

## INTRODUCTION

This poster describes the architecture of a new kind of digital repository service that includes components that run on desktop computers, designed to close the gap between Institutional Repositories (IRs) and the day-to-day electronic work environment used by researchers, and to address the too-often heard cry from repository managers of “we built it but they didn't come”.

The team at the Australian Digital Futures Institute are working with researchers to provide software that can:

- **index and expose** the research data content on their hard disks
- **extract metadata** from files
- **automatically process data** according to highly configurable workflows including producing web-ready renditions of research objects including:
  - documents,
  - domain specific data visualizations (such as chemical molecules), and
  - converting video and images so that they may be easily previewed.

## System Overview

At the individual researcher level, The Fascinator provides **search, tagging and organising services** to assist with **personal digital resource management**. By itself this would add value to the research process by aiding researchers in dealing with a large number of digital resources, but allowing research to “leave” the individual's desktop and be shared is critical in contemporary research.

To encourage the researcher to engage in this networked environment, features for building and contributing to collaborative information environments need to be easily accessible via a variety of models, such as allowing researchers to **publish both documents and data** from their desktop to outputs such as journals, blogs and project websites.

In addition to the collaborative aspect of publishing, we have begun to provide support for **collaborative environments** that provide the ability to direct data to other repositories for a range of services such as team analysis initially using the SWORD protocol. These repositories will include team-based repositories for data sharing and organization, and institutional repositories.

The architecture is inspired by the success of consumer software in two ways; the way entertainment programs organize content via faceted browse and search interfaces using embedded metadata, and the way photographic software allows content to be grouped into collections and pushed to online services, which are essentially repositories.

## Key development requirements

The key development requirements of the system are that it:

1. Work cross-platform on Windows, Mac and Linux platforms.
2. Work alongside other tools, and be extensible to handle all different kinds of research data.
3. Be web-based as much as possible, showing data as being 'in the repository' and as part of the web, from the moment it is created. The idea is to develop plugins for various kinds of research data to provide a useful web view; at a minimum this is a summary of metadata but, if possible, an interactive visualization.
4. Provide collaboration functionality so that researchers may work across their communities; be it team members, supervisors, research participants or the general public.
5. Use standard protocols to interact with other systems, including OAI-PMH and SWORD for repository integration and the Atom Publishing Protocol for interaction with content management systems.
6. Allow for curation as required as data moves between sites that span private and public domains.
7. Present the researcher's data as Linked Data without requiring technical input from the researcher.

Figure 1: Data flow to other repositories

