RDA VP16 Regional Event - Skills

PRESENTED BY
Kathryn Unsworth, Liz Stokes, Matthias Liffers - ARDC,
Nick May - FAVeR, Mark Crowe - QCIF
In opening this session, I would like to acknowledge and celebrate the Bunurong people whose traditional lands I’m meeting with you from today. And I would like to pay my respects to their Elders both past, present and emerging. I extend that respect to Aboriginal and Torres Strait Islander peoples and Elders from other communities who may be joining us today.
Muted during presentations
Unmute if you’d like to talk

For questions & comments
use the Chat box

Turn on your video
if you have the bandwidth

#RDAPlenary
We welcome little people & pets

Come prepared to contribute
and have fun!
Please go to menti.com and use the code 54 21 86 7
RDA Virtual Plenary 16
VP16 Regional Event – Daily Sessions

• Monday, 23 November: COVID-19: Role of data in responding to COVID-19 pandemic. What can we do to respond better in the future?

• Tuesday, 24 November: Data versioning and data life cycle

• Wednesday, 25 November: FAIR

• Thursday, 26 November: Skills

• Friday, 27 November: Research Software

For more information and to register, please visit https://tinyurl.com/yy2fz4vh
THANK YOU VERY MUCH!

RDA
Email - enquiries@rd-alliance.org
Web - www.rd-alliance.org
Twitter - @resdatall
LinkedIn -
www.linkedin.com/in/ResearchDataAlliance
Slideshare -
http://www.slideshare.net/ResearchDataAlliance

Stefanie Kethers
Director of Operations, Research Data Alliance
Email – Stefanie.Kethers@ardc.edu.au
Skills session agenda

Converging/diverging skills challenges - eResearch Australasia, ARDC Skills Summit and RDA P16

Skills Summit 2020 - Working Groups

FAVeR - eResearch Services and Capabilities

Carpentries Consortium

Australian Skills Landscape - missing skills blitz (activity)
Converging skills challenges
Skills & Training - eResearch Australasia Conference

Presentations primarily focused on transitioning from face-to-face (in-person) to online training delivery

Challenges

- Interactivity
- Individual assistance
- Different challenges for different skills/tools training
- Dependencies on technical abilities and the levels of topic knowledge of participants
- Solving technical issues much harder online
- Some content for in-person training doesn’t translate well to online
- People learn differently online

Innovative ways for mitigating the challenges...
Skills & Training - ARDC eResearch & Data Skills Summit

Presentations - local focus

Accelerating skills development in Data science and AI at scale

Astronomy Data And Computing Services - Upskilling the Australian astronomy community

Presenting MetoDHology – an open platform to support cross-disciplinary digital research methodologies

Successful data training stories from NCI

Data Fluency: a community of practice supporting a digitally skilled workforce

eResearch Services and Capabilities

An open source textbook for research software engineering

Monash University - University of Queensland training partnership in Data science and AI

Developing an organisation-wide framework to transform and uplift data capabilities
Skills & Training - RDA Plenary 16

IG - Professionalising Data Stewardship: Challenges and Developments

IG - Engaging Researchers with Data: What’s Next

BoF - How to Improve Virtual Collaborative Meetings - Sharing Experiences, Resources and Stories

IG - Education and Training on Handling of Research Data: Creating and Connecting Data Training Ecosystems


BoF - Data Concierge - Extending the Services, Roles, and Support of Data Stewards for the Success of International Research Partnerships
Converging skills - common challenges

- Transitioning training from in-person to online
- Surveying for skills gaps and skills training overlaps
- Coordinated approaches - (Inter)National level training
- Co-development of training materials - training partnerships
- Repositories for open source trainings
- Emphasis on Train the Trainer
- Establishing and supporting a network of volunteer trainers/community of practice - sharing expertise
- Skills profiles, identifying core skills for specific groups, identifying not only skills gaps but gaps in levels of capability/competency, mapping learner journeys
- Professionalisation of the eResearch workforce
4 task forces

01
EOSC minimal skillset
Identify and prioritise OS and digital skills for EOSC

02
Options for organisational models for competence centers (blueprints)

03
EOSC in national strategies for digital skills

04
Specifications for training catalogue(s)

Ongoing work. Consolidated reports expected Nov - Dec 2020

Kuchma, I. (2020). EOSC Skills & Training presentation
Diverging skills challenges
Interesting points of difference

Training delivery pathways
Leveraging partnerships with commercial training providers
User-centric design for training - proactively engaging with research communities to improve training
Training trainers/facilitators - best practice session facilitation
FAIR training
Data Concierge
Others?

Discussion...
What our skills communities will be focusing on -

Skills Summit Working Groups

- Jointly Developed Training & Materials
- Train the Trainer
- National Training Registry/Calendar
- eResearch Support Roles and Careers
Jointly developed training & materials - Mark Crowe
Jointly Developed Training Working Group

Mark Crowe

RDA Regional event: skills session. 26 November 2020

#QCIF
Motivation

● **Cost**
  ○ Training is expensive to develop, both in time and money

● **Demand**
  ○ Training groups supporting researchers have broadly overlapping needs

● **Expertise**
  ○ Domain specialists may be located at different institutions
Objective

To improve the quality and quantity of new digital skills training available to the Australian research community by supporting collaboration around the development of training content.
Initial Goals

- Communication platform to connect development partners
- Identify training topics for co-development
- Review existing solutions for collaborative development
- Address IP and ownership issues
- Incentivise full participation
Progress

First meeting 10am (AEST) Friday December 4th

- To get involved, contact mark.crowe@qcif.edu.au
Thank you

Mark Crowe
mark.crowe@qcif.edu.au
Train the Trainers Working Group - Liz Stokes
Common goals

Enhance the impact of eResearch and data skills training
Meta-skills that increase trainers efforts:
  - Facilitation
  - Organisation
  - Materials
  - Sharing
  - Collaboration
Connect up current trainers
Prepare for future trainers
Look at formal certification
First steps

Skeleton user stories

Explore certification options eg UK HEA, Carpentries

#trainthetrainer on ENRICH slack

Establish public presence

Community of practice meetup for skillsharing
Break time

Back in 10 minutes
National Training Registry/Calendar

Working Group
National Training Registry: Goals

Produce a National Registry & Calendar of “live” training events from Institutions primarily serving Aus/NZ researchers and research support staff.

- Create a “minimum viable product”*
- Determine and use a mechanism / technology platform
- Establish a process to produce this product
- Establish an ongoing “home” for this product

*Includes:

- Minimum standards for submissions (consistency, findability, metadata (ISO19115 for basic schema))
- Agreed boundaries for first-round “training events” is “live” (virtual/in-person)
National Training Registry: Stretch Goals

Prioritisation of Stretch Goals will depend on the Working Group’s weightings as well as “interest” within the group/individuals to delve into an area.

- Build on the product’s sophistication/depth/breadth/usability
- Increase the sophistication of the product’s underpinning technology/ways to submit
- Increase boundaries for capturing “non-live” training events

Other considerations:

- Quality control
- Metrics
- Impact
National Training Registry: Audiences

- Trainers
- Researchers
- Research Support Professionals
- Managers of training organisations

Use cases (so far)

- Trainer:
  - How to benchmark ourselves
  - Trainer/Institute: Do we develop training ourselves or outsource it; prioritise based on what is available elsewhere.

- Researcher: Who offers what I’m after (where do I look) - central “listing”.
National Training Registry: Outcomes (early discussions)

- Registry of training/events - Aus/NZ
  - WHO
    - Tertiary Educational Institutions or training events primarily serving AUS/NZ researchers and research support staff
  - WHAT
    - “Live” in-person or virtual training “events” in the first instance - first deliverable
    - “Non-live” training - second deliverable
- Calendar of those training/events
  - Same as above
National Training Registry/Calendar: Minimum viable metadata

Has fields:

- Event name
- Event date
- Event location (broad indication or online)
- Link to event details on host website
- Who the event is for or limited to
- Level (awareness, introductory, intermediate, advanced)

Other fields/metadata:

- Concrete learning outcomes (ideally standardized? Is this even possible?)
National Training Registry: Future activity

● Collaborative working doc - https://tinyurl.com/y65qmw9k
● Communication channels - email at this point
● Next meeting - Week of December 14 - Doodle Poll - inviting TESS developer
● Duration of the WG - TBA
● Contact - Ann Backhaus (Pawsey) ann.backhaus@pawsey.org.au
Other conversations

FAVeR - eResearch Services and Capabilities - Nick May

Carpentries Consortium - Liz Stokes
Separate pdf
No slides
Australian Skills Landscape - missing skills blitz - Jam board - https://tinyurl.com/y2fdhxjm
Australia’s eResearch & Data Skills Landscape

Why develop a Skills Landscape?

To identify:
• Data skills
• Skills ARDC focuses on
• Skills others in the sector focus on
• Overlaps and gaps

Who is the Skills Landscape for?

• Skills training developers & trainers
• Data Stewards
• Researchers, but not directly

Points to note:
• Data Integrity happens because of good DG, FAIR & DM
• Data Sovereignty is captured under Gov policy & legislation
• Change management -> Cultural change – important, but a separate set of skills

Other considerations:
• Skills Landscape takes a course/unit view
• Slides 2 to 6 focus on a generalised identification of skills, not roles
• Slide 8 describes four key/generalised roles
• Roles are not exclusive, i.e. an individual could undertake more than one data role at a time
Communities for skills and workforce development

**Data Stewardship Skills**
- Data Governance
  - Policies & Standards
- FAIR Data Principles
  - FAIR Outputs
  - Discovery & Reuse
  - FAIR Technical Environment
- Data Management
  - Working with Data
  - Preserving Data

**Data Generation and Use Skills**
- Data Methods
- Data Infrastructures

**Research Software Engineering Skills**
- Coding for Research

**eRes & Data Infrastructure Management Skills**
- Establishing, Implementing and Operating Infrastructure
Data Governance

Data Governance Skills
The process of creating and complying with data standards and policies which manage the availability, usability, integrity, use, and security of data. Effective data governance ensures that data is consistent and trustworthy, doesn't get misused, and generates value for the data owner(s).

Policies & Standards

Institutional Policies
Funders & Publishers Policies
Government Policies/Legislation

Intellectual Property
Research Integrity
Trust Certification
FAIR Data Principles Skills

Skills that are useful to create and use FAIR (Findable, Accessible, Interoperable, and Reuseable) data outputs and infrastructures that enhance the ability of machines and people to find, access, and use or reuse data.

These skills help facilitate trust through improved transparency and reproducibility.

FAIR Outputs
- Metadata
- Persistent Identifiers
- Open Formats

Discovery & Reuse
- Repositories & Discovery Portals
- Standardised Communications Protocols
- Provenance
- Terms of Access
- Licensing

FAIR Tech Environments
- Vocabulary / Ontology Application
Data Management

Data Management Skills
The operational management and oversight of data assets to help provide users with high-quality data that is easily accessible in a manner consistent with the data governance framework, i.e. these are **tactical** skills.

**Working with Data**
- Planning for Data Management
- Cleaning & Validating
- Categorising & Classifying
- Structuring Data
- Workflows
- Handling Sensitive Data
- Accessing & Storing
- Applying Local Protocols
- Citing & Tracking
- Moving Data
- Data Versioning

**Preserving Data**
- Retention & Discovery Infrastructures
- Archiving & Publishing
- Appraising, Selecting & Disposing of Outputs
- Preparing & Packaging Outputs
- Managing over Long-term
Data Generation & Use Skills
Skills which are useful for researchers and other data generators to ensure their data is, at the outset, structured and managed in such a way as to facilitate (re)use, high quality, and reflection of impact.

Data Methods
- Data Analytics
- Collection & Capture
- Compilation, Derivation & Aggregation
- Simulation & Modelling
- Reproducibility & Replication
- Data Visualisation & Storytelling

Data Infrastructures
- Data Repositories
- Data Portals
- Platforms/Facilities/Resources
- Access Management
- Citation & Impact Tracking

Data Stewardship Skills
- Policies & Standards
- Working with Data
- FAIR Outputs
Data Generation & Use Skills

Skills which are useful for researchers and other data generators to ensure their data is, at the outset, structured and managed in such a way as to facilitate (re)use, high quality, and reflection of impact.

Data Analysis

A data lifecycle stage that involves the techniques that produce synthesized knowledge from organized data/information.

- Handling geospatial data
- Uncertainty propagation
- Bayesian techniques
- Statistical inference
- Spatial analysis
- Meta-analysis and systematic reviews
- Time-series analysis
- Scientific workflows
- Simulation modeling
- Advanced linear modelling
- Nonlinear modelling
- Machine Learning
Data Generation & Use Skills

Skills which are useful for researchers and other data generators to ensure their data is, at the outset, structured and managed in such a way as to facilitate (re)use, high quality, and reflection of impact.

Data Visualisation

Data visualisation is the graphical representation of information and data.

- Visual literacy and graphical principles
- Interactive visualisations
- 2D & 3D visualisation
- Visualisation services and libraries
- Visualisation tools
- Web visualization tools and techniques
Research Software Engineering Skills

- Software development practices
- Code testing for reliability
- Programming (scripted, command-line)
- Algorithm design and development
- Software Workflows
- Data structures and algorithms
- Code parallelisation
- Numerical stability
- Compute architectures - HPC and Cloud computing
- Version control
- Data modeling
- SQL and database design
- Technical specs - documentation
- Modular design for reusability
- Containerisation principles and tools
Wrap up
Thank you

CONTACT
Kathryn Unsworth
kathryn.unsworth@ardc.edu.au