

Skills for data-intensive research from the most basic to advanced, with a subset identified as necessary at all levels and ranging in complexity along the continuum. These skills were identified by a group of experts involved in data-intensive environmental research and its computational support, who convened under the umbrella of the ISEES project, 2013.

<http://isees.nceas.ucsb.edu/workforce/>

On a continuum:	basic:	intermediate:	advanced:
basic command line programming	computational literacy	data archiving	model interoperability
version control	directories and files	basic command line programming	numerical analysis
annotation, metadata, documentation	data wrangling – handling diverse “messy” data	data structures	parallelization - code, hardware
intellectual property understanding	data management best practices	diversity of algorithms	numerical stability
data enabled science skills	fundamental computing architecture	uncertainty analysis/model assumptions / error	verification (code)
	visual literacy	spatial analysis	cloud computing
	units and dimension analysis	exploratory data analysis	object-oriented design
	how to archive data	bug tracking	code that interacts with the web
	inter-disciplinary thinking	open science	algorithms - coding for “big data”
	data-enabled science skills	standards and tools	interoperability (package API)
	collaboration skills	knowledge of public repositories	unit testing
	pseudocode	scientific workflows	advanced metadata
	conceptual modeling	semantics, ontologies, and taxonomies, vocabularies	high performance computing
		software life cycle	hardware knowledge