

NAME AUTHORITY CONTROL FOR INSTITUTIONAL REPOSITORIES - SURVEY

Welcome

Welcome to this short survey from the Names Project, a JISC funded initiative to identify the requirements of a name authority service for use by repositories of research outputs in the United Kingdom. More details can be found at: <http://names.mimas.ac.uk/>.

The results of this survey will be used to inform the project team as to the needs of the repository community with respect to authority control, to develop the Names prototype, and to assess the feasibility of this service.

This three part survey should take about 15 minutes to complete. For further information about this survey please email Amanda Hill at amanda.hill@manchester.ac.uk.

Part 1 - About your institutional repository

Overview

1. Name of your institution: (Optional)
2. Name of the institutional repository: (Optional)
3. When was this repository set up? (Optional)
4. What software is the repository based on?
 - DSpace
 - EPrints
 - Fedora
 - Greenstone
 - PURE
 - Don't know
 - Other (please specify):
What version? (Optional)

Repository management

5. Who is responsible for the management/administration of the repository? (Select all that apply)
 - Library staff
 - Administration staff at the Institution
 - IT staff at the Institution
 - Faculty members
 - Don't know
 - Other (please specify):
6. How many FTE staff are responsible for creating or maintaining metadata in the repository?
 - < 1
 - 1
 - 2 - 5
 - 6 - 10
 - 10 +
 - Don't know

Repository content

7. What resources are held (or described) in the repository?

(Select all that apply)

- Articles
- Books/Book chapters
- Monographs
- Conference papers
- Patents
- Theses
- Presentations
- Digital material
- Special collections
- All of the above
- Other (please specify):

8. How many records are currently in the repository?

- < 100
- 100 - 200
- 201 - 500
- 501 - 1,000
- 1,001 - 2,000
- 2,001 - 5,000
- 5,001 - 10,000
- 10,000 +
- Don't know

9. How many author/contributor names are stored in the repository?

- < 100
- 100 - 200
- 201 - 500
- 501 - 1,000
- 1,001 - 2,000
- 2,001 - 5,000
- 5,001 - 10,000
- 10,000 +
- Don't know

10. How many resources are added to the repository in a given year?

- < 100
- 100 - 200
- 201 - 500
- 501 - 1,000
- 1,001 - 2,000
- 2,001 - 5,000
- 5,001 - 10,000
- 10,000 +
- Don't know

Depositing material

11. Who is entitled to submit resources to the repository?

(Select all that apply)

- IR managers/administrators
- Authors/contributors
- Faculty
- All of the above
- Other (please specify):

12. Which of the following metadata elements related to individuals are required upon submission of materials? (Optional) (Select all that apply)

- Affiliation
- Co-author(s)
- Date of birth
- Email address
- Family name
- Field(s) of interest
- Given name(s) or initials
- Occupation
- Phone number
- Post nominal initials (e.g. BA, PhD)
- Postal address
- Status (student, faculty member, administrator etc.)
- Terms of address (Prof., Dr., Ms, etc.)
- All of the above
- Other (*please specify*):

13. Are users required to register prior to depositing material?

- Yes
- No
- Don't know

If so, what information is required upon registering? (Select all that apply)

- Affiliation
- Date of birth
- Email address
- Family name
- Field(s) of interest
- Given name(s) or initials
- Occupation
- Phone number
- Post nominal initials (e.g. BA, PhD)
- Postal address
- Status (student, faculty member, administrator etc.)
- Terms of address (Prof., Dr., Ms, etc.)
- All of the above
- Other (*please specify*):

Part 2 - Authority control

Disambiguation (i.e. the process of distinguishing between similar or identical names)

14. Which of the following statements adequately describes authority control at your repository?

- All names are controlled/disambiguated
- Some names are controlled/disambiguated
- No names are controlled/disambiguated
- Don't know

Please comment (Optional)

15. Does the repository software support authority control?

- The software supports the level of authority control currently undertaken at this repository
- The software does not support the level of authority control currently undertaken at this repository
- The software does not support any authority control
- Don't know
- Other (please specify):

a. Does the software enable control of variant forms of names e.g. Joe Bloggs, J. Bloggs, J. P. Bloggs, Joseph Bloggs?

- Yes
- No
- Don't know

Please comment (Optional)

b. Does the software enable control over changes to names,(e.g. married/maiden names; current/former names of Departments)? (Optional)

- Yes
- No
- Don't know

Please comment (Optional)

c. Does the software prevent entry of duplicate names (e.g. can there be more than one instance of John Smith)? (Optional)

- Yes
- No
- Don't know

Please comment (Optional)

d. Does the software enable names to be merged or crossreferenced? (Optional)

- Yes
- No
- Don't know

Please comment (Optional)

e. Can authors/contributors update or correct their records? (Optional)

- Yes
- No
- Don't know

Please comment (Optional)

Identifiers

16. Are unique identifiers assigned to authors/contributors upon registration or submission?

- Yes
- No
- Don't know

Please comment (Optional)

17. Are any of the following external identifiers used for disambiguation purposes? (Optional)

(select all that apply)

- HR numbers
- Staff numbers
- Student numbers
- Email addresses
- All of the above
- None of the above
- Other (please specify):

Ingested records

18. Are citation records harvested/imported from external sources, such as Thomson Reuters World of Knowledge?

- Yes
- No
- Don't know

If so, are externally harvested records disambiguated prior to ingest?

- All names are disambiguated prior to ingest
- Some names are disambiguated prior to ingest
- Names are not disambiguated prior to ingest
- Names are disambiguated following ingest
- Don't know

Please comment (Optional)

19. Are externally harvested records matched with pre-existing identities? (Optional)

- All externally harvested records are matched with pre-existing identities
- Some externally harvested records are matched with pre-existing identities
- No externally harvested records are matched with pre-existing identities
- Don't know

If so, express the percentage of working time spent disambiguating records, where 100%= 1xFTE. (Optional)

- Less than 5%
- 5-10%
- 11-20%
- 21-30%
- 31-40%
- 41-50%
- More than 50%

Managing identities

20. Which (if any) of the following sources are used to disambiguate author names? (Optional)

(select all that apply)

- Internal directories
- People finder databases
- Co-author details
- Author CVs
- LC/NACO
- VIAF
- Other (please specify):

21. Which (if any) of the following issues have you encountered with regards to name control at your repository?

- Duplicate author/contributor names reducing the quality of search output
- Variant author/contributor names resulting in incorrect attribution of repository materials
- Lack of data preventing adequate disambiguation of author/contributor identities
- None of the above
- All of the above
- Other (please specify):

Part 3 - JISC Names Project

JISC Names Project

22. The JISC Names Project aims to provide a free service for the institutional repository community through its website and via an Application Programming Interface (API). An API facilitates interaction between different software programs and enables remote access to services. The Names API will allow for the automated disambiguation of individual and institutional names.

Additional services such as retrospective authority control work and large-scale de-duplication and disambiguation of repository content may also be provided, if there is a need for them. In order to gauge the demand for such services please rate the following based on potential usefulness:

	Please rate accordingly			
	Very useful	Useful	Not very useful	Useless
a. Bulk de-duplication of identities (i.e. merging and deleting all duplicate names)				
b. Bulk disambiguation service (i.e. matching all names using disambiguation algorithm)				
c. Retrospective cleanup services (e.g. assigning unique identifiers to old name headings)				
d. Bulk disambiguation of batch imports from publishers				

23. Which (if any) of these services would you be willing to pay for? (Optional)

(select all that apply)

- Bulk de-duplication of identities
- Bulk disambiguation service
- Retrospective cleanup services
- On-going services for new submissions/bulk harvesting
- All of the above
- None of the above

Please comment (Optional)

24. Would you be willing to discuss your answers further or to be a case study for the Names project?

(Optional)

- Yes
- No

If so, please provide your name and contact details. (Optional)

Thank you for taking the time to complete this survey.