

Academic Analytics at CU-Boulder

September 2013 Lou McClelland, Planning Budget and Analysis – IR

Academic Analytics (AA) is a private firm providing data on faculty scholarly activity to PhD-granting universities in the US and UK. Founded in 2005 by then-graduate dean Lawrence Martin of Stoney Brook University, as of July 2013 AA had 35 AAU and 90 total subscribers plus over 70 employees working from locations throughout the US and UK. Website: <http://www.academicanalytics.com/>

The AA database links over 300,000 individual faculty members in almost 400 PhD-granting institutions both to a common disciplinary categorization and to comparable records of journal articles, citations, books, conference proceedings, federal research grants, and honorific awards. These records are collected by AA itself from publishers, granting agencies, and awarding bodies.

AA's data and comparisons emphasize

- recent scholarly activity – generally in the prior four to ten years – rather than cumulative activity over a scholar's entire career;
- past scholarly activity rather than stated interests and plans;
- comparisons within disciplines, to accommodate different disciplinary patterns in scholarly activity;
- annual updates;
- comparable collection and summarization methods across disciplines and institutions.

The AA comparative database is released annually early in the calendar year based on lists of faculty active roughly 15 months prior – e.g., January 2014 release for faculty as of November 1 2012. Standard comparisons are based on 4 years of articles and proceedings, 9-10 years of books, 5 years of grants, and as many as 50 years of honorific awards. The comparative database is released to paying clients only in the AA online portal and in custom files. Comparisons are at multiple levels: departments, PhD programs, individual faculty members within disciplines, broad fields, and entire schools. Comparisons list volume (e.g., number of books), percentages of faculty with a type of activity, and per-capita measures plus an index combining standing relative to the discipline or broad field on multiple measures.

AA's larger database adds records of scholarly activity data daily – for example, articles from a new issue of *Science* are collected immediately upon publication. The database also holds records for many publications and books as early as 2000, and for article, book, and grant types not included in the comparative database – e.g., book reviews and training grants.

In March 2013 CU-Boulder contracted with AA for our second five-year subscription, extending to February 2018. Our "master license" subscription gives us

- Matching: AA's pledge to match their records of scholarly activity to all CU-Boulder faculty members we send. Our contract allows for us to send 2,350 individuals each year. Prior to September 2013 AA had matched only those listed as holding some level of "professor" title.

- Current and past detail data: Access to and permission to use *and publish* all AA information on CU-Boulder individuals (e.g., lists of publications, honors), departments, PhD programs, and groups of individuals such as research institute faculty. This access includes CU-Boulder aggregate and detail data (e.g., individual journal article title-authors-publication specifications) from the comparative database plus detail data on earlier and more recent scholarly activity from the larger AA database.
- Comparative data: Access to comparative data at other institutions for our own internal use.
 - Access can be through the portal or custom files. For example, for 2013 we plan custom files to allow comparisons to specified groups of peer departments that cross disciplinary lines, with differentiations not now possible in the portal. *Status: Modification of files received in 2012; CU-Boulder input needed.*
 - Market share: The portal also displays selected discipline-level data comparing our department or program to all others on articles in specified journals, and grant funding from specified agencies. Downloads from the portal present more information. E.g., 2004-2011 the UCB aerospace engineering faculty published most frequently in three journals, *Journal of Geophysical Research*, *Journal of Guidance Control and Dynamics*, and *Geophysical Research Letters*. Our 30 faculty members (2% of all in the discipline) were responsible for roughly two-thirds of all articles *by aerospace faculty anywhere* in the two geophysics journals, for 12% of the articles *by aerospace faculty anywhere* in the third.
 - We may not make public other schools' data or precise relative standings, or those of individuals at other schools, in a document or discussion or presentation that may be covered by the press and made public. We can do so in internal documents, including those made available to external reviewers. We can in public documents list, e.g., the top 10 (or 20 or whatever) schools on a measure, or the 10+ with a measure value close to ours.
- Access to AA custom reporting and consultation, and to additional AA products before they become generally available. Products and special reports available or in the works now:
 - Semantic similarity: Facility to select and list schools, broad fields at schools, disciplines at schools, and individuals, both within CU-Boulder and elsewhere, based on the "semantic similarity" of the titles and abstracts of their published work to a text input by the user. This is currently called "expert picker," but that understates its potential utility. *Status: Beta. Possible: custom files with "semantic signatures" of individuals.*
 - Following PhD recipients: Data on the scholarly activity of CU-Boulder PhD recipients (from as long as 50 years ago) found on the AA database. *Status: Beta. UCB work proceeding.*
 - Data feeds: Development of a regular, frequent (e.g., monthly), automated feed of recently-collected AA data to CU-Boulder Faculty Affairs for use in FIS and VIVO. *Status: In beta at some schools; to be arranged at UCB.*
 - Collaboration: Files of person-to-person and unit-to-unit links based on co-authorship, perhaps citations, semantic similarity of work, or other communalities, for use in creating network diagrams for understanding past and potential collaborations. *Status: Beta. Files not yet received.*

- Custom comparative files: AA will work with us to design and create custom files to improve the utility of comparative data by such measures as mixing departments and PhD programs, defining other aggregations of individuals at UCB and elsewhere (such as research institutes), mixing disciplines, limiting on faculty-member characteristics, calculating market shares for journals or granting agencies regardless of discipline, and more. *Status: Available, limited by our and AA's time, creativity, and knowledge of potential utility for clients.*
- Two visits of AA staff to CU-Boulder, upon request
- CU-Boulder obligations to acknowledge AA as the source of data, follow the non-disclosure rules, manage access to the AA portal, do beta-testing, and clear public statements with AA staff.

Faculty lists

- AA gets lists of faculty members from subscribers and from web and catalog listings for PhD programs, in all cases seeking faculty members whose positions carry an "expectation of scholarly activity."
- At CU-Boulder we cast a wide net in identifying CU-Boulder faculty to list, because we have many non-tenure-stream faculty with scholarly responsibilities and activity and because we can use custom files and filters as necessary to limit their participation in comparisons.
- We list all tenure-stream individuals (including those currently in administrative positions), plus others in job titles of assistant-associate-full-distinguished professor in adjoint, clinical, and research lines (but not adjunct or visiting), plus others in the two job titles research associate (RA) and senior research associate (SRA).
 - Adjoints are typically employees from the federal labs active in Boulder campus research and teaching. Example: Eric Cornell
 - SRA and RA titles are included based on their significant participation as PI's on Boulder campus contracts and grants (see e.g. <http://www.colorado.edu/pba/facstaff/research/jobcode11.htm>, showing one-quarter of SRA/RA's as PI's with \$67 million in grants).
- Next issue, by 11-1-13, any others to add to the 2,098 already sent for as-of November 1 2012. Possibilities:
 - Museum curator adjoint, N 32, jobcode 1424, and/or museum associate, N 12, jobcode 1427, both unpaid. One-quarter list addresses outside Colorado. Most have only one appointment.
 - Three individuals omitted in error
- AA generally associates each individual in a year with one or more departments and PhD programs, rank (professor, associate, assistant), tenure-stream status and tenure status, administrator flag, and a categorization differentiating tenure-stream, clinical, research, libraries, and other faculty types. AA also collects year and school of the terminal degree.

How it works at CU-Boulder

- Provost and CFO split the cost of the contract, with payments due each November 1 to 2016. Planning, Budget and Analysis (PBA) handles logistics.
- Provost and CFO also funded one 40-50% hourly position to be UCB AA liaison and facilitate use. That's Lou McClelland after retirement as IR director 2013.
- PBA-IR submits and reviews the faculty lists. To date only Lou and Rob Stubbs of PBA-IR are authorized to access the AA portal.

- Lou has been on the AA advisory board since it formed in 2008, along with Lydia Snover of MIT and Julie Carpenter-Hubin of Ohio State, joined by David Jamieson-Drake of Duke about 2011 and Michael Monticino of the U. of North Texas in 2013. The board meets 2-3 times per year in person, occasionally by phone, and members work directly with AA officers on new products, needs, issues, and feedback. CU-Boulder got a discounted contract rate to acknowledge this work.
- The only prior CU-Boulder use of AA data was in November 2012
 - PBA-IR delivered comparative data for all departments through a CU-Boulder website accessible only to those completing a non-disclosure agreement.
 - This was announced only to units whose Academic Review and Planning self-studies were due soon.
 - Description of process, product, reactions, lessons posted from <http://www.colorado.edu/pba/peer/acanalytics.htm>
 - Main bottom lines
 - Discipline/peer mix-ups were crippling in some cases
 - Users need more flexibility and tutoring
 - Faculty Affairs (and the Academic Review and Planning process) has no desire to require units to use AA data at all, or for each unit to use the same way
 - What we delivered used a fraction of what we've bought
 - Portal access for units is a mixed bag – risks, time
- Now I am working on
 - Understanding what we bought
 - Understanding and developing my role
 - Understanding who the UCB clients are (*more on this below*) and what AA data will serve what clients
 - Handling incoming files, with feedback to AA – e.g., PhD tracking
 - Checking what we've sent
 - Continuing service to Academic Review and Planning by building on 2012 posting but modifying and expanding and working 1-1 with units
 - Tracking and communicating UCB needs, issues to AA
 - Setting up an internal advisory/user/client board
- UCB clients
 - Academic Review and Planning
 - Academic units in upcoming Academic Review and Planning cycle, with issues
 - Libraries – We send Library faculty for matches but there are no standard comparisons for them; library science (which is different) is not even a discipline in the AA taxonomy.
 - Journalism/Mass Communication – Taxonomy issues, differentiation from Communication, fit into disciplines
 - Business – Peer and discipline mix, level of differentiation (“business” vs. marketing, finance, etc.), PhD vs. master’s and MBA’s
 - Education - Peer and discipline mix, level of differentiation (“education” vs. curriculum, research methods, etc.), PhD vs. master’s
 - Law – No corresponding taxonomy entry
 - Museum- No corresponding taxonomy entry; affiliated TTT faculty are from several departments
 - From 2012 – SLHS
 - Plan to focus on Business, Education, and SLHS first, working with units to define useful files/data

- Others (not in order of priority!)
 - FIS and VIVO through Faculty Affairs, for individual person/publication data and data feeds
 - For reviews, profiles, new program proposals, analysis of individuals, collaborations, PhD tracking, more
 - A&S dean, divisional deans
 - Engineering dean
 - Graduate school
 - Department chairs in A&S and Engineering
 - Deans of other schools/colleges
 - Research institutes
 - For the whole campus
 - Libraries
 - U Communication – e.g., total volumes, where we fit
 - VC Research
 - Provost
 - PBA?

Standard comparisons through the AA portal are best for

- Departments 1-1 with PhD programs both at UCB and in peers – *History, not environmental studies or women and gender studies*
- Departments where the PhD program plays a major role both at UCB and at peers – *Not Business, Law*
- Departments/PhD programs that fit clearly into one and only one discipline at only one AA taxonomic level both at UCB and in peers – *History, not SLHS, not Communications, not Education*
- Departments where non-TTT play an insignificant role in scholarly activity, both at UCB and at peers – *Art and Art History, not Chem/Biochem or MCDB or Physics or Psychology*
- Departments whose scholarly activity is captured by articles, citations, conference proceedings, books, federal grants, and honors, both at UCB and at peers – *Not Ethnic Studies, Music, Theatre/Dance*
- Departments whose sponsored research is predominantly federal both at UCB and at peers – *Not Electrical/Computer Engineering, Sociology*
- Departments with at least 10 faculty members – *Not Linguistics, French/Italian*
- Departments with a large proportion of faculty members at least 5 years post PhD both at UCB and at peers – *Maybe not Asian Languages and Civilizations*
- Departments where book chapters are not a significant part of scholarly productivity either at UCB or at peers – *Many humanities disciplines? Book chapters are coming from AA but have not been delivered yet*
- Departments emphasizing journal articles that are happy to count publications in all journals the same, and are happy with AA journal coverage – *A facility to weight by impact factor, or select journals, may be coming from AA; custom files could do this. AA readily accepts suggestions for additional journals, honors, etc.*
- Departments emphasizing honors that are happy . . . as above.

Comparisons through the AA portal with modified peer groups (modified in the portal) can help with

- Departments in disciplines with departments qualitatively different from one another in ways affecting scholarly activity patterns – *E.g., Anthropology (with physical-cultural split), Psychology (with proportion clinical varying from zero to 100%)*

- Departments who meet the criteria for standard comparisons above but want comparisons to a subset of institutions represented in the discipline. *The modified peer group facility can handle this now, with a planned attack.*

Very long term possibilities requiring AA action

- Federal labs in even though don't grant PhD's
- Everything through the portal
- Send and track all our PhD graduates
- More!