

Utilitaire d'aide à la recherche de bulletins

L'**utilitaire de recherche automatisée de bulletins** pour le serveur de données du Service Météorologique du Canada est disponible à l'adresse: http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl

Ce puissant utilitaire a plusieurs paramètres et peut être utilisé de plusieurs manières. Les recherches se font à l'aide de requêtes dans les chaînes de caractères des bulletins. La syntaxe est:

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?variable1=valeur1&variable2=valeur2

Il peut parfois se passer beaucoup de temps, plusieurs minutes, avant que l'utilitaire ne réponde à une requête.

La recherche peut se faire en utilisant des expressions régulières ou encore de manière plus intuitive. Les pour et contre des deux approches seront discutés ci-après, suivi par des exemples de l'approche intuitive utilisant les requêtes par chaînes de caractères. Pour de plus amples détails concernant les expressions régulières en PERL voir :

- <http://perldoc.perl.org/perlre.html>
- http://www.journaldunet.com/developpeur/tutoriel/cgi/010315cgi_expreg.shtml

Note: la commande de recherche a cependant une faiblesse en ce qui concerne les recherches par dates. À l'occasion, les bulletins ne sont pas placés dans les bons répertoires de dates ce qui fait qu'ils ne seront pas trouvés lors de la recherche. Par exemple, un bulletin émis en fin de journée pourrait être placé dans le répertoire du jour suivant, et alors sera introuvable lors de la recherche. Des situations similaires peuvent se produire en fonction de l'heure du bulletin et dans quel répertoire horaire il a été placé. Élargir votre période de recherche pour inclure les jours ou heures adjacentes à ce que vous désirez facilitera votre recherche dans ces cas particuliers.

Contents

- 1 À propos de ce document
- 2 Exemples de recherche par chaînes de caractères
- 3 Liste des champs de recherche disponibles
 - 3.1 Champs universels alphanumériques
 - 3.2 Champs universels numériques
 - 3.3 Champs alphanumériques optionnels
 - 3.4 Champs numériques optionnels
- 4 Autres variables disponibles pour requêtes par chaînes de caractères
 - 4.1 Format de sortie de données (format)
 - 4.2 Type de sortie de données (output)
 - 4.3 Technique de recherche (search)
- 5 Interface intuitive vs. expressions régulières PERL
 - 5.1 Les champs numériques n'ont pas besoin d'être complété avec des zéros
 - 5.1.1 Interface intuitive
 - 5.1.2 Expression régulière
 - 5.2 Facilité à effectuer des recherches avec « ou »
 - 5.2.1 Interface intuitive
 - 5.2.2 Expression régulière
 - 5.3 Recherche entre des bornes facilité
 - 5.3.1 Interface intuitive
 - 5.3.2 Expression régulière
 - 5.4 Facilité à utiliser les opérateur « plus grand/petit que »
 - 5.4.1 Interface intuitive
 - 5.4.2 Expression régulière
 - 5.5 Facilité à inverser les recherches
 - 5.5.1 Interface intuitive
 - 5.5.2 Expression régulière
 - 5.6 Combiner les inversions avec des bornes est facilité
 - 5.6.1 Interface intuitive
 - 5.6.2 Expression régulière
 - 5.7 Combiner les inversions avec des valeurs multiples est facilité
 - 5.7.1 Interface intuitive
 - 5.7.2 Expression régulière

1 À propos de ce document

Version du document: Version 1.2 en date du 2014-09-24

La version la plus récente de ce document peut être téléchargée à cette adresse:

<http://dd.meteo.gc.ca/bulletins/doc/>

Pour plus d'information à propos des bulletins, veuillez consulter:

http://dd.meteo.gc.ca/bulletins/doc/README_bulletins.txt

2 Exemples de recherche par chaînes de caractères

Ci-bas se trouvent des exemples de requêtes par chaînes de caractères démontrant les possibilités et la flexibilité qu'offre ce système en utilisant l'approche intuitive (c.-à-d. la valeur du paramètre «*search*» est mis par défaut à «*intuit*»).

- Pour trouver le bulletins WBCN04 pour le jour courant:

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=WB&location=CN&header=04

- Pour trouver tous les bulletins du 26 février, 2006 pour le produit FP provenant du bureau émetteur CWAO

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?year=2006&month=2&day=26&product=fp&issuer=cwao

- Pour trouver tous les bulletins pour aujourd'hui pour les produits SA venant du bureau émetteur approprié et pour le site CN

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=sa&location=cn

- Pour trouver tous les bulletins pour aujourd'hui pour les produits SA, FP, et WW

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=sa,fp,ww

- Pour trouver tous les bulletins pour aujourd'hui qui ne sont pas les bulletins SA, FP, et WW

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=!sa,fp,ww

- Pour trouver tous les bulletins pour aujourd'hui émis par tous les centres émetteurs autres que CWAO ou DRRN pour la station CYYZ

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?issuer=!cwao,drn&station=cyyz

- Pour trouver tous les bulletins pour aujourd'hui pour les bulletins SA émis avant midi par le centre émetteur EGGY

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=sa&issuer=eggy&hour=<11

- Pour trouver tous les bulletins d'aujourd'hui pour tout produit ayant un code de correction

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?correction=!none

- Pour trouver tous les bulletins d'aujourd'hui pour tout produit ayant un code de correction mais sans code de station

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?correction=!none&station=none

- Pour trouver tous les bulletins d'aujourd'hui pour tout produit provenant du centre émetteur CWAO

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?issuer=cwao

- Pour trouver tous les bulletins d'aujourd'hui pour le produit SA venant du centre émetteur approprié dont le code d'en-tête est plus grand que 50

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=sa&header=>51

- Pour trouver tous les bulletins d'aujourd'hui dont le nom de produit se termine par A ou Z pour les stations qui ne débutent pas par un C et dont le code d'en-tête se situe entre 35 and 72 et émis après midi

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?product=*a,*z&station=!c*&header=3572&hour=>13

- Pour trouver tous les bulletins d'aujourd'hui pour les stations dont le nom de code débute par C et se termine par R, et ceux débutant par V

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?station=C*R,V*

- Pour trouver tous les bulletins d'aujourd'hui provenant de centres émetteurs dont le nom ne débute pas par C et se termine par L, et ceux se terminant par Z

http://dd.meteo.gc.ca/cgi-bin/bulletin_search.pl?issuer=!C*L,*Z

3 Liste des champs de recherche disponibles

Il y a plusieurs champs de recherche disponibles avec l'utilitaire. Certains sont numériques, d'autres alphanumériques, certains obligatoires dans des bulletins et d'autres optionnels. Les champs optionnels peuvent avoir des contrôles d'existence ou de non-existence, les champs alphanumériques peuvent avoir des contrôles de sous-chaînes de caractères, les champs numériques des contrôles d'intervalles de valeur, et tous les champs peuvent prendre une valeur négatives et être regroupés ensemble avec un point d'exclamation et séparées par des virgules pour les différentes valeurs. Voir ci-après une liste des champs possibles (avec le nom de paramètre à être utilisé dans la chaîne de caractères) incluant des exemples de leurs valeurs ainsi que de leur valeur par défaut.

3.1 Champs universels alphanumériques

Variable	Description	Valeur par défaut
product	Code de produit. Exemples de code: SA, AA, etc.	Tout produit valide
location	Code d'emplacement. Exemples de code: CN, TS, etc.	Tout produit valide
issuer	Code du centre émetteur. Exemples de code: CWAO, CYPL, etc.	Si programmé: centre émetteur par défaut pour le produit Si non-programmé: tout centre émetteur valide

3.2 Champs universels numériques

Variable	Description	Valeur par défaut
----------	-------------	-------------------

year	Année d'émission du bulletin	Année courante
month	Mois d'émission du bulletin	Mois courant
day	Jour d'émission du bulletin	Jour courant
hour	Heure d'émission du bulletin	Toute heure valide
minute	Minute d'émission du bulletin	Toute minute valide

3.3 Champs alphanumériques optionnels

Variable	Description	Valeur par défaut
station	Code de station. Exemples de code: CYYG, WZB, CWKW, etc.	N'importe quelle station valide (ou même aucune station)
correction	Code de correction Exemples de code: CCA, AAA, etc.	N'importe quel code de station (ou même aucun code)

3.4 Champs numériques optionnels

Variable	Description	Valeur par défaut
header	Code d'en-tête. Exemples de code: 15, 31, etc.	Toute valeur d'en-tête valide (ou même aucun en-tête)

4 Autres variables disponibles pour requêtes par chaînes de caractères

Il y a d'autres variables disponibles pour les requêtes par chaînes de caractères, ces variables permettent de choisir parmi les différentes méthodes de sortie de données qu'offre la commande de recherche.

4.1 Format de sortie de données (*format*)

Description: L'option *raw* produit une nouvelle ligne de texte séparée ou un contenu de bulletin ou un emplacement de fichier dépendamment de la valeur du paramètre de sortie (*output*). L'option *HTML* produit une liste non-ordonnée d'emplacements de fichiers ou contenu de bulletins dépendamment de la valeur du paramètre de sortie (*output*).

Options disponibles: Texte brut (*raw*) et liste HTML (*html*).

Valeur par défaut: texte brut (*raw*).

4.2 Type de sortie de données (*output*)

Description: Fournit l'emplacement exact du bulletin lorsque programmé à *name* ou le contenu du bulletin lorsque programmé à *content*.

Options disponibles: emplacement du bulletin (*name*) et contenu du bulletin (*content*).

Valeur par défaut: contenu du bulletin (*content*).

4.3 Technique de recherche (*search*)

Description: Quand programmé à *regex* aucune devinette n'est fait pour vous, cela supposant que vous avez entré une expression régulière par vous-même.

Options disponibles: interface intuitive (*intuit*) et expressions régulières (*regex*). *Intuit* prépare quant à lui vos paramètres de recherche pour vous, tel que décrit plus bas.

Valeur par défaut: interface intuitive (*intuit*).

5 Interface intuitive vs. expressions régulières PERL

La syntaxe utilisée par l'interface intuitive vise à simplifier la recherche pour les usagers non-familiers avec la complexité de l'utilisation des expressions régulières. Ceci crée une approche plus intuitive pour la recherche de champs mais peut porter à confusion pour les usagers habitués à utiliser les expressions régulières.

Par exemple, tandis que dans «*regex*» on utilise «*.**» pour désigner zéro ou plus d'un certain caractère, avec le point décimal désignant tout caractère et l'astérisque désignant zéro ou plus du caractère le précédant (dans ce cas un caractère spécial signifiant tout caractère), l'interface intuitive par défaut utilise plutôt «***» pour désigner zéro ou plus du caractère approprié du champ recherché.

Cela veut dire qu'un champ qui n'accepte que les lettres de l'alphabet sera automatiquement limité à ne rechercher que les lettres de l'alphabet contenu dans l'interface par défaut; de la même manière il sera aussi limité par la longueur acceptable de la chaîne de caractère. Si par exemple vous vouliez faire une recherche pour trouver tout bulletin avec un code de produit débutant par un «*S*», vous programmeriez la variable dans votre chaîne de caractère à «*product=S**» et l'interface par défaut convertirait cela dans une commande sécuritaire et confinée de caractères d'expressions usuelles qui pourront être utilisés pour la recherche. En raison de cette simplification les recherches pour des caractères répétitifs ne sont pas possibles dans l'interface par défaut; si de telles fonctionnalités (où «*S*» désigne zéro ou plus de *S* consécutifs) vous sont nécessaires alors vous devrez utiliser les expressions régulières.

L'interface intuitive est le mode par défaut de l'utilitaire de recherche, mais il peut être changé au mode d'expressions régulières de PERL pour les usagers plus familiers avec cette syntaxe. Plus d'information sur les expressions usuelles PERL se trouvent à ces adresses: <http://perldoc.perl.org/perlre.html> ou http://www.journaldunet.com/developpeur/tutoriel/cgi/010315cgi_expreg.shtml

Les expressions régulières Perl sont activées en ajoutant une variable dans la chaîne de caractères «*search*» avec une valeur «*regex*». C'est-à-dire «*search=regex*». Pour l'interface intuitive il faut programmer la variable de recherché à «*intuit*» (ex: «*search=intuit*»).

Les recherches par sous-chaînes de caractères (incluant préfixes et suffixes) sont traitées par trois caractères spéciaux empruntés des expressions régulières, bien que leur signification aient été quelque peu ajustée. Tel que montré plus haut, l'usage de l'astérisque suppose une correspondance avec un caractère particulier et il en va de même pour le point décimal et le point d'interrogation. Le point décimal conserve la même signification que dans les expressions régulière «*regex*» comme étant n'importe quel caractère unique valide, permettant un meilleur contrôle sur la longueur des valeurs correspondantes. Si vous vouliez par exemple vous assurer que les codes d'émissions de vos bulletins soient de cinq caractères de longueur (plutôt que quatre comme typiquement) vous pourriez programmer la variable de votre chaîne de caractères à «*issuier=.....*» ce qui permettrait de trouver les

centres émetteurs correspondants à 5 caractères mais tout en étant limité par les valeurs possibles pour ce champ (dans ce cas autant les lettres que les chiffres).

Une des méthodes de recherche où l'interface intuitive surpasse les expressions régulières est pour la comparaison numérique. Les expressions régulières sont basées sur des caractères, de telle sorte que les comparaisons numériques sont prohibitivement complexes; l'interface intuitive permet des recherches numériques à l'intérieur (et à l'extérieur) de bornes de même que des comparaisons d'égalité (plus grand que, égal, plus petit que). Une des limitations est qu'il ne supporte que les valeurs de 0 à 99; cependant, étant donné que ces valeurs sont les seules que peuvent prendre les champs numériques dans les bulletins qui peuvent être recherchés, c'est un compromis acceptable.

De plus, il est possible de faire une recherche sur l'opposé de n'importe lequel de ces champs dans l'interface intuitive simplement en plaçant un point d'exclamation «!», ce qui a pour effet d'inverser la recherche courante. Toutes les méthodes discutées avant peuvent être inversées à l'aide du point d'exclamation. L'inversion de phrases (plutôt que d'un seul caractère) dans les expressions régulières n'est pas simple et la syntaxe peut être problématique.

Vous trouverez ci-bas des exemples démontrant les avantages de l'interface intuitive.

5.1 Les champs numériques n'ont pas besoin d'être complétés avec des zéros

5.1.1 Interface intuitive

```
?hour=2
```

5.1.2 Expression régulière

```
?hour=02
```

5.2 Facilité à effectuer des recherches avec « ou »

5.2.1 Interface intuitive

```
?hour=0,6,12,18
```

5.2.2 Expression régulière

```
?hour=(00|06|12|18)
```

5.3 Recherche entre des bornes facilitée

5.3.1 Interface intuitive

```
?header=1352
```

5.3.2 Expression régulière

```
?header=(1[39]|[24][09]|5[02])
```

5.4 Facilité à utiliser les opérateur « plus grand/petit que »

5.4.1 Interface intuitive

?hour=>13

5.4.2 Expression régulière

?hour=(1[39]|2[03])

5.5 Facilité à inverser les recherches

5.5.1 Interface intuitive

?issuer=!CYUL

5.5.2 Expression régulière

?issuer=(?!CYUL).*

5.6 Combiner les inversions avec des bornes est facilité

5.6.1 Interface intuitive

?header=!1520

5.6.2 Expression régulière

?header=(?!1[59]|20).*

5.7 Combiner les inversions avec des valeurs multiples est facilité

5.7.1 Interface intuitive

?product=!SA,AA,WW,WF

5.7.2 Expression régulière

?product=(?!SA|AA|WW|WF).*