Who owns faculty data? Fairness and transparency in UCLA’s new academic HR system

Chloe Reynolds, University of California, Los Angeles
Heather Small, University of California, Los Angeles

Abstract
Beginning in 2015, Opus will be the information system of record for faculty activities at the University of California, Los Angeles (UCLA). Opus will serve as both a profile system, storing data about faculty work, and as a workflow and approval engine for the promotion and tenure process. Opus leverages institutional master data wherever possible to collect data about faculty activity. However, repurposing institutional data collected for purposes not related to academic review necessitates allowing data subjects (UCLA faculty), to contextualize and reframe the data for the review process. Collecting, displaying and storing these augmented records (master data with manually added metadata from faculty) has forced the project team to grapple with questions regarding fairness and transparency to both data subjects and to data consumers. How can we hold to “good design” and usability practices, while faithfully representing the inherent “messiness” of the data? How does the context in which the data was collected impact repurposing the data for academic review? What does it mean to “own” faculty data? This paper outlines our attempts to address these questions, noting the tradeoffs and limitations of the selected solutions.

Keywords: Information Architecture, User Interface Design, Data Governance, Data Literacy, Fair Information Practices
Copyright: Copyright is held by the author(s).
Acknowledgements: We would like to thank our Project Director, Meg Buzzi and the entire Opus Development Team; Professor Fred Zimmerman and the other members of our Steering Committee; Kent Wada, UCLA Chief Privacy Officer; Jim Davis, Vice Provost for Information Technology; Carole Goldberg, Vice Chancellor, Academic Personnel, and all of the faculty and staff who have helped us understand the myriad issues involved in creating a faculty information system.
Contact: creynolds@it.ucla.edu, hsmall@it.ucla.edu

1 Introduction

Beginning in 2015, Opus will be the information system of record for aggregated information about the employment, teaching, research, and service activities of UCLA faculty. Opus will serve as both a profile system, storing data about faculty work and activities, and as a workflow engine for the promotion and tenure process. The system will collect data from disparate sources and store them within a secure electronic profile, provide a paperless review process, and offer robust reporting capabilities based on the underlying data. This replaces UCLA’s current, largely paper dossiers and academic review process.

In order to improve consistency of the data, Opus leverages institutional master data wherever possible. While this process provides more consistency, repurposing institutional data collected for purposes unrelated to academic review may impose new interpretations, removed from the context of the faculty themselves and from the “real world” activities that these data represent. In order to “recontextualize” the data, Opus will allow UCLA faculty, to annotate, augment, and reframe the data for the review process.

Collecting, displaying and storing these augmented records (institutional master data with manually added metadata from “non-authoritative” sources) has forced the project team to grapple with the following questions:

• When data is an amalgam of repurposed institutional data, distanced from the data subject and original context, and when metadata is added to reintroduce the human experience behind the “facts”, how do we provide space for the various “truths” or stories of the data?

• What does it mean to be a data steward or system of record for data about and often originally reported by faculty - whose voice is authoritative or privileged? In which context?

• How can we balance the data needs of university administrators, with faculty concerns about misuse and privacy?
• From a usability perspective, how do we represent the inherent messiness of the data in a way that people can understand, and interpret?

2 Background

2.1 The Dossier

When UCLA faculty undergo the review process, they are required to submit a summary of their activities, which includes their educational background, UCLA employment history, grants, scholarly and creative works, teaching and advising, and university and public service. Traditionally, data has been collected and documented ad hoc, by the individual faculty member or their department. Each act of collection, selection, and formatting imposes layers of interpretation that shape how downstream viewers will interpret the data. The lack of a standard practice has meant that similar items (like serving on an editorial board) have been placed under different categories by different departments, or even by different individuals within a department. Though data is often collected from the same institutional sources, there is not a rigorous process to ensure that the same data point or data state is collected across departments or individuals. One example we encountered was that there was no standard for when enrollment figures were captured for review purposes. While one department might capture enrollment at the first week (the peak enrollment figure), another might capture it at the eighth week (the lowest enrollment figure). While seemingly mundane, this diversity in data collection practices leads to a situation in which central reviewers cannot be certain that they are evaluating faculty based on the same data, putting the university at risk. Opus, UCLA’s upcoming faculty information system, attempts to standardize this process.

2.2 Opus: A faculty information system

While Opus will provide a workflow engine, routing the review files through the peer review and approval process, the heart of the Opus system is the data that it aggregates and integrates from many disparate on- and off-campus sources (as well as from faculty themselves) to build more holistic representations of individual faculty members’ work. In this way, Opus attempts to counter the ad hoc data collection process outlined above, providing a consistent view of the data. Returning to our enrollment count example, Opus will capture the Registrar’s data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses. The work of tracking down stewards’ data at the end of the third week of the academic quarter for all courses.

Opus will create a secure, online profile for each faculty member, pre-populated with courses taught, grants and awards, graduate advisees, etc. from the sources discussed above. We won’t go into much detail about the profile (the entire set of gathered dossier information) in this paper, because our paper focuses on the academic review and reporting functions of the Opus system. Briefly, however, it is helpful to know that for the purposes of academic review (UCLA’s promotion & tenure process), the relevant data from the faculty member’s Opus profile will be extracted and formatted into a standardized academic review template for organizing, and documenting faculty activities. The design of the reporting system is well underway but has yet to be finalized.

Historically, the Academic Personnel Office (APO) provided a paper academic review dossier template, but school, departmental and individual interpretations of that template and its guidelines led to differing styles of data organization and compilation. That meant that dossiers became, long, complex, messy, and extremely difficult to view and interpret, especially when comparing faculty within or across disciplines during fair pay evaluation efforts. Part of the complexity of the Opus project is the magnitude of the data in a dossier itself, and the variety of interpretations and representations across disciplines and across campus organizations, in addition to the sheer volume of annual reviews. As part of preparing for Opus, a new template format was agreed upon by the faculty and administrative staff of Opus Steering Committee, which includes representatives from APO. This new template, being electronic, having built in validation mechanisms and detailed help text, will increase the clarity of the intended template usage and interpretation, not completely of course, but far more than our past paper template was able to do.
In addition to providing data for academic review, Opus must also provide useful information to multiple stakeholders whose needs are often overlapping, but also frequently conflicting. A large set of our stakeholders are faculty members. Opus will hold dossier information for roughly 15,000 faculty and other academic appointees such as librarians, project scientists, researchers, etc. Additionally, Opus must provide reporting capabilities for administrators such as department chairs, deans, various vice chancellors and administrative departments such as Development, Government and Community Relations, and those who report to the system governance boards, state and federal government, the public (via public information requests), and so on. Ease of access to consolidated university-wide reporting will save not only frustration and errors created by manual data collection, but will also aid decision-making and save time and money university-wide. System design can help university administrators monitor and communicate issues that are otherwise hard to locate due to shared leadership in complex organizations (Birnbaum, 1989). Faculty and administrators are often at odds in how they want faculty data to be represented, accessed, used, and maintained (Arreola, 2000, xxi).

2.2.1 Project Governance

The project was initiated as a response to a report by a joint committee comprised of faculty and administrative leaders. Following the launch of the project, a steering committee made up of faculty and administration representatives was formed to supervise and provide direction on project features, goals, and policy issues (see Steering Committee in Figure 1 below). Additionally, the Opus Team has participated in Data Governance Task Force, and a faculty-led Workgroup on the Access and Use of Faculty Data (see Figure 1 below). The team has also engaged dozens of campus data stewards, and UC-wide committees on identity management and faculty recruitment (not shown).
2.2.2 Our Role

We are two former iSchool students and current members of the IT Team that is building the Opus system. In some ways, our relation to Opus is irrelevant, as the issues that surfaced would exist regardless of our role. Yet our roles are at the same time central to the topic at hand, in terms of clarifying
our contribution and methodology, as well as highlighting the decisions that are frequently left to development teams.

Our functions on the team consist of performing detailed data, policy, and process analysis, and transforming those data, process, and design requirements into specifications for the developers to build. In many ways we feel as if we have a foot in both the academic and business worlds; on the one hand we examine the issues our project raises with a critical, academic eye - on the other hand we are faced with the persistent pressures, deadlines, and compromises that exist in any development project. We want to put this balance on display, to pull the curtain back on how these competing pressures get negotiated in a project, and more uniquely, a project at an academic institution.

In exploring the deep, project-specific issues above, we ultimately explore the broader issues of organizational communication and decision-making in a complex and decentralized environment. As we discuss the issues above, the questions our team has struggled with include:

- When do we raise an issue that may have implications beyond the technical?
- How can we move forward with technical development while policy issues are still being defined and negotiated - how do we balance delivery with deliberation?
- How can we engage decision-makers in the often-mundane details of system design? How can we better highlight the policy implications that often result from these micro decisions?
- We realized that many design decisions have policy impacts. How can we engage policy-makers to make these often mundane decisions rather than leaving them to the technical team

As Drucker (2009) notes in a discussion of scholarly research infrastructure, “the task of modeling an environment for scholarship...is not a responsibility that can be offloaded onto libraries or technical staffs. I cannot say this strongly or clearly enough: The design of digital tools for scholarship is an intellectual responsibility, not a technical task.” While Opus is not an environment for scholarship per se, we would argue that the spirit of this argument applies to Opus as system tasked with representing that work for the purposes of academic advancement and university reporting. In exposing our work, we hope to share lessons learned, and receive feedback and suggestions from the academic and technical communities.

Opus is in the process of beta-test launching a minimal set of functionality - much of the data that will eventually be collected has not yet been collected. We therefore present our preliminary results – our methodology and outcomes to date, which will of course be influenced by the feedback we receive in these early release cycles. In the Spring of 2015, readers and any other members of the public will be able to access http://opus.ucla.it.edu/. Those without a UCLA affiliation will have limited access to information beyond the public landing page. However, the public can click the “About” link in the header and the Privacy link in the footer, for the official version of the project's background.

3 Issues and Preliminary Findings

3.1 Data and Truth: Creating space for multiple narratives

The creation of Opus has advanced university discussions about data privacy, data ownership, and data governance. Faculty feel they should have the power to control how data about their activities is used (or reused, or not used), who can access it, and which data can be viewed by which audiences. A primary concern centers on how the data can be amalgamated and repurposed to tell new stories for malicious or misleading purposes. A related concern is that the data will be used to make comparisons about faculty productivity without considering context such as the individual’s field of study, workload, service load, etc. Boyd argues that transparency in itself is not enough if those in power can shape the interpretation of the data to tell a story that supports their interest: “information is power...[but] spinning the interpretation is even more powerful” (boyd, 2010).

While we knew we needed to use “authoritative” (read institutional) data sources, we did not want to privilege the institutional narrative above the faculty members’ own narratives about the data or record. In designing the Opus system, the Opus team worked with the UCLA Board on Privacy and Data Protection,
as well as with Opus’ faculty-led Steering Committee, to ensure that fair information principles were embedded in the design and representation of the data. These principles include displaying information about the source of the data, allowing faculty to submit data correction requests to data stewards, giving faculty control over how their information is displayed to the public (including other faculty members), creating processes for responding to claims of misuse and for reviewing data access requests to ensure they are driven by legitimate use cases (UCLA Statement on Privacy and Data Protection, 2011).

3.1.1 Methodology

When faculty, reviewers, or campus administrators view a record in Opus (for example, a record documenting recent grant activity), they see an information icon (ℹ️) next to the record. When the viewer hovers over the information icon, an information box pops up and displays information about the data. See Figures 3 and 4 below. In Figure 3, the “Instructions” would contain relevant data definitions as well as instructions and links to appropriate policy. Hovering over the ℹ️ next to “Information Incorrect?” on Figure 3 would display the name of the data steward, the source system, the data definition according to the source, the population covered by the source, and the date when the source began collecting the data (Figure 4). In exposing this information directly to the user, Opus exposes the institutional view of the data.

![Figure 3: Data as Displayed in Opus](image-url)
Additionally, where faculty feel data is erroneous, they may click on “Information Incorrect?” and submit the error resolution request directly to the data steward (see Figures 3 and 5). Faculty may also flag data as “contested” (Figure 5). Contested data will be represented as such to reviewers and those who have reporting rights.

Figure 4: Data Source Information Box

Figure 5: Data Error Resolution Request Screen
The process for data error resolution differs by data type. Information entered into Opus directly by faculty themselves will remain editable by faculty. This applies both to entirely self-reported records, such as an advising record about a student advised at another institution, as well as to data added by faculty to records supplied by institutional sources. For university service records, for example, Opus will collect the name of the committee, and the dates of service from UCLA’s Academic Senate system. This entry is a record from an institutional data source. Nevertheless, faculty may then augment this data with fields such as description, level of contribution, links or attachments, and/or comments. Faculty would have the ability to directly update any of the latter fields. For fields that come from the Senate, however, the faculty would have to correct an error via a request to the source system (see Figure 5). It is possible to request data corrections electronically via Opus, or the faculty may opt to contact the system owners outside of Opus if they prefer. Either way, Book of Record data gets corrected at the source, and then corrections are fed into Opus during the regular (usually daily) feeds. Figure 6 depicts this process. A “trusted source” is the source that Opus is accessing (such as a data warehouse copy of the original data), whereas the Book of Record source is the source at which data corrections must be made. An “AA” is an Academic Appointee, such as a faculty member, librarian, researcher, etc.

![Opus Data Flow Diagram](image_url)

Figure 6: Error Resolution Data Flow Diagram

Finally, the information added to institutional data by the faculty provides additional insight into the real world experience behind the data point or record. In this way we add faculty members’ own interpretations to the institutional one, creating a new hybridized record where the sources of truth include both institutional and faculty perspectives. This issue is significantly integrated with issues of data representation, which is discussed in its own section as Section 3.4. See screenshots in Figures 7 and 8 in that section for samples of these faculty-augmentable fields.
In determining which fields to add, and to understand the types of experience and context the faculty wished to add, we elicited faculty input via hosting user engagement sessions, faculty town hall meetings, and speaking at departmental meetings. As the faculty begin to use Opus we anticipate the need for additional improvements and hope that the current design contributes to the conversation.

3.2 Data Ownership: What does it mean to be an authoritative source?

From a functional perspective, the value of the Opus depends on a consistent set of expectations about data fidelity, security, and access (Wada, 2014). Data fidelity and security is addressed for component data by their respective authoritative book of record offices (for many types of data), and augmented by an Opus-defined error resolution process that faculty can use to make corrections at the source. For other types of data, such as data about the academic review process (the dates on which a case was approved by each reviewer, etc.), the Opus system itself is an authoritative source.

While stewardship of some data seems relatively clear cut, for example few would dispute that the Registrar is the steward of course data, the line between ownership, stewardship, and accountability can get quite blurry. Agre states that data requires transparency about ownership, arguing however, “to speak of ‘owning’ data glosses a variety of things, from trade secrets to research ethics to contractual constraints on its use” (1994). While we have already outlined our efforts at making data ownership transparent, we feel that further reflection on what makes a source “authoritative” is warranted given the current drive towards enterprise data management and standardization. Consider the following examples:

A. Campus organizations often first learn about faculty activities from faculty themselves, for example, the Office of Contracts and Grants Administration (ORA) receives a grant application from a faculty member which provides the names of investigators or collaborators, the grant title, the funding request, etc. The university then administers the grant, and the information provided by the faculty member is augmented and stored in a university owned system, which now becomes the “authoritative” source for the data. Faculty no longer have direct access to edit this information but must instead submit updates or corrections to the ORA office, which will then has the authority to approve or deny updates.

B. Let us return to our enrollment data example. For multi-instructor classes, one office calculates enrollment for each instructor based on an algorithm involving section numbers (of cross-listed courses) and relative ranking of status among instructors. The calculation is designed for the purpose of showing how tuition funding was allocated to various departments; it does not answer the question of how many students were in the room or how much actual effort each instructor expended. For a 200-student course with 3 instructors, one instructor might get enrollment credit for 190 students while another gets credit for 10 students and the third receives no credit for enrollment at all. Sometimes the instructor credited with the brunt of enrollment by this algorithm is a figurehead overseeing the class who never actually attends the class, writes an exam, or grades an assignment. None of these scenarios are well-reflected by the existing enrollment figures, and one can assume that for purposes of academic review, all 3 instructors would wish to be credited for the full 200 students enrolled in the course – after all the work it took to read papers and prepare for lectures was based on the number of people in the room, not the number the university reported to the state. Nevertheless, allowing faculty themselves to report unverified enrollment opens the university to lawsuits of unfair promotion practices. Legality aside, this places the burden of determining which number is “fair” on each individual faculty member, advantaging those who would mark the higher of two possible counts when any gray area exists - over those who would mark the lower number out of politeness, or fear of overstating, or confusion over the definition of the criteria of what counts as an enrolled student.

We seek to find a way to tell both the individual and institutional stories of the data without privileging one voice over the others. Initially, we will categorize data records as either self-reported or as coming from a campus system. As Opus allows faculty to both augment records that came from authoritative sources and to add new records that were not captured by institutional systems of record, we need to distinguish for viewers of the data between self-reported data and data that comes from campus systems. See Figures 7 and 8 below for examples of self-reported data that may be added to records that originally came from a campus book of record. This means there are three kinds of records in Opus: 1) records that are entirely owned by campus books of record, such as the campus identification numbers for faculty
and staff; 2) book of record data records that have been augmented with self-reported data, and 3) records that are entirely self-reported. In practice, self-reported data are often viewed as less accurate, suspect because they cannot necessarily be independently validated, and yet it is really often the individuals themselves who can speak with the most “authority” about the event or work.

3.2.1 Methodology / Contribution

One of our most important lessons learned is that who is “authoritative” is context-specific, rather than enterprise-specific; rather than choosing a single authoritative source for a specific type of data, the enterprise may instead require multiple “authorities” depending upon the purpose for reuse. Our team’s first step was to identify data stewards and systems of record for the data we needed to collect. With the sponsorship of the Office of Information Technology, and the Academic Personnel Office, a meeting of campus data owners was convened to discuss issues of ownership, outline a methodology for identifying data stewards, and to gain support for the process and work involved in identifying the correct data for Opus’s purposes. It was agreed that Opus would be the pilot project for a larger initiative to identify campus data stewards.

Following these discussions, we met with each relevant office on campus (Registrar, Research Administration, etc.). Our approach was to understand the story and the journey of the data, mapping the flow of data through university offices and systems. We worked with data stewards to understand why each data element was created, how it was created, who used the data, for what purposes, and what meaning was embedded in it. In this way we were able to follow the data and ensure not only that was the data called the same thing (e.g. enrolled students at the end of the third week of the quarter), but that the interpretation was appropriate (all students in the room as opposed to only matriculated students in only a particular section).

Next we worked with data source system teams to get access to the data. In doing so, additional questions of interpretation, and understanding each office’s “language” surfaced, often requiring clarification from Opus stakeholders about the specific information they were after. Finally we will be working with each of these offices to develop the error resolution process discussed in section 3.1.

3.2.2 Outcomes

We continue to work with these offices to resolve questions and ensure we collect the most appropriate data for the project. We have also begun to outline the responsibilities of data stewardship at UCLA, identifying a need for a shared central repository of data dictionaries and data models, a list of campus identified data stewards, guidelines regarding what constitutes valid data, and identifying an approach to determine access to data that is based on the type of information represented. The UCLA Taskforce on Data Governance is utilizing this documentation to further its own work towards campus data governance. As with our attempts to design a system that allows multiple narratives to emerge from the data, we anticipate that once Opus is fully released, it will serve as a catalyst for ongoing discussions about data ownership.

3.3 Balancing Data Access and Privacy

Another aspect of ownership is control over how data is represented, accessed, and reused. A useful distinction here is Eschenfelder and Johnson’s “access control” versus “use control” (2014). While the Opus Team and the Data Governance Group are nearing agreement on our access control mechanisms, we anticipate future work on use controls – even if a campus office has legitimate access to the data in Opus, how are they permitted to use the data? Will Opus only display the data in aggregate to this user? How can we create a policy around data use that is flexible enough to be applied to the myriad uses people may have for Opus data, without being overly prescriptive or predictive of those uses? As Borgman argues, stakeholders often have overlapping and/or competing interests and incentives around how the data are collected, used, and interpreted (2013).
In addition to inheriting access controls from source systems, Opus may also need to address data stewards’ own concerns about how the data may be interpreted. The Office of Instructional Development (OID) is responsible for collecting student evaluations of courses. OID has requested that if evaluation data is used in Opus for the purpose of academic review, that Opus allow OID to add a disclaimer about the purpose for which the data were collected, and the problems that could arise with reinterpretation and reuse.

### 3.3.1 Methodology/contribution

Initially, we presented on Opus to the UCLA Board on Privacy and Data Protection. The group raised a number of important questions and we convened a shared group made up of members from the Privacy Board, the Opus Steering Committee, and the Opus Team. In laying the groundwork for policy discussions, we (the authors) first reviewed relevant university, UC system, state, and federal policy to understand which data already fell under existing protections and which were expected to be public, given the university’s status as a public institution. We looked as well at any precedents for how to balance the faculty’s right to control how information about them is represented, with the decision-making needs of the university that rely on this data. In many ways the process thus far has been an act of discovery, particularly for faculty, bringing to light the fact that these data are often already stored in various shadow systems, and used without faculty consent – in both policy and practice.

The analysis of existing policy was then brought to our Access and Use of Faculty Data Workgroup, as well as to the Opus Steering Committee. In addition to faculty, these groups included representatives from the campus records office, general counsel, and other administrative units that have an interest in using Opus data for reporting purposes. Based on the Opus team’s analysis, the workgroups helped to outline data visibility standards for the project. While much progress was made towards understanding which data would be “public” there is still much to be discussed regarding how data may be used by entities internal to UCLA.

### 3.3.2 Outcomes

As a result of groundwork laid by the Opus project in concert with the Office of Information Technology, UCLA has charged a Data Governance workgroup to outline guiding principles for managing access and usage of campus data whether about students, faculty, or staff.

One of the governance group’s realms of responsibility is data access. While these conversations are ongoing, our preliminary solution has been to allow faculty to control the data displayed on their public profiles. We have opted to show a minimum of data by default (name, title, department), and to never allow other data, such as data about students to be displayed (e.g. the names of advisees) on the public-facing profiles. For all other records, the decision to make the record public is at the discretion of each faculty member. For administrators who desire access to the data, we follow existing data access policy which states that those who have a legitimate business case to use the data will be granted access. We anticipate working further with the Data Governance workgroup to describe “legitimate business cases,” to formalize a process for requesting and evaluating requests to access, and to address any additional concerns that may arise.

For intellectual property purposes, we have agreed not to allow public or reporting access to information about grant proposals or patent applications (awarded patents and grants, on the other hand, may be shared publicly). Faculty may opt to share this information in the context of academic review, but will be restricted from making it “public” via their profile pages. We have also agreed to work with various offices to insert help text, disclaimers and links to additional information that can aid in interpretation. We will attempt to follow Borgman’s guidance to bring competing interests and interpretations to light by creating detailed documentation about our process of data collection and reuse, policies that adhere to data, and access constraints in a help section available to all Opus users. We are also engaging in a communications strategy exposing these ideas to faculty and the campus at large.
3.4 Data Representation

In her presentation, “Transparency is not enough”, boyd claims that transparency alone does not create an informed citizens, and when an end in itself, can actually create unintended consequences (2010). When providing access to data, boyd wants us to think through not just the issue of transparency, but interpretation and to recognize the need for data literacy. In Opus, though we realize our data schema, and interface are themselves interpretive acts, we strive to allow our users to understand the context of the data, and to allow the faculty to augment and add context to data about them. One major challenge for the project team has been the display of metadata in a way that allows us to be transparent, without being overwhelming or misleading.

3.4.1 The Representation of Multiple Truths

We discussed in section 3.1 the importance of allowing multiple truths. The problem that follows from that conclusion is the need to display the source of each truth to the viewers in a way that is both transparent and easily understood.

3.4.1.1 Methodology

To allow faculty a vehicle to speak to these issues within the dossier itself, we enabled faculty to contextualize data via multiple comments/notes fields, as well as more specific free-text fields, a best practice recommended in Theall (2005). For example, manually-entered Service entries can be described via not only an Additional Comments field, but also through the Description, and Level of Contribution fields, linking to other content, and uploading files that provide additional information (Figure 7). It is possible to annotate one’s case at the overall dossier level (Figure 8) as well.

Figure 7: Record-Level Notes and other Contextualizing Fields
Narrative areas of the dossier (not shown) also lend voice to context: the Self Statement, Contributions to Diversity, and Teaching and Curriculum Development.

### 3.4.2 From Information to Data Points

Annotation opportunities aside, some faculty have concerns about the use of quantitative data to measure their performance in the first place. Take the contentious idea of publication and author influence rankings, for example. Dennis et al. recommend that review committees lower the threshold expected of junior faculty for publications in elite journals, in the Information Systems field. Cameron, on the other hand, takes a stronger stance, summarizing that the use of bibliometric information in tenure promotion evaluations can be “misleading and prejudicial.” Limitations include the narrow time-lag between the work itself and later citations of it which emphasizes certain types of work, a preference towards disciplines with many journals, incomplete selection of journals or publication type included, bias toward larger journal size and English language journals, and more (2005). Shockingly, 90% of journal articles are never cited, according to Meho (2007). Like academic review, Van Raan also suggests that using such data for university administration decisions is equally ill-advised (2005). Others agree with the disadvantages of current bibliometric data, but call for improving the formula underneath the metric (e.g. Meho, 2007) rather than avoiding bibliometrics altogether. Then again, others have found bibliometrics to be correlated with peer review judgment (e.g. Van Raan, 2006). UCLA faculty have expressed worries similar in spirit about quantifying, and thereby making rank-able, aspects of faculty performance as well. On the organizational side of the coin, it is exactly the monitoring of numeric measures that yields performance improvements (e.g. Arreola, 2004; Becker & Huselid, 1998). A final group urges those consuming the data – particularly the group of faculty conducting a review their colleague - to be responsible for actively evaluating the value of data appropriately in its context (e.g. Theall, 2005). There is another debate about the use of student evaluations in faculty performance reviews. One can imagine a faculty could feel similarly about dossier data on any other topic as well.

#### 3.4.2.1 Methodology

In the case of Opus, the decision was made by a faculty-led Steering Committee and the Academic Personnel Office to collect faculty performance metrics. While the new decision makes a few changes to the prior template, the brunt of the metrics gathered in Opus are those of long-standing university policy. In short, we recognize concerns with reducing diverse and complicated work to standardized data points.
The decision whether or not to use metrics, we believe, belongs with the faculty and administration, rather than the IT team. As the IT team, we have attempted to implement a balanced approach, given the charge to use specific metrics. We are open to suggestions from the community on further improvements.

In a few circumstances, we have requested new, more appropriate, metrics be created and hosted by an existing data steward, and/or brokered the negotiation ourselves for a shared understanding of data items critical to Opus. One example of a new metric was our request that the Registrar begin recording an enrollment snapshot of both matriculated and non-matriculated students at the end of the third week.

3.4.3 You say categories, I say categorization

Another issue that surfaced was the mapping of controlled vocabularies. Opus has determined a controlled vocabulary for publication and scholarly or creative work types. When importing publications from another campus system, which uses its own controlled vocabulary for publication types, we were faced with unclear or outright incompatible options for mapping one type to another.

3.4.3.1 Methodology

Our decision is to map the straightforward types (source system’s publication type of “Book” translates to Opus’ “Book or Catalog” type), also map the likely correct mappings to their Opus counterpart, and lastly map the incompatible types to a temporary type of “Unknown”, which the user will need to correct at the time of import. In addition, during the import process, faculty will see an explanation of our mapping process and our request for them to verify the mappings done. Figure 6 shows our idea of how this would appear to the faculty.

![Figure 9: Prototype of Import Verification Screen](image-url)
3.4.4 Shared Terminology Across Diverse Cultures

Differences in controlled vocabularies are part of a larger topic, namely (pun intended) that of shared terminology. Differences in definitions for the same term, and use of different terms for the same idea across campus exposed a need for shared institutional data definitions. One term with different meanings across UCLA is “awards”. To some, “awards” implies honorary awards, such as a Nobel Prize, while to others “awards” signifies awarded grants.

3.4.4.1 Methodology

Where semantic differences have been discovered, Opus selected neutral terms. Instead of “awards”, Opus has one section called “Honors” and another section called “Grants.” The Opus glossary will be accessible via the application as its own document, as well as in-line where the terms are used. Definitions of terms, especially potentially contentious definitions, have been arrived at through discussion and negotiation with data stewards for enterprise definitions, and with departmental administrators regarding local usage of terms “in the wild”, a rather elaborate effort - as is often the case (Glushko et al., 2007). Opus will provide its consolidated list of definitions to the data governance board as a starting point for comparison. Within Opus itself, we display help text defining the terms we use, making those definitions easily available at the time and place in which the faculty and reviewers will encounter them.

4 Conclusion

Fairness, transparency and data ownership are not clear-cut concepts. Even when a vision is shared among stakeholders, implementing that vision via system design adds its own constraints. We have wrestled with these issues on many fronts while designing Opus and proposed our best efforts at mitigating them. As with any complex issue, each possible solution has tradeoffs, particularly considering the number of parties and moving parts involved in a project of this size. Further work for our specific project will entail responding to upcoming user testing and, later, to feedback after early roll-out, and then wider roll-out. Our purpose has been to expose some of the issues faced by a technical team developing in a university environment; each topic we discussed deserves additional research. Wider work could propose best practices regarding combined profile and review data and/or aggregated plus manually-augmented data in public university setting. Another path for expansion on this work would be an analysis detailing the tradeoffs for possible solutions.
References


**Table of Figures**

- Figure 1. Opus Project Structure ............................................................................................................. 4
- Figure 2. Opus Project Workgroups ........................................................................................................ 4
- Figure 3: Data as Displayed in Opus ......................................................................................................... 6
- Figure 4: Data Source Information Box ................................................................................................... 7
- Figure 5: Data Error Resolution Request Screen ..................................................................................... 7
- Figure 6: Error Resolution Data Flow Diagram ....................................................................................... 8
- Figure 7: Record-Level Notes and other Contextualizing Fields ............................................................. 12
- Figure 8: Case-Level Notes ..................................................................................................................... 13
- Figure 9: Prototype of Import Verification Screen ..................................................................................... 14