

Exposing and Exploring Academic Expertise with Virtu

Luanne Freund, Kristof Kessler, Michael Huggett, Edie Rasmussen

University of British Columbia, Canada

<http://www.diigubc.ca/virtu>

virtu@diigubc.ca

The Virtu academic expertise finder system is built on the Apache Solr platform using data from the Mendeley social network and bibliographic management system. Virtu takes a task-based approach to expertise, exposing and giving the user control over dimensions of expertise that can be applied to a variety of expert-finding tasks. The search interface supports information interaction and exploration through a number of browsing and filtering tools, including facets and sliders.

Searching for Experts

A search box retrieves candidate profiles based on an Expertise Score, defined as the average of —

- **Knowledge:** the sum of number of publications weighted by the popularity of the publication venue (40%), years of experience (40%), and the number of disciplines in which the user has published (20%)
- **Reputation:** the sum of number of readers weighted by the readers' academic rank (30%), the average number of co-authors per paper (30%), the number of contacts (20%), and the number of groups joined weighted by group size (20%). The user also receives a 20% bonus to their score if they have served as editor of a publication
- **Profile Completeness:** the sum of the size of the user's research interests (40%) and of the user's biographical information section (60%).

Filtering the Best of the Best

The results set of candidates can then be filtered along the dimensions of expertise, which are based on a set of basic measures —

	Evidence
Subject Knowledge	Number of academic publications (of types Book, Book Section, Conference Proceedings, Encyclopedia Article, Journal Article, Thesis)
Applied Knowledge	Number of publications in all non-academic publication types.
Experience	Number of years since date of first publication.
Reputation	<ul style="list-style-type: none"> - Number of readers in same main discipline as author - Publication outlets, ranked by number of readers - Number of professorial versus non-professorial readers of the user's publications
Connectedness	<ul style="list-style-type: none"> - Number of contacts - Number of co-authors per publication - Number of groups joined - Total number of members in groups joined
Multi-disciplinarity	<ul style="list-style-type: none"> - Number of disciplines - Number of Sub Disciplines - Percent of total readers from disciplines other than the author's

User Interface

The components are -

1 Query input: A drop down box offers the option of searching for a person, for expertise in an area, or for all matches (default).

2 Discipline Facets: The main disciplines listed in profiles are presented as hierarchical facets, grouped into broad categories at the top level. Checkboxes allow the searcher to select one or more disciplines as limits on a keyword query, or to browse through profiles within a given discipline. The discipline facets are dynamically updated for each query and are combined as a Boolean *OR* search if multiple disciplines are selected.

3 Expertise Sliders: The six expertise dimensions are manipulated directly using dual-control sliders that set lower and upper bounds on each measure. The underlying values for each dimension are calculated as a linear combination of basic measures and mapped to the range 0 to 100. The median value of each dimension is set to the mid-point of the slider. Search results are updated as soon as a slider handle is released.

<http://www.diigubc.ca/virtu>
virtu@diigubc.ca

The screenshot displays the virtu1 search interface. At the top, a search bar contains the text 'ecological design' and a 'GO' button. Below the search bar, there are two profile cards. The first card is for 'Marcus Blake' and the second is for 'Brandon Tearse'. Each card includes a profile picture, name, publication activity, main discipline, sub-disciplines, and a list of top publications. The interface also features a 'Reset' button and 'Results 1 to 3 of 3' indicator. The search results are displayed in a grid layout with numbered callouts (1-5) highlighting specific components: 1. Query input (search bar), 2. Discipline Facets (checkboxes for Formal Sciences and Computer And Information Science), 3. Expertise Sliders (six sliders for Subject Knowledge, Applied Knowledge, Experience, Reputation, Connectedness, and Multidisciplinarity), 4. Profile (index card for Brandon Tearse), and 5. Find Similar and Related (links for finding similar and related profiles).

4 Profile: Search results are displayed as index cards that include basic profile information such as name, publication activity, main discipline, sub disciplines derived from publications, the user's top three publications, and an image and title if available in the dataset. Links to the user's Mendeley page and the pages of the publications are also provided. Each profile shows the user's expertise score components as six horizontal bars representing the score for each dimension.

5 Find Similar and Related: Two links provided in each index card trigger prototype queries based on the selected profile of interest. *Find similar* retrieves profiles in the same discipline that have similar values on the expertise measures (+/-10 %), by adjusting the values of the Expertise Sliders. *Find related* uses data on co-authors, connections and groups to retrieve profiles that are most closely linked to the profile of interest.