Case study - The Alan Turing Institute - Data Science Teams

Background

In the ever-evolving landscape of data science, specialised roles have become integral to supporting interdisciplinary collaboration and guiding the adoption of open, reproducible, inclusive and ethical practices. Our project aims to address the gap in the policy landscape that would standardise and strengthen specialised research infrastructure roles alongside the traditional roles in data science. The goal is to professionalise dedicated specialised as well as general roles for ensuring high-quality ethical research at institutional and national levels.

To do this we need to find out about how current data science teams work in different sectors across the UK - what roles do they have, how do they define these roles, what skills do the roles need and how do they work as an effective research team. We are building organisational case studies to better understand working in data science.

Link to project GitHub repo: https://github.com/alan-turing-institute/professionalising-data-science-roles

Target Groups & Links

- Research Community Managers
- Research Application Managers
- Research Software Engineers
- Data Wranglers
- Project/Programme Managers
- PI/Project Leads
- Postdoctoral researcher/Research Fellow
- Research/Applied Ethicist

- Participatory scientist

Resources

- Alliance of Data Science Professionals - https://allianceforadatascienceprofessionals.co.uk//standards
- Detailed standards: https://allianceforadatascienceprofessionals.co.uk/documents/AfDSP_Standards_June22.pdf

Steps for case study

1. Collate names for your role at organisation
2. Write a short definition for your role/s
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above - this can be simplified or detailed.
4. Are there different roles in your team? How are these different?
5. Are there different levels of roles? What makes them different levels?
6. Consider how you have worked with other teams at Turing,
7. Which skills are essential and which ones cross over with other roles?
8. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Research Community Managers (still work in progress)

1. Collate names for your role at organisation
Research Community Manager
Senior Research Community Manager

2. Write a short definition for your role/s

3. Make a framework of skills for your role/s - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Research Community Managers have five main skills areas, A to E.

- A and B are Core Community responsibilities.
- C to E are shared responsibilities that cross over with other teams working with RCMs.

Skill Area	**Evidential Requirements**	**Types of evidence**
A. Communications *This skill relates to communicating within a project/programme to the research community and disseminating research to external audiences.*		Essential ● Creating and sharing guidelines and templates from communications (example to be shared in The Turing Way) ● Communication planning and content development ● Writing community documents ● Curation of project resources to share across projects (institute wide) - via The Turing Way ● Communication channel and social media management guidelines ● Speaking & presentation about the project and their work Desirable ● Editing research and external communication content ● Community specific documentation, impact and assessment ● Social Media management and monitoring ● Graphics & branding of the community work ● Program and event planning (steering and decision-making) ● Community-level governance, inreach and outreach process ● Translation and contextualisation (non-technical docs or supporting the multi-language comms)

B. Engagement *This skill relates to engagement throughout a research project - initial stakeholder mapping, onboarding researchers and other contributors and creating engagement opportunities, events and training.*		Essential ● Creating and sharing engagement and technical skill documents (examples to be shared in The Turing Way) ● Workshops for knowledge share and training/skill building ● Mentoring opportunities (upskilling and onboarding community members on visible roles as well as connecting them with other projects at the Turing) ● Community Policies and their alignment with organisational policies ● Recognition and visibility of community members (rewards, incentive alignments and support) ● Onboarding diverse community leaders (EDIA embedding in research) ● Community survey and reporting (community engagement, health and infrastructure) ● Creating collaborative, community events and opportunities for discussions ● Open collaboration and participatory approaches ● Reproducibility skills (Bidirectional conduit of The Turing Way concepts) Desirable ● Scientific engagement with researchers and (volunteer) onboarding ● Roadmap for community engagement ● Short term and long term engagement strategy (stakeholder mapping and planning)
C. Strategic development		Essential ● Embedding open science in the project strategy (a process built using this reference: Leonelli, S. (2018). https://ore.exeter.ac.uk/repository/handle/10871/31337) ● Institutional policy & guidelines ● Institutional impact metric and templates for reporting

		● Consulting other communities - creating opportunities Desirable ● Project and programme strategy ● Project Roadmap ● Identifying growth opportunity ● Contributor, user, stakeholder and public advocacy ● Funding, scoping the growth and implementing change ● New staff recruitment ● Org level impact and reporting
D. Technical skills		Essential ● Open Research Objects / Outputs / Communications / Licensing guidance/upskilling and support ● Project set up and design ● Technical training and upskilling ● Open Science support and planning (open source, open access, open data) Desirable ● Research Data analysis ● Data visualisation ● Data Management & Stewardship ● Open source software/ code/ infrastructure (alongside RAM) ● Platform/ product management (alongside RAM) ● Web, UI, UX design ● Content management and Archiving ● Technical Support / Code Review (alongside RAM) ● Reproducible workflows/research process
E. Project/Programme Management		Essential ● Risk identification, delivering outputs ● Record-keeping and reporting

		Desirable ● Operational planning & implementation ● Time management ● Project specific documentation, impact and assessment ● Producing ethics documentation ● Event planning and logistics ● Financial and budget management ● Meeting planning, calendar management and facilitation ● Recruitment support ● Project Organogram, project-level governance and project stakeholder engagement ● Contract management, legal and data protection ● Updating CRM records for project

4. Consider how you have worked with other teams at Turing - can you give a short example
   - Which skills were essential for your role and which ones cross over with other roles at your organisation?

5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Research Application Managers

1. Collate names for your role at organisation

2. Write a short definition for your role/s
3. Make a framework of skills for your role/s - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing - can you give a short example
   - Which skills were essential for your role and which ones cross over with other roles at your organisation?

5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Research Software Engineers

1. Collate names for your role at organisation
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Data Wranglers

1. Collate names for your role at organisation
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Research Project/Programme Managers

1. Collate names for your role at organisation

Alexandra Araujo Alvarez
Arielle Bennett
Batool Almarzouq
David Sarmiento
+ more

2. Write a short definition for your role

3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?

- Five different roles: Research Project Administrator, Research Project/Programme Coordinator, Research Project Manager, Programme Manager, Senior Programme Manager

- Are there different levels of roles? What makes them different levels?

Skill Area	**Evidential Requirements**	**Types of evidence**
	-	

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Project Leads/PIs

1. Collate names for your role at organisation
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?

- Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Postdoctoral Researchers/Associates/Fellows

1. Collate names for your role
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Research Ethicist

1. Collate names for your role
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	Evidential Requirements	Types of evidence

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.

Participatory scientist

1. Collate names for your role
2. Write a short definition for your role
3. Make a framework of skills for your role - see Alliance of Data Science Professionals above.
   - Are there different roles in your team? How are these different?
   - Are there different levels of roles? What makes them different levels?

Skill Area	**Evidential Requirements**	**Types of evidence**

4. Consider how you have worked with other teams at Turing,
   - Which skills are essential and which ones cross over with other roles?
5. Build a Turing wide case study that includes these roles and explains how we are working in Data Science Teams at Turing.