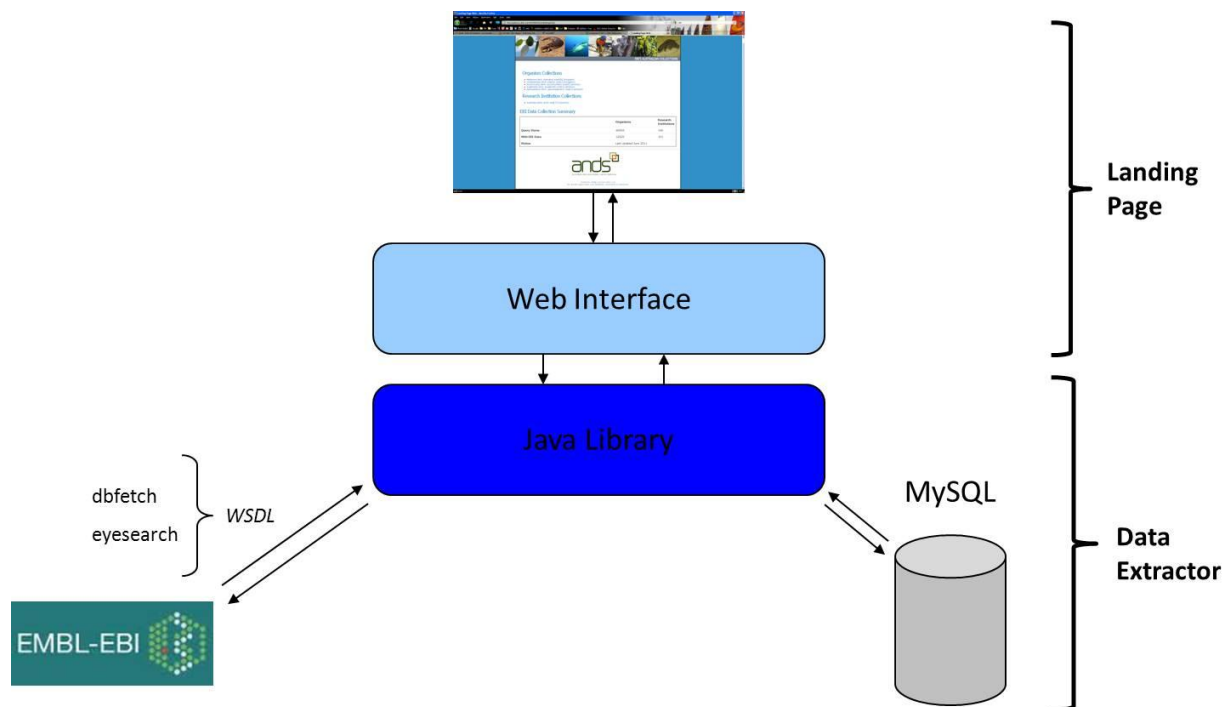


## Software architecture

The EBI-RDA linkage system is comprised of two main parts (Figure 1):

1. **Data extractor:** this system is a standalone Java Program (executable jar) that pulls query items from a local data store (MySQL database) and queries them (using web services) against the Uniprot and ENA databases made available by EBI.
2. **Landing page:** a website to allow visual display of the data. This was developed using JSP/Servlets and displays the relevant data that is stored in the MySQL database.



**Figure 1: Software architecture overview**

A more complete depiction of the software interfaces used is displayed in Figure 2:

- Public data repositories (Atlas of Living Australia, National Library of Australia) are used to generate lists of potential “Australian” query items (species and research institutions).
- These query items are used to query the European Nucleotide Archive (DNA sequences) and Uniprot (protein sequences) databases at EBI via web services.

- These results are stored in the Australian data collection repository (MySQL database)
- The data in the Australian data collection repository is visualised using the landing page.
- Data in the Australian data collection repository is used to create collections which are submitted to the RDA. These collections can be discoverable by interested parties, which links from the RDA to the landing page back to the original data stored in EBI databases.

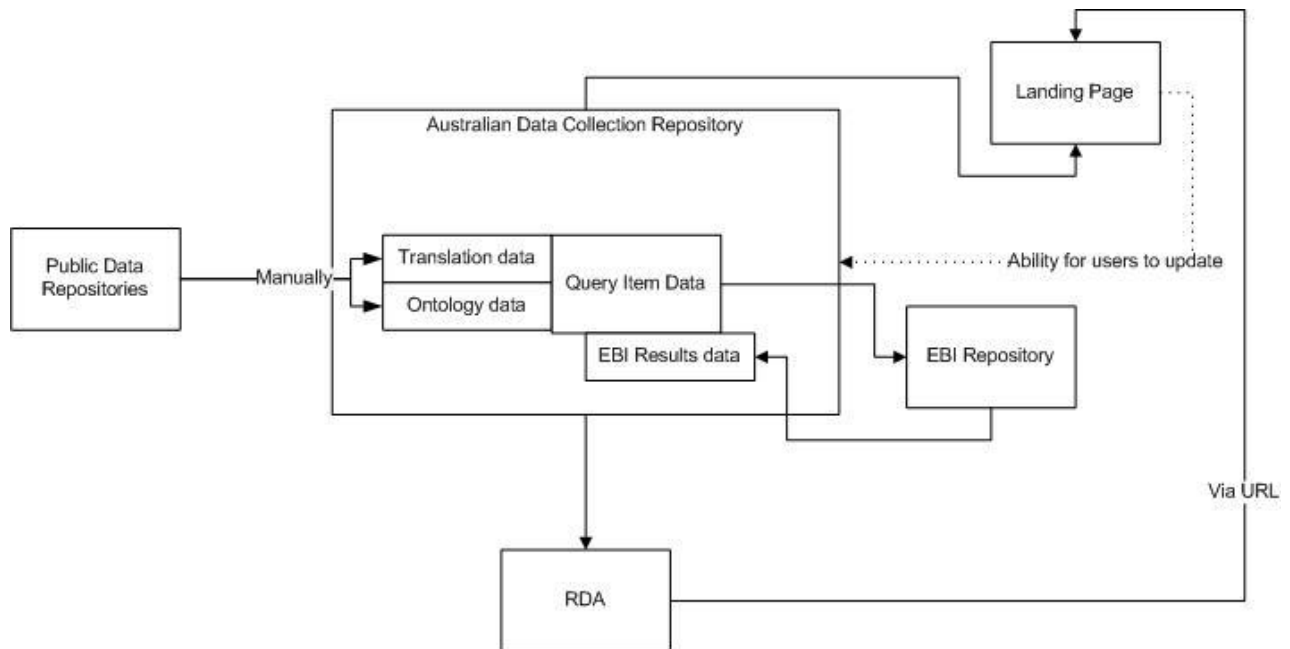


Figure 2: Software interfaces