Challenges and opportunities in timely and efficient delivery of IT for eResearch projects

- **Lahiru Ariyasinghe** and David Eyers
- eResearch NZ 2020

Motivation for this presentation

- The University of Otago is undergoing an internal restructuring that includes IT and eResearch support

- UoO should enjoy improved eResearch potential
  - … but there are risks involving complex stakeholder groups
  - Many departments are losing local technical support
  - Transitions so far have emphasised capability over timeliness

- We want to share thoughts on opportunities and risks

An approximate definition of 'eResearch'

eResearch covers a spectrum spanning two extremes:

- **Big ticket items**—direct impact; narrow scope
  - Equipment—e.g., HPC; large-scale storage; custom hardware
  - Funded projects—e.g., HPC provisioned in a Marsden Grant
  - 'Superstar' researchers—justifying support expenses is easy

- **Overall capability lift**—indirect impact; broad scope
  - The Carpentries (Software, Data, Library)
  - Efficiency of research management software and systems

Challenges seen in UoO eResearch projects

- Challenges involving funding:
  - Projects may face **costs** outside the scope of research grants
  - Capital expenditure versus operating expenditure
    - DIY clusters (**CapEx**) versus cloud computing (**OpEx**)

- Challenges involving personal expectations:
  - Academics and professional staff have very different routines
    - Academics may underestimate **technical difficulty**
    - Professional staff may overestimate **academic responsiveness**

Challenges seen in UoO eResearch projects

- Project management challenges:
  - Estimation of **projects' timelines**
  - Estimation of **response times** for resolving IT issues encountered
  - **Tracking issues** across different teams—e.g., across departmental and central IT, researchers, NRENs, etc.
  - **Prioritisation** and opportunity costs—e.g. the mechanisms that can support escalation of issues in an efficient manner

- **Sustainability**: support after grant funding finishes…
  - Different severities, e.g., no maintenance versus vanishing

Developments: cloud computing

- **Cloud computing** can perform eResearch off campus
  - University infrastructure and IT not on researchers' critical path
    - University IT + NRENs can provide cloud tenancies
    - University IT cloud bursting hybrid services an even better option
  - Technology on offer is increasingly HPC-like: InfiniBand, GPUs…
  - Generous grants for academic research and teaching

- There are plenty of technical challenges and risks:
  - **Data migration**—transfer speeds / latency
  - **Jurisdictional concerts**—data sovereignty

Developments: desktop supercomputing

- Cloud is centralising but other trends are decentralising
  - Has been a perpetual pendulum in terms of technology

- GPGPUs: Machine learning / AI workloads
  - Gravitating to devices on researchers' computers
  - Fiddly to manage as HPC, since devices don't share well
  - May be freely available to academics (!)

- Data processing: shift to streaming data processing

Developments: DevOps tooling

- DevOps tooling for reproducible research: e.g.,
  - **Containers** (Docker, Singularity, etc.) capture software envs.
  - Researchers take a central role in software they develop / use
  - **Infrastructure as Code** (IaC) helps snapshot deployment needs

- Difficulties: ridiculously high pace of change in tools and only emerging methodologies for sustainability
  - Promise: e.g., Kubernetes leading container orchestration tools

Developments: delegated administration

- More university systems are offering self-service
  - (Still pales in comparison to workflows online in the cloud.)
  - …arguably compliance within Uni. sector uniquely challenging
    - Difficult to manage information that staff need to know
    - Staff often have high autonomy: hard to train and manage

- Better availability of APIs is emerging
  - Commercial vendors pressed by common practice
  - … although many older University systems lack APIs

Research Software Engineers (RSEs)

- RSEs span University IT and research groups
  - Key to embed staff to ensure responsive project progress
    - Documenting all requests to central org. unit? Slow and expensive!
  - Needs a career path: not a rolling sequence of contracts
    - Currently do not have an official job title or role at UoO
    - High potential for UoO research groups to pool resources
    - Existing UoO 'RSEs' are often successful postdoc researchers too

- For efficiency, need to trust RSEs, and IT-authorise them
  - … although given limited central IT resources, how to prioritise?

Assessing potential eResearch success

- eResearch resources (all types) are limited:
  - thus want evidence-based, accountable resource allocation
    - … and appreciate that approach will evolve dynamically
  - Centralisation helps enumerate projects
    - … but understanding their edges is difficult—complex environment

- Need shared methodologies for cost estimations:
  - Weight against estimated impact?
  - Points system to prioritise and effect sharing?
    - e.g., to balance peaks in need against average needs

Changes at University of Otago

- UoO ITS gaining a new unit for research & teaching IT
  - Planned to link to existing UoO eResearch advocacy groups
  - … plus push to centrally coordinate dept. research IT

- UoO University Library's Research Support Unit
  - changing role of libraries within university research

- UoO ITS investigating cloud tenancies
  - Research and teaching can be added to admin transitions

Conclusion … to be continued?

- IT restructuring at UoO will affect eResearch activities:
  - combined with Cloud + DevOps tools + RSEs, future is bright
  - … but there are risks to timely delivery of IT needs to RSEs

- Some UoO shadow IT is being moved into central ITS
  - … but original need must be addressed or shadow IT will return

- Feedback and other thoughts are invited!