



# Université du Québec à Chicoutimi

Cours 8INF803 : Bases de données réparties  
Travail 2 - Exercise 1

AL HADDAN Ayoub  
DELEFORGE Thomas

## Exercice 1 : La guérison

Dans cet exercice, nous avons réalisé un Crawler en java qui nous a permis de récupérer toute la liste des sorts pour chaque créature du bestiaire sur le site paizo.com. Pour pouvoir le faire, nous avons d'abord téléchargé les 5 pages.

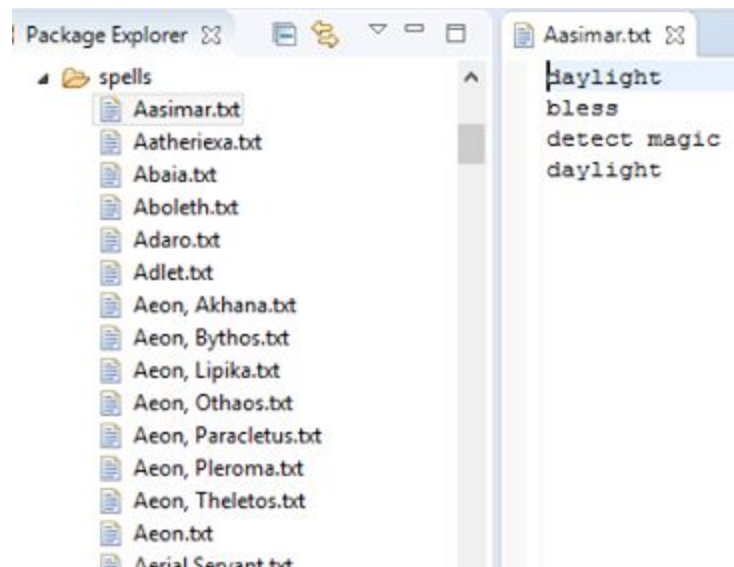
```
Main (1) [Java Application] C:\Program Files\Java\jre1.8.0_152\bin\javaw.exe (25 a
Page 1 downloaded with success !
Page 2 downloaded with success !
Page 3 downloaded with success !
Page 4 downloaded with success !
Page 5 downloaded with success !
```

Nous avons ensuite analysé les pages 5 pages téléchargées pour ne laisser que les noms des monstres et leur url, nous avons enregistré les résultats dans 5 fichiers monster1.txt, monster2.txt, monster3.txt, monster4.txt et monster5.txt, voici un extrait du fichier monster1.txt :

```
monster1.txt
Aasimar:aasimar.html#aasimar
Aboleth:aboleth.html#aboleth
Angel:angel.html#angel
Angel, Astral Deva:angel.html#angel-astral-deva
Angel, Planetar:angel.html#angel-planetar
Angel, Solar:angel.html#angel-solar
Animated Object:animatedObject.html#animated-object
Ankheg:ankheg.html#ankheg
Ant, Giant:ant.html#ant-giant
Army Ant Swarm:ant.html#army-ant-swarm
Ape:ape.html
Ape, Dire:ape.html#ape-dire
Ape, Gorilla:ape.html#ape-gorilla
Archon:archon.html#archon
Archon, Hound:archon.html#archon-hound
Archon, Lantern:archon.html#archon-lantern
Archon, Trumpet:archon.html#archon-trumpet
Assassin Vine:assassinVine.html#assassin-vine
Azata:azata.html#azata
Azata, Bralani:azata.html#azata-bralani
Azata, Ghaele:azata.html#azata-ghaele
Azata, Lillend:azata.html#azata-lillend
```

*NomMonstre:UrlMonstre*

On va ensuite accéder au lien de chacune des créatures pour pouvoir récupérer la liste des spells qu'elle peut pratiquer. On va stocker la liste des spells pour chaque créature dans un fichier au format « nom\_creature.txt », comme vous pouvez le voir sur l'image suivante. Le but est d'établir une liste de spells facilement accessible pour chaque créature. Dans notre cas on va Crawler tous les spells pour chaque créature.



On peut voir la structure de chaque fichier ici, avec le nom de la créature en nom de fichier et la liste des spells qu'elle possède. Ici l'exemple montre la liste des spell que le Crawler a trouvé pour la créature « Aasimar ».

```

Main (1) [Java Application] C:\Program Files\Java\jre1.8.0_152\bin\javaw.exe (25 a
Page 1 downloaded with success !
Page 2 downloaded with success !
Page 3 downloaded with success !
Page 4 downloaded with success !
Page 5 downloaded with success !
Page 1 cleaned with success !
spell 0 Aasimar downloaded with success !
spell 1 Aboleth downloaded with success !
spell 2 Angel downloaded with success !
spell 3 Angel, Astral Deva downloaded with success !
spell 4 Angel, Planetary downloaded with success !
spell 5 Angel, Solar downloaded with success !
spell 6 Animated Object downloaded with success !
spell 7 Ankheg downloaded with success !
spell 8 Ant, Giant downloaded with success !
spell 9 Army Ant Swarm downloaded with success !
spell 10 Ape downloaded with success !
spell 11 Ape, Dire downloaded with success !
spell 12 Ape, Gorilla downloaded with success !
spell 13 Archon downloaded with success !
spell 14 Archon, Hound downloaded with success !
spell 15 Archon, Lantern downloaded with success !
spell 16 Archon, Trumpet downloaded with success !
spell 17 Assassin Vine downloaded with success !

```

Aperçu console de l'affichage lors du téléchargement de la liste des spells pour chaque créature depuis le site paizo.com.

Maintenant que nous avons la liste des sorts par créature, nous allons construire les Rdd contenant la liste de créatures par sort, de telle sorte à créer une Batch View permettant à Pito de se sortir rapidement de la galère dans laquelle il se trouve.

```

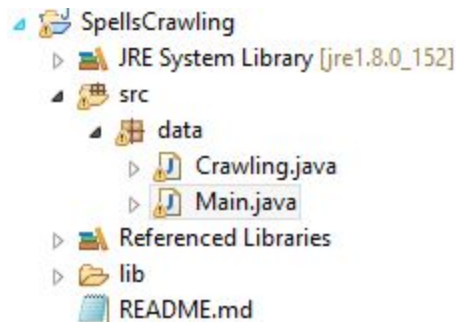
(greater magic fang,Set(Animal Lord))
(shout,Set(Giant: Ocean))
(banishment,Set(Metallic Dragon: Silver, Devourer, Aeon, Archon: Trumpet, Azata: Ghaele, Ae
(stone tell,Set(Korred))
(greater shadow evocation,Set(Shen, Rakshasa: Maharaja, Rakshasa))
(entropic shield,Set(Sceaduinar, Outer Dragon: Vortex))
(cognitive block ,Set(Psychemental))
(explosive runes,Set(Azata: Raelis))
(chill metal,Set(Winter Hag, Esoteric Dragon: Etheric, Taxidermic Creature, Empyrean Lord:
(Control water,Set(Imperial Dragon: Sea))
(crushing despair,Set(Kyton: Ostiarius, Kyton, Sangoi, Seilenos, Demon: Incubus, Devil: Apc
(continual flame,Set(Angel: Planetary, Phoenix, Outer Dragon: Solar))

```

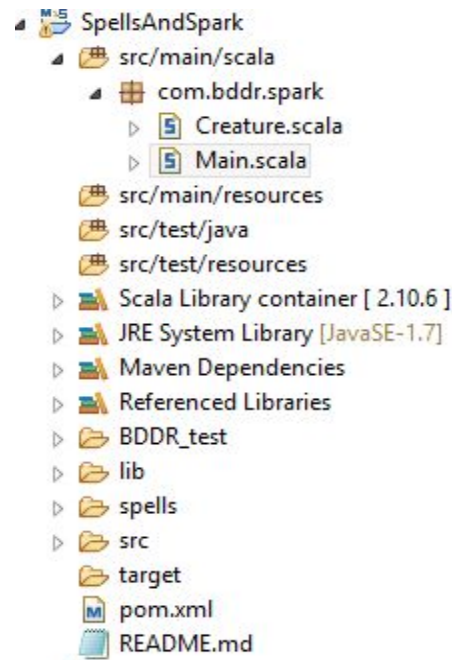
Voilà la sortie en console lors de la création des Rdd, on a un ensemble (Key, Value) telle que la Key corresponde au nom du spell et que la Value soit un array comprenant la liste des créatures pouvant utiliser ce spell particulier. Chaque nom de créature est séparé par une virgule.

## Structure du projet :

Pour la partie crawling nous l'avons fait en Java :



Nous avons ensuite exporter le jar de ce projet et nous l'avons utilisé dans le projet scala pour trouver les batchs view, la structure du projet scala :



*Le jar du projet "SpellsCrawling" se trouve dans le fichier lib et la classe Crawling.java est utilisée dans le Main.scala*