CSIRO’s eResearch Service
Provisioning in 2019

CSIRO Scientific Computing

Justin Baker, John Zic | October 2019
Australia’s National Science Agency

- 5500 talented staff
- $1 billion+ budget
- Working with over 2800+ industry partners
- 57 sites across Australia
- Top 1% of global research agencies
- Each year CSIRO contributes $4.5 billion+ to the economy
# IMT Scientific Computing Services

*Providing computational expertise that enables great science*

## CSIRO Researchers/CSIRO Research Projects

<table>
<thead>
<tr>
<th>Digital Science Consultancy</th>
<th><strong>Embedding Computational Science Expertise</strong></th>
<th>Capability Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning &amp; Provisioning; Analytics; Science domain computing and data expertise: Bioinformatics; Astronomy; Chemistry; Earth Sciences; Synchrotron Science;</td>
<td>Driven by requirements from research projects and strategic initiatives Data Analytics and Visualisation Research Software Engineering Technical Solutions Data Pipelines and Modelling</td>
<td>Digital Learning Python; Using CSIRO HPC; ML, AI and Data Analytics; Software Engineering</td>
</tr>
<tr>
<td><strong>Concierge</strong></td>
<td><strong>Service delivery and partnerships</strong> Software services</td>
<td>Outreach Conferences - C3DIS+RDA 2020 Student and Internships Technical Leaders Drop-in expert sessions</td>
</tr>
</tbody>
</table>

## IMT Scientific Computing Services

<table>
<thead>
<tr>
<th>Scientific Software Licensing</th>
<th>License management</th>
<th><strong>Research Data</strong></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>High Performance &amp; Scalable Compute Platforms</th>
<th><strong>Partner Facilities</strong></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Service delivery and partnerships Software services</th>
<th>Partnership advisor and liaison</th>
</tr>
</thead>
</table>

## Scientific Software Licensing

- License management

## Digital Science Consultancy

- Planning & Provisioning
- Analytics
- Science domain computing and data expertise: Bioinformatics; Astronomy; Chemistry; Earth Sciences; Synchrotron Science
- Concierge

## Embedding Computational Science Expertise

- Driven by requirements from research projects and strategic initiatives
- Data Analytics and Visualisation
- Research Software Engineering
- Technical Solutions
- Data Pipelines and Modelling

## Capability Development

- Digital Learning: Python; Using CSIRO HPC; ML, AI and Data Analytics; Software Engineering
- Outreach: Conferences - C3DIS+RDA 2020; Student and Internships; Technical Leaders; Drop-in expert sessions

## Research Data

## High Performance & Scalable Compute Platforms

## Partner Facilities
# Scientific Computing Services - Expertise

<table>
<thead>
<tr>
<th>Expertise</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data analytics and visualisation</td>
<td>Domain specific algorithms and software</td>
</tr>
<tr>
<td>Science workflows – data and compute centric</td>
<td>Science data management</td>
</tr>
<tr>
<td>High performance scientific applications</td>
<td>Solution Design</td>
</tr>
<tr>
<td>Simulation and modelling</td>
<td>Specialised full stack web development</td>
</tr>
<tr>
<td>Research Software Engineering</td>
<td>...</td>
</tr>
</tbody>
</table>

*Total of 42 staff in 4 Teams*
eResearch Collaboration Projects

Delivering Scientific Computing expertise into CSIRO research

- Competitive process – RFP
- Researchers submit ~ 60-80 projects proposals per annum
- ~ 40 projects accepted this cycle

- Projects
  - run for a 6 month cycle
  - 0.2FTE allocated per project and can be applied in various ways
### eRCP Proposals

<table>
<thead>
<tr>
<th>Activity / Project Title</th>
<th>Proposer</th>
<th>Jira #</th>
<th>Round</th>
<th>Status</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mquest - improving a visual quality control tool for ocean temperature data</td>
<td>Cowley, Rebecca (O&amp;A, Hobart)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
<tr>
<td>Tools for defining chemical crystallisation space</td>
<td>Newman, Janet (Manufacturing, Parkville)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
<tr>
<td>AI for image analysis</td>
<td>Newman, Janet (Manufacturing, Parkville)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
<tr>
<td>Facilitating access to Sentinel 1/2 satellite data</td>
<td>Poulsen, Brett (Energy, Pullenvale)</td>
<td>ERRFP-652</td>
<td>2018 July - Dec</td>
<td>Application Received</td>
<td>VIEW PROPOSAL</td>
</tr>
<tr>
<td>Real Time Soft Matter Modeling for Autonomous Design</td>
<td>Hockings, Nick (Data61, Pullenvale)</td>
<td>ERRFP-651</td>
<td>2018 July - Dec</td>
<td>Application Received</td>
<td>VIEW PROPOSAL</td>
</tr>
<tr>
<td>Development of image analysis algorithms for automated tag reading</td>
<td>Wang, Dadong (Data61, Marsfield)</td>
<td>ERRFP-650</td>
<td>2018 July - Dec</td>
<td>Application Received</td>
<td>VIEW PROPOSAL</td>
</tr>
<tr>
<td>Biomass Quality Database</td>
<td>Roberts, Daniel (Energy, Pullenvale)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
<tr>
<td>Immersive Marine Sensor Data Analytics</td>
<td>Engelke, Ulrich (Data61, Sandy Bay)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
<tr>
<td>Processing electronic health records using deep learning</td>
<td>Karimi, Sarvnaz (Data61, Marsfield)</td>
<td>2018 July - Dec</td>
<td>Awaiting Endorser</td>
<td>VIEW PROPOSAL</td>
<td></td>
</tr>
</tbody>
</table>
eRCP per BU
H2 2019
Example Project 1 - Freight transit across Australia

TraNSIT is a transport logistics tool developed by CSIRO that maps millions vehicle trips between production and domestic and export markets. This visualisation uses Deck.gl to animate traffic flows for tens of thousands of freight trips across Australia, providing valuable input to inform infrastructure investment and regulatory decisions.

Data Analytics and Visualisation team: Louise Ord (IM&T, South Eveleigh)
TraNSIT team: Adam McKeown (L&W, Cairns) and Artiom Bondarenco (L&W, Dutton Park)
Example Project 2
Bayesian Network for understanding systems risk

Shiny dashboard enables clients to interact with expert knowledge-driven Bayesian network – allows model parameter selection and scenario analysis without the need to interact with the underlying model.

Data Analytics and Visualisation team: Louise Ord (IM&T, South Eveleigh)
Modelling team: Justine Murray (H&B, Dutton Park) and Bill Venables (Data61, Dutton Park)
Challenges

Resourcing:

• Allocation complexity:
  Staff/skills ↔ projects
• Science priorities, mix of BUs

Variable recognition of contributions to project outcomes

Occasional project failures

Skills shortage
eResearch Collaboration Projects - Future Plans

Ongoing process improvement
• project proposals, endorsement, submission
• project execution, close, reporting

Engaging RSE movement
• Communication of capabilities
• Capability champions

More flexible scope

Dropping the “e”!
Thank you

Justin Baker
Team Leader, Data Analytics and Visualisation
T +61 3 9545 2012
E justin.baker@csiro.au

Dr John Zic
Executive Manager, Scientific Computing Services
T +61 2 9490 5682
E john.zic@csiro.au

eRCP Case Studies