Research IT Day 2018 - Session Descriptions

8:00 - 8:45	**Networking and Breakfast**			
8:45 - 9:00	**Welcome (Mendola)**			
9:00 - 9:45	2 sessions Tableau Training 9:00 - 10:15 and 10:30 - 11:45	Intro to Amazon Web Services (Circe Tsui)	REDCap Project Design to Deployment (Sean Mann)	Demos, Tables, Snacks, and Drinks
10:00 - 10:45		AWS Elastic Compute Cloud (EC2), Spot Instances, HPC, and High Throughput (James Anderson)	Protecting your Data (Brad Sanford)	
11:00 - 11:45		AWS Machine Learning and Big Data analysis (Vadim Omeltchenko)	Dimensions - Funding Intelligence for Research Organizations (Tom Lickiss)	
Noon - 1:00	**Lunch**			
1:00 - 1:45	AWS Training	Using R in an AWS Environment - an interactive demo (Circe Tsui)	Clinical data for research: Children, EHC, Grady, and the VA (Jeff Weaver)	Demos, Tables, Snacks, and Drinks
2:00 - 2:45		Intro to Tableau (Tyler Morris / Patrick Maloney)	Data Management Plans (Jen Doty / Jeremy Kupsco)	
3:00 - 3:45		Intro to LabKey (Wayne Harris)	3D Visualizations in Research (Arya Basu)	

- **Introduction to Emory AWS Service – Circe Tsui (9:00 – 9:45)**
  - Learn how Amazon Web Services (AWS) can expand the options for analysis in your research. AWS is a cloud services platform, offering compute power, storage, content delivery and other functionality. A basic overview of AWS services will be given, successful use cases discussed, and an R Studio analysis will be demonstrated.

- **AWS Elastic Compute Cloud (EC2), Spot Instances, HPC, and High Throughput - James Anderson (AWS) (10:00 – 10:45)**
  - Access to unprecedented compute power and storage are drivers for cloud adoption, but the learning curve to use cloud-based services can be steep, and costly, if you don't plan accordingly. Learn the difference between on-demand, spot, and reserved instances at AWS and when to use each to optimize costs. Also, learn how Clemson University has leveraged Amazon Web Services to develop a Dynamic AWS High Performance Compute (HPC) Cluster, to provide easy access to scalable HPC compute capacity to process large workloads for students and faculty.

- **AWS Machine Learning and Big Data analysis - Vadim Omeltchenko (AWS) (11:00 – 11:45)**
  - AWS enables researchers to build data lakes (genomic, health care, social, etc) on top of Amazon Simple Storage Service (S3) where information from multiple disparate sources can be merged. Users can search, retrieve and analyze data while decoupling storage from compute. Joining, exploring and mining large datasets is made possible by the AWS Big Data services such as Athena or Elastic MapReduce (EMR). Customers can also use the Artificial Intelligence and Machine Learning

(AI/ML) services offered by AWS to build, train and host models with all levels of sophistication – from regression and classifiers, to deep neural networks.

- **Using R in an AWS environment - an interactive demo - Circe Tsui (1:00 – 1:45)**
    - Learn how to use AWS for R analysis. The session will provide an online demo of analyzing data with an R server.

- **Introduction to Tableau – Tyler Morris / Patrick Maloney (2:00 – 2:45)**
    - Interested in how Tableau can be used in your role? Please join us for a brief overview of how the platform works, what makes it unique, and how you can get started. Tableau will discuss how they are helping researchers better see and understand their data, as well as a couple specific use cases from your peers at Emory.

- **Introduction to Labkey – Wayne Harris (3:00 – 3:45)**
    - Learn about the LabKey solution to organize, query, review and share your lab's data

- **REDCap Project Design to Deployment – Sean Mann (9:00 – 9:45)**
    - Please join us for an interactive presentation that will walk you through the initial project design to deployment of a production ready research data collection tool. The session will cover: traditional and longitudinal studies; repeatable instruments and events; survey distribution; data access groups for multi-site research; project design method.

- **Protect Your Data – Brad Sanford (10:00 – 10:45)**
    - Learn about new and emerging threats and how best to protect your data from Emory University's Chief Information Security Officer

- **Dimensions - Funding Intelligence for Research Organizations – Tom Lickiss (11:00 – 11:45)**
    - Dimensions is a new data source available to Emory that contains data on publications, grants, patents, and clinical trials. In this session, we will explore ways to leverage this data to find collaborators with a similar scope of interest and analyze the global funding landscape in a detailed fashion.

- **Clinical Data for Research: Children, EHC, Grady, and the VA – Jeff Weaver (1:00 – 1:45)**
    - We will be focusing on the types of data that exists in the Emory Clinical Data warehouse and the process for requesting data for individual

projects. We also will give some examples of projects we have completed. Representatives from Children's Healthcare of Atlanta and Grady Health Systems will be there to help answer questions.

- **Data Management Plans – Jennifer Doty and Jeremy Kupsco (2:00 – 2:45)**
    - Does your grant proposal require a data management plan? Learn about resources to help you craft a comprehensive and realistic plan for managing your research data from soup to nuts.

- **3D Visualization in Research – Ayra Basu (3:00 – 3:45)**
    - Our discussion in this presentation will primarily revolve around the emerging trends in our community and its relevant applications in pedagogy including but not limited to Virtual Reality, Augmented Reality, etc.