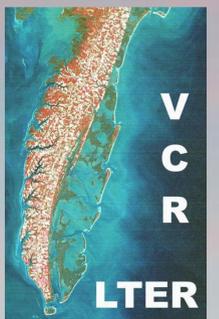


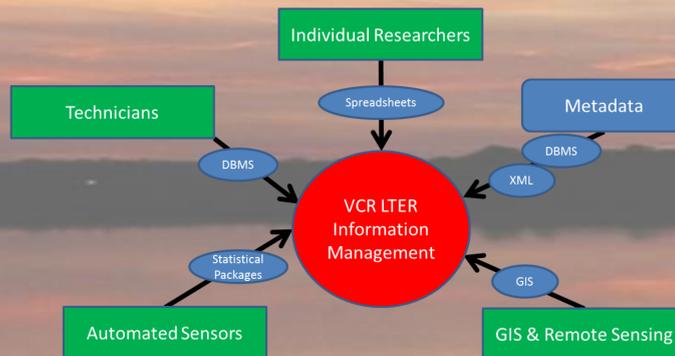
# Information Management at the Virginia Coast Reserve Long-Term Ecological Research Project

John Porter, University of Virginia



**ABSTRACT:**  
The Virginia Coast Reserve Long-Term Ecological Research (VCR/LTER) project was initiated in 1986 with a large multi-institutional set of researchers collecting a wide array of data types. The information system has evolved from a single PC to a multi-server LAMP (Linux, Apache web server, MySQL database, PHP language) system managing nearly 200 datasets and approximately 200 GB of raw data. Technologies used include automated workflows and field wireless networks for streaming data, a web-enabled meta-database which is used to generate Ecological Metadata Language (EML) metadata, an EML-based data catalog, and web services which produce user-ready programs for analysis. Automation is widely employed to allow a relatively small staff (~1 FTE) to manage and update a wide array of data products. Underlying software tools include the Drupal content management system, MySQL databases, and PHP, PERL, Python, SAS, SPSS and R scripts. In addition to data, the VCR/LTER web page hosts a wide array of additional information resources including research descriptions, proposals, personnel and publication lists, images of research sites and activities, and interactive maps. A key to making the system operate has been the selection of technologies that remain stable over long-time periods so that automated applications and workflows require relatively little maintenance. Similarly, archival data is typically maintained as text files (most frequently comma-separated-value files), in order to assure that data will remain readable far into the future.

Data Holdings	
Datasets	188
Data Objects	343
Data Volume	230 GB



The system incorporates a wide variety of data types, each of which will typically require different types of processing.

## Data Archiving

### Data Archived

To increase long-term usability, data are stored, whenever feasible as text files, typically comma-separated-value (CSV) files. For specialized data (GIS, image, sound) data are stored in one or more popular formats that are widely used by scientists.

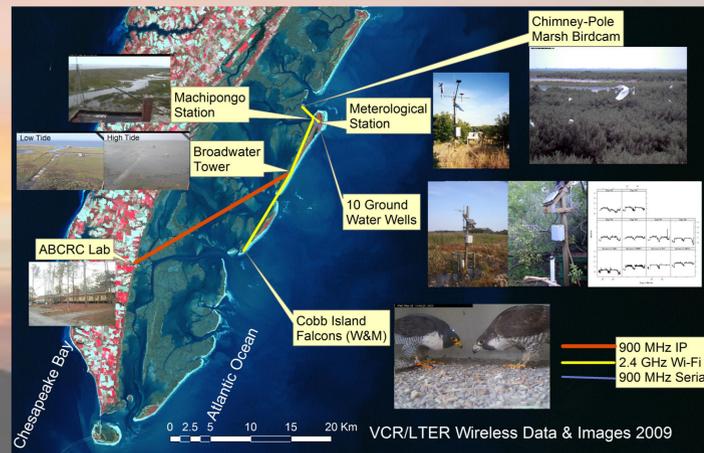
### Long-Term Data

81 datasets have a duration of 5 years or more, 66 of a decade or more, and 37 of 30 years or more. The longest dataset is over 300 years long (Hurricane record).

### Metadata Input

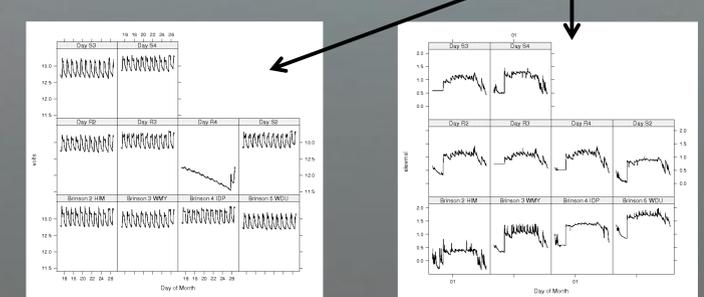
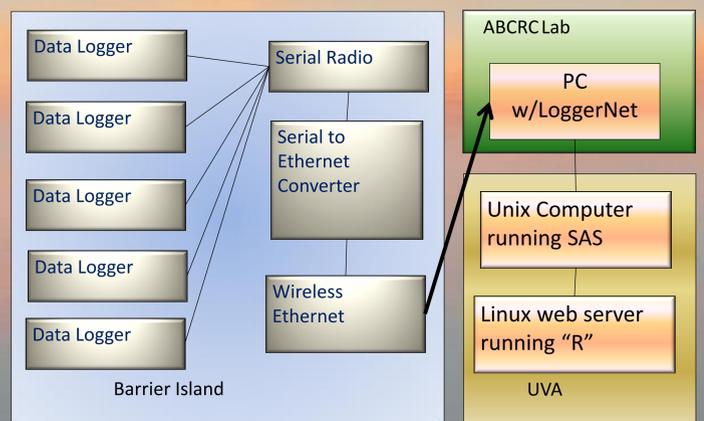
Web forms are used to create a metadatabase, which is subsequently used to generate standard Ecological Metadata Language (EML) metadata.

## Sensors & Workflows



### VCR/LTER Wireless Sensor Network

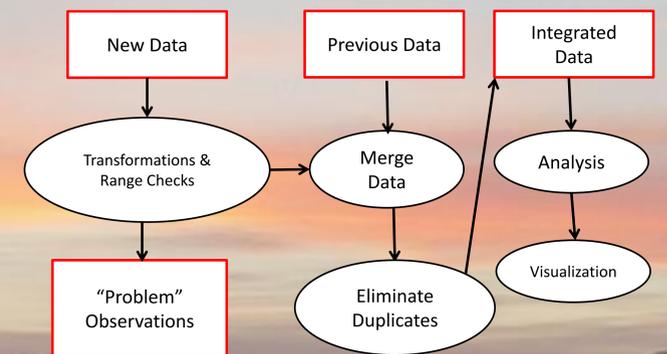
The VCR/LTER Wireless Sensor Network spans over 30 km and includes more than 15 sensor systems.



### Sample Workflow & Products

Ten groundwater wells on Hog Island produce data that is processed to create raw data files and graphical outputs for quality control and quality assurance.

<http://www.VCRLTER.virginia.edu>



Generic Workflow for Sensor Data

## Data Sharing

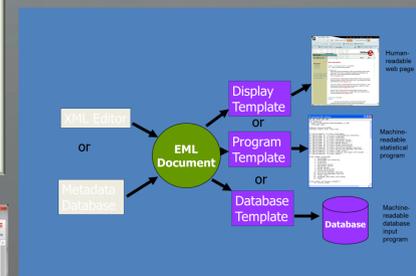
Year	# of Dataset Downloads
2011	860
2012	1,760
2013	1,656
2014	5,375

### Data Sharing

Data is provided via the VCR/LTER data catalog and via the LTER Data Portal (<https://portal.lternet.edu>)

### Code Generation

The VCR/LTER provides web services to translate EML metadata into statistical programs for analyzing data that are used by the LTER Data Portal



### Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1237733. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).