



Facilitate Open Science Training for European Research

Open access and research data management:

Horizon 2020 and beyond

University College Cork, April 14th & 15th 2015

Lessons learned developing the Digital Repository of Ireland

Dermot Frost
Trinity College Dublin

Talk Outline

- DRI Introduction
- Policy
- Procedure
- Platform
- Partners
- People
- Summary



Digital Repository of Ireland

- The Digital Repository of Ireland is an interactive trusted digital repository for contemporary and historical, social and cultural data held by Irish institutions
- €5.2M grant from HEA PRTL I
- 6 academic partners
 - Royal Irish Academy Dublin Institute of Technology
 - Trinity College Dublin NUI Galway
 - Maynooth University NCAD
- Public launch June 2015 at DPASSH

DRI Vision Statement

- The Digital Repository of Ireland links together and preserves the rich data held by Irish institutions, providing a central internet access point and interactive multimedia tools.
- DRI is an educational resource used by the public, students and scholars, and will provide preservation and access services for our stakeholders and partners.
- DRI acts as a focal point for the development of national guidelines and policy for digital preservation and access.

DRI Project Structure

- Four project strands and ten work packages
- Strand 1 – Management (RIA led)
 - WP1 – Long term planning
 - WP2 – Project management
- Strand 2 – Context (MU led)
 - WP3 – Requirements Analysis
 - WP4 – Policy and Guidelines

DRI Project Structure

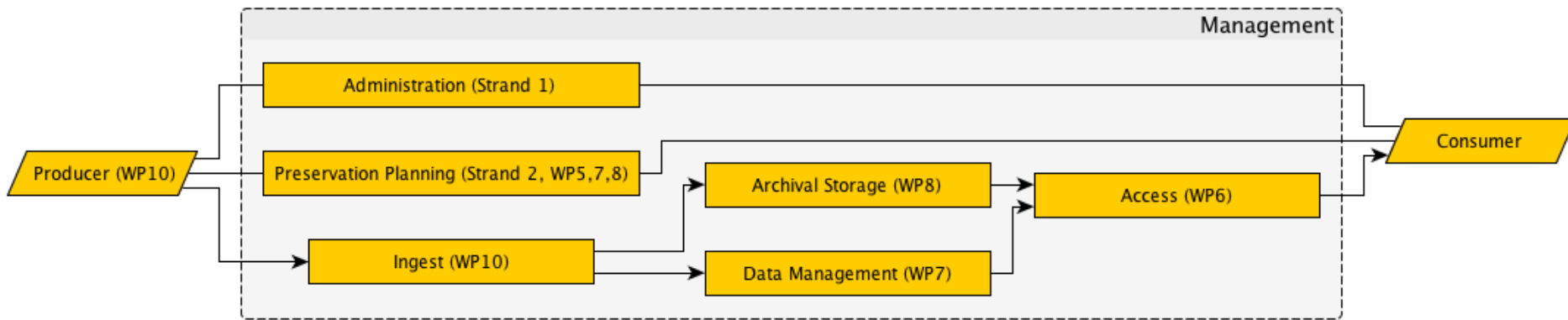
- Strand 3 – Design and Implementation (TCD led)
 - WP5 – Architecture
 - WP6 – User Interfaces
 - WP7 – Data Management
 - WP8 – Storage
- Strand 4 – Rollout (RIA led)
 - WP9 – User support, Training and Advocacy
 - WP10 – Demonstrator projects

Policy

- Trusted digital repository status achieved in 3 ways
 - Data Seal of Approval
 - ISO 16363 - Audit and certification of trustworthy digital repositories
 - De-facto trust within the community
- Strong emphasis on policy and audit
- Policies need to be based on reality and common sense
- Input from all parts of the project
 - Ideal policy may not be technically feasible
 - Desired technical solution may not be legal

Policy

- DRI follows the OAIS model
- Strands and Workpackages map onto the model



Policy

- LESSONS LEARNED

- Policies need to be clear
- Policies need to be public
- Policies need to be regularly reviewed
- Policies may need input and drafting with legal expertise

- Engage with your community
- Consensus approach better than top down



Process

- DRI built by a distributed team
- Strong requirements gathering and analysis phase
- These inform the design and implementation phase
- Agile methodology
 - Daily stand-up calls
 - Regular face to face meetings
 - Code sprints
- Executable specifications and automated tests
 - Rspec & Cucumber



Process

- LESSONS LEARNED
 - Communication is vital
 - There are (almost) no stupid questions
 - Agile works but requires buy-in from management
 - Requirements need to be complete



Platform

- DRI is a green-field repository
- No legacy systems or data to migrate
- Could start from scratch - 'vim repository.php'
- Better to select best of breed components and then customize as appropriate



Platform



Platform

- Open Nebula as a private cloud platform
- Fedora Commons as core repository software
 - Currently upgrading to FC4 before launch
- Hydra framework for building on top of Fedora
- Apache Solr industry leading search engine
- Blacklight provides the discovery interface
- Shibboleth for federated identity management
- Ceph scalable, robust, parallel data storage
- Ansible for configuration management



Platform

- LESSONS LEARNED

- Quality open source solutions are out there
- Get involved with the relevant open source community

- Talk to your tech transfer people

- Mixing technologies (Java, Ruby, SQL) not an issue
- IRC is not dead



Partners

- A repository needs data to justify its existence
- Need depositors to provide this data
- DRI has spent a lot of time working to build relationships with the community
- Stakeholder advisory group (met yesterday!) 40 people from diverse groups including
 - Raidió Teilifís Éireann
 - National Museum of Ireland
 - Dept of Arts, Heritage and the Gaeltacht
 - UCC Library
 - IBM

Partners

- DRI is well connected and represented
 - Europeana
 - Research Data Alliance
 - DARIAH
 - European Data Forum
 - Horizon 2020 Expert Advisory Group on European Research (e-) Infrastructures
 - National Steering Committee on Open Access
 - Open Repositories Conference

Partners

- LESSONS LEARNED

- Engage with your stakeholders
- Take comments and suggestions on board
- Don't become a slave to the stakeholders

- Build a reputation for getting stuff done



People

- None of this could be done without talented and dedicated staff
- DRI is truly an inter-disciplinary affair
 - Librarians
 - Archivists
 - Software developers
 - Systems architects
 - Photographers
 - Historians
 - Geographers
 - Social Scientists



People

- Full details on <http://dri.ie/dri-team>



People

- LESSONS LEARNED

- Language barrier (tech vs non-tech) is not trivial to overcome
- Technical people are hard to hold on to in the public sector
- Given freedom and a challenge, good people do great work



Summary

- The Digital Repository of Ireland is Ireland's trusted digital repository for humanities and social science data
- Building it has been challenging but a lot of fun and immensely rewarding
- Nothing beats engaging with others to exchange ideas

Open access and research data management: Horizon 2020 and beyond

This event was funded by FOSTER through the European Union's Seventh Framework Programme

<http://www.fosteropenscience.eu>

and organised by

- University College Cork <http://www.ucc.ie>
- Teagasc <http://www.teagasc.ie>
- Repository Network Ireland <http://rni.wikispaces.com>

