FACE-IT: Earth science workflows made easy with Globus and Galaxy technologies
(Provide more capability for more people at lower cost by delivering “Science as a Service”)

Raffaele Montella\textsuperscript{1,2}, Alison Brizius\textsuperscript{2}, Joshua Elliott\textsuperscript{2}, David Kelly\textsuperscript{2}, Ravi Madduri\textsuperscript{2,3}, Ketan Maheshwari\textsuperscript{3}, Cheryl Porter\textsuperscript{4}, Peter Vilter\textsuperscript{2}, Michael Wilde\textsuperscript{2}, Wei Xiong\textsuperscript{4}, Meng Zhang\textsuperscript{4} and Ian Foster\textsuperscript{2,3,5}

\textsuperscript{1}Department of Science and Technologies, University of Naples Parthenope, Naples, ITALY;
\textsuperscript{2}Computation Institute, Argonne National Laboratory and University of Chicago, Chicago, Illinois, USA;
\textsuperscript{3}Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, USA;
\textsuperscript{4}University of Florida, Department of Agricultural and Biological Engineering, Gainsville, Florida, USA;
\textsuperscript{5}Department of Computer Science, University of Chicago, Chicago, Illinois, USA;

\texttt{faceit-portal.org} \hspace{1cm} \texttt{usefaceit.org}
Facing real problems with Information Technology

...less magic wands, more magic minds...

What’s in a name...
No buzzword
Real things!
An open playground for the next generation of earth system scientists

The user profile...
Scientists
Experts of their fields
Limited programming skills
Complex experiments

Data + Workflows = Results

Effective and efficient solutions to real problems
Experts in design and abstraction

Built on widely used Galaxy, Globus, and Swift systems

Development-experts (in wizardry)...

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology

faceit-portal.org
People create services (data or functions) ... which I discover (& decide whether to use) ... & compose to create a new function ... & then publish as a new service.

→ I find “someone else” to host services, so I don’t have to become an expert in operating services & computers!
→ I hope that this “someone else” can manage security, reliability, scalability, ...
### The Science Stack

<table>
<thead>
<tr>
<th></th>
<th><strong>SaaS</strong></th>
<th><strong>PaaS</strong></th>
<th><strong>IaaS</strong></th>
</tr>
</thead>
</table>
| **Galaxy** | - Interactive execution  
- Creation, Execution, Sharing, Discovering Workflows | | |
| **Globus** | - Data management  
- Identity Management | | |
| **AWS** | - HTCondor, Chef, EC2, EBS, S3, SNS, NEWT  
- Spot, Route 53, Cloud Formation | | |

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
Teaching Galaxy to speak Earth Science

- Datatypes
- Tools
- Tool parameters
- Aggregated datatypes
- Data providers
- Visualizers

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
Teaching Galaxy to speak Earth Science

Step ONE of 6: earth system datatypes

- Color scheme

- Regular Galaxy
- New datatypes for earth sciences
- Extensions of general datatypes
- FACE-IT applications datatypes
- Other FACE-IT apps datatypes
- data.Data
- data.Text
- tabular.Tabular
- xml.Generic_Xml
- netcdf.NetCDF
- generic_json.GenericJson
- acmo.Acmo
- netcdf.NCML
- enhanced_xml.EnhancedXml
- geojson.GeoJson
- geojson.CrsJson
- geojson.GeometryJson
- gcm.GCM
- ecmwfera40.EMWFERA40
- enhanced_xml.DsRef (RAFT)

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
Teaching Galaxy to speak Earth Science ("Globusally")

Step TWO of 6: new tools

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
**Tool parameters:**
Define the user interface elements for a tool

- Regular tool parameters wrap text fields, radio buttons and drop down lists.
- Custom tool parameters for Globus, OpenDap, date peaking and feature selection on maps.

**Teaching Galaxy to speak Earth Science**

---

**FACE-IT:** A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
**Dataset References:**
XML based datatype grouping references to different datasets in the same history.

- The regular Galaxy works on single file datasets or composite file datasets.
- Acts as a ‘struct’ or an ‘array’ or a mix of both.
- Supports schemas and translators.

**Used when:**
- A tool consumes and/or produces a variable number of datasets
- The tool is implemented using a Swift script working in parallel

**DsRef (EnhancedXML)**

**Teaching Galaxy to speak Earth Science**

**Step FOUR of 6:**
aggregated datatypes (RAFT*)
**Data providers:**
software components interfacing the datasets with the web browser.

- They provide data as array of JSON objects
- Key/Values, Columnar, custom
- Implemented in Datatype classes

**Web Browser**

**Data Providers**

*Web page...*  
...dynamically generated...

*.form Mako template*  
(mix of server side python code with client side web technologies)

**Galaxy Instance**
**Teaching Galaxy to speak Earth Science**

**Step SIX of 6 (!): map visualizers**

- **Visualizers**: client-side software components for interactive data visualization

- **Quasi-GIS**

- **Map**: Visualizes vector data produced as GeoJson objects by a data provider

- **Wms** (World Map Server): Visualizes raster data from NetCDF datatypes.
• wrf-model.org
• +30k people community
• 150 countries
• high computing demanding

Weather Research and Forecast @AWS: a (real) application.

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
Weather Research and Forecast @AWS

Weather Forecast on Chicago area (333m)

- WRF
- 4 nested domains
- 9km to 333m

FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology
FACE-IT: A Framework to Advance Climate, Economic, and Impact Investigations with Information Technology

- NetCDF Scavenger
- Indexing / Ingestion
- Tiling
- Spatial / metadata query
- Selection
- Usage

NetCDF scavenging, discovery and provisioning

Breaking the fourth wall

WORK IN PROGRESS
Face-IT Galaxy is a creative playground for the next generation of earth scientists powered by Globus for data movement and more.

http://www.faceit-portal.org

Propose your application, write your code and share it!

Spin-off projects: extreme weather simulations in the Bay of Napoli, IT (UniParthenope)