
In preparation for the Workshop on Scholarly Databases & Data Integration on August 30th and 31st, 2006 at the School of Library and Information Science at Indiana University in Bloomington, Indiana, we ask you to provide a brief bio and your input to the questions below. We plan to make your input available via the workshop web page before the meeting to ease introduction of participants. We will also use your feedback to more effectively structure the workshop. Thank you for your time.

Biography (about 250 words)
Please provide a photo of yourself and a link to your home page and relevant sites and publications.
Image URL: [https://ivl.slis.indiana.edu/images/people/sumeet-ambre.jpg](https://ivl.slis.indiana.edu/images/people/sumeet-ambre.jpg)
[https://ivl.slis.indiana.edu](https://ivl.slis.indiana.edu)

General Questions
What is your main interest in attending the workshop?
- To learn more about the ways large datasets are organized effectively for efficient data retrieval.
- Methods by which different datasets can be integrated with each other.
- Learn more about open source DB Schemas available for specific domains.

What is your main interest in ‘data sharing’ and/or ‘data integration’?
- To learn more about data warehousing.

Scholarly Datasets and Data Integration Challenges
Please provide information on the datasets you serve/use in your research.

Dataset/Project Description
- COS, Medline, PNAS, JCR, USPAT, NSF, NIH / IVCDB interface: Develop a common interface which can be used to query all the above datasets.

Preservation – what processes or procedures are in place to keep data useable over time?

Provenance – what processes or procedures are in place to insure that the data contributed to the project is authentic and authoritative?

Privacy Policy – what processes or procedures are in place to insure the privacy of contributors and/or end researchers?

Data formats / types – what type of data is stored and in what format?

Database Technology – what database management system is used (Oracle, MySQL, etc)?

- PostgreSQL

Storage Technology – what hardware is used to store the data?

Backup Strategy – with what frequency is the data backed up? Are copies kept off site? Do you have a disaster recovery plan?

Partners – what other organizations participate in this project?

Funding / Ownership – which organization owns this project, and how is this project funded?

What data integration issues are you facing/addressing in your research? Please address policy and technical level.

What data integration solutions do you use/have you developed?
Please provide references/links/pointers to relevant work, papers, and efforts.

Please send the completed document by **August 22, 2006** to Katy Borner <katy@indiana.edu> and Stacy T. Kowalczyk <skowalcz@indiana.edu>, using the subject header ‘ScholarlyDB&I’.