May the 4th be with you! Happy Star Wars Day!

This graphic also from a recent Symposium on Human Dynamics Research that may be of interest:
A Dark Side to Data-Centric Geography? Where are the Reward Systems?
Panel
AAG Annual Meeting, San Francisco, April 1, 2016 - http://dusk.geo.orst.edu/aag16-darkside.html

Main Sponsors: AAG GIScience & Systems, Cyberinfrastructure, and Remote Sensing Specialty Groups
Related Sponsor: AAG Public/Private Affinity Group

Note that you can rate this session in the AAG meeting app by clicking on the clipboard icon to the left of the session description

Panel will discuss an apparent disconnect in academia where skills in research computing and programming are still not properly rewarded, and what we can do about it. By not properly rewarding computing and coding, publishing of data, etc. in the same way that we reward traditional research, teaching and service, are we unwittingly driving a host of promising researchers away from the academic community?

Dan Goldberg, TAMU: “As someone coming to a love of geography following 3 degrees in computer science (bs, ms, and phd) this is my daily struggle toward tenure in a dept of geography. I routinely need to explain to my colleagues why computing is part of geography and how what I do is not simply widgetology.”
Esri and ODIP

Dawn Wright
Esri Chief Scientist
Affiliate Professor, Oregon State U.
The standard view of GIS but GIS is not longer about just eyeballs on a map, it’s looking at the invisible rubber bands of mathematical manipulation of these different layers. It’s about the coupling of the appropriate data (interoperable), analysis and manipulation, and compelling design to effectively communicate the results, and when appropriate, the possible models and future scenarios.
Strengthen science base at Esri … Establish Esri as a *member* of the scientific community …
Across the entire Esri organization we launched an ocean GIS initiative in 2012, which is strong, ongoing, and including a comprehensive strategic plan, a new oceans/maritime resource center, and we’re working with a growing list of terrific partners.
Esri part of a data alliance with Amazon and Microsoft within the NOAA Big Data project.
We at Esri can do science too, and these are some of the recent contributions that we’ve made TO science.
Clockwise from upper left (based on a slide prepared for the FedGIS conference in Washington, DC, hence the call-outs to US federal government agencies):

**ELU – Ecological Land Units**
The Ecological Land Units (ELU) project is a collaboration between Esri and Dr. Roger Sayre of the USGS. It was officially launched last December at the ACES 2014: A Community on Ecosystem Services meeting in Washington, DC. For the ELU, we essentially undertook a massive biophysical stratification of the planet at a finest yet-attempted spatial resolution (250 m) to produce a first ever map of distinct physical environments and their associated land cover. We also offer a concept for delineating ecologically meaningful regions that is essentially both classification-neutral and data-driven. Our intent is to provide scientific support for planning and management (including as an important variable for GIS geodesign models and apps), and to enable understanding of impacts to ecosystems from climate change and other disturbances; hence for valuation of ecosystem SERVICES. In this way, we also offer fulfillment of one of the main recommendations of the White House PCAST report on sustainable environmental capital.

**EMU – Ecological Marine Units**
Following on from the ELU project, Esri is collaborating with over 10 partners including the USGS, NOAA, and NatureServe, in developing a standardized, robust, and practical ecosystems classification and map for all the world’s oceans, completely in 3D. Officially commissioned by The Group on Earth Observations (GEO). To inform MPA design, marine spatial planning, biodiversity observation and conservation.
Group on Earth Observations MOU
One of the most comprehensive efforts in place to monitor the entire face of the Earth is a group of over 140 governments and leading international organizations (GEO) seeking to establish a fully functioning Global Earth Observation System of Systems (GEOSS). Esri is working with the National Research Council of Italy to integrate public scientific content from ArcGIS Online into the GEOSS data ecosystem, while encouraging the Esri GIS community to participate as contributors and users of GEOSS.

Geoss.maps.arcgis.com

Improved Cloud Access to Imagery: LERC/MRF
Esri and NASA have collaborated to improve access to imagery and raster data stored in the cloud using a combination of two technologies, Meta Raster Format (MRF) and Limited Error Raster Compression (LERC). MRF is an Open raster format originally designed at the NASA JPL to optimize web access to rasters. LERC is a highly efficient algorithm that provides fast lossless and controlled lossy compression of image and raster data, and is especially suitable for geospatial applications. Esri was recently awarded a US Patent for LERC. LERC is integrated into NASA’s MRF support within GDAL (Geographic Data Abstraction Library), one of the most widely used open source projects in our industry. Esri’s LERC and its integration into MRF will enable our users to significantly reduce storage costs for enterprise image management solutions, and it will likely find its way into virtually every mapping/GIS application stack in use. Both are open sourced under an Apache 2 license.


ALSO OF NOTE REGARDING HOW Esri has improved its engagement with NASA: Esri helped with NASA JPL wizard Hook Hua’s review/feedback on his Advanced Rapid Imaging & Analysis for Monitoring Hazards (ARIA-MH) testbed project for the Federation of Earth Science Information Partners (ESIP). Esri has released its Spatio-temporal Big Data Store with the 10.4 release for managing large collections, ArcGIS now supports SAP HANA as an enterprise GDB, and Esri is set to release the GeoAnalytics extension for parallelizing processing of big data later this year. NASA still has plenty of passes to UC in San Diego at the end of June through the ELA, so Tripp Corbett has invited any NASA PIs who would like to attend on the Big Data track.

Open Water Data Initiative
Esri is both a participant and stakeholder in the Open Water Data Initiative, as well as a partner in the National Flood Interoperability Experiment, developing spatial data processing and visualization tools. These help Federal agencies make the data as useful and accessible as possible (e.g., for other Federal agencies, state and local gov’t, water mgmt. authorities, agriculture sector, insurance sector, investment sector, etc).

The National Water Model is a new project that will be released by NOAA in June. It is a project to forecast stream flow (cubic feet per second of water in streams). Currently
they do this at about 3600 locations around the country. This new project will forecast at 2.7 million locations, a 700x increase. How they get from 3600 to 2.7 million is a combination of GIS and modeling which Esri helped develop with NOAA, NCAR, Univ of Texas, and others last year. The GP tools for that will be available on GitHub soon. Esri is collaborating with NOAA to support them with this, led by Jeff Donze, tracked by Steve Kopp.

**R – ArcGIS Community**

R (aka the R Project for Statistical Computing) is the world’s fastest growing environment for statistical computing. We are working with ArcGIS and R users worldwide to develop a community to promote learning, sharing, and collaboration. This community will include a repository of free, open source, R scripts, geoprocessing tools, and tutorials.


R-arcgis.github.io

**A whole suite of apps for Citizen Science – esriurl.com/citizens**

**Scientific Python**

To further strengthen the link between GIS and science, we have fully integrated ArcGIS with SciPy, a Python-based ecosystem of open-source software for mathematics, science, and engineering. By integrating the ENTIRE STACK of SciPy modules with ArcGIS we have made developing scientific and technical geoprocessing tools and scripts easier and more efficient.


**Three New Science Books from Esri Press**

Features the peer-reviewed work of leading researchers and managers for advancing science through GIS, as well as how to incorporate spatial thinking and GIS technology into research design and analysis.

esripress.esri.com and esriurl.com/ocnso1ns
Welcome to the R – ArcGIS Community

Combine the power of ArcGIS and R to solve your spatial problems.

The R – ArcGIS Community is a community-driven collection of R, open-source projects making it easier and faster for R users to work with ArcGIS data, and ArcGIS users to leverage the analytics capabilities of R.

**r-bridge-install**

**Python**

Install the R-ArcGIS Tools

![GitHub](https://raw.githubusercontent.com/ArcGIS/r-bridge/master/R-bridge/)

---

**r-bridge**

**C++**

Bridge library to connect ArcGIS and R, including `arcgisForR` library

![GitHub](https://raw.githubusercontent.com/ArcGIS/r-bridge/master/R-bridge/)

---

**r-sample-tools**

**R**

Sample tools illustrating R usage in geoprocessing scripts

![GitHub](https://raw.githubusercontent.com/ArcGIS/r-bridge/master/R-bridge/)
Analytics....
We aim to register large data stores / data sets with ArcGIS Server, then distribute analysis across a cluster of machines for parallel processing.

**Performance example:** buffer 8.2 million points or thousands of polygons in a little over a minute,

Coming: ~250,000 writes to disk per second across 5 nodes.

Many frameworks/technologies exist for distributing computation:

- E.g., Hadoop, MapReduce, Spark
  - **Spark:** processes distributed data in memory; Supports MapReduce programming model
  - Includes additional framework level distributed algorithms

ArcGIS Server integrates these technologies on a cluster to solve analytic problems.
Includes scripts, a working toolbox, sample data, and step-by-step workflows in Word and pdf form to help scientists work with a variety of multidimensional datasets in GIS, including netCDF.
Just Making Data & Code “Available” NOT Enough!
Share workflows and use cases

Landsat 8 Script

This is a ZIP archive containing Python scripts and documentation for building and configuring mosaic datasets containing imagery from the Landsat 8 satellite. This script requires ArcGIS Desktop 10.2, and must also have installed the Landsat 8 patch available for download here [link]. Data from Landsat 8 can be downloaded using CoLDR or EarthExplorer.

esriurl.com/landsat8
The Living Atlas of the World is presented as a use case of a data framework seeking to cut effectively across multiple spatial and temporal scales.
The Living Atlas of the World is presented as a use case of a data framework seeking to cut effectively across multiple spatial and temporal scales. The Living Atlas was first created in 2014 to make authoritative geographic information accessible via hosted cloud services so that users could more quickly address scientific and societal problems and decisions at spatial scales ranging from a small study area to the entire globe, while using a range of interactive map functions to tell engaging narratives along the way (aka “story maps”).
Marine Conservation Institute layers based on data in PLOSOne paper on Global Habitat Suitability for Cold Water Corals
EMU explorer
http://chrismahlke.github.io/emu/
Username: emuuser
Password: 24clusters
Interoperability....
Along these lines we certainly want to engender interoperability and crosswalking among approaches, such as the examples here with the **Python Scientific Computing Environment**, or simple integration with a host of scientific tools and libraries.
Sensor Web

With standardisation:
Standards need to be implemented only once for each source and consumer
⇒ less integration efforts needed
IMPORTANT SLIDE!!!!
Including over 200 imagery or sensors!!
And we are not dictating on providing open data access
The basis for geospatial literacy
Twitter reference: https://twitter.com/heidiklaine/status/704135333034590208

Providing an Open Platform for Collaboration and Innovation

These are the 6 pillars of “open science” advocated around the world.
ODIP Prototypes 1, 3
netCDF standardization
big data, cloud services
workflows
Thank You
esriurl.com/scicomm and esriurl.com/oceandata

Dawn Wright
dwright@esri.com
Twitter: @deepseadawn
Parallelized batch analytics on tabular, vector, raster, and imagery datasets (big and standard data)

**Performance example:** buffer 8.2 million points or thousands of polygons in a little over a minute,

**NetCDF output is supported on …**

- **Aggregate Points** – Point in polygon aggregation (spatial join)
- **Aggregate by Cell** – simple aggregation
- **Summarize Nearby** – spatial join where left dataset and a distance are used to inflate the bounding box. A distance calculation is then employed (no drivetime/distance).
- **Summarize Within** – a spatial join that is a generalization of Aggregate Points
- **Find Existing Locations** – an attribute query followed by/with a spatial query (?); the attribute portion is SparkSQL, the spatial relationship is: intersections, within distance, contains, and within.
- **Find Similar Locations** – non-spatial similarity search (SparkSQL).
- **Calculate Density** – cell-based spatial aggregation by attribute with a weighted neighborhood/distance-based summarization.
- **Find Hot Spots** – very similar to Calculate Density but with the Getis-Ord Gi* statistic. This can be cell-based (point observations) or polygon-based (polygon source data).
- **Create Buffers** – simple spatial buffer generation across the dataset (the Java geometry library supports Euclidean and Geodetic buffer generation). This is done.
- **Extract Data** – a simple region query with a specified output format.
- **Field (RASTER) Calculator** – apply a Scala or SQL expression to each feature, one expression per field.