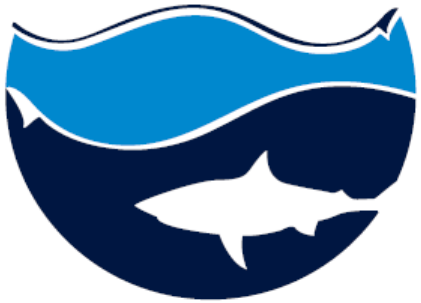
DATANAUTS
2016 CAPSTONE SCHEMATIC REPORT
SHARK SHARE GLOBAL



SHARK SHARE GLOBAL

Heather Moore, Elizabeth MacCready, Molly Ostheller

Table of Contents

- Abstract
- Table of Contents
- Who is Shark Share?
- What is the problem?
- Requirements Analysis
- Solution – Technical
- Solution – Design
- Solution – Process
- Solution – Results
- Impact of Shark Share Global
- Next Steps
- Conclusion

Abstract:

Sharks are enormously important to our ecosystem, yet poorly understood. Researchers seek to understand and study them but are hampered by siloing, wasted research opportunities, and underdeveloped collaborative networks. Shark Share Global is a website and database that facilitates sample sharing and collaboration between researchers around the world. It is custom-built to encompass the needs of researchers- bringing the technology solution as close to their existing processes of sample cataloguing as possible- to ensure a low barrier to adoption. Open only to accredited shark and ray researchers, it is a simple, elegant solution to a sprawling problem, introducing a formal research collaborative tool where none have existed before.

Who is Shark Share?

SHARK SHARE GLOBAL

Mission

- Shark Share Global was founded in order to create an international shark and ray tissue database to increase the efficiency of field work and facilitate researcher collaboration by providing a central repository for researchers to list surplus tissue samples that other researchers can use. The organization has recently gained NGO status during our time working with them.

Founders

- Shark Share Global was founded by Madeline Green and Lauren Meyer, both PhD students. Lauren is an ecotoxicologist at Flinders University, and Madeline is a shark geneticist at University of Tasmania. Both founders are located in South Australia- as is Shark Share.
- Madeline and Lauren created Shark Share Global in order to address pressing problems facing shark and ray research collaboration.

Madeline Green

Lauren Meyer

Photo credit: Save Our Seas Foundation saveourseas.com/project/shark-share-global/

Who Supports Shark Share?

Community Support

- The database is surrounded by a healthy community, and they hope that their site will be used by Elasmobranch scientists and researchers all over the world. Shark Share Global is a standalone organization, but has many other research allies and supporters associated with it as well, including the Save Our Seas Foundation and Collaborative Research Networks Ltd, and Shark References, which provided the taxonomy list for the database.

Leveraging People

- Shark Share Global has already begun to reach out and determine need and desire for the database solution that we have created. They have outside funding in the form of grants, crowdsourced funding, and pledges of support from many different groups from within the research community.

Financial Sponsorship

- Funding for Shark Share Global is provided by a generous grant from

save our seas foundation

Centre for Sustainable Tropical Fisheries & Aquaculture

shark-references

Bibliography database, typedatabase and species descriptions of living and fossil Elasmobranchs

Marine Biodiversity Hub

National Environmental Science Programme

RED LIST Guiding Conservation for 50 Years

Manta TRUST

IN PARTNERSHIP WITH THE SAVE OUR SEAS FOUNDATION

What is the Problem?

- Shark researchers have an esoteric, inefficient method for collaborating on projects and sharing samples to prevent them from going to waste.
- Researchers rely solely on spreadsheets, emails, and cold-calling in order to get samples and support for projects outside the purview of their own university or lab community.
- 47% of shark species are data deficient, of the ones we know about, a third of those are threatened by human activity.
- Poorly understood aquatic ecosystems lead to devastating effects on seafood and sea product industries. As apex predators, sharks are great regulators of environments.
- The ocean is the last 'great frontier' on the Earth, which means that research is vital to understanding our planet, but also difficult to come by.
- New researchers, especially PhD students, have no network outside of their research institution by which to leverage samples and help with research. This leaves them with only emailing and cold-calling other experts in their subdomain around the world- an awkward approach often prone to failure.
- Sample data are stored in isolated spreadsheets, and labs are often unequipped for the open sharing of data.

Requirements Analysis- Functional

Website
Functional Requirements

Search by species, and by part
- *Search by keyword*
- *Nomenclature should be by scientific name only*
- *Searches focused around species > sample type > location hierarchy of need*

Submit available samples into database
- *Batches/sample sets can be uploaded at once, not line-by-line*

Submit request for available tissue
- *Button next to available sample batches*
- *'I'm looking for' feature in profile/landing page*

Accessible from multiple places around the world

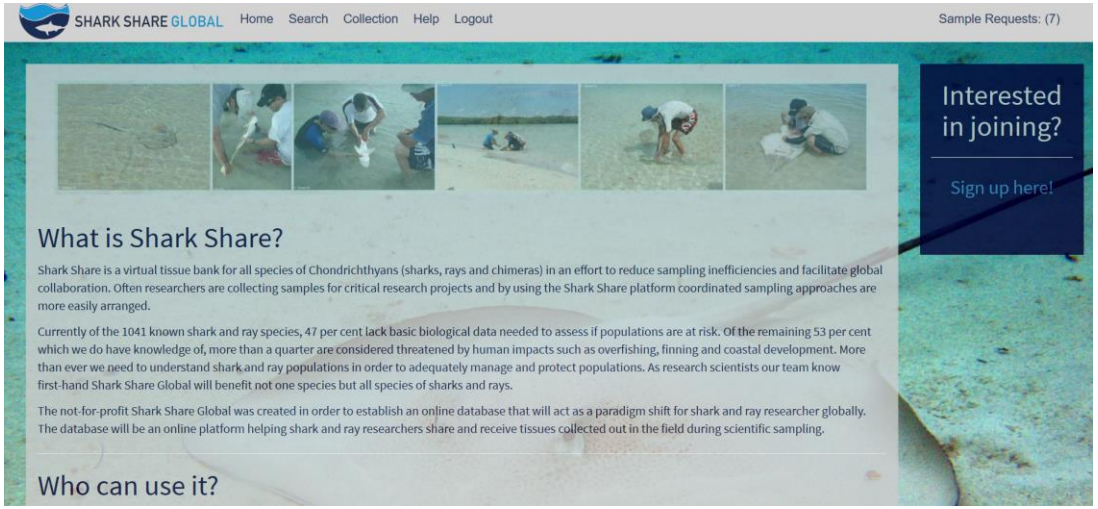
Support further interactions by researchers via email

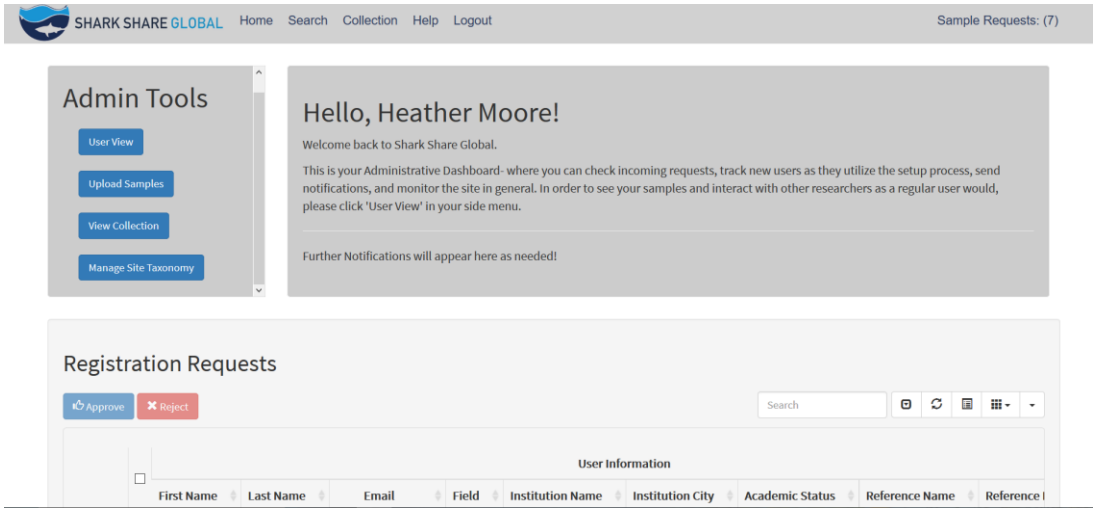
Must remove (or provide process for removal) of claimed or expired sample

Will utilize SQL queries to execute retrieval

Will work in the relational database format

Will be built in html/css/js/d3js for front-end design





Heather

Requirements Analysis- Non-Functional

Non-functional

How it works/How it should behave

Accepts records of samples in batch format

- *Supports user copy-pasting from records*
- *Multiple records can be entered this way (see 1b)*

Have user sign-up that contains:

- *Name (req)*
- *Institute affiliation (req)*
- *Email address (req)*
- *Reference contact (req)*
- *Research Statement (req)*
- *Phone number contact (opt)*

Explains why Shark Share needs registration requirements on sign-up

Contains page of terms and conditions for the website

- *New users must accept this during sign-up process*

System of manual approval by Shark Share administrators

Users cannot browse contents until account has been approved

Contain landing page after successful login

Have a focus on identifying the owner of the sample at every step of the process to smooth attribution concerns

Have profile creation screen/UN/PW/profile edit options

- *Must set up profile before sharing or requesting samples.*
- *Contains:*
  - *Name (req)*
  - *Institution (req)*

Request form for samples that contains:

- *Research intent (to avoid overlap)*
- *Use agreement terms*
- *Who bears cost of shipping*
- *Request template comes up in an email popup which facilitates transition to personal email communication.*
- *Requires explanation/walkthrough of process*
- *Location and contact data for sample provider*

Contains links page

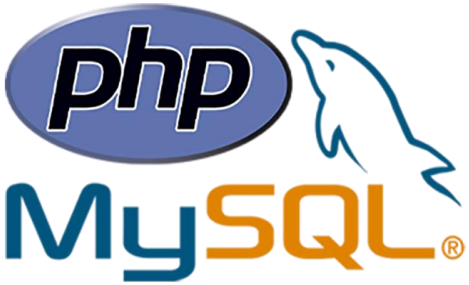
- *IUCN assignment page link*
- *Shipping help and information links*
- *Shark references and sponsors links*
- *Allows for browsing and searching of samples*

Will have a controlled vocabulary for user use

Must be easy to use, easily navigable, and user-friendly for people with some computer literacy, but not necessarily advanced users.

- *Must follow when possible existing processes for elasmobranch information retrieval and upload.*

Solution - Technical



- The solution is a custom web app built using the LAMP stack
  - It is built using CodeIgniter, a lightweight PHP framework
  - And leverages libraries like PHPExcel to extract and transform data
  - The site is styled using the Bootstrap framework, as well as custom css work
- The database is a MySQL database hosted on AWS RDS
  - Hosting in the cloud improves performance worldwide by distributing service, maximizing efficiency for a dispersed, global audience
  - As a result, the site has increased resilience against service loss and greater agility
  - This provides lower initial costs for our sponsor, and low ongoing maintenance, with minimal effort for the two administrators.
  - Amazon's Elastic Beanstalk Service also contributes technical help to our sponsor, who do not have deep technical experience, by resolving for them issues of scaling and elasticity automatically while allowing them to maintain a custom database that has the features they need
- Emphasis on security and usability
  - Search leverages controlled vocabularies to improve precision and recall
  - Valid login credentials necessary to view site
  - Form data are protected from Cross-Site Scripting (XSS) attacks

Solution - Design

Database

Global accessibility and good speed and performance at that scale were needed to encourage adoption of the database, but the solution also needed to be maintained by just two scientists with occasional IT assistance. Additionally, it needed to be cost -efficient, as Shark Share is a newly founded NGO. To solve this, an open source solution database solution hosted in the cloud was chosen. This ensured the scalability needs were met, but the only costs would be hosting and occasional maintenance in the future.

Web Portal

To access the database, a custom web portal was developed for uploading, browsing, discovering, and deleting samples, as well as making contact with other researchers. To encourage adoption we designed this website in accordance with best-practice UX principles and modeled as many processes for interacting with the website with existing practices used by researchers, These practices reduced the amount of learning new users would need to do to start using the site.

Metadata

The site metadata needed to be versatile, well-defined and usable. We worked with Shark Share to determine what metadata should be included as attributes in the database, as well how users would likely use it for search functions. From there we developed the search fields as well as controlled vocabulary lists for fields like locations. measurements, and the Linnaean shark taxonomy.

Security

Finally, the app also needed sufficient security as its access is restricted to researchers vetted and approved by Shark Share. Each user must provide a personal reference and the institution they are associated with and will be manually approved by Shark Share. We also wanted to protect it from potential deletion other hacking concerns. To do this, standard security measures such as hashed passwords, filtering user input to prevent SQL injections, and other methods mentioned in the technical solution above were implemented.

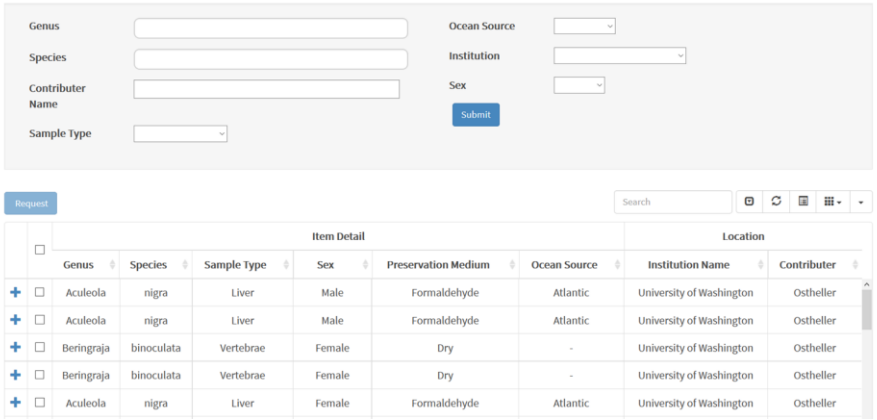
Solution - Process

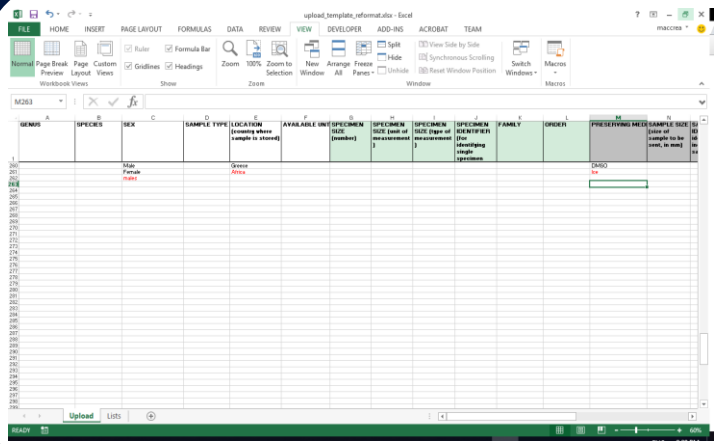
- Our first step was to make an analysis of existing work from past consultants who scoped out the process, but were not selected to go forward.
- We then had a series of conferences with our sponsor to determine where the previous work had been correct, where it needed improvement, and what needed to be discarded.
- As a result of these efforts, we created a requirements analysis that laid out the expectations for what the project needed to be successful.
- We proceeded from there to lay out scheduling for website creation, as well as wireframes and design documentation. While we did not have the time to do a high-fidelity prototype, we used other methods to get a clearer idea of what the sponsor wanted.
- For the metadata, we also made a controlled vocabulary list that is updatable as the field of shark and ray research changes. We controlled vocabulary both in the upload process and through a series of checks in the database itself.
- In order to make things as simple as possible, we mirrored broadly-used data storage techniques in the field in our data upload procedures, instead of making researchers learn a new interface.
- To balance the need to settle distrust and ensure correct academic attribution, we created a multi-check process where users agree to levels of attribution prior to joining and sharing.
- We test with volunteers from the shark and ray community identified through our sponsors to reinforce data, method, and design decisions.
- Before launch, the project will be hosted on AWS RDB services for global distributed use, with help in place for the sponsor to aid growth.

Solution – Notable Features

Search

Researchers can search for samples using multiple facets. The most commonly searched fields will be genus and species of the shark, but users can also search by ocean, institution, contributor, sex, or sample type. Search results return filtered rows that give basic information about each sample, which can be expanded to show optional fields by clicking the plus sign to the left of each row.





Upload Sheet

An excel template was created for database uploads that highlights any data violating the controlled vocabularies and checks for field validity so researchers can simply cut and paste sample information, make a few minor spelling adjustments, and upload hundreds of samples to the database quickly and easily.

Tagged Alerts

Alerts about relevant samples will display on the user's dashboard based on tag words selected by the user.

Impact of Shark Share

Shark Share has the potential to impact many areas of shark research.

Environmental

Sharing samples will reduce the need to harvest more sharks from the ocean to do research, as the ones that are taken can be used more efficiently. This will reduce researchers' impact on these populations. Additionally, more efficient science can be done which should help reduce the data deficiency on many of these species at a more rapid rate.

Research

By being able to share samples across the world, researchers will be able to do more efficient and better science because they will have much easier access to potential samples needed to test their theories. This will also serve to reduce unnecessary field work, saving many scientist significant time and money on their already tight grant budgets.

Collaboration and Networks

New scientists and those from developing nations will be better able to grow their professional networks and gain access to connections in similar research areas that otherwise can take many years and good luck to develop.

Next Steps

- Shark Share Global goes live in July 2016 at the American Elasmobranch Society's yearly conference in New Orleans, LA.
- The Datanauts will transition the site over at the beginning of July prior to this event. The work completed after graduation will be compensated with a stipend from funding through the sponsor.
- Shark Share Global will continue to mount their social media campaign to gain wider awareness, and tap sponsors of their own that have been waiting for the service for some time.
- Shark Share Global will also promote the product at several further conferences throughout the summer, and at other events over the next year.
- In the future, features provided by AWS' Elastic Beanstalk, as well as further recommendations by the Datanauts team, will guide the technological development and growth of the product as, we hope, it grows ever larger. This includes either scaling up in size or moving to another service that allows for greater discounting to NGOs as they expand to larger data storage needs.

Conclusion

In summary, Shark Share is a global platform to facilitate sample sharing and collaboration between shark researchers. It was designed to scale easily and have good performance and speed at a global level. It is also cost-effective and low maintenance so it can be maintained by Shark Share with minimal IT assistance. With this solution we hope to support increased research efficiency, lower harvesting impact on shark populations, and foster greater collaboration among new and established researchers.

Photo Credit [all photos]: Green, M. and Meyer, L. *Shark Share Global, sharkshareglobal.org accessed: 5/14/2016, Problem Statistics:* 47% of sharks and rays are at risk: Dulvy, Nicholas K., et al. "Extinction risk and conservation of the world's sharks and rays." Elife 3 (2014): e00590.. Our Sponsor Graphics courtesy of: saveourseas.com/projects/shark-share