

# MPI Workshop Abstract | May 2024

**Project:** One Algorithm to Rule Them All

## Problem Description

The Graduate College at University of Delaware oversees several initiatives and programs aimed at advancing the academic and research success as well as personal and professional development of our 4,200 graduate students and postdoctoral trainees. One of our signature programs is an alumni mentoring program called Grad Leveraging Engaged Alumni Program (Grad LEAP) (Figure 1). In Grad LEAP, we recruit volunteer alumni to mentor current graduate students with the hope that the alumni mentor will add to the constellation of advisors in the students' support network. Grad LEAP mentors have provided guidance to their mentee(s) on a variety of academic, professional, and personal topics. Interested participants complete a matching survey to share a range of information that will be used in the matching process. For example, they can share their academic program, career interests/current career field, lived experiences, and what they look for in a mentor or mentee (Table 1). Mentors and mentees are paired using a proprietary matching algorithm created by the education technology vendor we contracted with. The algorithm proposed matches by weighing mentee's top matching criteria and their goals for joining the program (Figure 2). The proposed matches are reviewed by the Graduate College's Grad LEAP administrators who determine if a match is approved or requires rematching. For approved matches, the mentor-mentee pairs are notified, and can start meeting in-person or virtually. For matches requiring rematching, Grad LEAP administrators can (a) suggest specific names of mentors from the matching pool to be matched with the student or (b) indicate the administrators' preferences for matching. The matching process concludes when all prospective mentees in the program are matched. It is possible that a small number of

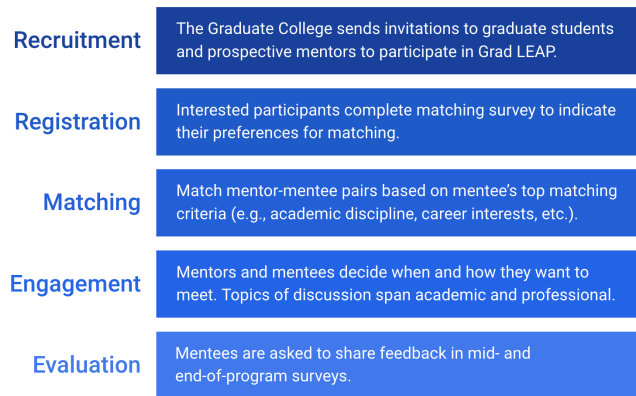


Figure 1. Overview of Grad LEAP

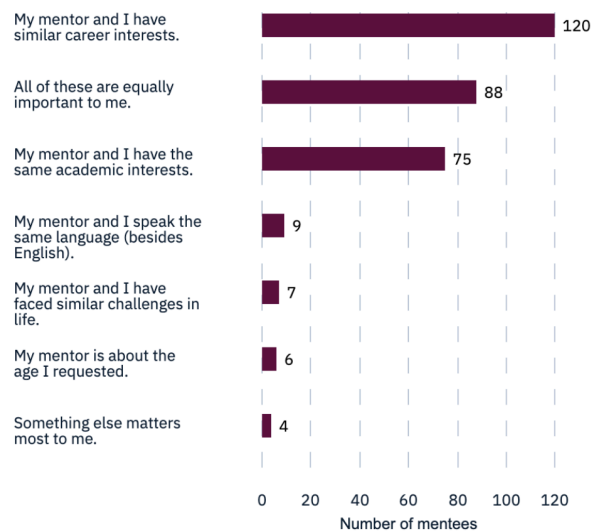


Figure 2. Mentees' top matching criteria as indicated in the matching survey.

mentors who signed up will not be matched. Mentors and mentees can also request to exit the program or be rematched at any time.

Table 1. Matching survey questions and sample responses.

Question	Response
What is your field of study?	Biological Sciences
What career fields are you interested in?	Management, business, & finance; Health & medicine; Teaching, college or university level
Do you identify with any of these terms?	LGBTQIA+, Neither of my parents or guardians completed graduate-level education
What challenges have you experienced?	Supporting family financially, Growing up in a single-parent household, Working and studying at the same time,
What language(s) do you speak?	English
What would you like to discuss with your mentor?	Learn about career options
What are your top mentor preferences?	My mentor and I have the same career interests.
Is there anything we should know when searching for your mentor?	As a international student, I like to be matched with a mentor who was also an international student so that they can guide me in International student related questions, too.

As we grew Grad LEAP from 200 to 300 mentoring pairs, we found that the proprietary matching algorithm was yielding a success rate of approximately 30%. Meaning, only 30% of the proposed matches met our criteria of being good matches. Consequently, we had to manually match the rest of the cohort by hand. We define “good matches” as matches that reflect the preferences of the mentors

**Table 2.** Example of proposed matches from the proprietary algorithm. The first column shows categories of data from matching surveys. The second column represents a proposed match between Rosa (mentee) and Sam (mentor) that needs to be rematched because of misalignments in academic discipline and career interests. The third column represents a proposed match between Alex (mentee) and Yang (mentor) that was approved because they are from closely related academic disciplines, share similar career interests and lived experiences (indicated in the “identity” column).

Student Alias	Rosa	Alex
<b>Student Matching Priority</b>	My mentor and I have the same academic interests.	My mentor and I have similar career interests.
<b>What do you hope to gain from joining the program?</b>	Receive internship, career, or job search advice	Receive internship, career, or job search advice
<b>Student Academic Program</b>	Computational Science & Engineering	Energy & Environmental Policy
<b>Student Career Interest</b>	Computers & technology, Data Science	Renewable energy
<b>Student Identity</b>	Being an international student	Changing advisors or degree programs, Unsure about career
Mentor Alias	Sam	Yang
<b>Mentor Academic Program</b>	Languages, Literatures, and Cultures	International Development
<b>Mentor Career Interest</b>	Teaching, K - 12, Applied Linguistics	Law and Public Policy, Government, Renewable energy
<b>Mentor Identity</b>	Being an international student, English is not my first language, Neither of my parents or guardians completed a two or four-year college	Unsure about career

and mentees based on the responses they provided in matching surveys (Table 2). For example, Student A from the molecular biology program shares that she is interested in research careers with the federal government and her goal of participating in Grad LEAP is to seek career advice. Based on this information, we determine characteristics of an ideal mentor for this student to be (1) sharing similar academic interests such as graduating from the same or closely-related degree program and (2) the mentor is currently working in a federal government agency or has in the past. Using the existing proprietary algorithm, we would find that mentees like Student A are proposed to be matched with mentors from an unrelated degree program (e.g. humanities, not biological sciences) and who are currently working in unrelated career fields (e.g. a curator in an art museum, not a laboratory researcher).

We hope that MPI workshop participants will be able to help create a superior algorithm for our mentor-mentee matching process so that we can more effectively and efficiently identify optimal mentors for graduate students participating in the program. Here are our criteria for an effective algorithm.

1. **Data input.** The matching survey data will be collected via google form and exported into Excel spreadsheets. The algorithm must be compatible with input data in this format (a file for mentor survey data and a file for mentee survey data). The capacity to automatically recognize the survey question and rename/translate the category will be important. As shown in Table 1 and Table 2, data from the mentee's matching survey question "Do you identify with any of these terms?" were exported into the "mentee's identity" column while data from the mentor's matching survey question "Do you identify with any of these terms?" were exported into the "mentor's identity" column.
2. **Data output.** The output on the algorithm should also be in an Excel spreadsheet format with column headings for sorting and reviewing purposes. Each proposed match must have a (proposed) match number. For example, from Table 2, Alex and Yang will be match number 1. Additionally, each participant must have an automatically generated participant identification number. For example, Rosa's ID is 100, Sam's ID is 101, and so on.
3. **Confidence score and rationale.** For each proposed match, please define and generate a confidence score so that the quality of each proposed match can be reviewed by Grad LEAP administrators. For example, the range can be 1 to 10 with the proposed matches containing the largest number of matching elements (i.e. a good match) receiving higher scores. In the case of Alex and Yang's match above, the match would receive a confidence score of 9 while Rosa and Sam's match would receive a confidence score of 1. In addition to the confidence score, please include in the output a column containing matching rationale data for each proposed match. In the example of Alex and Yang's match, this column will contain information such as "academic match", "career match", and "identity match".

As you develop this winning algorithm, please remember the following important considerations and be sure to address them during your presentation and discussed in the written report.

- **Definition of closely-related academic disciplines.** What is the definition of closely related academic discipline for matching purposes? How would we teach closely-related academic disciplines to the algorithm? A list of University of Delaware's current degree programs will be provided to assist.

- **Definition of “mentor’s career fields” and “mentee’s careers of interest”.** The matching survey data can be confusing in these two overlapping categories. As such, it will be important to consider and weigh this element carefully. For example, a career teaching in the K-12 level is different from a career in “Teaching, college or university level”. Likewise, we wonder how relevant careers in “Engineering” are to “Cybersecurity” and “Robotics”. Should there be such a distinction? Should we remove broad career fields such as engineering, health & medicine, life sciences and focus on narrower, more specific careers or job titles instead? Relatedly, there are instances where a prospective mentor with a degree in the social sciences may have experience in data science. In this case, how would the algorithm decide if this prospective mentor should be matched with a mentee interested in a data science career but from a STEM discipline or to a mentor in social sciences interested in other careers (assuming there is not a mentor from the social sciences who are also interested in data science careers). A list of career fields and career interests will be provided to assist. We welcome suggestions on how to improve these answer options to assist with the matching process.
- **The maximum number of mentees per mentor as a matching constraint.** Each prospective mentor will be able to indicate the maximum number of mentees they are willing to mentor in the matching survey. Please ensure that the algorithm takes into account this constraint during the matching process. We will not want to match 4 mentees to a mentor who has indicated she is willing to mentor 2 students. However, matching 1 mentee to a mentor who is willing to mentor 3 is acceptable.
- **Proposed match revisions.** In the cases where rematching is needed, what would be the steps for the Grad LEAP administrators to do so? How would we exclude “approved matches” from the matching process? Would each rematching pair need to be done individually or would they be aggregated into a pool for reconsideration?
- **Data collection considerations.** Are we asking for too much or too little information? Are there better criteria we should be asking? Are there published research we can refer to? What is the best approach for weighing various preferences in the matching survey data?

The MPI participants will have access to the following data files.

1. Mentor matching survey questions and answer options.
2. Mentee matching survey questions and answer options.
3. Anonymous data file from mentor matching survey.
4. Anonymous data file from mentee matching survey.
5. A list of current degree programs.
6. Examples of good matches and matches requiring rematching.

**Thank you! We look forward to working with you.**