Journal of Informetrics 12 (2018) 736–756

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Regular article

The more multidisciplinary the better? – The prevalence and interdisciplinarity of research collaborations in multidisciplinary institutions

Zhiya Zuo, Kang Zhao*

University of Iowa, United States

ARTICLE INFO

Article history:
Received 15 February 2018
Received in revised form 15 June 2018
Accepted 16 June 2018

Keywords:
Multidisciplinarity
Interdisciplinarity
Collaboration
Network analysis
Text mining

ABSTRACT

Scientific research is increasingly relying on collaborations to address complex real-world problems. Many researchers, policymakers, and administrators consider a multidisciplinary environment an important factor for fostering research collaborations, especially interdisciplinary ones that involve researchers from different disciplines. However, it remains unknown whether a higher level of multidisciplinarity within an academic institution is associated with internal collaborations that are more prevalent and more interdisciplinary. Analyzing 90,000 publications by 2500 faculty members in over 100 academic institutions from three multidisciplinary areas, information, public policy, and neuroscience, we investigated the connection between multidisciplinarity and research collaborations. Based on social network analysis and text mining, our analysis suggests that more multidisciplinary institutions are not necessarily more collaborative, although they do feature collaborations that are more interdisciplinary. Our findings provide implications for academic administrators and policymakers to promote research collaborations and interdisciplinarity in academic institutions.

1. Introduction

Knowledge creation in the scientific community depends heavily on collaborations (Dong, Ma, Shen, & Wang, 2017; Wuchty, Jones, & Uzzi, 2007). Scientific research as a whole has become more collaborative, as evidenced by increasing multi-authored papers (Adams, 2012; King, 2012; Regalado, 1995). Collaborative research not only becomes more prevalent, but also tends to produce papers with better quality with respect to citations (Bu et al., 2018; He, Geng, & Campbell-Hunt, 2009; Wuchty et al., 2007) and help researchers increase productivity as measured by number of publications (Bu et al., 2018; Lee & Bozeman, 2005; Petersen, 2015). As Popper (1962) pointed out, *"We are not students of disciplines but students of problems. And problem may cut across the borders of any subject matter or discipline."* Different from the conventional collaborations in which researchers work only with peers with similar educational backgrounds or expertise, scientists now often form diverse collaborative teams to investigate novel and difficult problems that need to be addressed with an interdisciplinary approach (Derrick, Falk-Krzesinski, Roberts, & Olson, 2012; Ledford, 2015; NSF, 2005). Indeed, research collaboration not

* Corresponding author.
E-mail addresses: zhiya-zuo@uiowa.edu (Z. Zuo), kang-zhao@uiowa.edu (K. Zhao).

https://doi.org/10.1016/j.joi.2018.06.006
1751-1577/

only connects tangible entities such as researchers, organizations, and countries. Further, it contributes to the increasingly blurry conceptual borders between academic disciplines. The importance of interdisciplinary research in major scientific advances has been widely recognized (Derrick et al., 2012; NSF, 2005; Van Hartesveldt & Giordan, 2008). Interdisciplinarity occurred not only in emerging areas such as HIV study (Adams & Light, 2014), nanotechnology (Wang, Notten, & Surpatean, 2013), and astrobiology (Gowanlock & Gazan, 2013), but also in traditional fields, such as physics (Pan, Sinha, Kaski, & Saramäki, 2012; Sinatra, Wang, Deville, Song, & Barabási, 2016) and applied math (Xie, Duan, Ouyang, & Zhang, 2015). In fact, science, as a whole, has become more interdisciplinary (Porter & Rafols, 2009), and interdisciplinary research also has higher long-term impact measured by citation counts (Van Noorden, 2015).

Interdisciplinary research can be conducted by a sole investigator. However, typical interdisciplinary practice would involve people with disparate backgrounds (NSF, 2005). On one hand, collaboration of researchers with diverse expertise can facilitate integration of knowledge from different areas. On the other hand, interdisciplinary research invites and even demands collaboration to resolve the underlying complexity. Investigation on interdisciplinary collaboration, nonetheless, is still at an early stage that lacks in-depth analysis. Two common ways to report collaborations that are interdisciplinary are (i) co-authorship (Bordons, Zulueta, Romero, & Barrigón, 1999; Qiu, 1992; Schummer, 2004) (more interdisciplinary if authors come from different department); (ii) self-report survey (Cummings & Kiesler, 2005, 2008; van Rijnsoever & Hessels, 2011; Woolley, Sánchez-Barrioluengo, Turpin, & Marceau, 2015). However, the former suffers from arbitrary disciplinary classification, whereas the latter inevitably introduces individual biased understanding of what interdisciplinarity means.

In university systems, the smallest units of academic operation are at department, school, or college levels. Inspection on collaborative patterns as well as dynamics, however, is often made at the level of university, field, and beyond (Bu, Ding, Liang, & Murray, 2017; Dong et al., 2017; Jones, Wuchty, & Uzzi, 2008), belittling the importance of (interdisciplinary) collaborations between researchers within the same academic department. In fact, intra-organizational research collaborations within an academic institution are still very important. Cummings and Kiesler (2005) pointed out that distance not only yielded unwanted costs due to the need of researchers getting together, but lowered the productivity of interdisciplinary projects. From the perspective of gender equity and hiring, departments are the basic organizational units that allocate human resources and shape career prospects (Clauset, Arbesman, & Larremore, 2015; Su, Johnson, & Bozeman, 2015). Indeed, invincible assets such as consistency of reward and evaluation systems and convenience of in-person communications can hardly, if not impossible, be obtained when collaborating with people outside researchers' home institutions. On the other hand, the current formation of departments is no longer restricted to branches of science. Emerging areas such as information (Zuo, Zhao, & Eichmann, 2017) have pushed the movement of building multidisciplinary institutions that have similar functionality as traditional well-established departments (e.g., physics). These institutions gather people with diverse backgrounds and hence offer opportunities for colleagues to walk on the boundaries of disciplines, leading to interdisciplinary research, without needing to seek external collaborators. However, there is little study to show how these institutions perform in stimulating both the prevalence and interdisciplinarity of collaboration. A systematic examination of these new types of institutions' functionality contributes to a deeper understanding of their benefits and costs, and therefore providing empirical evidence for university departmental structures and future science policies.

There are many ways to promote more collaborations and interdisciplinary collaborations within an academic institution, such as organizational culture and promotion/tenure policies. One way our study is particularly interested in is to create a multidisciplinary environment with researchers from various domains, so that they have more opportunities to form research teams, especially interdisciplinary ones driven by complex problem-oriented research (Van Hartesveldt & Giordan, 2008). Multidisciplinarity, a concept closely related to and easily confused with interdisciplinarity, is about the co-existence of multiple disciplines, whereas interdisciplinarity focuses more on the integration of knowledge from several disciplines into research endeavors (Derrick et al., 2012; Wagner et al., 2011). An academic institution with faculty members from many different disciplines, for example, have a higher level of multidisciplinarity compared to others whose faculty members have very similar background. Interdisciplinarity, on the other hand, is achieved when people with different expertise actually collaborate with each other. In fact, diverse teams were found to be associated with high productivity (Stvilia et al., 2011). While research has argued that having a diverse group of researchers within the same organization may help enhance team performance (Salazar, Lant, Fiore, & Salas, 2012), it is unclear whether, at the very first place, collaborations will arise in a multidisciplinary institution. This can be challenging due to the heterogeneous nature of different disciplines (Jackson, Joshi, & Erhardt, 2003; van Knippenberg, van Ginkel, & Homan, 2013).

From a unique organizational perspective, this study examined the connection between multidisciplinarity of an institution and the prevalence, as well as the interdisciplinarity of collaborations within the institution, with datasets collected from three disparate disciplines – information, policy, and neuroscience[1] (Appendix A). Specifically, we utilized social network analysis and text mining techniques to address two research questions: First, *do more collaborations occur when an academic institution has a more multidisciplinary environment?* Second, *do collaborations that are more interdisciplinary emerge in a more multidisciplinary environment?* While it may seem intuitive to have an affirmative answer to the second research question, the heterogeneous disciplinary boundaries may also hinder such collaborations (Disis & Slattery, 2010; Yegros-Yegros, Rafols, & D'Este, 2015), due to coordination costs and team management. Specifically, it may take more time as well

[1] Faculty data used in this study is available at https://data.mendeley.com/datasets/6c2p7r6p6y/.

as costs for people with different expertise to work together. In a multidisciplinary environment, it is still possible that people are more willing to collaborate with researchers who are similar, according to the homophily principle (McPherson, Smith-Lovin, & Cook, 2001). Answers to these questions can help research institutions, funding agencies, and other policy makers better understand and more effectively promote collaborations, especially interdisciplinary collaborations.

2. Related work

In organizational studies, collaborations are commonly recognized as beneficial, because they can pool more resources and expertise from a team. Teamwork is, for example, an essential component in achieving high performance (Baker, Day, & Salas, 2006). Using data from various industries, Zhao, Huo, Selen, and Yeung (2011) found strong and positive impact of internal collaboration on a company's performance. In addition, various studies have revealed the importance and value of diversity in organizations or teams (Erhardt, Werbel, & Shrader, 2003; Gibbons, 2010; Lash & Zhao, 2016; Siciliano, 1996; Wilde, 2010). Admittedly, as diversity can be based on many different individual or group attributes, such as race, gender, sexual orientation, and national origins (Shore et al., 2009), there are different effects on organizational performance depending on the specific dimension of diversity (De Abreu Dos Reis, Sastre Castillo, & Roig Dobón, 2007; van Knippenberg et al., 2013).

For scientific research, multidisciplinarity, which reflects the diversity in researchers' expertise and knowledge, is beneficial in several ways. It has been revealed that diversity in educational backgrounds exerts positive influence on team success (Harrison, Price, Gavin, & Florey, 2002; Jackson et al., 2003; van Knippenberg et al., 2013). Besides, multidisciplinarity can generate novel ideas at the intersection of disciplinary knowledge, enhance collaborations (Salazar et al., 2012), and increase research productivities (Porac et al., 2004; Stvilia et al., 2011).

Despite the benefits, however, collaborations in a multidisciplinary environment face challenges. For example, a team has to deal with problems such as being in middle ground, choice of collaborators, and the evaluation of research impact (Whitfield, 2008). Coordination cost, team management, and infrastructure can be obstacles for multidisciplinary collaborations (Disis & Slattery, 2010Yegros-Yegros et al., 2015). Thus, it is still an open question whether a higher level of multidisciplinarity within an academic institution can help more prevalent research collaborations.

Compared to multidisciplinarity, interdisciplinarity is a subtler concept that assesses the integration of knowledge from various fields (Wagner et al., 2011). In the literature, there are two types of nonexclusive approaches to quantify interdisciplinarity (Rafols & Meyer, 2010): top-down and bottom-up. Top-down approaches utilize predefined disciplinary categories for journals and analyze their interrelationships via citation records (co-citation and bibliography coupling) or co-authorship. Porter, Cohen, David Roessner, and Perreault (2007), for example, proposed an integration score for both research papers and researchers using Web of Knowledge Subject Categories. The basic idea is that a paper's integration score gets higher if it cites papers from a more diverse set of Subject Categories. From the perspective of collaborations, Bordons et al. (1999) considered the degree of interdisciplinarity as being proportional to the number of authors' disciplines/departments associated with a document. Sayama and Akaishi (2012) developed a method that quantifies the interdisciplinarity at both researcher and topic levels with a set of manually selected authors and keywords using web search engines. Using the disciplinary classification of the Proceedings of the National Academy of Sciences, Xie et al. (2015) employed network metrics to evaluate the role of applied mathematics in interdisciplinary research at paper level. Although simple and straightforward, such methods suffers from the arbitrary classification scheme, and fails to capture the context of citations and the extent to which a citation influenced the citing paper. Besides, it is impossible to assess the knowledge integration among authors of a paper simply via their affiliations. Bottom-up approaches investigate interdisciplinarity from emerging clusters. Existing methods include keywords analysis, text clustering, topic models, network analysis, etc. (Gowanlock & Gazan, 2013; Leydesdorff, 2007; Nichols, 2014; Rafols & Meyer, 2010; Wang et al., 2013; Xie et al., 2015). However, interdisciplinarity measurements based only on network analysis (Leydesdorff, 2007; Rafols & Meyer, 2010; Xie et al., 2015) are limited to topological structure, without considering nodal attributes. Being too fine-grained, the analysis of words and phrase occurrences (Gowanlock & Gazan, 2013; Wang et al., 2013) has problems with synonyms and phrases that may correspond to more than one research area (e.g., the word network can be about social network analysis or computer network design). By contrast, topic modeling techniques provide an abstract representation of what a document is about with the proper granularity.

The contributions of this paper are four-fold: *First*, we clearly distinguished the concepts of multi- and interdisciplinarity at institutional level. While the two terms have been discussed in the past (Wagner et al., 2011), past research looks at a single aspect, belittling the subtle relationship between mutli- and interdisciplinarity. Moreover, as Wagner et al. (2011) pointed out, most of the past research focus on interdisciplinarity at article or journal levels. There is a lack of clarification and definition of institutional level multi- and interdisciplinarity. A side-by-side investigation between these two closely related yet different concepts in the context of academic institutions can shed light on the structure and operationalization of disciplines, as well as higher education systems. Specifically, we quantified multidisciplinarity of an institution by its intellectual compositions based on faculty educational backgrounds and research interests. Interdisciplinarity, on the other hand, is measured by the extent to which co-authors differ in their research topics. The more different co-authors are, the more interdisciplinary this collaboration tie is. *Second*, our novel measures of multidisciplinarity of an institution and interdisciplinarity of research collaborations are based on individuals' research interests derived by applying topic models to publication data. Although topic modeling has been adopted in the study of multidisciplinarity, the approach used by Nichols (2014) still relied on a discipline taxonomy from the National Science Foundation. Such a taxonomy, along with the detailed topic modeling results, are not publicly available though. Bordons et al. (1999) also tried to measure interdisciplinarity based on research interests, but they only used researchers' affiliations to approximate such interests. While educational backgrounds, departmental/institutional affiliations, and subject categories of journals (Qiu, 1992; van Rijnsoever & Hessels, 2011; Wang, Thijs, & Glänzel, 2015) can, to some extent, reflect a researcher's expertise, they are too coarse-grained and can be inaccurate. Many disciplines, such as information science, and journals, such as Science, already cover highly diverse research areas. *Third*, by analyzing the difference among collaborators' research interests before the tie was formed, our proposed measure of collaboration interdisciplinarity can better capture the degree of interdisciplinarity of a collaboration tie even though a collaborator's research interest may have changed over time. *Finally*, we are the first to evaluate the effects of multidisciplinarity on the prevalence and interdisciplinarity of collaborations in academic institutions. Our findings have policy implications for administrators at academic institutions and funding agencies.

3. Methods

3.1. Measuring institutional multidisciplinarity

For each institution, we used two measures: educational multidisciplinarity based on faculty members' educational backgrounds, and research multidisciplinarity based on faculty members' research interests extracted from their publications.

3.1.1. Educational multidisciplinarity

Educational background has been used extensively to approximate faculty members' research directions and expertise (Wiggins & Sawyer, 2012; Zhu, Yan, & Song, 2016). Although arbitrary, disciplinary classification schemes make it easy and straightforward to quantify multidisciplinarity. Therefore, we first measured the level of multidisciplinarity using this traditional approach by categorizing faculty members' doctorate degrees (Table A.4). Following Wiggins and Sawyer (2012), we calculated Shannon Entropy (Shannon, 1948) to evaluate variety. Specifically, educational multidisciplinarity of an institution is defined as $MD_{edu} = -\sum_i p_i log\, p_i$, where p_i is the proportion of faculty members whose doctoral programs belong to category i. The higher the entropy value is for an institution, the more multidisciplinary the institution is in terms of faculty educational background.

3.1.2. Research multidisciplinarity

Albeit straightforward and convenient, educational multidisciplinarity was based on a top-down approach that utilizes predefined categories of disciplines. Such arbitrary classifications of educational backgrounds can be inaccurate proxies for faculty members' actual research interests. For example, a computer science PhD focusing on health informatics may be more similar to a public health PhD also working in this area, than to another computer science PhD who studies programming languages. Moreover, it is very difficult, if not impossible, to quantify the distances among different discipline categories.

Besides educational backgrounds, we proposed research multidisciplinarity, which measures the level of multidisciplinarity based on how diverse faculty members' research interests were prior to joining their current institutions. We focused on research interests before their current positions because hiring a faculty member is an important and deliberate decision. It is also one of the keys to create a multidisciplinary environment. When an institution makes a decision on whom to hire, each candidate is only represented by her previous research. After joining an institution, collaborations with peers could potentially influence a faculty member's research directions and confound the quantification of research multidisciplinarity from publication data.

While the year in which a faculty joined her current institution can be inferred from affiliation changes in her publication records, we still need to capture a faculty member's research interests over time, so that we can get her research interests before joining the current institution. We decided to adopt topic modeling techniques, which is used to extract latent topics from texts of faculty members' publications. Specifically, we used latent Dirichlet allocation (Blei, Ng, & Jordan, 2003) for this study. In the outcome of LDA, each topic is represented by a distribution over words, and each document (i.e., paper, in our case) has a probabilistic distribution over topics. Specifically, we collected a faculty member's publications up to a certain year, and used the average topic distribution of these papers to represent the faculty member's research interests till that year (i.e., before her appointment in the current institution). We ran three LDA models, one for each of the three areas. In other words, we have one LDA model fitted by papers of information schools, one for policy schools, and one for neuroscience departments. Titles and abstracts of papers retrieved from Scopus were preprocessed (stop words removal and stemming) before being fed to LDA as inputs. The number of topics was simply set to 20, because we were more interested in the differences in topic distributions among faculty members than the topics themselves. Different numbers (e.g., 30, 40, and 50) have been tried and the eventual results are consistent (Fig. B.1). Summaries of LDA results for the three disciplines are shown in Table B.1. The diverse topic distributions among researchers in the three disciplines again reflect the three areas' multidisciplinarity.

With faculty members' annual topic distributions, we applied cosine distance to measure the level of research multidisciplinarity within an institution. We define research multidisciplinarity for an institution as the average pairwise distance between its faculty members' topic distributions: $MD_{res} = \frac{1}{N}\sum_{i \neq j} dist(T_i, T_j)$, where N is the number of possible faculty pairs within the institution; T_i is faculty member i's topic distribution before her current position; $dist(\cdot)$ is the pairwise cosine distance function. The higher the average cosine distance is, the more multidisciplinary the school is on faculty research interests.

3.2. Intra-institutional collaborations

3.2.1. Prevalence of collaboration

We measure collaboration prevalence within an institution with (i) the average percentage of internally collaborative papers (ICPs) for each institution and (ii) co-authorship network connectedness. We considered a paper to be an internally collaborative paper if its co-authors include more than one faculty member from the same institution. A higher average percentage of ICPs indicates more prevalent collaborations between faculty members within an institution.

We built an unweighted and undirected co-authorship network for each institution, in which each node represents a faculty member and there will be an edge between two nodes if they have one or more co-authored publications. Internal co-authorship networks can help us understand collaborative relationships among faculty members within one institution and how knowledge and expertise are exchanged among faculty members. The more connected the network is, the more prevalent collaborations are within an institution. In this research, we used co-authorship network density as the measure of connectedness. It is defined as the ratio between the observed and possible numbers of co-authored faculty pairs. A higher co-authorship network density indicates higher prevalence of collaborations in the corresponding institution.

3.2.2. Interdisciplinarity of collaboration

Edges in co-authorship networks clearly manifest collaborative relationships between faculty members. However, some collaborations may be between those whose research interests are similar. As mentioned earlier, we believe that interdisciplinary collaborations are those between faculty members with different research interests. The more divergent two connected faculty members' research topics are, the more interdisciplinary their collaboration is. Such diversity at the dyadic level can be measured by assortative mixing patterns of co-authorship networks. Assortativity is the tendency of nodes to connect to similar others in a network (Newman, 2002; Zhao, Ngamassi, Yen, Maitland, & Tapia, 2010). In our study, assortativity of a co-authorship network is the likelihood of faculty members with similar topic distributions to co-author papers. To make it more intuitive, we adopted the opposite of assortativity, disassortativity, which quantifies the extent to which dissimilar faculty members collaborate with each other. The more disassortative an internal co-authorship network is, the more interdisciplinary the corresponding institution's collaborations are. To calculate disassortativity, we specified each faculty member's topic distribution as nodal attributes. Specifically, we capture research topics of two faculty members before their first collaborative paper – after they co-author a paper, their research topics inevitably get closer to each other than before due to the co-authored paper.

Topic distributions take the form of vectors. However, the original calculation of assortativity by default considers nodal attributes as scalar (e.g., degree), which would be infeasible for this study. Instead, we adopted a method proposed in Zhang and Pelechrinis (2014): for each edge, we calculate the cosine distance between nodal attributes (i.e., authors' topic distributions prior to formation of this edge.) The disassortativity of a network is then the average cosine distance over all edges in the network. It is worth noting that we will be focusing on institutions whose internal co-authorship network has at least one edge (i.e., 25 out of 26 information schools; 47 out of 66 policy schools, and all neuroscience departments). After all, it is meaningless to evaluate the interdisciplinarity of internal collaborations when there is no such collaboration within the institution.

4. Results

4.1. Multidisciplinarity

Fig. 1A shows the distributions of educational multidisciplinarity among individual institutions. All three areas have left skewed distributions (skewness is -0.72, -0.16, and -0.52 for information, policy, and neuroscience, respectively). The Information School at University of Washington has the highest educational multidisciplinarity; the Information School at Georgia Institute of Technology is the lowest, with 84% of its faculty members earning doctoral degrees from computing related disciplines. Among policy schools, the University of Delaware is the most multidisciplinary in faculty educational background, whereas the University of Illinois at Springfield is the least, with 87.5% of its faculty members graduating from public policy programs. Department of Neuroscience at Brown University has the highest educational multidisciplinarity in neuroscience. Having 92% faculty members with degrees in life or social sciences, the entropy of Oregon Health and Science University's Neuroscience Department is the lowest. Please note that we are not trying to compare educational multidisciplinary of the three areas with each other because of (i) differences in classification schemes of educational backgrounds and (ii) conceptual distances between categories.

In distributions of research multidisciplinarity (Fig. 1B), information (skewness -0.51) and neuroscience (skewness -1.20) both have long left tails, with the majority of individual institutions with above-average research multidisciplinarity and a few with relatively low multidisciplinarity. Policy schools have a right-skewed distribution (skewness 0.48). Similar to educational multidisciplinarity, it is meaningless to compare research multidisciplinarity of the three areas with each other

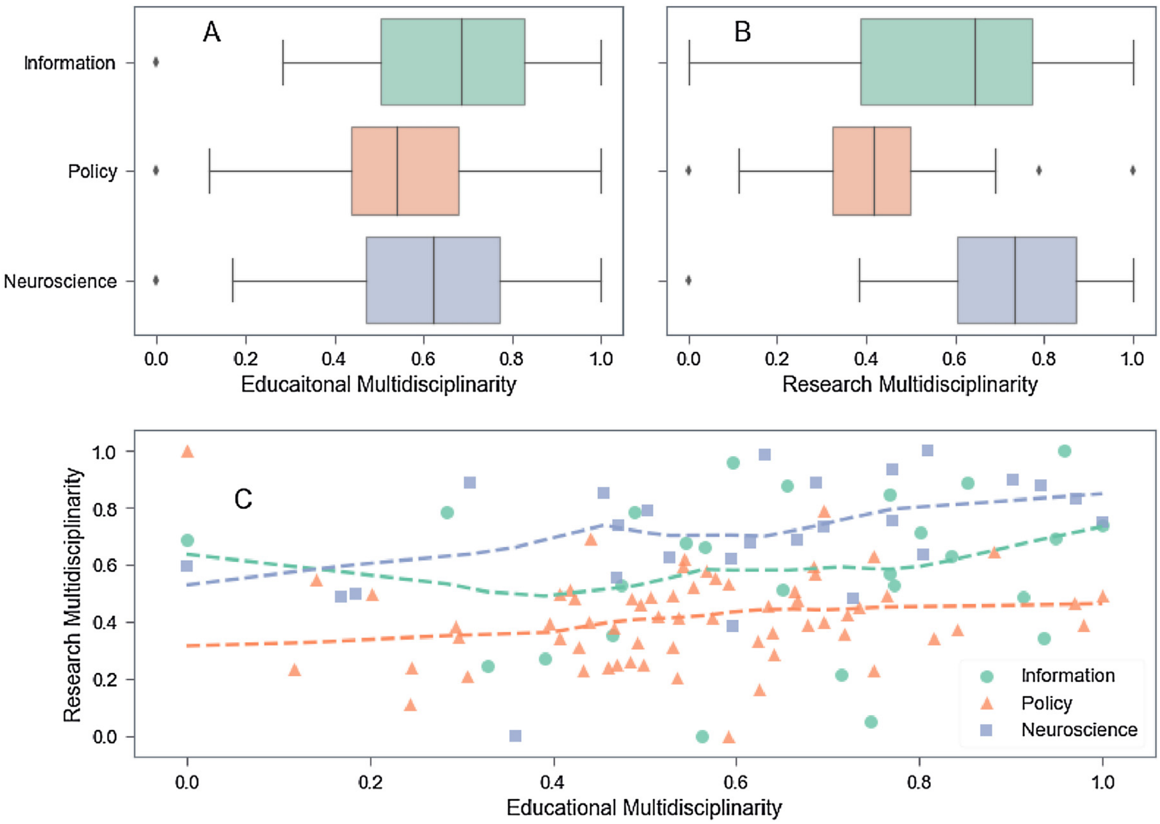

Fig. 1. Boxplots of educational multidisciplinarities (A), research multidisciplinarities (B), and scatter plot of research vs. educational multidisciplinarity (C). Dashed lines in C are fitted by local regression (Cleveland, 1979). In A and B, the upper and lower bounds represent 75% and 25% percentiles, respectively; the middle bars are medians. Note that educational multidisciplinarities in C are plotted in the same range [0,1] by min–max scaling for each discipline respectively.

because research topics are intrinsically heterogeneous due to both the difference between the areas and the separate topic extraction processes.

In addition, a Pearson correlation test confirms that research multidisciplinarity based on publication profiles indeed capture institutional multidisciplinarity from a different perspective than educational multidisciplinarity (Fig. 1C). The two measures have little correlation for information ($r^2 = 0.13$) and policy ($r^2 = 0.03$), and small for neuroscience ($r^2 = 0.20$).

4.2. Collaboration prevalence and multidisciplinarity

The average percentage of ICPs for Information Schools is 5.80%, with the highest being 13.10% (Department of Information Systems at University of Maryland Baltimore County). Among policy schools, the average percentage of ICPs is 2.89%, with School of Public Service at Old Dominion University having the highest percentage (17.54%). Faculty members at the Department of Neuroscience at Temple University have the highest of average percentage of ICPs (33.01%), whilst the average percentage across 26 neuroscience departments is 8.79%. Co-authorship network densities follow a very similar distribution with ICPs (Fig. 2A and B).

The distribution of both measures (Fig. 2) suggests that (i) neuroscience on average exhibit the highest collaboration propensity, partly because neuroscience has deep roots in biological sciences, the vanguard of larger research teams (Pavlidis, Petersen, & Semendeferi, 2014); (ii) policy, on the other hand, tends to favor solo author publications. For all three areas, distributions for both prevalence measures have positive skewness values. There are moderate or high correlations between the average percentages of ICPs and co-authorship network (Fig. 2C): $r^2 = 0.27$ for information, $r^2 = 0.71$ for policy, and $r^2 = 0.22$ for neuroscience.

Across the three areas, there are negative correlations between co-authorship network density and educational multidisciplinarity (Fig. 3). For policy and neuroscience, research multidisciplinarity are also negatively correlated with both collaboration prevalence measures. All these suggest that a more multidisciplinary institution may not be more collaborative. In fact, the heterogeneity that comes with a highly multidisciplinary institution may not be beneficial for promoting intra-institutional collaboration between colleagues.

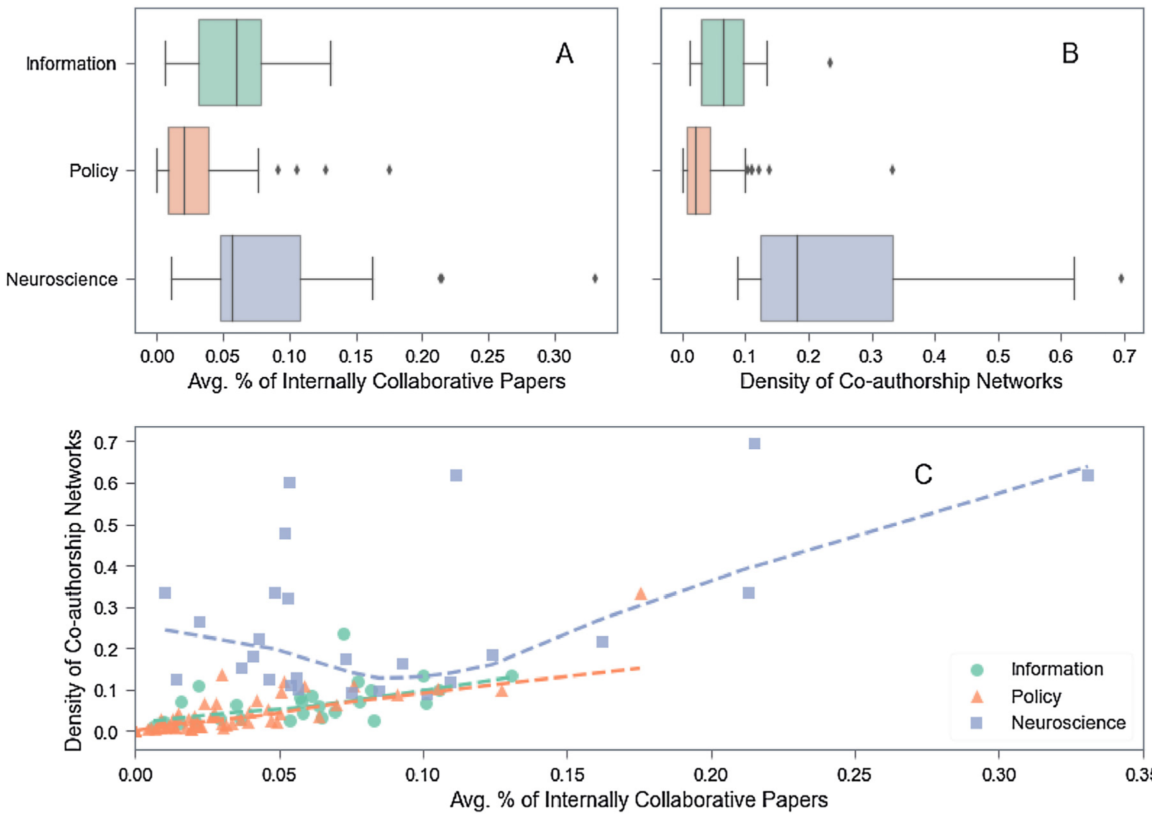

Fig. 2. Collaboration prevalence of the three areas: Avg. % of ICPs (A) and co-authorship network density boxplots (B); scatter plot of avg. % of ICPs vs co-authorship network density (CA-DEN) (C). See Fig. 1 for figure aesthetic.

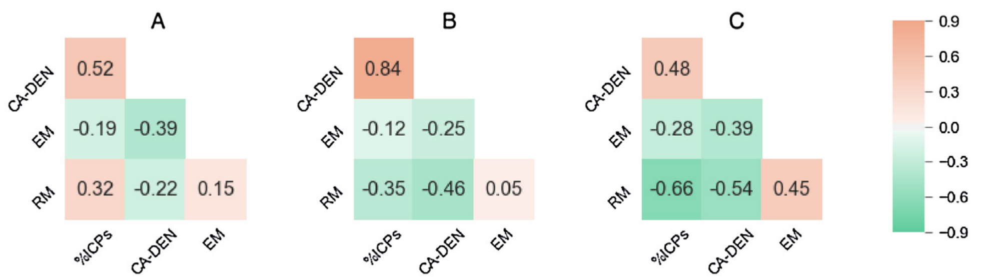

Fig. 3. Pearson correlations between multidisciplinarity and collaboration prevalence for information (A), policy (B), and neuroscience (C).

Meanwhile, other institutional factors can potentially affect the prevalence of internal collaborations. Therefore, we applied hierarchical regression models, where control variables mentioned below are entered first and then multidisciplinarity serving as the study variables will be entered last. For each type of institutions, we ran regression models with the average percentage of ICPs and co-authorship network density serving as dependent variables respectively. Independent variables include educational and research multidisciplinarity. We further incorporated three control variables: (i) size of an institution (measured by the number of faculty members in the school/department); (ii) average research productivity at an institution (i.e., average number of publications per faculty member within the school/department); (iii) home university's Carnegie Classification (Indiana University Center for Postsecondary Research, 2015). For information school, we ran six regression models–three sets of independent variables (educational multidisciplinarity only, research multidisciplinarity only, and both multidisciplinarity measures) with each of the two dependent variables. The same settings were applied to policy schools and neuroscience departments.

Fig. 4 plots the standardized coefficients and the corresponding confidence intervals – a standard deviation increase in each of the dependent variable lead to a standard deviation increase in multidisciplinarity, controlling for the linear effects of the others. This facilitates the prediction power comparison across different variables which were originally in various scales.

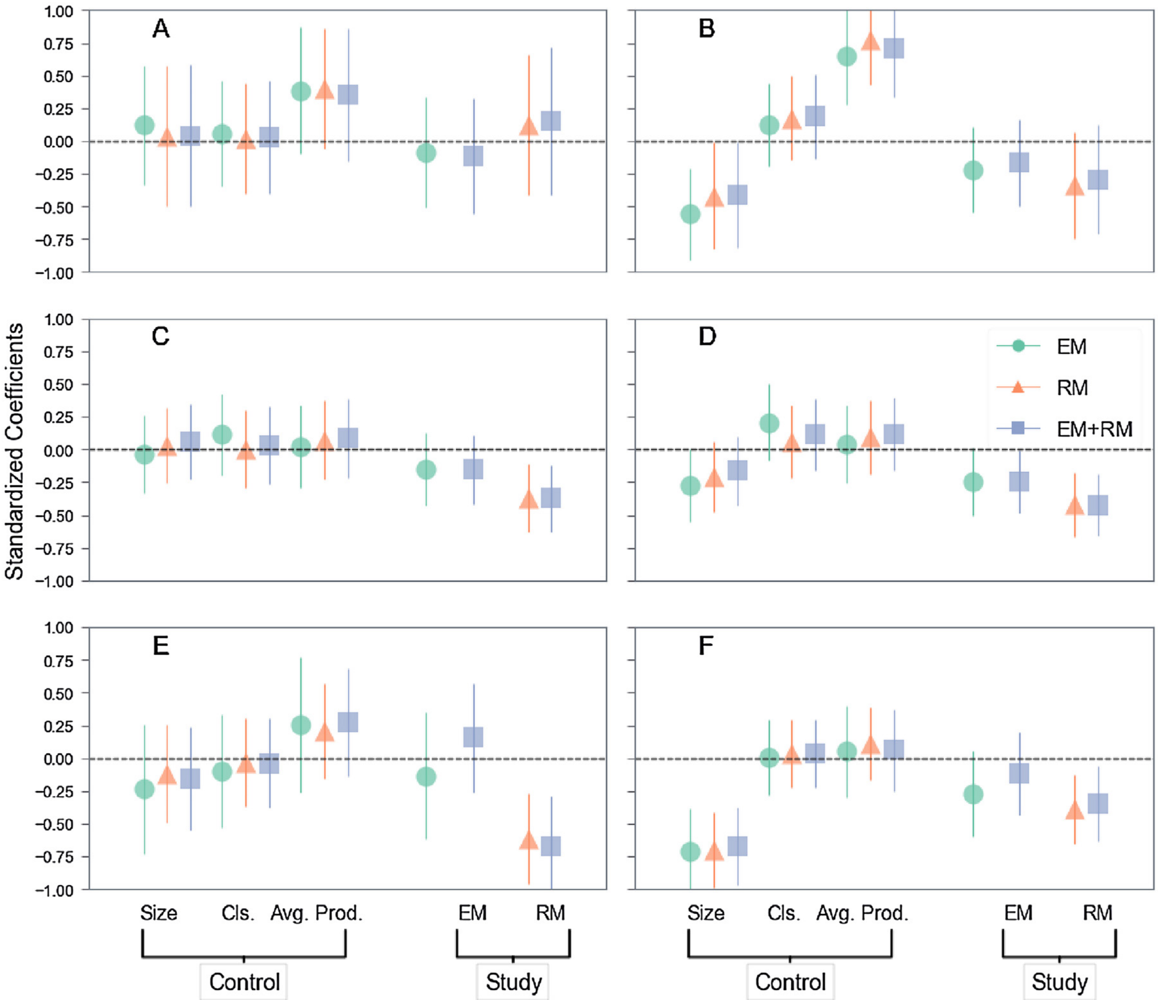

Fig. 4. Standardized coefficients of regression models for the relationship between multidisciplinarity (educational, EM and research, RM) and collaboration prevalence (the dependent variable, measured by the average percentage of ICPs in first column, and density of co-authorship networks in the second one), controlling for institutional size (Size), Carnegie classification of the home university (Cls.), and average faculty productivity (Avg. Prod.). A and B are for information; C and D for policy; E and F for neuroscience. Each bar shows a 95% confidence interval, while the filled marker is the point estimate.

Results from these models are consistent for the three types of institutions: Educational multidisciplinarity is still not informative for either the average percentage of ICPs or co-authorship network density. Meanwhile, internal collaboration prevalence in policy schools and neuroscience departments will likely decrease as research multidisciplinarity increases (Fig. 4C–F). The effects of control variables and multidisciplinarity measures are consistent no matter which measure of collaboration prevalence serves as the dependent variable. For example, across the three areas, the size of a school/department is a (marginally) strong factor that negatively affects co-authorship network density, which may be an artifact of the definition of density. More detailed discussions on the regression analyses are in Appendix D.

4.3. Collaboration interdisciplinarity and multidisciplinarity

To explore the relationship between institutional multidisciplinarity and collaboration interdisciplinarity, we also started with Pearson correlations (Fig. 5). Across the three disciplines, the level of internal collaboration interdisciplinarity is positively correlated with research multidisciplinarity. In the subsequent regression analysis (Fig. 6), we controlled for institutional size, Carnegie classification, average productivity, and co-authorship network density (one of the collaboration prevalence measures). Consistent across the three areas, research multidisciplinarity has a strong and positive relationship with the interdisciplinarity of collaborations. It is worth noting that the inclusion of research multidisciplinarity uniquely contributes to 50% of the variance in the model of information schools. Educational multidisciplinarity, on the other hand, has little effects on research interdisciplinarity.

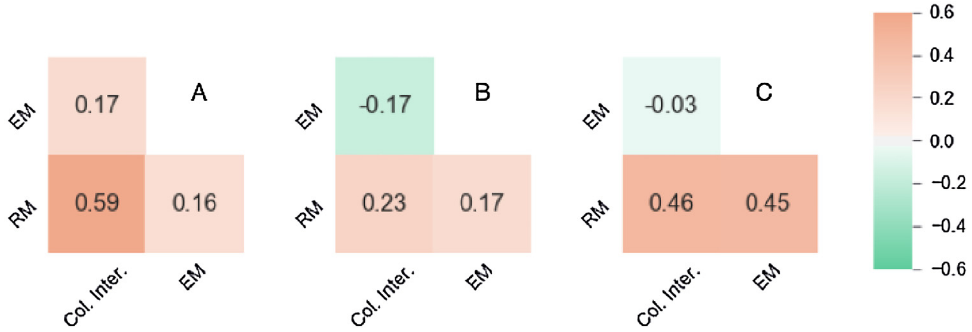

Fig. 5. Pearson correlations between multidisciplinarity and collaboration interalence for information (A), policy (B), and neuroscience (C).

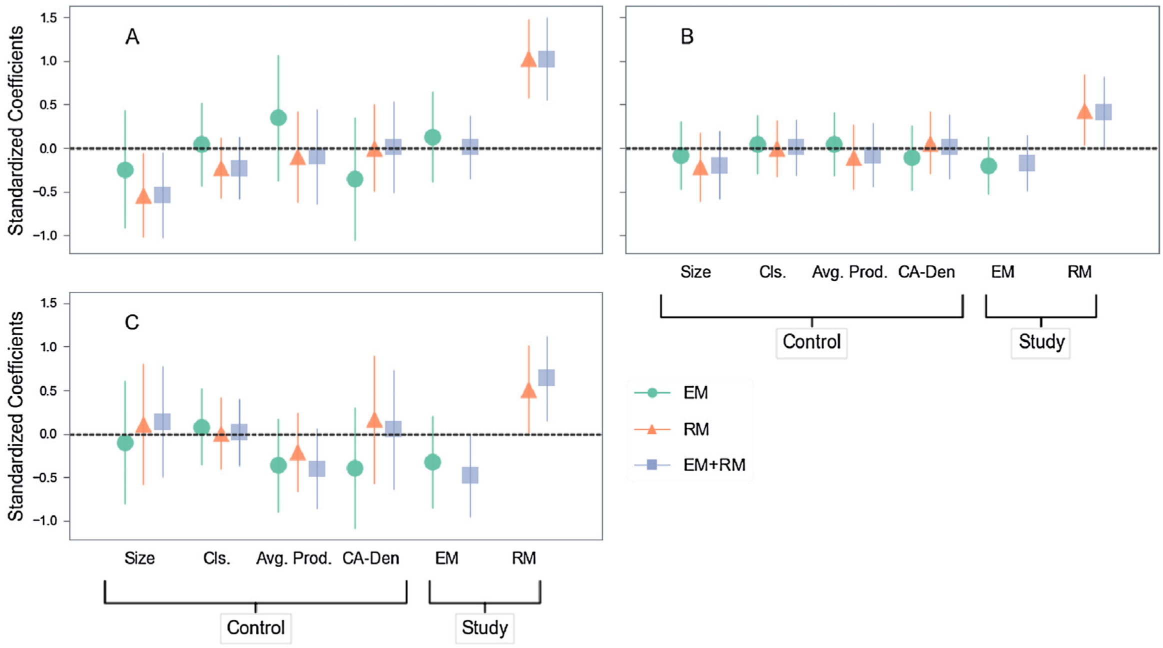

Fig. 6. Standardized coefficients of simultaneous regression models for the relationship between multidisciplinarity (EM and RM) and collaboration interdisciplinarity (the dependent variable), controlling for Size, Cls., Avg. Prod., and CA-DEN. A is for information; B for policy; and C for neuroscience. See Fig. 4 for figure aesthetic.

5. Discussion and conclusion

Using social network analysis and text mining techniques, this research investigated the relationship between institutional multidisciplinarity and collaborations. We quantified institutional multidisciplinarity using both faculty educational background and research topics derived from their publications. For collaborations within an institution, we measured both the prevalence and interdisciplinarity of collaborations using network metrics and publication topics. Analyzing data from information schools, policy schools, and neuroscience departments in the U.S., we found that the three disciplines are indeed different in faculty educational backgrounds, publishing venues, and research topics. Despite the fundamental differences in research foci, our analysis revealed very similar and interesting relationships between institutional multidisciplinarity and internal collaborations. Specifically, there is a paradox of multidisciplinarity and research collaborations, brought by the heterogeneity of diverse faculty composition. On one hand, higher levels of multidisciplinarity do not necessarily mean a more collaborative institution. On the other hand, research multidisciplinarity helps to foster interdisciplinarity – research collaborations within a more multidisciplinary institution are more likely to occur between those with different research interests. At the same time, educational multidisciplinarity, which based on disciplines of faculty members' doctoral degrees, is not informative for either the prevalence or the interdisciplinarity of collaborations. In other words, in emerging multidisciplinary institutions such as those studied in this research, a researcher's educational background alone can no longer accurately reflect her research interests.

Our results have implications for research policies, especially for academic and research administrators. Specifically, just creating an academic institution with a diverse faculty body is not sufficient to foster more internal research collaborations.

We suggest that for a multidisciplinary institution, more coordination, management, and incentives may be necessary to fully exploit the potential benefits of multidisciplinary diversity in stimulating more intra-organizational research collaborations that span disciplinary boundaries. As our analyses revealed, when collaborations do occur within institutions that have a highly diverse faculty body, they are expected to have high levels of interdisciplinarity.

Although we focused on three academic areas in this study, our approach can also be applied to research beyond them. For example, the novel measurement of research interdisciplinarity based on topic modeling and co-authorship networks can be extended to study the interdisciplinarity of papers in citation networks (e.g., a paper that cited papers with divergent topic coverages tend to be more interdisciplinary). In addition, our approach can be applied to investigate the role of diversity in expertise or knowledge in intra-organizational collaborations in for-profit or non-for-profit organizations outside academia.

Admittedly, this research has limitations. First, the study analyzed empirical data using regression and correlation analyses. Causal relationships between multidisciplinarity and research collaborations are beyond the scope of this paper. Second, by using co-authorship as the only proxy for research collaborations, we may miss other types of collaborations that did not lead to publications, such as brainstorming sessions and un-successful grant applications (Katz & Martin, 1997). Further, despite the broad coverage of publications by Scopus, there are other published papers that are not indexed by Scopus. Third, we analyzed the interdisciplinarity of an institution through the lens of collaborations, although an institution can also improve its research interdisciplinarity by hiring faculty members whose research is by itself already interdisciplinary without collaborating with colleagues in the same institution. Finally, we note that our regression analyses inevitably suffer from specification errors, as variables are not an exhaustive list of factors that may pose influence on collaboration practice. In particular, faculty members' decisions to collaborate with (diverse) colleagues is a complex process. Factors including personality, monetary and non-monetary reward systems, and institutional policies, may also affect the final collaboration outcome. However, collecting such data is challenging, especially at a large scale across different areas.

Looking ahead, our findings in this study also point to several intriguing future research directions. Among the three areas analyzed in this study, we found that information is different from policy and neuroscience in several ways. For example, the level of research multidisciplinarity in information schools by itself accounted for remarkably 50% of the variance in collaboration interdisciplinarity. With less than 3 decades of history, information schools may still be establishing its unique identity from the commonly recognized root disciplines of library, computer science, and management sciences (Zuo et al., 2017). A further investigation into such differences will shed lights on the fundamental drivers of interdisciplinary research in contemporary multidisciplinary academic institutions. Besides, it will be interesting to include well-established disciplines (e.g., computer science) in the same analyses, which can serve as a baseline to facilitate the interpretation of regression results on multidisciplinary institutions. In addition, while the prevalence and interdisciplinarity of collaborations are important, research impact (Wang et al., 2015; Yegros-Yegros et al., 2015) is what many stakeholders care about – *Do more multidisciplinary academic institutions also tend to produce research with higher impact?* With large-scale electronic databases, faculty members' digital footprints will enable systematic data-driven analyses on the processes of scientific discoveries (Clauset, Larremore, & Sinatra, 2017), including collaboration patterns and interdisciplinary research. Future studies answering these questions could provide valuable insights for policymakers, funding agencies, as well as academic institutions.

Authors contributions

Zhiya Zuo: Conceived and designed the analysis; Collected the data; Performed the analysis; Wrote the paper.
Kang Zhao: Conceived and designed the analysis; Performed the analysis; Wrote the paper.

Appendix A. Data collection

We collected faculty data from formal academic institutions in three different and multidisciplinary areas: (i) information; (ii) policy; (iii) neuroscience. The three areas have roots in applied sciences, social sciences, and biological sciences, respectively, and can thus provide a more comprehensive view of the connection between multidisciplinarity and collaborations in different academic institutions. Specifically, 27 information schools were listed on *ischools.org* in the member directory in 2014[2] (Table A.1); 64 policy schools were retrieved from *NASPA - Student Affairs Administrators in Higher Education*'s list of universities that offer doctoral programs[3] in 2016 (Table A.2); 26 neuroscience departments were obtained from a list of neuroscience training program directory by *Society for Neuroscience*[4] in 2016 (Table A.3). During the data collection, we excluded academic programs that have no dedicated faculty members (i.e., all faculty members have main appointments at other departments), because there is no formal institutional or organizational structure for such programs. Institutions that do not offer doctoral programs are also removed. We did not include those outside the U.S. to eliminate the effects of possible differences in faculty ranks and higher education systems, and to avoid publications written in languages other than English. When a school includes academic units that are not related to the domain, only the related units are included in

[2] We identified 27 information schools but removed the one at University of California, Berkeley due to data quality issues.
[3] http://www.naspaa.org/students/graduate/doctoral.xlsx.
[4] http://www.sfn.org/Careers-and-Training/Training-Program-Directory.

Table A.1
List of 26 information schools.

University	Academic unit
Carnegie Mellon University	Heinz College
Drexel University	College of Computing &Informatics (only the former College of Information Science &Technology)
Florida State University	School of Information
Georgia Inst. of Tech.	School of Interactive Computing
University of Illinois at Urbana-Champaign	Grad School of Library and Information Science
Indiana University	School of Informatics and Computing
University of Kentucky	Library and information science
University of Maryland at College Park	College of Information Studies
Michigan State University	Dept. of Media and Information
University of Michigan	School of Information
University of Missouri	School of Information Science &Learning Technologies
Penn State University	College of Information Sciences &Technology
University of Pittsburgh	School of Information Sciences
Rutgers University	Library and Information Science Dept.
Simmons College	School of Library and Information Science
Syracuse University	School of Information Studies
University of Tennessee at Knoxville	School of Information Sciences
University of Texas at Austin	School of Information
University of California at Irvine	Dept. of Informatics
University of California at Los Angeles	Dept. of Information Studies
University of Maryland, Baltimore County	Dept. of Information Systems
University of North Carolina at Chapel Hill	School of Information and Library Science
University of North Texas	Dept. of Library &Information Sciences
University of Wisconsin at Milwaukee	School of Information Studies
University of Washington	Information School
University of Wisconsin at Madison	School of Library &Information Studies

Table A.2
List of 64 policy schools.

University	Academic unit
American University	Dept. of Public Administration &Policy
Arizona State University	School of Public Affairs
Boise State University	Dept. of Public Policy &Administration
Brandeis University	Heller School for Social Policy &Management
Carnegie Mellon University	School of Public Policy &Management
City University of New York	Dept. of Public Management
Cleveland State University	Maxine Goodman Levin College of Urban Affairs
Columbia University	School of International &Public Affairs
Duke University	Sanford School of Public Policy
Florida Atlantic University	School of Public Administration
Florida State University	Askew School of Public Administration &Policy
George Mason University	School of Policy, Government, &International Affairs
George Washington University	School of Public Policy &Public Administration
Georgia Institute of Technology	School of Public Policy
Georgia State University	Dept. of Public Management &Policy
Harvard University	Kennedy School of Government
Indiana University at Bloomington	School of Public &Environmental Affairs
Jackson State University	Dept. of Public Policy &Administration
Mississippi State University	Dept. of Political Science &Public Administration
New York University	Robert F. Wagner Graduate School of Public Service
North Carolina State University	Dept. of Public Administration
Northeastern University	School of Public Policy &Urban Affairs
Northern Illinois University	Dept. of Public Admin
Ohio State University	John Glenn College of Public Affairs
Old Dominion University	School of Public Service
Penn State University at Harrisburg	School of Public Affairs
Portland State University	Division of Public Administration
Princeton University	Woodrow Wilson School of Public &International Affairs
Rutgers University at New Brunswick	Edward J. Bloustein School of Planning &Public Policy
Rutgers University at Newark	School of Public Affairs &Administration
Southern University	Nelson Mandela School of Public Policy &Urban Affairs
State University of New York at Albany	Dept. of Public Administration &Policy
Syracuse University	Maxwell School of Citizenship &Public Affairs
Tennessee State University	Dept. of Public Administration
The New School	Milano School of International Affairs, Management, &Urban Policy
University of Arizona	School of Government &Public Policy
University of Baltimore	School of Public &International Affairs Faculty

Table A.2 (*Continued*)

University	Academic unit
University of California at Berkeley	Goldman School of Public Policy
University of Central Florida	School of Public Administration
University of Chicago	Harris School of Public Policy Studies
University of Colorado at Denver	School of Public Affairs
University of Delaware	School of Public Policy &Administration
University of Georgia	Dept. of Public Administration &Policy
University of Illinois at Chicago	Dept. of Public Administration
University of Illinois at Springfield	College of Public Affairs &Administration
University of Kansas	School of Public Affairs &Administration
University of Kentucky	Martin School of Public Policy &Administration
University of Louisville	Urban &Public Affairs
University of Maryland at Baltimore County	School of Public Policy
University of Maryland at College Park	School of Public Policy
University of Massachusetts at Boston	Dept. of Public Policy &Public Affairs
University of Michigan	School of Public Policy
University of Minnesota at Twin Cities	Humphrey School of Public Affairs
University of Missouri at Columbia	Truman School of Public Affairs
University of Missouri at Kansas City	Dept. of Public Affairs
University of Nebraska at Omaha	School of Public Administration
University of Nevada at Las Vegas	School of Environmental &Public Affairs
University of North Texas	Dept. of Public Administration
University of Pittsburgh	Graduate School of Public &International Affairs
University of Southern California	Price School of Public Policy
University of Texas at Arlington	Dept. of Public Affairs
University of Texas at Austin	Lyndon B. Johnson School of Public Affairs
University of Texas at Dallas	School of Economic, Political &Policy Sciences
University of Washington	Daniel J. Evans School of Public Policy &Governance
Virginia Polytechnic Institute &State University	School of Public &International Affairs
Western Michigan University	School of Public Affairs &Administration

Table A.3
List of 26 neuroscience departments.

University	Academic unit
Baylor College of Medicine	Department of Neuroscience
Brown University	Graduate Department of Neuroscience
Georgia State University	Neuroscience Institute
Icahn School of Medicine at Mount Sinai	Department of Neuroscience
University of Texas At Austin	Department of Neuroscience
Oregon Health and Science University	Department of Behavioral Neuroscience
Johns Hopkins University	Department of Neuroscience
Thomas Jefferson University	Department of Neuroscience
University of California at San Diego	Department of Neurosciences
University of Minnesota	Department of Neuroscience
University of Pennsylvania	Department of Neuroscience
University of Pittsburgh	Department of Neuroscience
University of Rochester	Department of Neuroscience
University of Wisconsin at Madison	Department of Neuroscience
Albert Einstein College of Medicine	Dominick P. Purpura Department of Neuroscience
Columbia University	Department of Neuroscience
Mayo Clinic	Department of Neuroscience
Medical University of South Carolina	Department of Neuroscience
Ohio State University	Department of Neuroscience
Rosalind Franklin University	Department of Neuroscience
Stanford University	Department of Neurology &Neurological Sciences
Temple University	Department of Neuroscience
University of Florida	Department of Neuroscience
University of New Mexico	Department of Neurosciences
Washington University	Department of Neuroscience
Yale University	Department of Neuroscience

our data collection. For example, the information school at the University of California at Los Angeles is the graduate school of education and information studies. In this case, we only included the department of information studies in the dataset, whereas the department of education was excluded. While these institutions are selected based on accreditation, they are representative and the major players of the three disciplines.

For all core faculty members (full-time tenured or tenure-track) in these institutions, we collected their educational backgrounds (including PhD programs and degree-granting institutions) from their personal or schools websites. Those with titles such as emeritus, adjunct, or visiting professors were not included. Our final dataset includes 708 faculty mem-

Table A.4
Disciplinary classifications for information, policy, and neuroscience.

Category	Included area(s)
Information schools	
Communication	Media and Mass Communication; Journalism
Computing	Computer Science; Electrical Engineering; Mathematics; Computer Engineering
Education	Education; Learning Technology
Humanities	History; English; Philosophy; Literature; Music; Geography; Art; Anthropology
Information	Information Science; Information Studies; Information Transfer; Informatics
Library	Library Science; Information and Library Science
Management &Policy	Business Administration; Management; Policy; Economics; City &Regional Planning; Public Administration
Science &Engineering	Life Sciences; Physics; Statistics; Engineering (not Electrical); Biology
Social &Behavioral	Psychology; Sociology; Law; Social Sciences; Linguistics; Political Science; Government
Policy schools	
Policy	Interdisciplinary Evaluation; Sociomedical Sciences; Sustainable Development; Health/Environmental Policy; Public Policy/Administration/Affairs/Health/Service
Business	Organizational Behavior; Operations Research; Management; Business Administration; Finance; Accounting
Planning	Urban Studies/Affairs; Regional/Urban/City planning
Humanities	Cultural Studies; Architecture; Anthropology; History; English; Philosophy; Literature
Social &Behavioral	Human Resource; Communication and Journalism; Industrial Relations; Demography; Geography; Economics; International affairs/relations; Government; Psychology; Law (including Criminology); Education; Sociology; Politics; Conflict
Science &Engineering	Agriculture; Linguistics; Ecology; Computing; Physics; Chemistry; Statistics; Library and Information Science
Medicine	Medicine; Toxicology
Neuroscience departments	
Formal	Mathematics; Applied Mathematics
Physical	(Bio)Physics; (Bio)Chemistry; Experimental Physics; Plasma Physics
Life	Biology; Genetics; Anatomy; Developmental Biology; Microbiology; Toxicology; Zoology; Physiology; Cell and Molecular Biology; Immunology; Biomedical Science
Health	Medicine; Veterinary; Pharmacy; Pathology
Social	Anthropology; Psychology; Education
Neuroscience	Neuroscience
Engineering	Computer Science/Engineering; Biomedical Engineering; Electrical Engineering; Bioengineering; Biomaterial engineering
Interdisciplinary	Behavioral Sciences; Cognitive Science; Complex Systems; Brain Sciences; Movement science; Biological Cybernetics; Kinesiology

bers for information schools, 1537 for policy Schools, and 620 for neuroscience departments. Program names are highly unstructured. Therefore, we applied disciplinary classifications (Table A.4; left column in Fig. A.1) to individual programs to unify doctoral programs. Based on the taxonomy proposed by Zuo et al. (2017), doctoral programs where information schools faculty members received their doctorates were classified into nine categories: communication, computing, education, humanities, information, library, management & policy, science & engineering, social & behavioral. Following a similar scheme, we categorized the doctoral programs of policy school scholars into social & behavioral, policy, planning, science & engineering, business, humanities, medicine, and other. For faculty members in neuroscience department, the categories are formal science, physical science, life science, health science, social science, neuroscience, engineering, and interdisciplinary science.[5] It is worth noting that some policy school and neuroscience department faculty members graduated from doctoral programs with names spanning more than one categories. In this case, we allow multiple assignments – a faculty members PhD program may be assigned to different categories. For example, if a policy professor graduated from a program named *urban affairs & public policy*, we would categorize her doctoral program into both planning and policy.

Research collaborations among researchers can occur in different ways, but co-authoring papers has been used as a powerful and valid proxy in previous studies (Katz & Martin, 1997; Newman, 2004; Qiu, 1992). We used Scopus as the source for faculty members' publication data (right column in Fig. A.1). For information school faculty members, many tend to publish in conference proceedings. Policy school faculty members, besides journals, are very likely to publish books/book series. Researchers in neuroscience departments mainly publish their papers in journals. Titles, abstracts, authors' names and affiliations, as well as publication dates of each faculty member's papers were retrieved from Scopus APIs,[6] based on author name and affiliation. Finally, we gathered 23,758 publications for information, 24,903 for policy, and 45,550 for neuroscience.

[5] https://en.wikipedia.org/wiki/Science.

[6] The data of (i) information school was downloaded between November 2014 and May 2015; (ii) policy school was downloaded between February and April 2016; (iii) neuroscience department was downloaded between November 2016 and February 2017 from Scopus API via http://api.elsevier.com and http://www.scopus.comhttp://www.scopus.com.

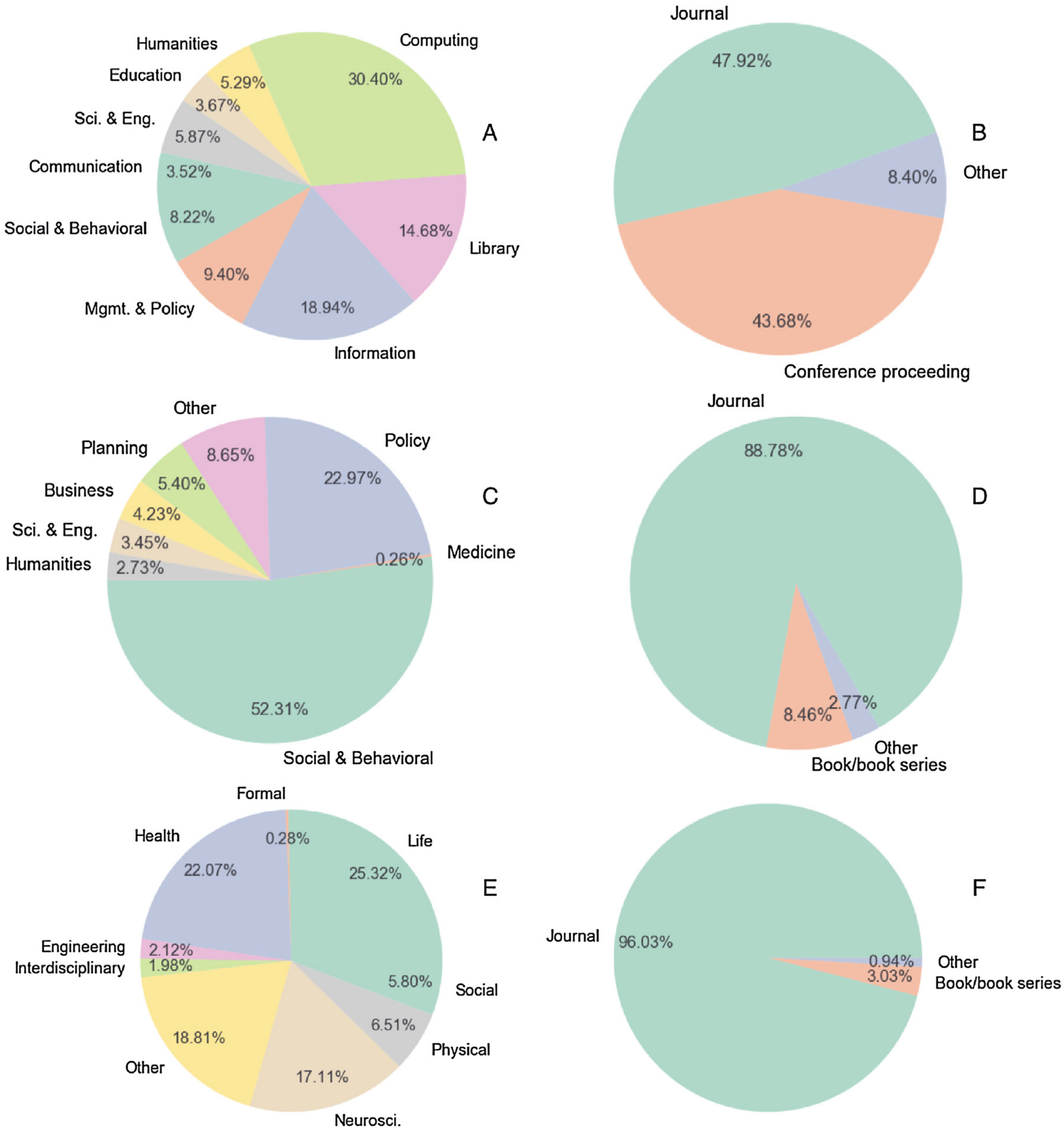

Fig. A.1. Distributions of doctorate programs (left) and publication types (right) for information (A and B), policy (C and D), and neuroscience (E and F)).

Appendix B. Topic modeling results

We applied LDA on faculty publication data for information, policy, and neuroscience separately, using Matlab Topic Modeling Toolbox.[7] We followed Griffiths and Steyvers (2004) and set the hyperparameters to $\alpha = \frac{50}{K}$; $\beta = 0.1$, where α and β are symmetric Dirichlet priors for document-topic and topic-word multinomials, θ and ϕ respectively; K is the number of topics. The results were retrieved from a single Gibbs sampler after 1000 iterations. We tried $K \in \{20, 30, 40, 50\}$ and calculated research multidisciplinarity. Across the three disciplines, pairwise Pearson correlations between research multidisciplinaritiese range from 0.87 to 0.99, implying consistent results independent of K (Fig. B.1). We pick $K = 20$ and show the topical interpretations along with the top keywords (Table B.1).

[7] http://psiexp.ss.uci.edu/research/programs_data/toolbox.htm.

Table B.1
Topic distribution ($K = 20$).

Topic interpretation	Proportion	Representative keywords
Information schools		
IT for Collaboration and Communication	7.31%	Information; technology; communication; practice; collaboration
Software and System Engineering	5.91%	Design; system; develop; software; process
Information Privacy and Policy	5.83%	Privacy; policy; govern; market; internet
Social Networks and Media	5.62%	Social; community; online; media; network
Machine Learning and Data mining	5.48%	Measure; perform; test; predict; data
Information Retrieval and Recommendation	5.28%	Inform; user; search; web; query
Computing Infrastructure	5.18%	Application; system; service; compute; distribute
Cybersecurity and Networks	5.17%	Network; secure; scheme; attack; node
Digital Library and Library Science	5.14%	Library; digit; public; author; collect
User Interface and Experience	5.05%	User; design; interface; interact; mobile
Text mining	4.86%	Document; retrieve; text; term; topic
Algorithms	4.75%	Algorithm; optimal; time; space; efficiency
Data Storage and Visualization	4.66%	Data; visual; analysis; inform; collect
Education and Learning Technology	4.63%	Learn; student; education; compute; school
Robotics and Cognitive Systems	4.59%	Robot; human; agent; game; behavior
Health Informatics	4.56%	Health; patient; care; medic; inform
Programming Languages	4.50%	Program; language; type; function; structure
Spatial and Multimedia Data Analytics	4.22%	Image; locate; spatial; video; object
Bioinformatics	3.69%	Sequence; protein; gene; genomics; structure
Others	3.58%	Simulate; energy; measure; process; structure
Policy schools		
Info and Tech Management	6.59%	System; communication; develop; manage; inform
Social Policy	6.31%	Social; policy; theory; issue; science
Public Admin and Government	6.00%	Public; govern; manage; organ; perform
Data Analysis Methods	5.97%	Data; estimate; measure; analysis; variable
Patient Studies	5.70%	Patient; hospital; cancer; treatment; associate
Family Studies	5.50%	Children; age; family; child; associate
Social Studies	5.34%	Study; survey; social; behavior; individual
Healthcare Policy and Management	5.14%	Health; care; service; insurance; cost
International Relationship	5.10%	Country; intern; global; develop; economics
Finance and Econ	5.03%	Market; price; capital; finance; rate
Labor	4.96%	Employ; change; income; increase; immigration
Politics and Legislation	4.80%	Policy; politics; public; federal; govern
Tech Innovation	4.65%	Technology; network; innovation; knowledge; develop
Social Welfare	4.58%	Cost; program; effect; benefit; tax
Ecology	4.41%	Concentrate; water; measure; lake; organ
Environmental Policy	4.37%	Environment; energy; climate; change; emission
City Management and Urban Policy	4.25%	City; urban; house; neighborhood; region
Education	3.91%	School; education; student; black; racial
Epidemiology and Public Health	3.83%	Vaccine; population; effect; species; infect
Crime and Drug Abuse	3.55%	Treatment; drug; crime; violence; HIV
Neuroscience departments		
Seizure disorder	6.43%	Patient; clinic; study; treatment; seizure
Visual &Movement	5.92%	Visual; movement; response; active; direct
Memory	5.49%	Memory; age; cognitive; impair; brain
Brain Function	5.46%	System; function; brain; develop; mechanics
Dopamine	5.27%	Effect; behavior; dopamine; rat; drug
Peptide-Membrane Interaction	5.26%	Protein; membrane; domain; bind; peptide
Protein Regulation	5.20%	Active; protein; regulation; signal; express
Alzheimer	5.20%	Disease; Alzheimer; mice; mutate; protein
Neural Coding	5.18%	Response; neuron; frequency; active; cell
Axon Development	5.13%	Cell; neuron; axon; develop; growth
Astrocyte	5.05%	Induce; cell; increase; astrocyte; active
Circadian Rhythm	4.88%	Rat; male; female; increase; day
Cerebral Cortex	4.74%	Cell; neuron; layer; cortex; distribute
Synaptic Plasticity	4.70%	Neuron; synaptic; synapse; active; plasticity
Immune &HIV	4.69%	Cell; express; gene; mRNA; human
Neuroimaging	4.57%	Image; data; measure; analysis; base
GABA Receptor	4.48%	Receptor; GABA; effect; bind; subunit
Membrane Depolarization	4.38%	Channel; current; cell; calcium; active
Alcohol	4.05%	Ethanol; mice; gene; genetic; alcohol
Spinal Cord	3.92%	Neuron; spinal; cord; nerve; nucleus

Fig. B.1. Pearson correlation between research multidisciplinarity in information, policy, and neuroscience with respect to different *K*'s.

Appendix C. Caveats for multi- and inter-disciplinarity measures

We introduced three measures: two for multidisciplinarity and one for interdisciplinarity. Educational multidisciplinarity quantifies the extent to which faculty members in the same institutions have doctorate degrees from different areas, whereas research multidisciplinarity looks at topical diversities mined from publication papers. Interdisciplinarity, instead of looking at faculty composition, gauges knowledge integration from papers co-authored by faculty colleagues. It is possible that an

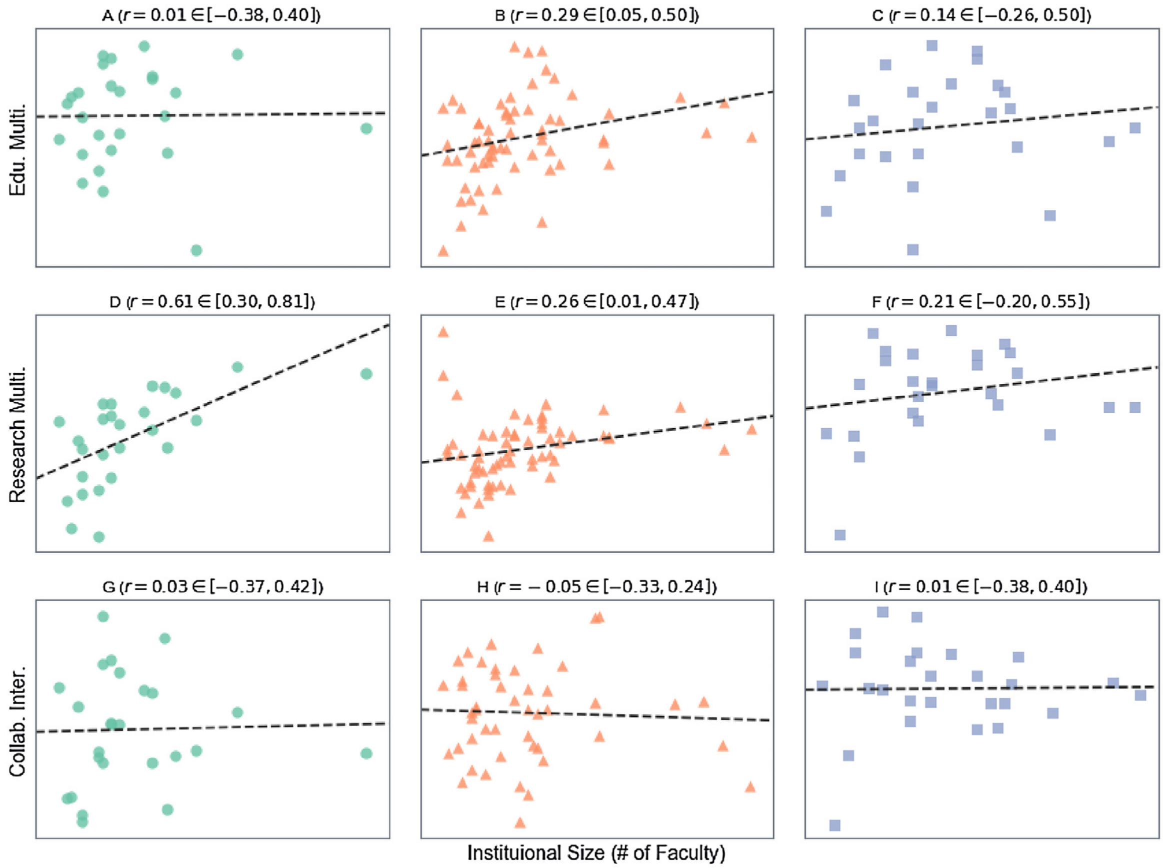

Fig. C.1. Scatter plots between institutional size and (i) educational multidisciplinarity (the first row), (ii) research multidisciplinarity (the second row), and (iii) collaboration interdisciplinarity (the third row). The three columns are for information (A and D), public policy (B and E), and neuroscience (C and F), respectively. Each marker is an institution: (i) dot for information; (ii) triangle for policy; and (iii) square for neuroscience. The dashed straight line is fitted by least square.

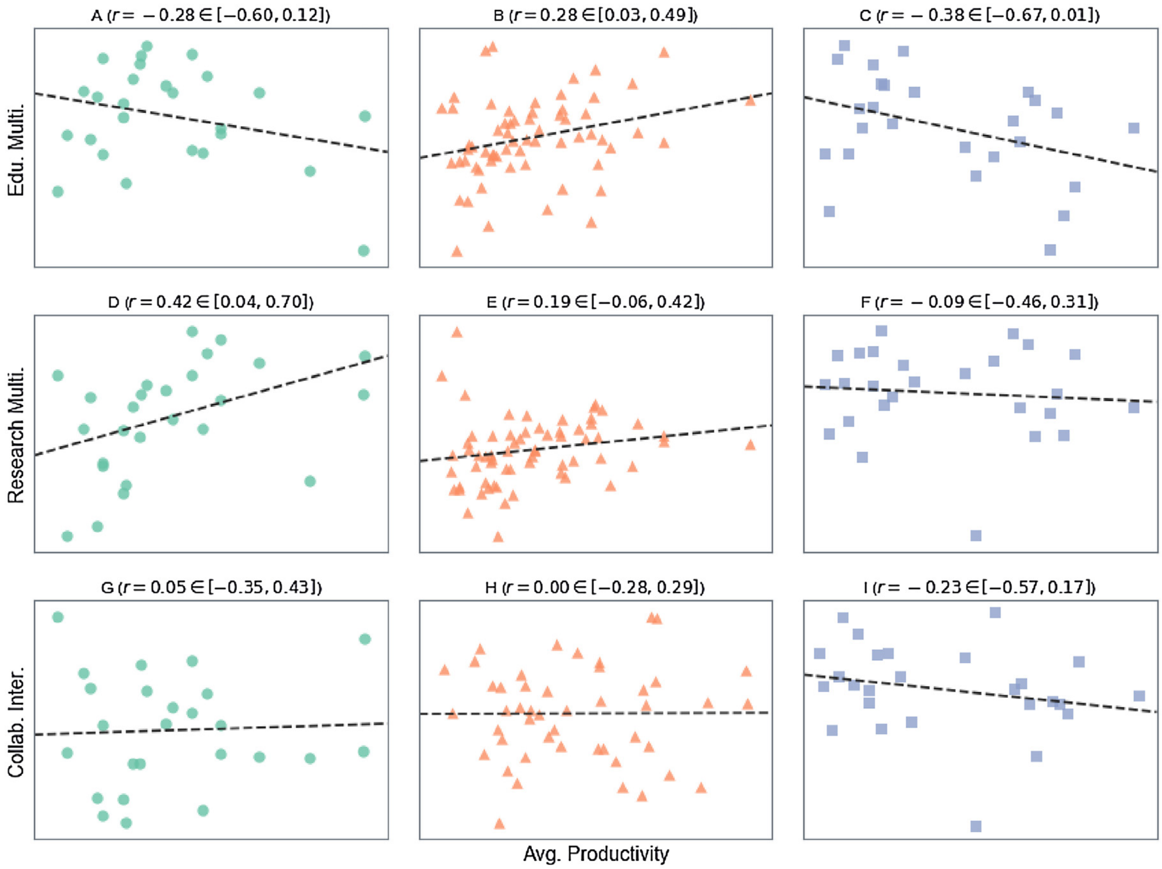

Fig. C.2. Scatter plots between institutional productivity and (i) educational multidisciplinarity (the first row), (ii) research multidisciplinarity (the second row), and (iii) collaboration interdisciplinarity (the third row). The three columns are for information (A and D), public policy (B and E), and neuroscience (C and F), respectively.

institution of a smaller size or lower productivity may be benefit from how we measure multi- and inter-disciplinarities, due to the limited number of graduate programs represented and research topics among faculty members.

In this appendix, we show evidence that our proposed measures are unbiased, with respect to institutional size (Fig. C.1) and productivity (Fig. C.2). Our correlation results suggest that there is generally no strong correlation, implying that our proposed measures are robust for institutions with various sizes and productivities. The only exception is information school, where there is a correlation of 0.61 between research multidisciplinarity and institutional sizes (Fig. C.1D). This echoes with our previous discussion in Section 5 – information schools manifest a slightly different collaboration atmosphere compared to policy schools and neuroscience departments. More exploration is needed to investigate why information schools are different from the other two areas.

Appendix D. Regression analysis and results

After controlling for institution size, research intensity, and average productivity, educational multidisciplinarity has little effect on either measure of collaboration prevalence (Fig. D.1A–F). The corresponding partial correlations are weak across three areas. The proportionate variance in collaboration prevalence uniquely contributed by educational multidisciplinarity takes small values (ranging from 0 to 6.12%).

On the other hand, while there is little correlation between research multidisciplinarity and either collaboration prevalence in information schools (Fig. D.1G and J), research multidisciplinarity has strong and negative effects on both the average percentage of ICPs and co-authorship network density within homogenous groups of institution size, Carnegie classification, and average productivity in public policy schools and neuroscience departments (Fig. D.1H, I, K, and L). Research multidisciplinarity is shown to account for a relatively good amount of variance (12% to 34%) in both average percentage of ICPs and co-authorship network density, except for information schools (0 and 6%). As for the interdisciplinarity of collaborations, the nearly flat-lined relationship between educational multidisciplinarity and collaboration interdisciplinarity (Fig. D.2A–C) indicate no association. Across information, public policy, and neuroscience, collaboration interdisciplinarity goes higher

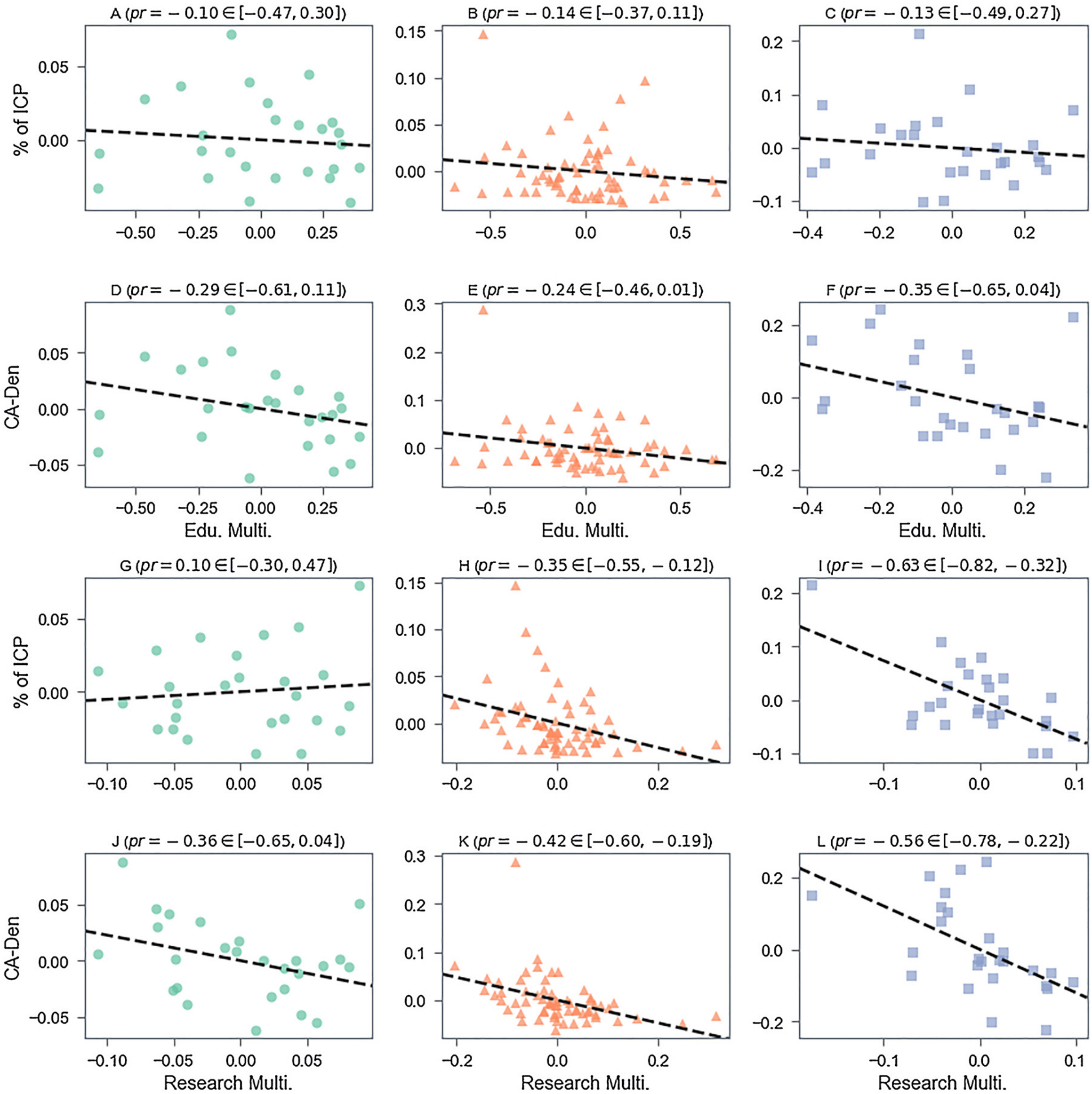

Fig. D.1. Scatter plots between residuals of collaboration prevalence and multidisciplinarity. The three columns are for information (A, D, G, and J), public policy (B, E, H, and K), and neuroscience (C, F, I, L), respectively. The four rows are the partial correlation for average % of ICPs vs. educational multidisciplinarity (A, B, and C); Co-authorship network density (CA-DEN) vs. educational multidisciplinarity (D, E, and F); average % of ICPs vs. research multidisciplinarity (G, H, and I); CA-DEN vs. research multidisciplinarity (J, K, and L), within homogeneous groups of institutional size, Carnegie classification, and average productivity. *pr* is partial correlations.

when the level of research multidisciplinarity increases (Fig. D.2D–F). With respect to the prediction power, educational multidisciplinarity merely contributes 1.4%, 3.6%, and 6.8% to the variance in collaboration interdisciplinarity in information schools, public policy schools, and neuroscience departments respectively, as opposed to 50%, 10.2%, and 16.5% for research multidisciplinarity. As a result, educational backgrounds have much less valuable information in explaining or predicting either measurement of collaborations. Research topics, which were mined from massive collections of documents using topic modeling algorithms, did a better job at accounting for collaboration characteristics in these three types of multidisciplinarity institutions.

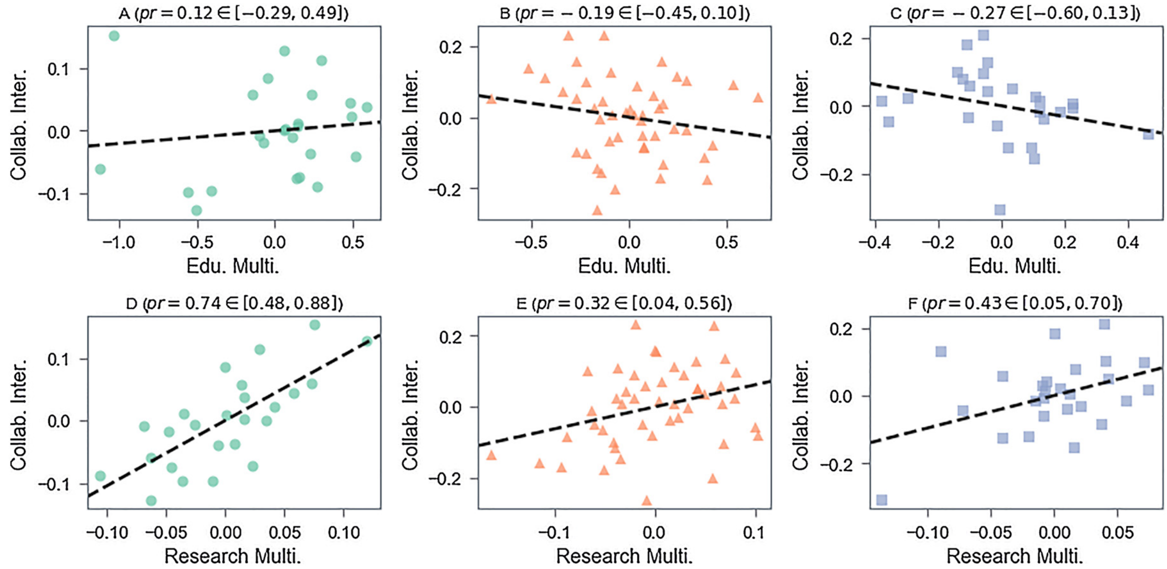

Fig. D.2. Scatter plots between residuals of collaboration interdisciplinarity and multidisciplinarity. The three columns are for information (A and D), public policy (B and E), and neuroscience (C and F), respectively. The two rows are the partial correlation for collaboration interdisciplinarity vs. educational multidisciplinarity (A, B, and C) and collaboration interdisciplinarity vs. research multidisciplinarity (D, E, and F), within homogeneous groups of institutional sizes, Carnegie classification, average productivity, and internal collaboration prevalence.

References

Adams, J. (2012). Collaborations: The rise of research networks. *Nature*, *490*, 335–336. http://dx.doi.org/10.1038/490335a

Adams, J., & Light, R. (2014). Mapping interdisciplinary fields: Efficiencies, gaps and redundancies in HIV/AIDS research. *PLOS ONE*, *9* http://dx.doi.org/10.1371/journal.pone.0115092

Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Services Research*, *41*, 1576–1598. http://dx.doi.org/10.1111/j.1475-6773.2006.00566.x arXiv:NIHMS150003

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. arXiv:1111.6189v1

Bordons, M., Zulueta, M. A., Romero, F., & Barrigón, S. (1999). Measuring interdisciplinary collaboration within a university: The effects of the multidisciplinary research programme. *Scientometrics*, *46*, 383–398. http://dx.doi.org/10.1007/BF02459599

Bu, Y., Ding, Y., Liang, X., & Murray, D. S. (2017). Understanding persistent scientific collaboration. *Journal of the Association for Information Science and Technology*, http://dx.doi.org/10.1002/asi.23966

Bu, Y., Ding, Y., Xu, J., Liang, X., Gao, G., & Zhao, Y. (2018). Understanding success through the diversity of collaborators and the milestone of career. *Journal of the Association for Information Science and Technology*, *69*, 87–97. http://dx.doi.org/10.1002/asi.23911

Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, *1*, e1400005. http://dx.doi.org/10.1126/sciadv.1400005

Clauset, A., Larremore, D. B., & Sinatra, R. (2017). Data-driven predictions in the science of science. *Science*, *355*, 477–480. http://dx.doi.org/10.1126/science.aal4217

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*, 829–836. http://dx.doi.org/10.1080/01621459.1979.10481038

Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science*, *35*, 703–722. http://dx.doi.org/10.1177/0306312705055535

Cummings, J. N., & Kiesler, S. (2008). Who collaborates successfully? In *Proceedings of the ACM 2008 conference on computer supported cooperative work – CSCW '08*. pp. 437. New York, New York, USA: ACM Press. http://dx.doi.org/10.1145/1460563.1460633

De Abreu Dos Reis, C. R., Sastre Castillo, M. Á., & Roig Dobón, S. (2007). Diversity and business performance: 50 years of research. *Service Business*, *1*, 257–274. http://dx.doi.org/10.1007/s11628-007-0022-2

Derrick, E. G., Falk-Krzesinski, H. J., Roberts, M. R., & Olson, S. (2012). *Facilitating interdisciplinary research and education: A practical guide*. http://www.aaas.org/cspsp/interdisciplinary/guide/

Disis, M. L., & Slattery, J. T. (2010). The road we must take: Multidisciplinary team science. *Science Translational Medicine*, *2*, 22cm9. http://dx.doi.org/10.1126/scitranslmed.3000421

Dong, Y., Ma, H., Shen, Z., & Wang, K. (2017). A century of science. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining – KDD '17*. pp. 1437–1446. New York, New York, USA: ACM Press. http://dx.doi.org/10.1145/3097983.3098016 arXiv:1704.05150

Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of director diversity and firm financial performance. *Corporate Governance*, *11*, 102–111. http://dx.doi.org/10.1111/1467-8683.00011

Gibbons, B. (2010). Celebrating diversity in organizations: An analysis of workplace initiatives. *International Journal of Knowledge, Culture and Change Management*, *9*, 27–37.

Gowanlock, M., & Gazan, R. (2013). Assessing researcher interdisciplinarity: A case study of the University of Hawaii NASA Astrobiology Institute. *Scientometrics*, *94*, 133–161. http://dx.doi.org/10.1007/s11192-012-0765-y arXiv:1204.5563

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*, 5228–5235.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal*, *45*, 1029–1045. http://dx.doi.org/10.2307/3069328

He, Z.-L., Geng, X.-S., & Campbell-Hunt, C. (2009). Research collaboration and research output: A longitudinal study of 65 biomedical scientists in a New Zealand University. *Research Policy*, *38*, 306–317. http://dx.doi.org/10.1016/j.respol.2008.11.011

Indiana University Center for Postsecondary Research. (2015). *The Carnegie Classification of Institutions of Higher Education* (2015th ed.). Bloomington.

Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent research on team and organizational diversity: SWOT analysis and implications. *Journal of Management*, *29*, 801–830. http://dx.doi.org/10.1016/S0149-2063_03_00080-1

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-University research teams: Shifting impact, geography, and stratification in science. *Science*, *322*, 1259–1262. http://dx.doi.org/10.1126/science.1158357
Katz, J., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*, 1–18. http://dx.doi.org/10.1016/S0048-7333(96)00917-1
King, C. (2012). *Multiauthor papers: Onward and upward*. http://archive.sciencewatch.com/newsletter/2012/201207/multiauthor_papers/
van Knippenberg, D., van Ginkel, W. P., & Homan, A. C. (2013). Diversity mindsets and the performance of diverse teams. *Organizational Behavior and Human Decision Processes*, *121*, 183–193. http://dx.doi.org/10.1016/j.obhdp.2013.03.003
Lash, M. T., & Zhao, K. (2016). Early predictions of movie success: The who, what, and when of profitability. *Journal of Management Information Systems*, *33*, 874–903. http://dx.doi.org/10.1080/07421222.2016.1243969 arXiv:1506.05382
Ledford, H. (2015). How to solve the world's biggest problems. *Nature*, *525*, 308–311. http://dx.doi.org/10.1038/525308a
Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, *35*, 673–702. http://dx.doi.org/10.1177/0306312705052359
Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinary of scientific journals. *Journal of the American Society for Information Science and Technology*, *58*, 1303–1319. http://dx.doi.org/10.1002/asi.20614 arXiv:0911.2719
McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444. http://dx.doi.org/10.1146/annurev.soc.27.1.415 arXiv:9702163
Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*, 208701. http://dx.doi.org/10.1103/PhysRevLett.89.208701
Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 5200-5. http://dx.doi.org/10.1073/pnas.0307545100
Nichols, L. G. (2014). A topic model approach to measuring interdisciplinarity at the National Science Foundation. *Scientometrics*, *100*, 741–754. http://dx.doi.org/10.1007/s11192-014-1319-2
NSF. (2005). *Facilitating interdisciplinary research*. Washington, DC: Technical Report National Academy of Sciences. http://dx.doi.org/10.17226/11153 http://www.nap.edu/catalog/11153
Pan, R. K., Sinha, S., Kaski, K., & Saramäki, J. (2012). The evolution of interdisciplinarity in physics research. *Scientific Reports*, *2* http://dx.doi.org/10.1038/srep00551 arXiv:1206.0108
Pavlidis, I., Petersen, A. M., & Semendeferi, I. (2014). Together we stand. *Nature Physics*, *10*, 700–702. http://dx.doi.org/10.1038/nphys3110
Petersen, A. M. (2015). Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, E4671–E4680. http://dx.doi.org/10.1073/pnas.1501444112
Popper, K. R. (1962). *Conjectures and refutations: The growth of scientific knowledge* (2nd ed.). New York: Basic Books Inc.
Porac, J. F., Wade, J. B., Fischer, H. M., Brown, J., Kanfer, A., & Bowker, G. (2004). Human capital heterogeneity, collaborative relationships, and publication patterns in a multidisciplinary scientific alliance: A comparative case study of two scientific teams. *Research Policy*, *33*, 661–678. http://dx.doi.org/10.1016/j.respol.2004.01.007
Porter, A., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, *81*, 719–745. http://dx.doi.org/10.1007/s11192-008-2197-2
Porter, A. L., Cohen, A. S., David Roessner, J., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics*, *72*, 117–147. http://dx.doi.org/10.1007/s11192-007-1700-5
Qiu, L. (1992). A study of interdisciplinary research collaboration. *Research Evaluation*, *2*, 169–175. http://dx.doi.org/10.1093/rev/2.3.169
Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, *82*, 263–287. http://dx.doi.org/10.1007/s11192-009-0041-y arXiv:0901.1380
Regalado, A. (1995). Multiauthor papers on the rise. *Science*, *268*, 25.
van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*, *40*, 463–472. http://dx.doi.org/10.1016/j.respol.2010.11.001
Salazar, M. R., Lant, T. K., Fiore, S. M., & Salas, E. (2012). Facilitating innovation in diverse science teams through integrative capacity. *Small Group Research*, *43*, 527–558. http://dx.doi.org/10.1177/1046496412453622
Sayama, H., & Akaishi, J. (2012). Characterizing interdisciplinarity of researchers and research topics using web search engines. *PLoS ONE*, *7*, e38747. http://dx.doi.org/10.1371/journal.pone.0038747 arXiv:1201.3592
Schummer, J. (2004). Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics*, *59*, 425–465. http://dx.doi.org/10.1023/B:SCIE.0000018542.71314.38
Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x. arXiv:9411012
Shore, L. M., Chung-Herrera, B. G., Dean, M. A., Ehrhart, K. H., Jung, D. I., Randel, A. E., & Singh, G. (2009). Diversity in organizations: Where are we now and where are we going? *Human Resource Management Review*, *19*, 117–133. http://dx.doi.org/10.1016/j.hrmr.2008.10.004 arXiv:arXiv:1011.1669v3
Siciliano, J. I. (1996). The relationship of board member diversity to organizational performance. *Journal of Business Ethics*, *15*, 1313–1320. http://dx.doi.org/10.1007/BF00411816
Sinatra, R., Wang, D., Deville, P., Song, C., & Barabási, A. L. (2016). Quantifying the evolution of individual scientific impact. *Science*, *354*, aaf5239. http://dx.doi.org/10.1126/science.aaf5239
Stvilia, B., Hinnant, C. C., Schindler, K., Worrall, A., Burnett, G., Burnett, K., Kazmer, M. M., & Marty, P. F. (2011). Composition of scientific teams and publication productivity at a national science lab. *Journal of the American Society for Information Science and Technology*, *62*, 270–283. http://dx.doi.org/10.1002/asi.21464
Su, X., Johnson, J., & Bozeman, B. (2015). Gender diversity strategy in academic departments: Exploring organizational determinants. *Higher Education*, *69*, 839–858. http://dx.doi.org/10.1007/s10734-014-9808-z
Van Hartesveldt, C., & Giordan, J. (2008). *Impact of Transformative Interdisciplinary Research and Graduate Education on Academic Institutions*. Technical Report May National Science Foundation. http://dx.doi.org/10.1001/jama.2010.987
Van Noorden, R. (2015). Interdisciplinary research by the numbers. *Nature*, *525*, 306–307. http://dx.doi.org/10.1038/525306a
Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., Rafols, I., & Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*, *5*, 14–26. http://dx.doi.org/10.1016/j.joi.2010.06.004
Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLOS ONE*, *10*, e0127298. http://dx.doi.org/10.1371/journal.pone.0127298
Wang, L., Notten, A., & Surpatean, A. (2013). Interdisciplinarity of nano research fields: A keyword mining approach. *Scientometrics*, *94*, 877–892. http://dx.doi.org/10.1007/s11192-012-0856-9
Whitfield, J. (2008). Collaboration: Group theory. *Nature*, *455*, 720–723. http://dx.doi.org/10.1038/455720a
Wiggins, A., & Sawyer, S. (2012). Intellectual diversity and the faculty composition of iSchools. *Journal of the American Society for Information Science and Technology*, *63*, 8–21. http://dx.doi.org/10.1002/asi.21619
Wilde, D. (2010). Personalities into. *Mechanical Engineering*, *132*.
Woolley, R., Sánchez-Barrioluengo, M., Turpin, T., & Marceau, J. (2015). Research collaboration in the social sciences: What factors are associated with disciplinary and interdisciplinary collaboration? *Science and Public Policy*, *42*, 567–582. http://dx.doi.org/10.1093/scipol/scu074
Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316.
Xie, Z., Duan, X., Ouyang, Z., & Zhang, P. (2015). Quantitative analysis of the interdisciplinarity of applied mathematics. *PLOS ONE*, 10. http://dx.doi.org/10.1371/journal.pone.0137424

Yegros-Yegros, A., Rafols, I., & D'Este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PLoS ONE*, *10*, e0135095. http://dx.doi.org/10.1371/journal.pone.0135095

Zhang, K., & Pelechrinis, K. (2014). Understanding spatial homophily. In *Proceedings of the 23rd international conference on World wide web – WWW '14* (pp. 271–282). http://dx.doi.org/10.1145/2566486.2567990

Zhao, K., Ngamassi, L. M., Yen, J., Maitland, C., & Tapia, A. (2010). *Assortativity patterns in multi-dimensional inter-organizational networks: A case study of the humanitarian relief sector*. pp. 265–272. *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6007) LNCS. http://dx.doi.org/10.1007/978-3-642-12079-4_33

Zhao, X., Huo, B., Selen, W., & Yeung, J. H. Y. (2011). The impact of internal integration and relationship commitment on external integration. *Journal of Operations Management*, *29*, 17–32. http://dx.doi.org/10.1016/j.jom.2010.04.004

Zhu, Y., Yan, E., & Song, M. (2016). Understanding the evolving academic landscape of library and information science through faculty hiring data. *Scientometrics*, *108*, 1461–1478. http://dx.doi.org/10.1007/s11192-016-2033-z

Zuo, Z., Zhao, K., & Eichmann, D. (2017). The state and evolution of U.S. iSchools: From talent acquisitions to research outcome. *Journal of the Association for Information Science and Technology*, *68*, 1266–1277. http://dx.doi.org/10.1002/asi.23751