

## PCC SCS Task Group on Privacy in Name Authority Records

Progress Report, June 3, 2024

After PoCo's approval of the [PCC Position Statement on Personal Data in Name Authority Records](#) in July 2023, the PCC Standing Committee on Standards (SCS) charged the [Task Group on Privacy in Name Authority Records](#) with developing guiding principles and best practices based on these principles and proposing revisions to LC/PCC Documentation as needed to support these best practices. This report aims to share with the community the progress made to date and outline the work that lies ahead.

### Establishment of principles

The Task Group met regularly to discuss privacy in NARs, with a view to achieving consensus on a set of principles, which would guide any recommendations for changes to practice and documentation. We began by reading a few background documents and critically looking at the PCC Position Statement (see above). We used this as a basis for a brainstorming session, the results of which were developed into a provisional statement of shared principles. In the following months, we iterated over these principles collaboratively until the group arrived at a consensus on the eleven principles outlined below.

The Guiding Principles are:

1. Privacy is not a supplemental feature of authority work, but an integral part of it. Catalogers derive data that identify and describe people, from public information sources, without prior consent. Therefore, from the outset, catalogers must incorporate a critical, privacy-aware mindset toward the creation and curation of authority data.
2. The management of authority data involves three relevant parties: "catalogers," who create and maintain the data and systems; "users," who make use of the data for research and discovery; and "subjects"--people (as well as organizations, etc.) who are identified and described by the data. Privacy-informed authority work seeks to maintain an effective balance among the goals and concerns of all three relevant parties.
3. Catalogers collect data about subjects to enable the identification, selection, and use of bibliographic works that are related to them in various ways. The mechanism by which these objectives are accomplished is the creation and maintenance of files of records, which consist of access points and related information that supports them. Ideally this system fosters a beneficial environment for catalogers, users, and subjects alike. However, the creation and management of an authority file is itself a form of compilation and publication, which has the potential for subsequent use by other parties and for purposes other than the support of bibliographic objectives, presenting a risk that the data in an authority file could be used to harm a subject.
4. Catalogers have an ethical and professional duty to reduce the risk of harm to subjects, by

practicing privacy-informed identification and description. The safety and well-being of people who are the subjects of authority data overrides bibliographic concerns.

5. The primary types of risk for subjects are:
  - a. adverse use of the data for purposes other than bibliographic management, in particular, the use of the data to harm or threaten a subject;
  - b. misrepresentation of a subject in a way that is incorrect, misleading or harmful.
6. Any piece of personal information about a person can potentially represent a privacy risk, and should be evaluated critically by catalogers before being included in authority data.
7. Some categories of information carry inherent risk. Recording this information in authority data should be avoided:
  - a. information about physical characteristics, e.g.: age, skin color, health condition, or observable physical characteristics (weight, height, etc.) of a subject;
  - b. information that assigns subjects to fundamental social categories, e.g.: gender, race, caste, socioeconomic status, criminal status, sexual orientation, or disability status of a subject;
  - c. information that would enable a malefactor to contact, stalk, harass, or impersonate, etc., the subject of an authority record, e.g., a subject's address or phone number.
8. Data that appears risky may be recorded if it is clear that the subject wishes this information to be known. Subjects of authority records may be associated with national, cultural, tribal, or ethnic groups; religions, professions, social and political movements; or sundry other affiliations. When these kinds of associations are used by subjects, in public sources of information, to describe themselves, they may be recorded to identify and describe those subjects in authority data. In these cases, the terminology that subjects use in public sources of information to describe their associations should be regarded as more relevant than externally provided descriptions, such as those selected from controlled vocabularies.
9. In addition to types of information, catalogers should evaluate the sources of information from which information about subjects is derived, to limit risks of harm to subjects.
10. Living people should be granted a significant amount of control over their own representation in authority data. However, changes to authority records must continue to accommodate the purpose of the authority file: to identify all bibliographic identities unambiguously.
11. While subjects of authority records may legitimately wish to be ignored completely, or "forgotten" by an authority file, the deletion of a bibliographic identity from the authority file does not solve the problem of it having been created, or of there being a bibliographic need for its subject to be identified and described in authority data. Catalogers should disassociate subjects from objectionable bibliographic identities instead of deleting authority data.

The exact wording and order of these principles remain subject to change but we believe the principles themselves provide a sound starting point for specific recommendations.

## Tasks in progress

The Task Group is writing a report to accompany any best practice recommendations and to provide background and justification for privacy-aware authority work.

The Group is drafting a set of specific practice recommendations for reducing the risk of harm to subjects of authority records. These recommendations will be made at the level of individual metadata elements and their corresponding MARC fields.

The Group is currently working on identifying specific areas of LC/PCC Documentation that are in need of revision.

## Tasks to be completed

The Task Group will discuss recommendations for guidelines on privacy data in existing NARs.

The Group will consult with the Standing Committee on Training (SCT) to draft or update examples, FAQs, or other resources to assist catalogers and authors/creators in understanding and implementing the guidelines and best practices.

## Time frame

- Date charged: November 1, 2023
- Date preliminary report submitted: March 5, 2024
- Date progress report submitted: June 3, 2024
- Date final report due: extended from June 3, 2024 to August 31, 2024

## Task Group membership

Ben Abrahamse, MIT Libraries (SCT Representative, Co-Chair)

Kathleen Burlingame, University of Pennsylvania (LD4 Ethics in Linked Data Affinity Group representative)

Keno Catabay, University of Colorado Boulder (ACDEI representative)

Laura Daniels, Cornell University Library

Michelle Gibeault, Smith College Libraries (ALA Intellectual Freedom Committee, Privacy Subcommittee representative)

Dina Herbert, National Archives and Records Administration (Feb. 2024 - )

Junghae Lee, University of Washington (SCS Representative, Co-Chair)

Tricia Mackenzie, Library of Congress (LC representative)

Alexandra Provo, New York University (LD4 Ethics in Linked Data Affinity Group Representative)

Michael Stewart, University of Delaware (PCC TG on Gender representative)

Anne Washington, OCLC (OCLC Representative, IMAC Liaison)

Thomas Whittaker, Indiana University Libraries

## **Past membership**

Jerry Simmons, National Archives and Records Administration (Nov. 2023 - Jan. 2024)