

Research data analysis: tools for SU researchers

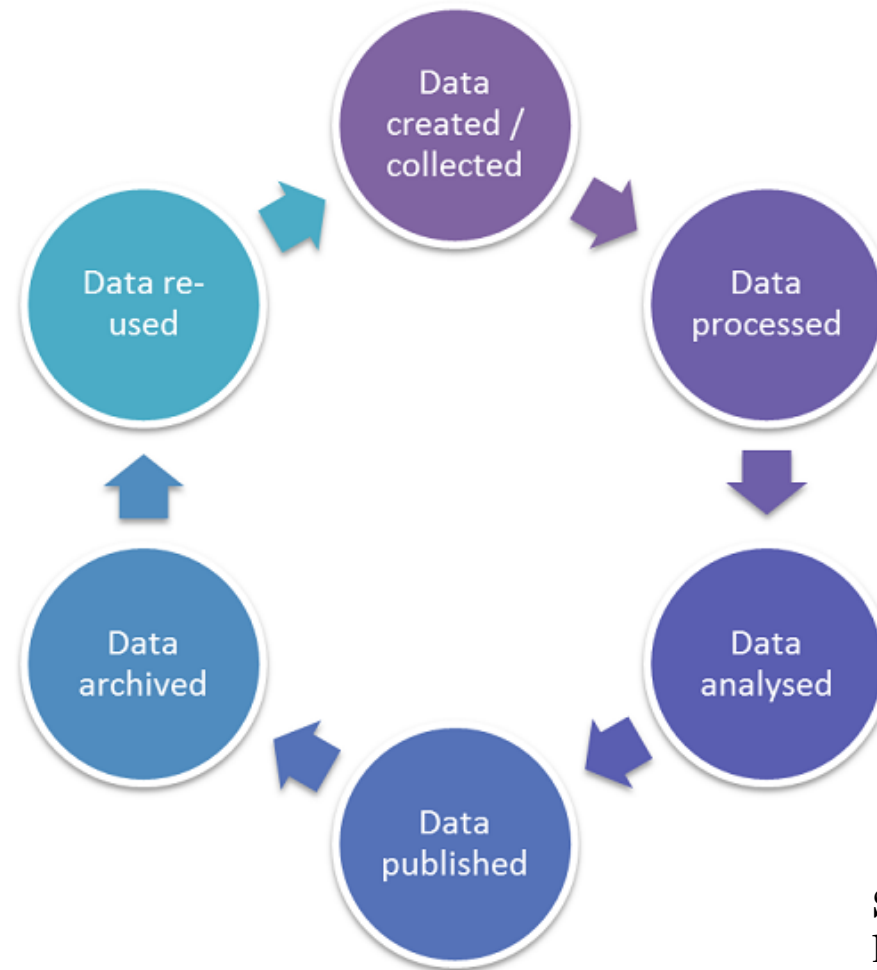
Hilda Kruger

SU IT

<http://infoteg.sun.ac.za/>

hkruger@sun.ac.za

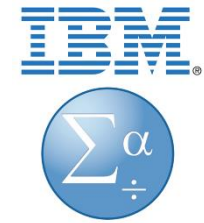
ICT for Research



Source:
<http://www.lancaster.ac.uk/library/rdm/wat-is-rdm/plan/data-lifecycle/>

Campus licenses

- Statistica, SPSS, SAS, Mathematica, Matlab
- Enquiries: software@sun.ac.za
- Requests for installation: help@sun.ac.za
- IT Hub
- Computer User Areas
- Matlab: <http://rga.sun.ac.za/firga/matlab.html>
 - CoP

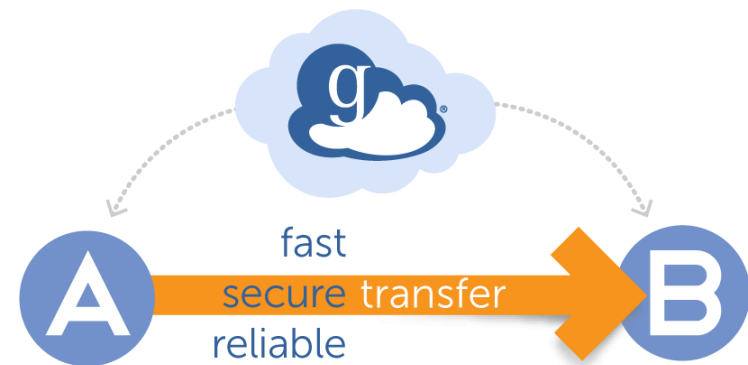


HPC and Linux systems

- Enquiries: Charl Möller | System Administrator: HPC and Linux systems | cwmoller@sun.ac.za
- “Help make it work”
- ± 250 tools
- R (Statistics, *omics), ADF (Computational Chemistry), Altair Hyperworks (Mechanical Engineering), ANSYS FLUENT (Mechanical Engineering), Gaussian (Computational Chemistry), Galaxy (*omics), SPAdes (*omics), Velvet (*omics)
- Training: classroom or bespoke

HPC and Linux systems

- Guidance
 1. log in to HPC front-end
 2. upload data
 3. develop analysis scripts
 4. submit jobs
 5. execute activities in the pipeline
 6. monitor
 7. download results
 8. post-process/visualise
- Centre for High Performance Computing (CHPC)
 - good practice
 - debugging/troubleshooting license and transfer issues
 - <https://www.chpc.ac.za/>
- Globus PoC!



Microsoft

- SQL Server R Services
- Microsoft SQL Server Analysis Services
- Predictive Analytics with Microsoft Azure Machine Learning
- Enquiries: help@sun.ac.za
















Skills development

- 2016: software carpentry workshop
 - Python/Shell/Git
 - To-do: <http://openrefine.org/>
 - Contact: hkruger@sun.ac.za
 - Also: ADA - <http://www0.sun.ac.za/ada/>



<http://www.datacarpentry.org/lessons/>

Lessons

Lesson	Site	Repository	Instructor Guide
Data Organization in Spreadsheets			
Data Cleaning with OpenRefine			
Data Management with SQL			
Data Analysis and Visualization in R			
Data Analysis and Visualization in Python			

Thank you for your time!

<http://infoteg.sun.ac.za/>