Larry L. Sautter Award Proposal 2022

Name of project:

Shoreline - GeoData platform for three UC Campus Libraries

Summary:

An open-source, cross-campus collaborative project to create an interactive map-based discovery platform for UC Libraries' geospatial data.

Submitted by:

Chrissy Rissmeyer, Director of Digital Library Development, UC Santa Barbara Library, crissmeyer@ucsb.edu

Project Leaders and Team Members:

Tom Brittnacher, UC Santa Barbara (Primary Product Owner)
Alexandra Dunn, UC Santa Barbara (Technical Lead)
Amy Work, UC San Diego (Product Owner for UC San Diego)
Thor Brickman, UC San Diego
Vivian Chu, UC San Diego
Matt Critchlow, UC San Diego
Jessica Hilt, UC San Diego
Dan LaSusa, UC San Diego
Ron Stanonik, UC San Diego
Michael Stuart, UC San Diego
David Hendlin, UC Santa Barbara
tamsin johnson, UC Santa Barbara
Margaret Kibi, UC Santa Barbara
Chrissy Rissmeyer, UC Santa Barbara
Kristian Allen, UCLA
Stephen Gurnick, UCLA
Zhiyuan Yao, UCLA

Project Narrative:

What is Shoreline?

Shoreline (branded publicly as GeoData) is a discovery platform for geospatial data built on GeoBlacklight. GeoBlacklight, the project, is an open-source collaboration with the goal of building a better way to discover, share, and access geospatial data online. GeoBlacklight, the software, is a powerful application that incorporates interactive maps to search for and preview geospatial data. Built as a Ruby on Rails gem, GeoBlacklight is based on the popular open-source discovery software Project Blacklight.

Why Shoreline?

Increasingly, students, faculty, and researchers draw on digital information identifying the geographic locations and characteristics of the places we live, travel, and research. This type of data (called geospatial data) grows more valuable when compared over time. For example, it can provide answers to basic questions such as, How have the streets in a given county changed over the past five or 10 years? or, How has affordable housing and land use changed in Santa Barbara since 2000?, or How much of the California coastline has eroded since 1990? The information derived from this data is useful on a variety of levels and interesting to students and researchers across disciplines.

The UC Santa Barbara, UC San Diego, and UCLA Libraries each license and acquire a number of geospatial datasets that are delivered as files for the libraries to redistribute to users. Historically, the path for accessing that data was cumbersome for the user and time-consuming for Library staff. After discussing their needs, UC Santa Barbara and UC San Diego decided that those needs were similar enough that pooling their resources and expertise would be more cost and time effective for the UC system than if they each tackled the problem alone.

To begin, the team reviewed existing spatial discovery, geoportals, and geographic information retrieval systems, assessing seven existing instances and evaluating 130 GeoBlacklight features. Working with a team of geospatial subject matter experts at each campus, a gap analysis was performed on an initial test deployment and the list of features constituting the minimum viable product (MVP) was created. Once development work began, we worked as a unified, cross-campus Agile development team in the same shared code base. The initial MVP, branded GeoData, was successfully deployed at each campus from that shared code base approximately 9 months after the project formally began.

However, this alone did not constitute our collaboration effort. We wanted to ensure that everyone who was interested in the project UC-wide felt invited to participate. We wanted other campuses (and other universities) to know what was working for us and what wasn't. We knew that not everyone could work on the same timeline as us, but might be interested in joining the effort in the future. We made our planning documents in Google Drive open to anyone interested. The code repository and documentation are publicly available via GitLab. We shared our sprint summaries and recorded demos on our project blog so that anyone interested could catch up on the current status. We ran all of our project communications out of the UC Tech Slack in a public channel (#surfliner). We presented at the UC-wide tech and library conferences (UC Tech, UCDLFx, UC Libraries Forum) and at geospatial data community conferences (ESRI, Spatial Discovery III, Geo4LibCamp) to get the word out about our project and to invite other UC campuses to participate.

Those outreach efforts paid off when the UCLA Library officially joined the Shoreline team in 2021. UCLA realized that they too could not build an effective geospatial discovery platform on their own, and saw Shoreline as a way to leverage expertise and contributions from other campuses while providing their own knowledge and time to the project. As a result, recent progress has been made at a more rapid pace and all three campuses have benefitted.

The Impact of Shoreline

Before the Shoreline project, library users were only able to access geospatial collections by coming in person and talking to staff members. If geospatial collections were put in the traditional library catalog, users had difficulty finding relevant data. (Text-based searching using geographic terms is rarely successful because users type in specific geographic study areas, but datasets are cataloged by their full geographic extents.) If a user was successful in finding a record for a relevant dataset, they would still need to travel to the library building to access the data with the help of staff. With GeoData, users can zoom into their study area and see all of the datasets held by the library in that area in the results list. Once they select a dataset from the list, a web page comes up with an interactive preview of the dataset as well as descriptive information. From that page, they can download the dataset directly without having to come into the library at all. Additionally, although each campus has its own instance of GeoData, metadata can be shared between campuses so that each campus' data can be discoverable and downloadable (if unrestricted) from any of the GeoData interfaces.

The impact of collaborating on Surfliner has gone beyond just the architecture/infrastructure. Because of the collaboration, each campus was forced to think about the system, geospatial data, and associated workflows and processes (metadata, ingest, etc.) through a different lens. The collaboration has spread beyond the immediate Shoreline project team to those who manage and describe geospatial data at each library. The addition of the third campus has pushed us to rethink our previous work such as making the ingest process more widely reusable for campuses with different underlying infrastructures. This collaboration among three campuses eventually makes Shoreline more usable to all UC campuses when they are ready to embark on this journey of making their geospatial data holdings more discoverable and accessible.

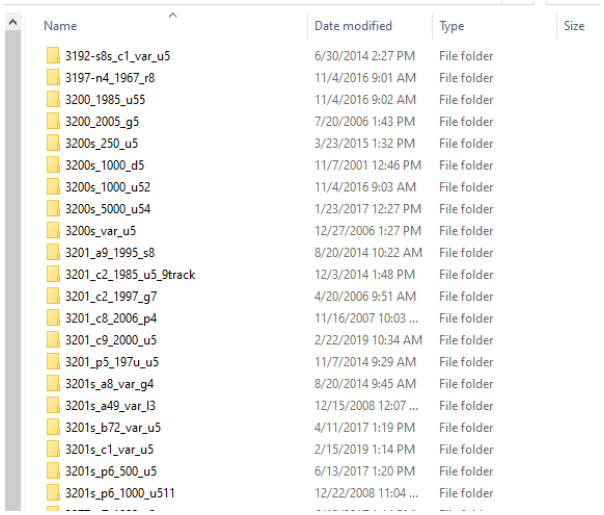
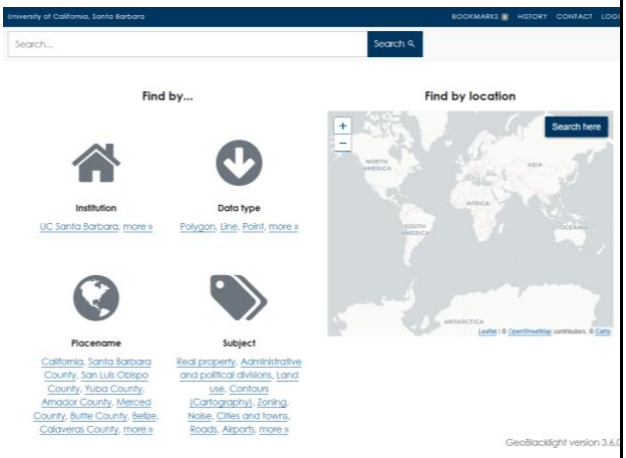
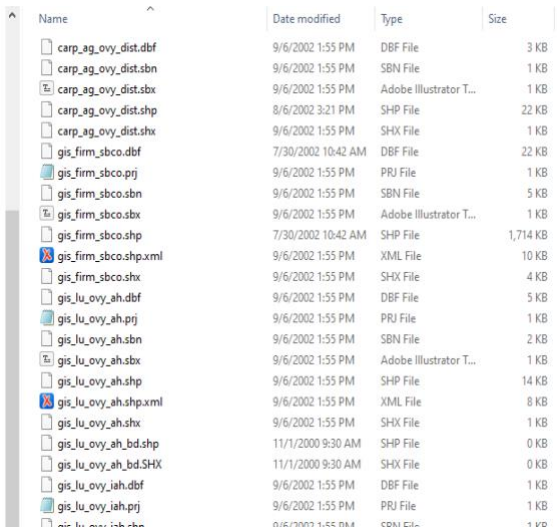
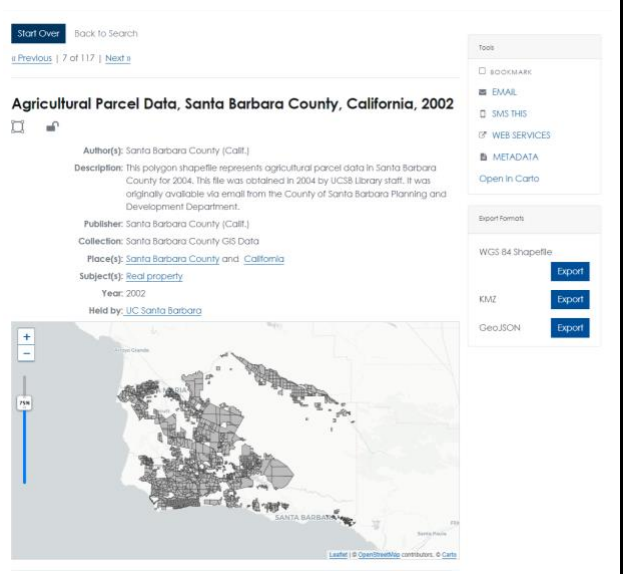
The Time Frame

Formal requirements gathering began in earnest in September 2019. Our first development sprint began in February 2020. The MVP was deployed in June 2020 at both UC Santa Barbara and UC San Diego. (UCLA did not formally join the collaboration until 2021.) The Shoreline team has continued to work on adding new features and improvements since the launch of the MVP. Recent work has been focused on adding support for additional data types and integrating the platform with our digital object management systems.

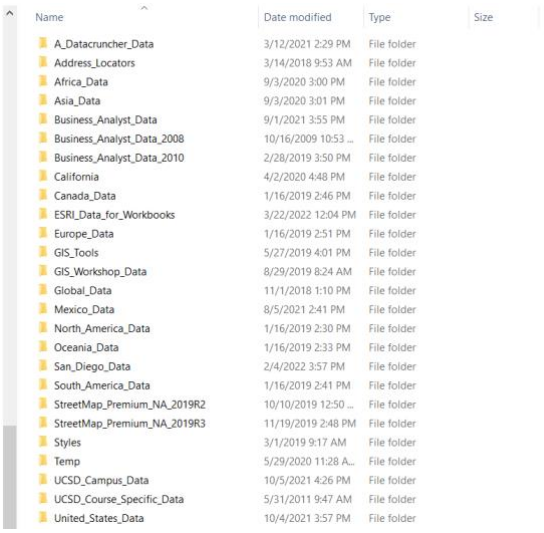
The Technology

- *GeoBlacklight* - an open-source Ruby on Rails software application for discovering geospatial content, including GIS datasets, web services, and digitized paper maps.
- *GeoServer* - an open-source server for sharing geospatial data
- *Gitlab* - an open-source code repository and collaborative software development platform. Shoreline code is available at https://gitlab.com/surfliner/surfliner/-/tree/trunk/shoreline
- *Helm/Kubernetes* - an open-source container orchestration system for automating application deployment, scaling, and management.

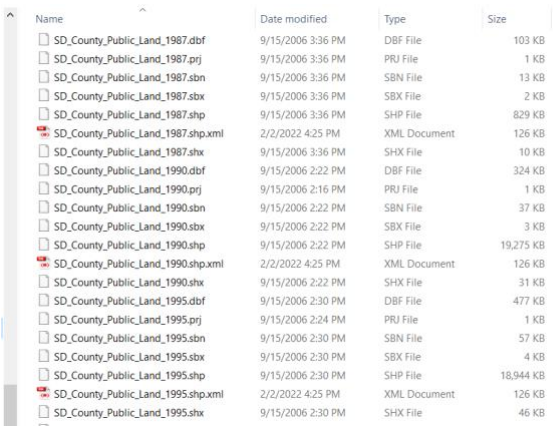
Examples and Screenshots:

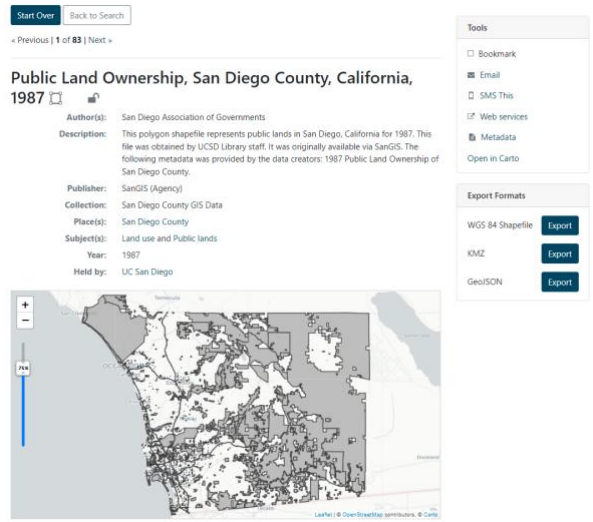
Before - UC Santa Barbara	After - UC Santa Barbara
	
	

Before - UC San Diego



Name	Date modified	Type	Size
A_Datacruncher_Data	3/12/2021 2:29 PM	File folder	
Address_Locators	3/14/2018 9:53 AM	File folder	
Africa_Data	9/3/2020 3:00 PM	File folder	
Asia_Data	9/3/2020 3:01 PM	File folder	
Business_Analyst_Data	9/1/2021 3:55 PM	File folder	
Business_Analyst_Data_2008	10/16/2009 10:53 ...	File folder	
Business_Analyst_Data_2010	2/28/2019 3:50 PM	File folder	
California	4/2/2020 4:48 PM	File folder	
Canada_Data	1/16/2019 2:46 PM	File folder	
ESRI_Data_for_Workbooks	3/22/2022 12:04 PM	File folder	
Europe_Data	1/16/2019 2:51 PM	File folder	
GIS_Tools	5/27/2019 4:01 PM	File folder	
GIS_Workshop_Data	8/29/2019 8:24 AM	File folder	
Global_Data	11/1/2018 1:10 PM	File folder	
Mexico_Data	8/5/2021 2:41 PM	File folder	
North_America_Data	1/16/2019 2:30 PM	File folder	
Oceania_Data	1/16/2019 2:33 PM	File folder	
San_Diego_Data	2/4/2022 3:57 PM	File folder	
South_America_Data	1/16/2019 2:41 PM	File folder	
StreetMap_Premium_NA_2019R2	10/10/2019 12:50 ...	File folder	
StreetMap_Premium_NA_2019R3	11/19/2019 2:48 PM	File folder	
Styles	3/1/2019 9:17 AM	File folder	
Temp	5/29/2020 11:28 A...	File folder	
UCSD_Campus_Data	10/5/2021 4:26 PM	File folder	
UCSD_Course_Specific_Data	5/31/2011 9:47 AM	File folder	
United_States_Data	10/4/2021 3:57 PM	File folder	



Name	Date modified	Type	Size
SD_County_Public_Land_1987.dbf	9/15/2006 3:36 PM	DBF File	103 KB
SD_County_Public_Land_1987.prj	9/15/2006 3:36 PM	PRJ File	1 KB
SD_County_Public_Land_1987.sbn	9/15/2006 3:36 PM	SBN File	13 KB
SD_County_Public_Land_1987.sbx	9/15/2006 3:36 PM	SBX File	2 KB
SD_County_Public_Land_1987.shp	9/15/2006 3:36 PM	SHP File	829 KB
SD_County_Public_Land_1987.shp.xml	2/2/2022 4:25 PM	XML Document	126 KB
SD_County_Public_Land_1987.shx	9/15/2006 3:36 PM	SHX File	10 KB
SD_County_Public_Land_1990.dbf	9/15/2006 2:22 PM	DBF File	324 KB
SD_County_Public_Land_1990.prj	9/15/2006 2:16 PM	PRJ File	1 KB
SD_County_Public_Land_1990.sbn	9/15/2006 2:22 PM	SBN File	37 KB
SD_County_Public_Land_1990.sbx	9/15/2006 2:22 PM	SBX File	3 KB
SD_County_Public_Land_1990.shp	9/15/2006 2:22 PM	SHP File	19,275 KB
SD_County_Public_Land_1990.shp.xml	2/2/2022 4:25 PM	XML Document	126 KB
SD_County_Public_Land_1990.shx	9/15/2006 2:22 PM	SHX File	31 KB
SD_County_Public_Land_1995.dbf	9/15/2006 2:30 PM	DBF File	477 KB
SD_County_Public_Land_1995.prj	9/15/2006 2:24 PM	PRJ File	1 KB
SD_County_Public_Land_1995.sbn	9/15/2006 2:30 PM	SBN File	57 KB
SD_County_Public_Land_1995.sbx	9/15/2006 2:30 PM	SBX File	4 KB
SD_County_Public_Land_1995.shp	9/15/2006 2:30 PM	SHP File	18,944 KB
SD_County_Public_Land_1995.shp.xml	2/2/2022 4:25 PM	XML Document	126 KB
SD_County_Public_Land_1995.shx	9/15/2006 2:30 PM	SHX File	46 KB

After - UC San Diego