<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Campus collaboration for Data Management</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Chan, FHF</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>JULAC Libraries Forum 2015, The Hong Kong Institute of Education, Hong Kong, 24 April 2015.</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2015</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/209960">http://hdl.handle.net/10722/209960</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Campus Collaboration for Data Management

Fred H. F. Chan
University of Hong Kong Libraries
Outline

• Research Data Management: What and Why?

• Building Blocks of Research Data Management
  – Policy
  – Services
  – Processes
  – Technologies

• Conclusion
Research Data Management: What?

- Research data management (RDM) is a concept used to describe the managing, sharing, and archiving of research data to make it more accessible to the broader research community.

- RDM provides an opportunity for a researcher to make plans to create a plan that will ensure that their data will be organized so that it can be shared with other researchers, and archived for long term preservation.

Research Data Management: Why?

- Librarians can assist researchers with their data management concerns by helping to assign metadata to describe research data; teaching data management best practices; providing guidance for transferring data into a data repository; and outlining a preservation plan to ensure long term access to data.

Starting Off RDM...

- Who owns the data?
- What requirements are imposed by others?
- Which data should be retained?
- For how long should data be maintained?
- How should digital data be preserved?
- Are there ethical considerations?
- How are data accessed?
- How open should the data be?
- How should the costs be borne?
- What alternatives to local data management exist?

Source: Erway (2013)
A Library-oriented RDM Model

Drivers: Why?
- Storage
- Security
- Preservation
- Compliance
- Quality
- Sharing
- Jurisdiction

RDM Programme Components: What?
- Strategies
- Policies
- Processes
- Guidelines
- Technologies
- Services

Influencing Factors: How?
- Acceptance
- Cultures
- Demand
- Incentives
- Roles
- Governance
- Politics
- Resources
- Projects
- Skills
- Communications
- Context

Stakeholders: Who?
- Library
- IT Services
- Academic Departments
- Senior University Managers
- Research Support Services
- Other Support Services

Researchers: Arts and Humanities Disciplines
Researchers: Social Sciences Disciplines
Researchers: Science Disciplines
Researchers: Engineering Disciplines
Researchers: Medicine and Health Disciplines
Researchers: Multi-Disciplinary Collaborations

Source: Pinfield, Cox, & Smith (2014)
RDM Policy: Who?

- Researchers, departments/faculties, divisions, central administrative units and service providers and, where appropriate, research sponsors and external collaborators, need to work in partnership to implement good practice and meet relevant legislative, research funder and regulatory requirements.

Source: University of Oxford (n.d.)
The University is responsible for:

• Providing access to **services and facilities** for the **storage**, **backup**, **deposit** and **retention** of research data and records that allow researchers to meet their requirements under this policy and those of the funders of their research;

• Providing researchers with access to **training, support** and **advice** in research data and records management;

• Providing the necessary **resources** to those operational units charged with the provision of these services, facilities and training.

Source: University of Oxford (n.d.)
RDM Services: Outreach

• **Work done:**

1. HKU Library Research Support Survey (Ithaka S+R), with a section on *Data Preservation and Management*

2. Received faculties’ feedback on a) *ownership/IP rights*, b) *nature and extent of data*, and c) *implementation/operations* of the drafted policy and implementation plan
RDM Services: Outreach

- **Partners**: Research Services, Faculties, Subject Librarians
- **Scope**: Research data creation, collection, and possibly sharing
- **Plans**:
  1. Approach researchers who expressed willingness to share their research data (☑ of data sharing in RGC grant proposals)
  2. Talk to individual researchers on their specific RDM needs, practices, and considerations

- **Useful reference**:
  1. Incentives for researchers and support stakeholders (Brown, Bruce, & Kernohan, 2015)
  2. Talking Points with Faculty (Tsang, 2015)
  3. Motivations and incentives for sharing research data (Van den Eynden & Bishop, 2014)
RDM Services: Training

- **Partners**: Graduate School, Research Services, Faculties, Instructional Librarians

- **Scope**:
  - Start with RPG students and staff, first round is voluntary
  - May later embed into existing research integrity/ethics training
  - Initial training/support will probably be more generic

- **Plans**:
  - Course(s) on data management planning, data discovery, publication and sharing
  - Individual data management consultation (or data curation interview)
Data Management Plan is important not just to the funder, but also the institution and the researchers.

**Work in progress:**
- Evaluation of DMPTool and DMPonline
- DMPTool: NTU Singapore has become the first non-US partner institution
- DMPonline: Already used by a number of non-UK institutions
- Further liaison needed to correspond DMP with data policy
- The need to customize:
  - Add own template(s) & questions to other templates
  - Add custom guidance, examples, suggested answers
  - Branding
• **Work done:**

1. Mapping between Dublin Core and [DataCite](https://www.datacite.org) Metadata Schema
2. Obtained a subject coding scheme [ACIRES](https://www.acires.org), for describing research data from Professor Chen Kuang-hua of NTU
   - Bilingual
   - Browsable with 3 degrees of hierarchy
   - Based on ANZSRC (Australian and New Zealand Standard Research Classification) published in 2008, with modifications
RDM Processes: Link With Researchers

- **Partners**: Research Services, Faculties, Technical Services Division

- **Future work:**
  1. Procedures for research data collection, storage & re-use
  2. Refine the DC-DataCite metadata scheme
     - mandatory, recommended, optional fields
  3. Consolidate the subject coding scheme with [RGC Research Field Areas](#)
  4. Explore the possibility to
     a. enrich the generic DC-DataCite metadata scheme with [subject-specific metadata](#) (see also [Research at risk by Jisc](#))
     b. map researchers’ research interests with the subject coding scheme
RDM Technologies: Data and Institutional Repository

- **Partners**: Research Services, Faculties, Technology Support Services Division

- **Scope**:
  - **HKU Scholars Hub** to host, preserve, and curate HKU data centrally
    - Create automated procedures for ingestion of data sets
    - Data sets as top level entities
    - Linking data sets to researchers, publications, grants, patents
    - DOI minting
Future work:
1. Decide on the costing model for data storage
2. Evaluate various data repository software & data-related features
3. In-house enhancement of Hub (currently DSpace-CRIS 4.1)
4. Get in touch with global registries (DataCite, re3data.org, etc.)

Projects to keep an eye on:
1. Research Data Repository Requirements and Features Review (Leahey et al., 2015)
2. Research Data Management Technical Infrastructure (Lewis, 2014)
3. Repository Platforms for Research Data, by Research Data Alliance
Conclusion

• RDM cannot be done alone, it’s important to leverage other parties’ skills and knowledge via collaboration.
• Given the large scope of RDM programme, starting with the generics may be a good idea.
• RDM offers opportunities to libraries to become important partners in the entire research lifecycle.
• However, disciplinary differences can be a major challenge.
• Policies, practices, and software development are rapidly changing, we can stay informed by participating in discussion groups, joining mailing lists, and attending webinars/training.
Acknowledgement

- The author thanks Mr. David Palmer for his advices and suggestions on this presentation.
- Template & design by Showeeet.com
- Archived at http://hub.hku.hk/handle/10722/209960
- This PPT & PDF are licensed under the Creative Commons Attribution “CC-BY” license. This license permits use of this work, so long as attribution is given. For more information about the license, visit http://creativecommons.org/licenses/by/3.0/hk/