



What does the next generation repository look like?

---

Paul Walk

 paul@paulwalk.net

 @paulwalk





3 cheers for the current generation of repositories!

---





**cheer #1:**

**proven technology,  
ubiquitous in our  
institutions**





cheer #2:

**strong community  
support**





**cheer #3:**

**distributed policy  
control**



*To position repositories as the foundation  
for a distributed, globally networked  
infrastructure for scholarly  
communication...*

the next generation repository....

---

1. **is connected**
2. **is talkative**
3. **is responsive**
4. **is busy (hard working!)**

# 1. the next generation repository is *connected*

---

- not just connected in a general sense, connected at **every** level:
- **outgoing:**
  - individual *content resources*
    - directly accessible on the network
  - individual *metadata records*
    - not just in batches
  - individual *users*
    - as part of a variety of professional and social networks
- **incoming:**
  - using all appropriate *global identifier systems*
  - allowing *overlay services* to interact with content
    - content mining
    - annotation services
    - etc.



*connected* means that:

repositories are **nodes** in the **network**

content items are **nodes** in the **network**

metadata records are **nodes** in the **network**

users are are **nodes** in the **network**

and the network is **The Web**

## 2. the next generation repository is *talkative*

---

- the contemporary repository is *reticent* - it waits to be interrogated
  - human agents: browsing and searching
  - software agents: downloading, harvesting and aggregating
    - but aggregation is a *tactic* (to quote Peter Burnhill)
- the next generation repository *talks* to the world
  - *publishing* events
  - *inviting* other systems to respond
  - reaching out to users with *targeted* content



### 3. the next generation repository is *responsive*

---

- while the current generation repository is *passive*, the next generation repository is *responsive*
- the next generation repository is an important component in a **distributed, globally networked infrastructure for scholarly communication**, so it:
  - *responds* to requests for content and metadata equally
  - *continuously improves* the information it has, adding value where it can:
    - responding to and supporting **annotation** and **peer review**
    - not just *allowing text-mining*, but *supporting* it and *benefitting* from the derived information

## 4. the next generation repository is *busy*

---

- the next generation repository is constantly *working*
  - actively *preserving* the content it holds
  - improving metadata by:
    - actively *interrogating* remote systems (powered by global PIDs)
    - supporting remote systems and users (*annotating* and *enhancing local metadata*)
  - *exchanging* information with other repositories
    - defining and managing collections
- supporting user workflows - providing and accepting data



# so what will it *look* like?

---

- no longer a **standalone system**
- ***ubiquitous***, plugged into research, teaching and preservation activity
- integral, foundational part of the **next generation of scholarly communications**:
  - fluid, user-mediated
  - providing *direct* access to content and data
  - part of the global ***infrastructure***

the *less clear* the boundaries of the repository, the *more useful* it will be

thanks for listening!

---

**connected**

**talkative**

**responsive**

**busy (hard working!)**