer plusieurs filtres qui sont appliqués avec la condition **ET**. Une attention spéciale doit être accordée au filtre Jours semaine. Pour obtenir les jours de fin de la semaine il faut saisir : 0,6. Pour le champ Confiance il faut saisir la limite inférieure de l'attribut [CONFIDENCE] sinon le champ reste vide. Pour chaque filtre, au bas de la page principale, il y a des informations et de courts exemples sur la façon de saisir les conditions de filtrage.



4. Cliquer sur le bouton <Exporter les données filtrées>

Il est possible d'afficher les vitesses harmoniques et arithmétiques en cliquant sur le bouton Afficher les vitesses. Les données de vitesse utilisées pour calculer et afficher les moyennes horaires sont filtrées en fonction des conditions saisies dans le formulaire Définition du filtre.

