

GROUP 3: GENOME BIOLOGY AND EVOLUTION

Jeremy Hui (jh564)

Alfred Zane Rajan Velladurai (ar2257)

Contributions:

Jeremy: Obtained all Articles as HTML files and downloaded them. Worked on the algorithm for extracting the 10 fields.

Alfred: Worked on the algorithm for extracting the 10 fields and wrote them into the text file. Also retrieved the data from the text file.

Challenges:

- Evading Captcha
- getURL function is not robust
- took long to get XPath expressions
- Writing the text file Row wise.