

Ensuring the Legacy Data for the Southern Right Whale is FAIR

Daureen Nesdill

daureen.nesdill@utah.edu

J. W. Marriott Library, University of Utah 295 South 1500 East Salt Lake City, UT 84112-0860 USA

Abstract

Worldwide the push is for research data to become FAIR: Findable, Accessible, Interoperable and Reusable. So what about legacy data? Vicki Rowntree, a Research Professor at the University of Utah, has been collecting behavioral and biological data of the Patagonia southern right whale since 1971 on over 3,000 individual whales. The dataset consists of over 84,000 slides of these whales for identification purposes, hand-drawn maps and a room full of file cabinets containing hand-written data sheets. Yes, she went digital when the world did and now has an out-of-date Microsoft Access database to add to the analog data. Other researchers have also been collecting longitudinal data on the southern right whale, *Eubalaena australis*, in Australia, New Zealand, Argentina, Brazil and South Africa. Obviously, the data were not collected and described following any standard procedure. Here at the University of Utah we are working towards bringing all the research together by hosting and standardizing the datasets. We will, in part, use some of the standardization procedure of the [Global Biodiversity Information Facility \(GBIF\)](#) to facilitate adding location data to their database. We are proposing to build a Web platform for accessing the data and tools to evaluate and analyze the data. This talk will be about our work and the Patagonia Right Whale.

Keywords: *Eubalaena australis*, Southern Right Whale, Southern Oceans, legacy data, FAIR, collaboration, climate change.

Initially naturalists studied species of animals by collecting specimens, noting the location and taking measurements. During the late 1960s and early 1970s the research shifted towards collecting information on individuals within a population of a specific species. Reproduction, diet, communication, and behavior were recorded. From this information, the social structure of the group could be determined. Jane Goodall, L. David Mech, Roger Payne all became household names because of their work with chimps, wolves and whales, respectively. These long-term studies provide unique information about a population's distribution and habitat requirements and facilitates the monitoring of the effects of environmental change.

Roger Payne, the discover along with Scott McVay of whale songs, visited Península Valdés, Argentina in 1970 and noticed he could recognize individual southern right whales, *Eubalaena australis*, from the patterns of white markings on their heads (Payne et al., 1983). The following year he initiated a program of annual aerial photographic surveys of the population under the direction of Victoria Rowntree, who has continued the research until the present. Rowntree, a Research Professor at the University of Utah, is Director of the Patagonia Right Whale Project. She is also the Right Whale Program Director for [Ocean Alliance, Inc.](#), a 501(c)3 organization founded by Payne in 1971.

As Director of the Patagonia Right Whale Project, Rowntree has led colleagues and graduate students from the National University of Córdoba in the study the ecology, behavior and biology of the southern right whale at Península Valdés. The data collected for 48+ years consists of photograph negatives, slides, digital images, hand-drawn maps, data sheets and an MS ACCESS database (Table 1). In addition to being an irreplaceable collection, these data are the basis for an ongoing, international research program, [Instituto de Conservación de Ballenas](#), ICB that can continue years after Rowntree is gone. Since they are continuing Rowntree's work, they need access to all of the project's data in a digital format that is FAIR, i.e. Findable, Accessible Interoperable, and Reusable.

Southern right whales inhabit the southern oceans. The research conducted at Península Valdés now has a core data set comprising a catalog of over 3,200 known individual whales and records describing when, where, and under what circumstances whales have been sighted. From this research it is now known that the whales begin arriving at the Península Valdés in May. Calves are born in August (Best, 1994). The greatest number of whales move to the peninsula by late September early October. By December the whales have left for their summer feeding grounds in Antarctica (Payne, 1986; Payne et al., 1990). The mothers primarily fast on the nursery ground while calves nurse and mature (Taber and Thomas, 1982; Thomas and Taber, 1984).

The data set recorded at Península Valdés is significant because this population has been studied longer than any other population of right whales. The data includes the long reproductive histories of many individuals and the results show that calving rates may be sensitive to the conditions at the feeding grounds. A study correlated a reduction in the number of calves when higher sea surface temperatures occurred as a result of El Niño years (Leaper et al., 2006).

In addition, the Patagonia Right Whale Project has been able to document the rise in deaths of the southern right whales due to attacks by Kelp Gulls, *Larus dominicanus*, from 1970 until present. The attacks have increased dramatically, from an average of 2% in the 1970s to 99% in the 2000s and the attacks have expanded to occur in both of the bays of Península Valdés (Maron et al., 2015).

To collect data, each whale must be identified. The patterns of white markings on the heads of whales noticed by Payne are callosities. Callosities are irregular patches of thickened and keratinized tissues. The coloring of the callosities are a result of amphipod crustaceans or whale lice that inhabit the structures. The pattern of growth of callosities on the head and back of right whales are unique to each individual and are therefore a sort of fingerprint that allows researchers to identify individuals (Payne et al., 1983; Payne and Dorsey 1983). To collect data, Rowntree and her research group fly over the bays and as a whale is spotted take photos. Initially the photos were analog but became digital in 2005 when digital cameras became available (Table 1). In the past whales were observed by boat. At this point in time drones are being incorporated into the research. These identification photos are analyzed using a computer-based system designed and implemented by Lex Hiby and Phil Lovell in the 1970's. It is not a completely automated system, so it takes a bit of time to determine if the whale in question is a new or formerly identified individual (Hiby and Lovell, 2001).

DATA	Amount Collected and Format
Identification data	84,000 color slides 12,856 negatives Since 2005 75,536 digital images have been collected
Biology/ecology data	78 3-ring notebooks which have been digitized and are in pdf format Since 1980 data has been entered into an MS ACCESS database
Location data	Hand-drawn maps of sightings which all have been converted to ARC GIS format

Table 1. Data Collected on the Southern Right Whale at Peninsula Valdez.

In addition to taking photos of the sighted whales during a fly over, data was recorded by hand on data sheets. In 1980 a MS Access database was developed so data could be stored electronically (Table 1). Location data were also recorded and plotted on hand drawn maps.

In 2015 Rowntree approached the author requesting assistance in digitizing the slides and data so that the information would be accessible to members of ICB and other whale researchers. Harish Maringanti, Associate Dean for IT and Digital Collections at the Marriott Library was brought into the conversation and an internal seed grant was applied for and granted by the University's Office of the Vice President for Research. The goal of the grant was to determine best practices, workflows, cost, FTE, equipment and time required to digitize the slide collection. Several informatics issues such as data ownership rights, data reuse, etc. also needed to be addressed before data could be widely shared. Three years of slides were digitized. In addition, the 78 notebooks of data sheets were converted to pdf format and Rowntree was able to work with graduate students who were knowledgeable about ARC GIS software to convert the analog maps to digital format.

During her 48+ years studying the whales, Rowntree's research has been funded by private donations from nonprofits, the National Geographic Society, the National Marine Fisheries Service and the Marine Mammal Commission. For the present research U.S. National Science Foundation funding is being sought to develop a portal for storing, curating, preserving and providing access to the identification catalog and its software, and to the maps and the biological and ecological data. In addition, standards and best practices will be developed to facilitate the addition of data from other researchers studying different populations of the southern right whale. Researchers of whale populations in Australia, New Zealand, Brazil and South Africa are known to be interested in the project. The software will be open source so it can be adapted to other species.

The Hiby-Lovell identification software requires upgrading. A new procedure has been developed for the North Atlantic right whale, *Eubalaena glacialis*, by deepsense.ai, incorporating deep learning integrated with a fully automated workflow. Since the morphology of the North Atlantic species is similar to the southern species, a group of collaborators is now working with [Wildme](#) to expand the software (Bogucki, 2018).

Whale watching in Argentina is a major tourist attraction (Arias et al., 2018). The ICB provides information and training to the whale-watching community, including the whale-watch-boat operators. Boat-operators are providing whale watchers with information about the whales and in return the whale watchers are contributing their photos to ICB. An educational web site will be developed as a component to the platform to facilitate educating whale watchers and to provide a place for them to contribute their photos of the whales they encounter. This will add to the amount of location points for sightings of each whale.

For the resulting portal to be useful, standards must be established and followed. The [Global Biodiversity Information Facility](#), GBIF membership includes biodiversity researchers working with data librarians and data scientists to establish a worldwide portal of data on biodiversity. By collaborating with North American members, we will ensure our datasets are developed using the appropriate standards for organizing and applying metadata. Working with GBIF will increase exposure to our portal and inform additional researchers of our portal for preservation and providing access to their data.

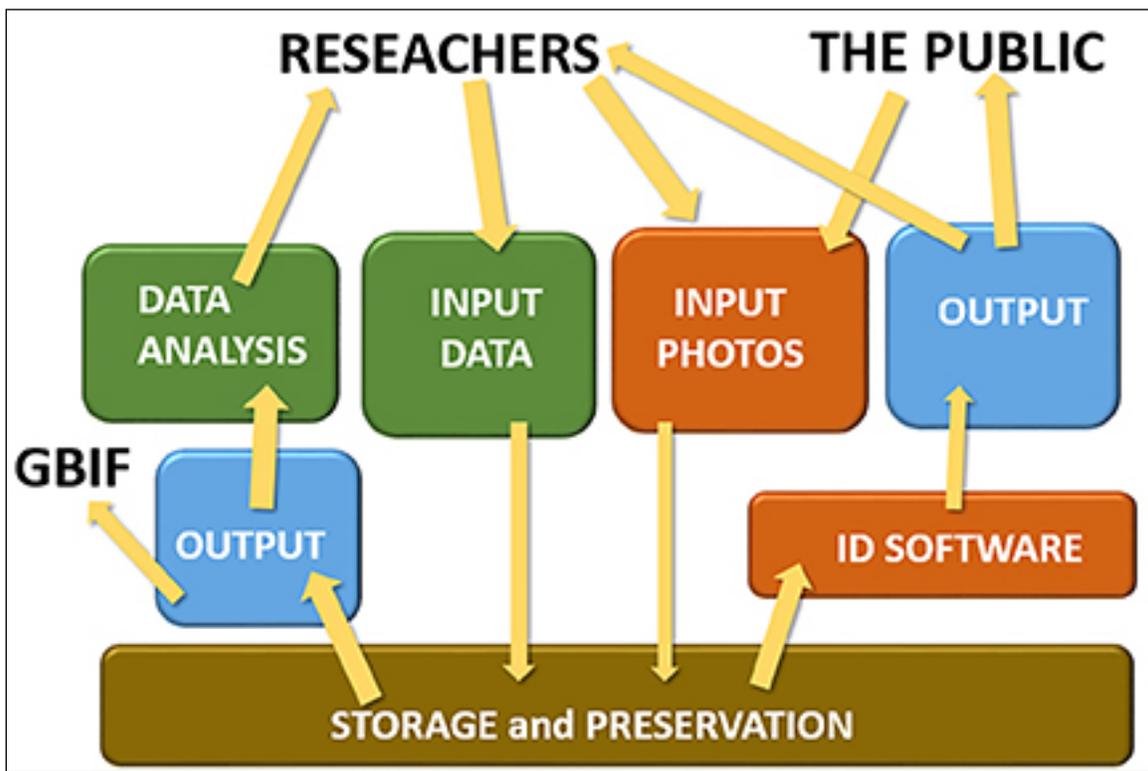


Figure 1. A schematic of the portal providing storage and preservation of data collected. Researchers can submit photos and data and use tools to select data for analysis and identification of whales. The public can submit photos and check to see if a whale they have spotted is in the system. GBIF has access to location data.

Research indicates that global climate change is leading to increasing ocean temperatures and is impacting the rate of ice loss of Antarctic glaciers leading to a rise in sea level (Robel, Seroussi and Roe,

2019). How will changes in the ocean environment affect the southern right whales? Having the longitudinal data collected on the Península Valdés population in a digital format and accessible along with the datasets from researchers studying additional populations in the southern oceans will potentially provide the information required to answer this question.

References

- Arias, T., Coscarella, M. A., Romero, M. A., Svendsen, G. M., Reinaldo, M. O., Curcio, N. S., Crespo, E. A., & González, R. A. C. (2018). Impact of whale-watching on southern right whale (*Eubalaena australis*) in Patagonia: Assessing the effects from its beginnings in the context of population growth. *Tourism Management Perspectives*, 27, 1-9. <https://doi.org/10.1016/j.tmp.2018.03.005>
- Best, P.B. (1994). Seasonality of reproduction and the length of gestation in southern right whales, *Eubalaena australis*. *Journal of Zoology, London*, 232(2), 175–189. <https://doi.org/10.1111/j.1469-7998.1994.tb01567.x>
- Bogucki, R., Marek Cygan, M., Khan, C. B., Klimek, M., Milczek, J. K., & Mucha, M. (2018). Applying deep learning to right whale photo identification. *Conservation Biology*, 33(3), 676–684. <https://doi.org/10.1111/cobi.13226>
- Hiby, L. & Lovell, P. (2001). A note on an automated system for matching the callosity patterns on aerial photographs of southern right whales. *Journal of Cetacean Research Management*, (Special Issue) 2, 291–295.
- Leaper R. I., Cooke J., Trathan P., Reid K., Rowntree V., & Payne, R. (2006). Global climate drives southern right whale (*Eubalaena australis*) population dynamics. *Biology Letters*, (2), 289-292. <https://doi.org/10.1098/rsbl.2005.0431>
- Marón C.F., Beltramino L., Di Martinao M., Chirife A., Seger J., Uhart M., Sironi M., & Rowntree, V. J. (2015). Increased wounding of southern right whale (*Eubalaena australis*) calves by Kelp Gulls (*Larus dominicanus*) at Península Valdés, Argentina. *PLoS ONE*, 10(10), e0139291 <https://doi.org/10.1371/journal.pone.0139291>
- Payne, R. (1986). Long term behavioural studies of the southern right whale (*Eubalaena australis*). *Report of the International Whaling Commission*, 10(special issue), 161-168.
- Payne, R. & Dorsey, E.M. (1983). Sexual dimorphism and aggressive use of callosities in right whales (*Eubalaena australis*). pp. 295-329. In R. Payne (ed.) *Communication and Behavior of Whales*. AAAS Selected Symposia Series, Westview Press, Boulder, CO. i-xii+643pp.
- Payne, R., Brazier, O., Dorsey, E., Perkins, J., Rowntree, V. and Titus, A. 1983. External features in southern right whales (*Eubalaena australis*) and their use in identifying individuals. In: Payne, R., editor. *Communication and Behavior of Whales*. AAAS Selected Symposia Series 76, Boulder, CO, Westview Press, p. 371-445
- Payne R., Rowntree V., Perkins J. S., Cooke J. G., & Lankester, K. (1990). Population size, trends, and reproductive parameters of right whales (*Eubalaena australis*) off Peninsula Valdés, Argentina. *Report of the International Whaling Commission*. (Special Issue 12), 271-278.
- Robel, A. A., Seroussi, H., & Roe, G. H. (2019). Marine ice sheet instability amplifies and skews uncertainty in projections of future sea-level rise. *PNAS*, 116 (30), 14887-14892. <https://doi.org/10.1073/pnas.1904822116>
- Taber S. & Thomas P. (1982). Calf development and mother–calf spatial relationships in southern right whales. *Animal Behaviour*, 30, 1072–1083. [https://doi.org/10.1016/S0003-3472\(82\)80197-8](https://doi.org/10.1016/S0003-3472(82)80197-8)
- Thomas. P. O. & Taber, S. M. (1984). Mother-infant interaction and behavioral development in southern right whales, *Eubalaena australis*. *Behaviour* 88(1/2), 42-60. <https://doi.org/10.1163/156853984X00470>