Fuglene ved de danske Fyr, 1883-1939

Birds at the Danish Lighthouses, 1883 – 1939

The digitization of "Fuglene ved de danske Fyr, 1895-1939" (UK: Birds at the Danish Lighthouses, 1895-1939) was initiated through a suggestion from the Thy National Park, Denmark (http://nationalparkthy.dk/). They wanted to use specific occurrence data from two lighthouses (Lodbjerg Fyr and Hanstholm Fyr) for an exhibition project. It would be a benefit to the Natural History Museum of Denmark to have DanBIF digitizes these volumes, namely, to preserve them for the future and provide online access for everyone. The digitization project eventually expanded after the discovery of more journal publications and thus resulted in the compilation of "Fuglene ved de danske Fyr, 1883 -1939" (UK: Birds at the Danish Lighthouses, 1883 – 1939).

Digitisation of the book volumes "Fuglene ved de danske Fyr, 1883 -1939" (UK: Birds at the Danish Lighthouses, 1883 – 1939) were carried out prior to compiling this dataset. The books were scanned individually using a PLUSTEK OpticBook A 300 book scanner and named as, for example, NHMD_LighthouseBirds_1883 and saved as a PDF file. These PDFs can be found at http://danbif.dk/links/litteratur.

The texts in the PDF files were then OCR-ed (Optical Character Recognition) using Adobe Acrobat. This is to enable the computer to recognize letter shapes in a scanned document, and turning them into digital text in order to copy, paste and edit as needed. Performing OCR to the scanned books was essential for this digitisation project as data were extracted by copying information from the PDF files and pasting them on a database file.

The database file chosen for this project was a simple Excel sheet with the following fields, mapped according to the Darwin Core terms: occurrenceID , locality (*Lighthouse name*), decimalLatitude, decimalLongitude, year, month, originalNameUsage (s*cientific name as appearing in the books*), identificationRemarks (*Danish names as appearing in the books*), vernacularName (*current Danish names*), scientificName, scientificNameAuthorship, individualCount (*observed individuals fallen by the lighthouse*), sex, lifestage, disposition (*individuals fallen that were sent to the Zoological Museum*) and eventRemarks (*notes by the lighthouse keepers, including weather conditions and wind description*).

The next process was opening the PDF files in Adobe Reader and performing a basic search for a particular lighthouse, for example, Lodbjerg Fyr, using the Find toolbar. Data that corresponded to this particular lighthouse were then copied from the PDF files and pasted in the correct fields in the Excel sheet (database file).

Most information were easily found from the books. However, there were a few fields that required some investigation. For vernacularName, scientificName and scientificNameAuthorship, http://allearter-databasen.dk/ was used to check for the current name, taxon, author and year. Thus the database contains both the legacy names used in the books and the names currently accepted in nomenclature. The last part of the digitization process was to give the records an occurrence ID following the template urn:NHMD:LighthouseBirds:01 with running numbers.

A total of 1212 records were generated from digitizing "Fuglene ved de danske Fyr": Birds at the Danish Lighthouses 1883 through 1939, 742 records from Lodbjerg Fyr and 470 records from Hanstholm Fyr. The complete dataset can be viewed through GBIF portal at http://www.gbif.org/dataset/ad331dcc-d0fa-4816-b1e6-d36f9f899c49. An appendix with description of the weather conditions and wind is available at http://danbif.dk/links/litteratur/Fyrfaldne-fugle_Vind_Forkortelser.pdf. The information provided at this particular link is in Danish. However, translation could be provided on demand. For translation request, kindly contact Isabel Calabuig, Node Manager of DanBIF via email, icalabuig@snm.ku.dk.