Pruebas unitarias

Las pruebas unitarias o test unitarios sirven para comprobar el correcto funcionamiento de unidades individuales de los programas informáticos. En el caso de Hércules-RPA, estas pruebas están centradas en el procesamiento de los datos que realiza cada uno de los cuatro procesos implicados en el proyecto. Los cuatro procesos tienen una estructura que se puede generalizar de la siguiente manera:

- Obtención de los datos.
- Procesamiento y tratamiento de los datos.
- Acción final.

Las pruebas unitarias realizadas para estos procesos están centradas en el procesamiento y tratamiento de los datos ya que esta es la parte crítica de cada uno de los procesos.

Proceso 1 : Segmentación y generación automática de boletines informativos

Este proceso es el encargado de la generación de un boletín informativo con los siguientes elementos: oferta tecnológica, artículos, propiedad intelectual, propiedad industrial, tesis doctorales, convocatorias, proyectos, contratos y noticias. Es por esto, que está compuesto de subprocesos que consultan la información, la procesan y crean un resumen para posteriormente enviar un correo electrónico con la información procesada de cada uno por estos subprocesos.

Las pruebas unitarias, por tanto, se centran en el procesamiento de los datos, es por esto que siguen la siguiente estructura:

- *Datos de entrada*: se han obtenido datos de entrada de prueba de los correspondientes subsistemas a los que se consulta la información para que la entrada de las pruebas sea inmutable.
- *Procesamiento de los datos*: para el procesamiento se hace uso de las funciones encargadas de realizar esta acción en cada uno de los subprocesos.
- *Comprobación de la salida*: esto nos garantiza que si el procesamiento de los datos cambia la salida debe de ser la misma, si esto falla, quiere decir que algo en el procesamiento de los datos ha cambiado.

Las pruebas de este proceso se dividen en siete ficheros. Se adjunta documento con el resultado de las pruebas realizadas:

Proceso 2 : Sexenios y acreditaciones

Este proceso está compuesto por dos subprocesos, la generación del informe de solicitud de un sexenio y la generación de un informe de solicitud de una acreditación. Las pruebas unitarias relacionadas con este proceso están orientadas a comprobar el correcto funcionamiento de:

- Procesamiento de los datos obtenidos.
- Evaluación o baremación de la producción científica.
- Inserción de información en un documento.
- Comprobación de funciones de formato.

Las pruebas unitarias realizadas para este proceso son las siguientes:

- Comprobaciones de funciones de formato:
  - test_period: Este test obtiene los periodos secuenciales en array dado en formato “2005-2020”.
  - test_period_formatter: Este test obtiene los periodos secuenciales separados por comas dado el formato “2005-2020”.
  - test_authors_list: Este test imprime una lista de autores separados por comas.
- Inserción de información en un documento:
  - test_print_researcher: Este test imprime en un documento la información de un investigador.
  - test_print_scientific_production: Este test imprime en un documento la producción científica de un investigador.
- Procesamiento de datos obtenidos y baremación de la producción científica para sexenios:
  - test_get_researcherInfo: Este test obtiene la información de un investigador.
  - Por cada comisión se realiza el procesamiento de los datos y la baremación de la producción científica, así se comprueba que si se modifica el proceso los resultados no cambian.
- Procesamiento de datos obtenidos y baremación de la producción científica para acreditaciones:

- Por cada comisión se realiza el procesamiento de los datos y la evaluación de la producción científica para los dos tipos de acreditaciones implementadas del programa ACADEMIA, titularidad y cátedra, así se comprueba que si se modifica el proceso los resultados no cambian.

Se adjunta documento con el resultado de las pruebas realizadas:

Proceso 3 : Automatización de procesos de gestiones documentales parametrizables

Este proceso se divide en tres subprocesos, para cada uno de ellos se han realizado las siguientes pruebas:

Extracción de Bases Reguladoras

Este subproceso es el encargado de la obtención de nuevas Bases Reguladoras a partir del Boletín Oficial del Estado (BOE), una vez obtenidas se enviará la información a la UGI y a la OTRI. Las pruebas realizadas para este proceso se centran en el procesamiento de los datos. Por lo que siguen la siguiente estructura:

- Extracción de información a partir de los documentos XML del BOE: para realizar esto se recorre el documento utilizando el módulo de tecnologías cognitivas de extracción de información de documentos XML y se obtienen los nodos que interesan para este proceso.
- Procesamiento de la información: una vez obtenidos los nodos, se procesa la información y se crean las entidades necesarias con la información obtenida.
- Comprobación de la información obtenida: a partir de la información procesada se realiza el filtrado de los resultados para obtener solo las Bases Reguladoras que son del interés de la UGI y la OTRI.

Extracción de Convocatorias

Este subproceso es el encargado de la extracción de convocatorias utilizando la Base de Datos Nacional de Subvenciones (BDNS) y Comisión Europea (Funding & tenders). La información obtenida de ambas fuentes se procesa y en el primer caso, se inserta en el subsistema Hércules-SGI y en el segundo, se envía por correo electrónico. Para comprobar el correcto funcionamiento de este proceso se han realizado los siguientes test:

- test_extract_call_info: Este test extrae la información de una convocatoria.
- test_extract_call_word_investigacion: Este test comprueba si existe la palabra investigación en un listado de convocatorias.
- test_extract_call_word_ayuda: Este test comprueba si existe la palabra ayuda en un listado de convocatorias.
- test_extract_call_word_desarrollo: Este test comprueba si existe la palabra desarrollo en un listado de convocatorias.
- test_extract_call_without_word: Este test comprueba si existe las palabras inventadas en un listado de convocatorias.
- test_extract_call_without_number: Este test comprueba si existe una convocatoria en el listado dado un número no valido.
- test_extract_call_with_number: Este test comprueba si existe una convocatoria en el listado dado un número válido.
- test_extract_call_number: Este test comprueba si existe una convocatoria en el listado con un número válido.
- test_extract_call_no_number: Este test comprueba si existe una convocatoria en el listado con un número válido y que no exista.
- test_extract_call_europa_dates_day: Este test comprueba si existe alguna convocatoria europea en un día dado.
- test_extract_call_europa_dates_week: Este test comprueba si existe alguna convocatoria europea en una semana dada.
- test_extract_call_europa_dates_half_month: Este test comprueba si existe alguna convocatoria europea en una quincena dada.
- test_extract_call_europa_dates_month: Este test comprueba si existe alguna convocatoria europea en un mes dado.
- test_extract_call_europa_dates_six_month: Este test comprueba si existe alguna convocatoria europea en seis meses dados.
- test_extract_call_europa_incorrect_dates: Este test comprueba si existe alguna convocatoria europea con fechas mal dadas.
- test_extract_call_europa_none_dates: Esta convocatoria comprueba si existe alguna convocatoria europea con ninguna fecha dada.
- test_extract_call_pdf_525644: Este test obtiene los documentos adjuntos a partir de un número de la BDNS.
- test_extract_call_pdf_607221: Este test obtiene los documentos adjuntos a partir de un número de la BDNS.
- test_extract_call_empty_pdf: Este test obtiene los documentos adjuntos a partir de un número de la BDNS incorrecto.
- test_extract_call_none_pdf: Este test obtiene los documentos adjuntos a partir de un None.

Extracción de Concesiones

Este subproceso es el encargado de la extracción de concesiones utilizando la Base de Datos Nacional de Subvenciones (BDNS) y Hércules-SGI. La información obtenida se procesa y se envía por correo electrónico. Para comprobar el correcto funcionamiento de este proceso se han realizado los siguientes test:

- test_get_valid_requests: Este test saca de una consulta guardada como fichero qué peticiones son válidas para procesar en el proceso de Concesiones.
- test_get_not_valid_requests: Este test saca de una consulta guardada como fichero qué peticiones no son válidas para el proceso Concesiones.
- test_get_valid_requests_bad_forms: Este test comprueba que la salida de la función get_valid_forms() metiendo un formulario erróneo.
- test_get_bdns_525644: Este test comprueba que se procese bien el sacar la información de la convocatoria con número 535644.
- test_get_bdns_644545: Este test comprueba que se procese bien el sacar la información de la convocatoria con número 644545.
- test_get_bdns_non_existent: Este test comprueba que no devuelva nada con la entrada de un número de la BDNS que no exista.
- test_get_bdns_bad_request: Este test comprueba que no devuelva nada con la entrada de un número de la BDNS erróneo.
- test_formatting_nifs: Método que formatea la entrada de NIFs para que te los devuelva en un array.
- test_formatting_nifs_none: Este test comprueba la entrada de NIFs cuando no viene ninguna entrada.
- test_formatting_nifs_bad: Este test comprueba la entrada de NIFs cuando viene una entrada vacía.
- test_formatting_nifs_bad_form_int: Este test comprueba la entrada de NIFs cuando viene mal formada por enteros.
- test_formatting_nifs_bad_form: Este test comprueba la entrada de NIFs cuando viene mal formada.
- test_get_persona_antonio: Este test comprueba dada una lista de concesiones si existe una entrada en concreto.
- test_get_persona_palma: Este test comprueba dada una lista de concesiones si existe una entrada en concreto.
- test_get_persona_no_awards: Este test comprueba si existe una concesión sin dar una entrada como parámetro.
- test_get_persona_bad_request: Este test comprueba si existe una concesión dando una entrada errónea como parámetro.
- test_get_persona_list_none: Este test comprueba si existe una concesión pasando una lista nula.
- test_get_persona_no_list: Este test comprueba si existe una concesión pasando una lista vacía.

Las pruebas realizadas para este proceso se dividen en tres ficheros, uno por cada subproceso mencionado anteriormente. Se adjunta documento con el resultado de las pruebas realizadas:

Proceso 4 : Sistemas Inteligentes de encaje entre convocatorias e investigadores

Este proceso realiza el encaje de convocatorias con los intereses de los investigadores. Para los datos de entrada de estos test se utilizan los siguientes ficheros:

- *testRecomendacionContenido.csv :* En este fichero se encuentran los datos de entrada con el que se evaluará el sistema de recomendación basado en contenido, es decir, teniendo en cuenta palabras clave o campos clave del histórico de convocatorias para tratar de hacer un match con la convocatoria a recomendar. Este fichero dispone de 101 entradas.
- *testRecomendacionFC.csv* : En este fichero se encuentran los datos de entrada con el que se evaluará el sistema de recomendación filtro colaborativo, es decir, teniendo en cuenta las puntuaciones que los investigadores han realizado a las áreas temáticas trata de buscar posibles relaciones con las convocatorias para poder determinar que convocatoria recomendar. Este fichero dispone de 379 entradas.

Las pruebas realizadas constan de cinco evaluaciones a cada sistema de recomendación y, por último, cinco evaluaciones al motor híbrido.

- Las distintas evaluaciones tratan de buscar diferentes recomendaciones modificando la convocatoria a recomendar y los parámetros de los sistemas de recomendación, por ejemplo, para la ejecución con mayor precisión de los sistemas de recomendación se necesita un mínimo de X puntuaciones o histórico de datos. Por tanto, con los test se obtiene diferentes resultados comprobando que el sistema se comprueba de la forma esperada y deseada.
- La evaluación del motor híbrido se realiza invocando al resto de sistemas de recomendación asignando un peso a cada uno para observar la importancia de cada uno.

Las pruebas de este proceso se encuentran en un único fichero. Se adjunta documento con el resultado de las pruebas realizadas:

process4...-42.html

Informe de cobertura

Para comprobar la cobertura de las pruebas unitarias mencionadas anteriormente se ha utilizado la herramienta Coverage de Python y se ha obtenido una cobertura del 58%. Se adjunta el informe de la ejecución de todos los test mencionados anteriormente:

Coverage report.pdf