

|  |  |
| --- | --- |
| **ID:** | **UC01** |
| **Objetivo:** | Extrair dados de interesse da fonte “DrugBank”. |
| **Requisitos:** | Estar conectado à internet. |
| **Atores:** | Usuário. |
| **Pré-condições:** | Os pacotes, linguagens, *frameworks* e bibliotecas usados pela ferramenta precisam estar instalados no ambiente.  Os dados de interesse precisam estar definidos no arquivo (JSON) de configurações/entradas da fonte. |
| **Pós-condições:** | As informações obtidas deverão estar salvas em arquivo(s) no disco. |
| **Fluxo Principal:** | **Scrap DrugBank**  1. O ator executa a ferramenta, através de linha de comando, com os parâmetros que definem a fonte “DrugBank”;  2. A ferramenta busca os dados de interesse (definidos em um arquivo configurável do tipo JSON);  3. A ferramenta gera arquivo(s) de saída (conforme os parâmetros informados) no disco com as informações encontradas. |
| **Fluxos Alternativos:** | Não há. |
| **Erros/Exceções:** | O tratamento de erros (como requisições inválidas e *timeouts*) é feito pelo *framework* utilizado (mensagens de erro podem ser exibidas para o usuário).  Em caso de falha(s) na execução as pós-condições não serão verdadeiras (os dados de interesse não serão obtidos). |
| **Mensagens:** | *Logs* de saída do *framework* utilizado. |
| **Regras de negócio:** | **[RN01]** O endereço de internet “<https://www.drugbank.ca/unearth/q>" deve aceitar requisições do tipo GET com os parâmetros “utf8”, “searcher” e “query” e os respectivos valores “✓”, “drugs” e “{variável\_de\_interesse}” (este último contendo a variável desejada) e deve ser utilizado para buscar os dados de interesse.  **[RN02]** O resultado da requisição anterior (**RN01**) deverá ser uma página HTML válida contendo um elemento do tipo “h1” e outro do tipo “table” com ID “drug-moldb-properties”, de onde os dados de interesse devem ser extraídos.  **[RN03]** O resultado da requisição feita pela **RN01** poderá ser uma página HTML válida contendo um elemento “a” dentro de um elemento “h2” que, por sua vez, deverá estar dentro de um elemento “div” com a classe CSS “search-result” atribuída. Neste caso, uma nova requisição do tipo GET deverá ser feita ao endereço contido no atributo “href” da referida *tag* “a” e o resultado obtido desta nova requisição deverá ser o descrito na **RN02**. |