What is PACE

Protected Analytics Computing Environment (PACE) is a highly protected network space where research users with approval can analyze and work with identifiable protected health information. The product will provide an individual workspace, “virtual laptop”, that is secure and compliant, yet flexible enough for researchers and educators. PACE replaces existing analytics services that range from personal computers to departmentally funded clusters.

Why PACE

Duke’s electronic health data and associated financial information are valuable assets desired by researchers and educators; however, obtaining sensitive data (from DEDUCE and other sources) and establishing a protected environment to perform analysis can often be a very difficult and time consuming process. As a result, these communities lose precious time and resources putting all the critical pieces together. PACE simplifies the effort of obtaining EHR data (Electronic Health Record - data from Duke Health enterprise data warehouse and Duke's MAESTRO Care (Epic-based) EHR system), creating a protected computing environment while providing access to rich computing resources for analytics.

Value

• A protected environment to interrogate sensitive data which minimizes risk of accidental disclosure
• An analytics platform with tools to perform the most common analyses that can be accessed from anywhere
• A platform that strengthens provenance via Duke Data Service and provides governance over data extraction (honest broker)
• A workspace that can be accessed by sponsored external researchers with Approved NetIDs

Capabilities

• Analytic and Statistical Capabilities Phase 1:
  • Standard Statistical Functions, Spatial Statistics
  • Static Graphic and Tabular Reporting
  • Duke Data Access to DEDUCE, REDCap
  • Modern Code Development Environment
• Analytic and Statistical Capabilities Phase 2:
  • Data Mining, Natural Language Processing
  • In-Memory High Performance Analytics, Visualization, and Machine Learning

Tools

• Analytic and Statistical Tools Phase 1:
  • Workstation class compute capacity
  • SAS, R, Python, ArcGIS
  • Jupyter, R Studio, SAS Studio, GIT
  • DEDUCE Data Query Tool
• Analytic and Statistical Tools Phase 2:
  • Enterprise Cluster Based capacity
  • SAS Viya, Apache SPARK, Microsoft R Open
  • Greenplum, NoSQL Systems, Oracle

For More Information Contact: PACE_Info@duke.edu
PACE Community

- PI or delegate can approve team member's access
- Team members will have individual Virtual Machines (Workspaces) to conduct research
- Researchers will have access to sensitive data via DEDUCE or other data sources
- From workspaces, researchers can use a variety of technology resources such as SAS Grid, and a rich set of analytical applications
- Software modules, code libraries, and GIS (ArcGIS) data can be pulled into workspaces from the Internet
- Finally, data will leave PACE through the engagement of an Honest Broker

For More Information Contact: PACE_Info@duke.edu
Key Stakeholders

Dr. William J. Fulkerson MD  Dr. Iain C. Sanderson, MD, BM, BCh
Dr. Jeffrey M. Ferranti MD  Dr. William (Billy) Willis
Dr. Larry Carin  Dr. Lesley Huntley Curtis
Dr. L. Ebony Boulware, MD  Dr. Stephen Blackwelder
Dr. Geeta Swamy, M.D.  Mr. Billy Newton, Jr.

Funding

This program is funded by Duke’s Clinical and Translational Science Award (CTSA) from the NIH and by Duke University Health System. The Health System’s primary interest in this project is the broad use of Duke Health System’s data for research and education coupled with analytics resources and enterprise class protection of data.

PACE Services

In short, PACE is a marketplace with upcoming services that will include:

<table>
<thead>
<tr>
<th>Service</th>
<th>Anticipate Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>OIT Containers via Singularity:</td>
<td>Currently Available</td>
</tr>
<tr>
<td>Whitelisted Resources</td>
<td>Currently Available</td>
</tr>
<tr>
<td>OIT GPU/Compute Resources</td>
<td>Currently Available</td>
</tr>
<tr>
<td>Oracle Exadata</td>
<td>Currently Available</td>
</tr>
<tr>
<td>PACE Data</td>
<td>Currently Available</td>
</tr>
<tr>
<td>Honest Broker</td>
<td>Currently Available</td>
</tr>
<tr>
<td>Isilon Storage</td>
<td>Currently Available</td>
</tr>
<tr>
<td>GitLab</td>
<td>Currently Available</td>
</tr>
<tr>
<td>Shib/DUO</td>
<td>Currently Available</td>
</tr>
<tr>
<td>OIT Containers Repository (Singularity/Docker)</td>
<td>2/1/2018</td>
</tr>
<tr>
<td>Microsoft Azure</td>
<td>2/1/2018</td>
</tr>
<tr>
<td>Radiology API (DICOM)</td>
<td>2/1/2018</td>
</tr>
<tr>
<td>SAS Grid</td>
<td>6/4/2018</td>
</tr>
<tr>
<td>Amazon WebServices</td>
<td>8/1/2018</td>
</tr>
<tr>
<td>Google</td>
<td>2/1/2019</td>
</tr>
</tbody>
</table>

For More Information Contact: PACE_Info@duke.edu