

Parler l'Idule !

L'Idule, le langage des IDU, est un langage Informatique pour accéder à des Données et les Utiliser.

Accès aux données :

Les données peuvent être accédées au travers de ressources. Une ressource est représentée :

- soit par une URI sous forme de chaîne de caractères,
- soit par une ressource interne définie par un nom (qui peut être vu comme un nom de variable).

La récupération d'un sous ensemble de données s'effectue par le mot clé « get ». Ce dernier doit prendre en paramètre :

- une liste de ressources
- une contrainte spécifiant le contenu recherché dans les ressources

Une contrainte est une expression booléenne utilisant les mots clés « contains » et « excludes » suivis d'une chaîne de caractères permettant de rechercher respectivement les données contenant et ne contenant pas la chaîne de caractères. Une contrainte pourra être composée de plusieurs contraintes liées par les mots clés « and » ou « or ».

Dans tous les cas, le résultat d'un « get » fournit une ressource interne qui doit absolument être nommée.

Manipulation de données :

La manipulation des données consiste en un calcul statistique sur les ressources internes. Le mot clé permettant de faire ce calcul est « stat ». Cette opération prend en paramètre

- un ensemble de ressources internes qui sont liées par les mot clés « union », « intersect » et « diff ».
- un ensemble de groupes de contraintes

Le résultat de cette opération est un dictionnaire d'une analyse statistique qui contiendra au minimum :

- le nombre d'éléments de l'ensemble des ressources internes en tenant compte des liens définis,
- pour chaque contrainte, le rapport entre le nombre d'éléments validant la contrainte par le nombre total d'éléments.

Exemples

Programme Idule initial :

get <http://www.machin-truc.org/page.html>

Programme Idule initial avec contrainte :

get <http://www.machin-truc.org/page.html> contains toto

Programme Idule initial avec groupe de contraintes :

```
get http://www.machin-truc.org/page.html contains toto and exclude titi or  
contains blabla
```

Programme Idule initial avec groupe de contraintes :

```
get http://www.machin-truc.org/page.html contains toto and exclude titi or  
contains blabla  
get http://www.bidule.org/page.html contains bli or contains blu  
and exclude blo
```

Programme Idule initial avec stats :

```
r1 = get http://www.machin-truc.org/page.html contains toto and exclude titi or  
contains blabla  
r2 = get http://www.bidule.org/page.html contains bli or contains blu  
and exclude blo  
stat r1 union r2 contains Trump and exclude Clinton contains Clinton and exclude  
Trump contains Trump and contains Clinton  
stat r1 intersect r2 contains Trump and contains Clinton
```

Réalisation :

Réalisez votre implémentation du langage Idule en :

- Python
 - Lexique en Python PLY (Lexer)
 - Grammaire et compilation en Python PLY (Parser)

Il vous est fortement conseillé d'implémenter la première partie « Accès aux données » petit à petit en traitant exemple après exemple, de les tester avant d'entamer la prise en compte de l'exemple suivant.

La téléchargement effectif des données et le calcul des statistiques seront à implémenter uniquement si votre lexer et parser sont opérationnels : valident les exemples donnés.