Directions for Professional Social Matching Systems

Thomas Olsson, Jukka Huhtamäki, Hannu Kärkkäinen

Key Insights

- Professional Social Matching (PSM) is an emergent and potentially very impactful area of social matching systems, building on recommender systems, decision-support systems, social network analysis, and machine learning.
- Considering the ethics of computationally influencing professional matching activities, such as team formation and networking, the current computational approaches for profiling, matching, and recommending actors need to be reconsidered.
- Future PSM systems should aim to enable unexpected encounters and social serendipity, incorporate a systemic perspective to the matching logic, help avoid human bias in decision-making, and identify optimal similarity-diversity trade-offs between actors.
- We argue that future PSM systems need to (a) be calibrated for different matching cases to cater for contextual dynamics, (b) be based on multi-dimensional analytics of profile and contextual data, (c) support meaningful cooperation between the end user and the matching system, and (d) feature proactive nudging to help the users avoid inherent human prejudice.

Abstract

Social matching, that is, the identification and facilitation of new social connections between people, is ubiquitous in professional life, yet little supported by computational systems. Taking a human-centric and systemic perspective, we review literature in computer science, management science, and social sciences to outline an emerging problem space and provide new design and research directions towards socially and ethically sustainable Professional Social Matching (PSM) systems. Our analysis indicates that many current approaches in recommender systems and social network analysis, when applied to PSM, could have detrimental long-term implications for collaboration practices and organizations' performance. For example, conventional mechanisms, such as optimizing for similarity and triadic closure, involve risks of strengthening the homophily bias and echo chambering. We call for diversity-enhancing yet contextually sensitive designs of future PSM systems, tapping on multi-dimensional analytics and meaningful cooperation between the user and the matching system. Our analysis of this emerging application area underlines the importance and urgency of rethinking the paradigmatic algorithmic solutions.

Introduction

Supporting human collaboration has been a central driver of the development of information and communication technology. A relatively recent approach to this end is *social matching*, referring to computational ways of identifying and facilitating new social connections between people [36]. Social matching is most often connected with partnering for leisurely and romantic relationships—in fact, the

most well-known social matching systems revolve around dating scenarios (e.g., Tinder[1]) or triggering opportunistic interactions with strangers (e.g., Happn[2]).

This paper focuses on *Professional Social Matching* (PSM), which we define as the matching of individuals or groups for vocational collaboration and co-creation of value. This covers organizational activities, including recruitment, headhunting, community building, and team formation within or across organizations as well as individually driven activities like mentoring, seeking advisory relationships, and general networking.

From a technological perspective, computer-supported PSM is based on computational approaches to profiling actors (organizations and individuals), modeling their qualities, analyzing their mutual social suitability and relevance, and presenting the recommendations to the users. For example, prescriptive data analytics [9] can utilize social network analysis (SNA) for explicating the social ties between actors and machine learning-based approaches to analyze their competences and interests and to identify suitable pairs of actors. The resulting computational system can manifest as proactive *people recommender systems*—or *social recommender systems* [15]—or other kinds of data-driven *decision-support systems* [31] for social matching. Conventional examples include recommendations of other professionals to follow in various social media services as well as more specific expert search systems [40].

This fusion of various analytical and algorithmic approaches for a broad and multi-faceted application area forms an intriguing, inter-disciplinary space for research and development of novel algorithms and information systems. This paper reviews relevant literature in three key scientific domains that present relevant understanding necessary to reach the potential of future computational systems (Figure 1).
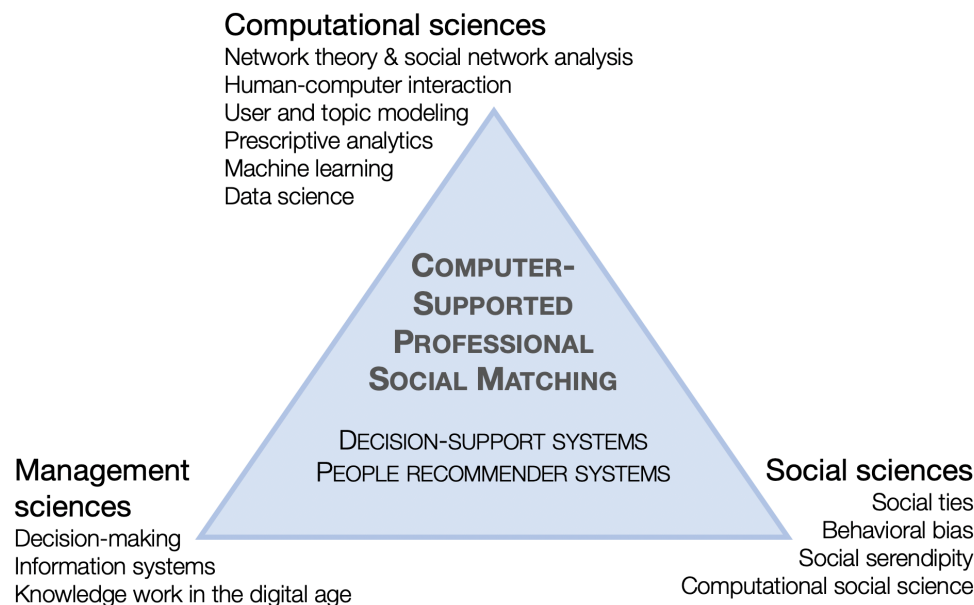
Computational sciences
Network theory & social network analysis
Human-computer interaction
User and topic modeling
Prescriptive analytics
Machine learning
Data science

COMPUTER-SUPPORTED PROFESSIONAL SOCIAL MATCHING

DECISION-SUPPORT SYSTEMS
PEOPLE RECOMMENDER SYSTEMS

Management sciences
Decision-making
Information systems
Knowledge work in the digital age

Social sciences
Social ties
Behavioral bias
Social serendipity
Computational social science

Figure 1. Overview of relevant scientific domains, concepts, and research areas to develop next-generation computer-supported Professional Social Matching.

---

[1] https://tinder.com/
[2] https://www.happn.com/en/

Our synthesis is based on two central yet paradoxical ideas. One the one hand, computational support for choosing collaborators seems beneficial because purely human-driven matching is prone to limitations and biases in human decision-making and the boundedly rational understanding of the breadth of alternatives. On the other hand, algorithmic matching involves risks of strengthening the human biases, for example, through biased training data [29][39], providing socially unacceptable or seemingly illogical recommendations or introducing unanticipated detrimental mechanisms to professional collaboration and social structures.

To address this problem of role differentiation between human and computational reasoning in PSM decisions, this paper builds on recent abundant discussion on the ethics of information technology (e.g., [29][33]). We review and problematize some premises in current paradigms of system design that might also seem intuitively appropriate when developing PSM systems. While the academic community has already acknowledged that simplistic aims and metrics, such as prediction accuracy in recommender systems, might misguide a research field at large [26][37], we follow a similar critical line of thinking on a broader scale. We underline the importance of area-specific application and systemic consideration when defining design goals and computational approaches.

Motivating the Computational Approach

Why is computational support necessary in professional social matching? Traditionally, the identification and choice of professional partners take place manually by individuals (e.g., matching an employer with a suitable employee, or choosing with which peers to collaborate). As long known in the science of cognition and reasoning, much of human decision-making is limited by the capacity of information processing (i.e., bounded rationality [34]), is inclined to be based on intuition, heuristics, and cognitive shortcuts [19], and strives for minimizing cognitive effort.

In choosing collaborators, human decision-making can result in tendencies like homophily, the preference of interacting with like-minded others [21], and leaning on existing social networks and a geographically limited pool of candidates. In social networking, people tend to strengthen existing social clusters (i.e., triadic closure [14]) rather than reach for unknown communities and individuals. Similarly, formation of working groups within organizations is often based on arbitrary choices even though the combination of people can significantly influence group productivity and satisfaction [35].

In other words, human biases can result in suboptimal collaboration and untapped co-creative potential, particularly in knowledge work and creative industries that demand cross-pollination of ideas and perspectives. It has been argued that fruitful collaboration and high innovation capability result from *complementary* viewpoints among a diverse group of actors [27]. Substantial research in management science and information systems hint that, in particular, activities that require divergent thinking in groups benefit from diversity [2] (e.g., startups pursuing innovations or corporation boards aiming at holistically optimal strategic decisions). Heterogeneity in terms of knowledge has also been found important to managerial performance and in particular to innovation performance [32].

While the jury is still out on the optimal balance of collaborators on the similarity–diversity continuum (e.g., [4]), we argue that current computational solutions strengthen similarity-seeking behavior. Furthermore, it is noteworthy that diversity is not only about differences in identity-related qualities like

gender, ethnicity, and personality types; it is also about *cognitive diversity* in terms of skills, competences, knowledge, and interests, and it is about *social diversity* in terms of social behavior and networks.

The need to identify optimal collaborators and fruitful skill combinations will increase in the future as co-creation chains tend to become increasingly complex and dynamism in work life is expected to increase. Examples of relevant trends and phenomena include increasing emphasis on ad hoc freelancer groups in creative industry, micro entrepreneurship, and piecework [3] as well as strongly interdependent actors in business and innovation ecosystems. The more interdependency and collaborative value creation there is on a systemic scale, the more important and impactful the decision-making on social matching becomes.

Opportunities for Computational Professional Social Matching

The question of how computational systems could support professional social matching in particular remains relatively unexplored when considering the breadth of collaborative professional activities. Prior research and development has addressed expert search (e.g., [40]) recommendations for inspiring individuals or organizations to follow in social media, and machine learning solutions for identifying suitable candidates in recruitment and headhunting (e.g., [11]). Recommending new collaborators has been explored in a few algorithmic experiments and user studies (e.g., [37]). However, this vein of research remains on the fringes of recommender systems research and is limited to the context of conferences (e.g., [8]) or analyzing publication data in academia (e.g., [30]). As for supporting team work, computational solutions have been explored in, for example, optimizing the organizing and managerial practices of a team [42]. However, the question of the ideal team composition remains underexplored.

At the same time, the new technical enablers and current trends in work-life allow for novel PSM solutions. The amount of data on people and organizations is increasing, which has paved the way for the rapid evolution of machine learning-based techniques. The aforementioned trends in leadership and organizing knowledge work welcome new forms of collaboration, and networking in general is highly valued by both organizations and individuals.

Consequently, we have recently witnessed the birth of commercial applications for browsing candidates for professional interaction (e.g., Brella[3], Grip[4], and Shapr[5]). Such applications tend to follow the so-called Tinder logic of user-based selection of seemingly interesting candidates based on simple profiles, often restricted to matching attendees at professional events. However, the simplistic and similarity-seeking nature of such applications calls for ethical consideration and expansion of the horizon in this area. We envision that future PSM systems could offer solutions on different levels of matching:

- Identifying optimal combinations of human characteristics and professional aims in certain professional activities (i.e., matching qualities and goals).
- Recommending partners for particular co-creative purposes, such as for business partnerships or mentoring relationships (i.e., matching individuals).

---

[3] www.brella.io
[4] https://grip.events/
[5] www.shapr.co

- Optimizing team formations for a project (i.e., matching multiple actors).
- Identifying suitably complementary actors for networked value creation (i.e., matching at ecosystem level).
- Balancing the supply and demand in the job market by suggesting dedicated trainings or new job openings (i.e., matching on societal level).

To provide an early framework of PSM activities, we identify three main tracks of social matching in professional life: one-to-one, one-to-many, and many-to-many (Figure 2). These categories display different levels of decision-making complexity and varying numbers of qualities to analyze. Specific matching cases further vary in terms of (a) *the temporal nature of the decision*: for example, long-term commitment in recruitment of a new employee vs. short-term perspective in forming a working group within an organization; (b) *intensity of collaboration*: for example, choosing a new group member for a co-creative project vs. choosing a member for a young company's advisory board; and (c) *probability and criticality of risk*: for example, low-probability but high-cost risk of making an unsuccessful recruitment vs. high-probability but low-cost risk of extending one's personal network with new individuals. The bottom part of Figure 2 relates particularly to this latter aspect of the cost of suboptimal matching.
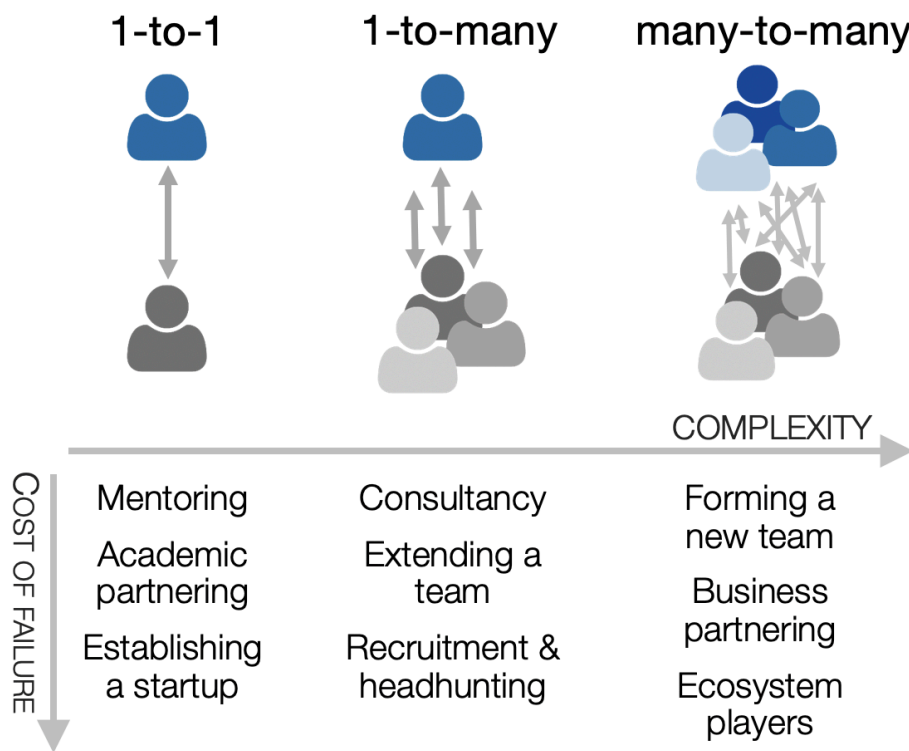

Figure 2. Main tracks of professional social matching, with examples of matching cases with different scales of cost upon a suboptimal matching decision. The complexity of matching decision-making increases as the number of actors increase.

These qualities set boundaries and requirements for information systems design, for example, in terms of the breadth of given alternatives (variety of matches), the need for transparency and explanatory capability of the logic behind recommendations, and the agency and role that a system has in the decision-making. In Figure 2, the recommendations related to increased complexity and cost of failure imply

higher risks and require increased explanatory power of recommendations from PSM systems. Deciding whom to meet at an event has low risks as the interaction would be short term and of low intensity; therefore, a user might be satisfied with only a few algorithmic recommendations and superficial reasoning behind the recommendations. However, matching for which actor to choose for close business collaboration poses higher risks and costs of failure. For the user to trust and follow the recommendations, this requires both more in-depth algorithmic reasoning and better explanation capabilities in the user interface.

Pitfalls in Computational PSM

A key argument of this paper is that the fundaments of state-of-the-art technologies and approaches that could be utilized to build PSM systems—for example, item recommenders, social network analysis, and machine learning—must be reconsidered in this application area. Directly applying the prevailing analysis or design patterns to PSM can introduce new risks with detrimental effects on the performance and collaboration practices of knowledge workers. The following critically reviews the suitability of common computational approaches.

First, it is imperative to realize that recommendation does not equal prediction [26], particularly in PSM. To truly enhance professional collaboration, recommender algorithms should not reproduce or strengthen the typical human bias. For example, in a recruitment system, using prior examples as training data for machine learning is expected to strengthen the demographic distribution that a certain organization or profession has traditionally had. As McNee et al. [26] pointed out, an accurate recommender engine might produce recommendations that are formally relevant as predictions, yet not very useful as recommendations (e.g., a well-connected person is recommended to most users).

Wallach gave a gender-related example of the same issue: "there is a substantial difference between a model that is 95% accurate because of noise and one that is 95% accurate because it performs perfectly for white men, but achieves only 50% accuracy when making predictions about women and minorities" [39]. Similarly, it would be straightforward to predict who someone might meet at an event based on their history of professional social encounters in similar situations. However, using that as a recommendation would strengthen their habitual behavior, which might be against their actual collaboration needs.

Second, social matching systems tend to look for maximal similarity; in fact, dating scenarios have particularly been found to benefit from emotional similarity [13]. A perfect match in dating services refers to the similarity of profiles, typically based on analyzing a simple user-defined, property-based profile content and clustering the pool of actors. This logic also seems to have affected the current vocational matching services: shared qualities tend to be highlighted in the user interface, and the brevity of profiles can lead the user to follow the natural tendency of seeking for similarity.

Third, from the perspective of social ties and networks, contemporary approaches for analyzing connections between individuals often utilize social network analysis and link prediction [24]. Following the triadic closure hypothesis, new ties are more likely to be formed between friends-of-friends or colleagues-of-colleagues [10], that is, between actors that share a strong connection. The triadic closure mechanism can, however, enforce echo chambers and increase polarization—the pitfalls of social networking services in the 2010s. In knowledge work, we argue that the narrowed thinking due to echo

chambers is bound to cause reduction of exposure to novel information and decrease divergent thinking and innovation capability.

Fourth, PSM system designers need to consider that good matches cannot be generalized across individuals. In content or item recommender systems, the same news article or product is often recommended to several users with corresponding consumption behavior and ratings—i.e., *collaborative filtering* [20]. However, in PSM, person A being a good partner for person B does not imply that A would also be a good match for person C, even if B and C had similar qualities. An optimal match in professional life is very case specific and determined by, among other things, the matched actors' current needs, interests, personality, and availability for collaboration. Matches can be generalized only across similar cases (e.g., matching for a similar professional activity). This means that narrowing down only to similar collaboration cases can lead to data sparsity and that the collaborative filtering approach would suffer from cold-start issues.

Finally, as professional activities are related to value creation for organizations—and, more broadly, communities and societies—PSM calls for a systemic perspective. For example, the same central and active individuals cannot practically be recommended to everybody (i.e., the Matthew Effect). The mechanism of preferential attachment [5] tends to lead to power-law distribution across the population. A system might recommend excessive collaboration opportunities for people that already have plenty of connections while undervaluing other criteria, such as urgency of the need for collaboration or the actors' practical capacities for exploring new collaboration opportunities.

New Design Directions

The discussed limitations in human decision-making and the pitfalls in the traditional computational approaches essentially imply that, in general, PSM is far from trivial bulk predictions based on masses. Neither the matching mechanisms in dating applications nor the analytics methods in contemporary content recommender systems and link prediction algorithms seem to match with the characteristics of PSM. To turn the focus from critique towards constructive thinking, we next propose general goals for the future generations of PSM systems. These are intended to advance this young research area towards directions that allow socially meaningful and ethically sustainable solutions.

Goal: Balance diversity and similarity

Comparing to leisurely social matching applications or content recommenders, we argue that PSM systems should employ a more diversity-enhancing approach that shifts the current aim for *convergence* towards *divergence*. While diversification has been identified as a relevant aim for recommender systems [23], matching people for professional collaboration makes a particularly strong case for this.

At the same time, neither optimizing for similarity nor diversity should be taken as a maxim: the optimum stands somewhere between these extremes. Successful collaboration requires mutual trust and shared working culture (i.e., similarity as a consolidating, convergent power) as well as openness for change and different perspectives (i.e., diversity as a divergent power). Similarity is useful when the aim is to validate ideas or methods and there is a need for agility in short-term collaboration. Strong common denominators can also open minds to appreciate individual differences. For example, diversity is beneficial in

developing novel ideas with the help of complementary perspectives, establishing long-term business ventures utilizing complementary social capital, or making well-informed strategic decisions based on diverse knowledge.

Identifying the optimal balance requires analysis and modeling of not only the pool of actors but also the characteristics of the intended collaboration and the collaboration context. Compared to leisurely activities, professional life is more dynamic in terms of interaction needs, interests, and resources at different times.

Goal: Enable experiences of social serendipity

We call for systems that help people make professional matching decisions with positive long-term benefit. We argue that the experience of *serendipity* is an indicator of successful knowledge work, making it a desirable goal for designing PSM systems. *Social serendipity* can be considered as a strong experience of both unexpectedness and instrumental benefit from social encounters and collaboration [25].

The unexpectedness can arise from recommendations outside one's conventional social circles and from others that have seemingly different qualities than oneself. However, turning such *chance* into serendipity calls for features that help the user to identify the possible benefits and gain advantage of the offered chance. The element of benefit can result from collaboration that provides value beyond what one could personally attain, for example, because of having complementary skills or knowledge or having formed an unparalleled team. Similar thinking can be identified in content recommenders where it is necessary to establish a certain level of familiarity for the user and, at the same time, support the discovery of interesting new content [22]. However, as social serendipity involves several actors, it is more challenging to design for it than for *information serendipity,* the more common aim of content recommender systems.

Goal: Support a systemic perspective in defining ideal matches

We argue that a matching algorithm should not only optimize matches for an individual user but also consider what is ideal on a systemic level, matching needs to be considered across the user population and social structures like organizations and teams. PSM systems introduce an opportunity to intervene the evolution of social networks to optimize for *systemic diversity*. While social link prediction is based on relationships on the local scale, the aim to enhance professional collaboration should be considered on the ecosystemic—even global—scale.

For example, in individual fields like research, a system should avoid reinforcing the processes of increasing homogeneity [21] and preferential attachment [5] by reproducing existing social network evolution mechanisms and recommending the same central actors to everyone. Instead, the recommendations need to be considered across the population (e.g., within a profession-based community). Organizations need to balance the workload across employees: the most likable and versatile individuals cannot practically contribute to all the groups or organizational actions.

Furthermore, the actors' interdependency demands bidirectional optimization: while person A could seem like a beneficial match to person B, person B might not perceive person A relevant enough. Matches that

are optimized only for an individual can create collaboration proposals with asymmetric benefit and thus remain inefficient or be rejected.

Goal: Help utilizing the existing social structures

We urge PSM developers to take advantage of existing social structures in organizations. First, existing weak or dormant ties (e.g., people who already know each other but lack understanding of all the mutual collaboration opportunities) can be re-introduced through recommendations. Second, from an organizational perspective, diversity-enhancing matching should enable making use of existing echo chambers by identifying potential pairs of actors that are able to connect and bridge the flow of knowledge between established clusters.

This objective differs fundamentally from the practices of straightforwardly predicting the likelihood of the formation of new social connections (cf. [24]). We point to Burt [7] who discussed social capital from two viewpoints—whether scarce or dense social networks produce social capital. Weak ties that gap structural holes can increase creativity and support the career development of those that occupy bridging roles. At the system level, these brokers serve as conduits of information flow. A related matching strategy that contributes to system-level diversity is identifying *tertius jungens* (or "third who joins"), that is, individuals that can serve as proxies in introducing "disconnected individuals or facilitating new coordination between connected individuals" [28]. Utilizing social proxies can be expected to facilitate building trust between the apparently diverse actors.
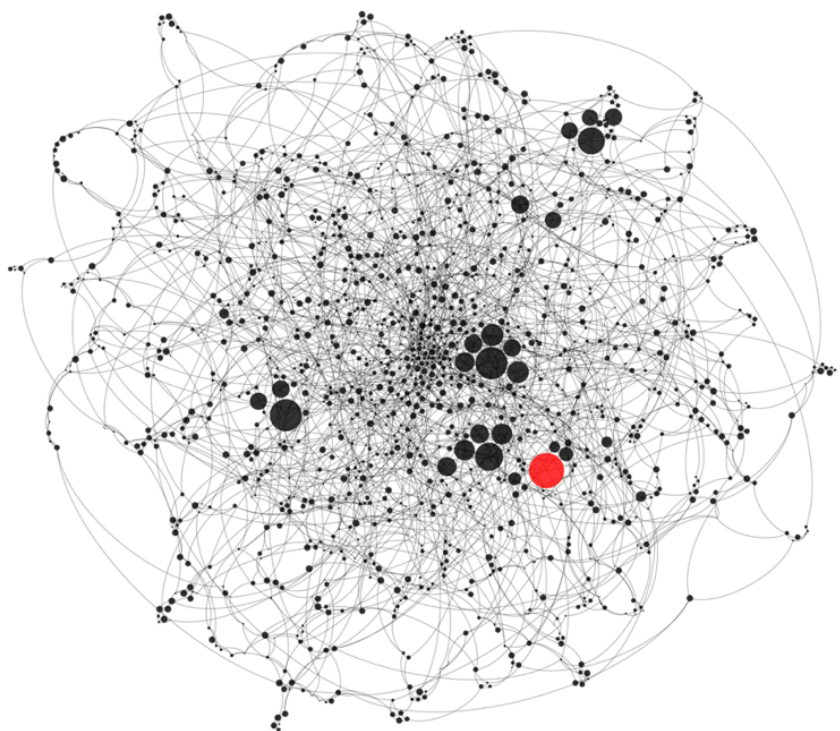

Figure 3. Illustrative example of matchmaking potential within a community. Nodes represent actors that are interconnected through collaboration, and node size represents similarity between actors' interests. Several existing clusters of actors similar to the active user (in red) are identifiable either computationally or manually through an interactive visualization.

PSM System Qualities and Future Directions

To address the aforementioned goals, we introduce a preliminary model of the analytical elements that we consider crucial for future PSM systems. Figure 4 outlines a high-level system architecture with key modules, inspired by [36], and their main functions. These are related to an analytics pipeline as well as general system qualities that need to be addressed across the presented modules. With this overview, we set visions for future PSM systems and underline research directions and requirements for future work.
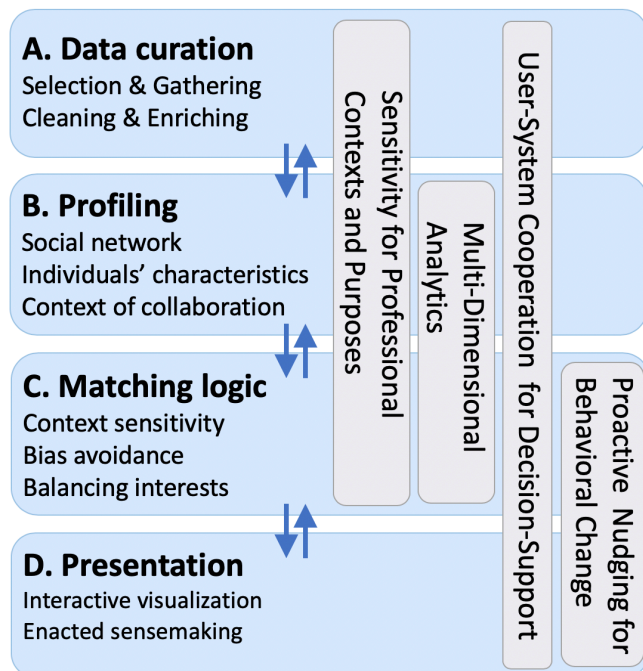

Figure 4. Key Modules and functions in a high-level system architecture and the System Qualities mapped to the different modules along the analytics pipeline. (A) Data curation refers to functions of gathering, cleaning, and managing data and enriching these with relevant semantic labels. (B) Profiling covers the functions for deriving insight from the data in order to truthfully present and compare different actors and matching cases. (C) Matching logic is about the algorithmic procedures for defining and predicting which social combinations would be most appropriate in a given matching case. (D) Presentation refers to the user interface and human-computer interaction solutions that support sensemaking, comparison, and well-informed decision-making.

Sensitivity for Professional Contexts and Purposes

Perhaps the most analytically challenging requirement is that systems should model not only the potential actors and their social structure but also the context of the intended professional relationship. The produced models should drive tailoring of the matching logic accordingly (e.g., changing priorities and weights of variables). The notions of context and context-awareness have been extensively discussed across the sciences, and various contextual variables have been investigated in relation to information systems in general and recommender systems in particular (e.g., location, culture, user personality, task context) [1]. In what follows, we focus on a few aspects that are particularly relevant when considering the matching logic.

The previously mentioned quest for a sweet spot between maximal diversity and similarity is already challenging per se. Furthermore, it is a moving, context-dependent target. The optimal degree of diversity depends on individuals and the targeted collaboration in question. For example, the end user's personality has been found to affect the readiness for accepting diversity of recommendations in content and item recommenders [41], and we expect that such inherent human factors also have an equally significant role in people recommenders. As for organizations, characteristics like openness, tolerance of difference, and the different purposes of collaboration presumably affect the sweet spot on the continuum.

Similarly, an understanding of the current social structures is needed to identify meaningful configurations of the matching logic. Our premise is that a typical structure is composed of clusters that are densely interconnected but loosely connected to other clusters. Consequently, we suggest that PSM developers draw from the diversity-bandwidth trade-off theory [4] to navigate the design space. Ties of different strengths (weak–strong) all have potential as conduits of novel information but in different ways in different domains of work. For example, the highly dynamic environment of a startup company highlights the importance of strong ties as high-bandwidth sources of relevant information. At the other end of the spectrum, an academic research group produces new knowledge over years rather than days. Therefore, the actors are able to invest in utilizing low-bandwidth weak ties due to their potential in serving non-redundant, novel knowledge.

In sum, many of the above-mentioned aspects represent largely unobservable contextual factors that are subjectively defined. This forms a cyclical relationship between context and collaboration where the collaboration activity gives rise to context and the context influences activity [1]. Moreover, the context forms as a mixture of several individuals' and organizations' contextual characteristics. Formalizing this contextual variance requires new methods and frameworks for conceptualizing PSM in more detail. This also means that the plethora of data needed to reach certain analytical targets is unpractical without new mechanisms of granting access to personal data. PSM development endeavors need to carefully consider which practically available data can best support the context-specific analytical goals and how to compensate missing data.

We suggest that there are two alternative design strategies to enable context sensitivity in PSM systems. *Reactive* design leaves the active user in charge of driving the system in a way that contextual requirements are met. *Proactive* design follows the logic of prescriptive analytics where the system suggests actions or even takes them on behalf of the active user. However, the proactive design introduces higher requirements for data availability and quality. This leads to another key quality, appropriate user-system cooperation, which we will discuss after reviewing the different analytical levels of PSM.

Multi-Dimensional Analytics

The above-mentioned contextuality and the goal of supporting systemic perspective call for multi-dimensional analysis with several parallel relevance algorithms and various data about the actors. Extant research on recommender systems shows that the perceived relevance of recommendations and the perceived usefulness of the recommender system increase when multiple different recommendation strategies are in parallel play. For example, Hupa et al. [18] discuss the advantages of multidimensional social-network recommenders to increase interdisciplinary collaboration. Tsai and Brusilovsky [37] used four different recommendation engines to support identifying new academic collaborators using

conference data (topic similarity, social similarity, interest similarity, and geographical distance) and showed that the users who are able and willing to use multiple engines in parallel receive more relevant recommendations.

Although the triangulation of these approaches can already help identification of relevant matches, we call for consideration of additional perspectives in the prescriptive analytics for people recommendations. For example, geographical distance needs to be considered, especially in long-term collaboration. Level of seniority (or expertise) affects the symmetry of benefit and trust: matching for mentoring or advisory relationships calls for high difference in expertise, while matching for a production team often demands relatively equal levels of overall seniority. The personalities and organization cultures need to be similar enough to enable matches that are sustainable in the long term (e.g., no conflicts due to drastically different ways of working or level of commitment).

Overall, developing analytics procedures and recommender engines for different perspectives (or *dimensions*) is a step towards so-called *hybrid recommender systems* [37]. At the same time, this introduces practical challenges. In addition to the axiomatic data availability issue, defining the logic in which the different analytical functions are combined in a context-sensitive manner requires deliberate research. Because PSM can potentially be affected by such a broad range of human features, all of them cannot practically be embedded in the algorithm design. We can only call for multidisciplinary research collaboration where social scientific understanding would support the identification of top-priority factors and formalizing this vast space.

User-System Cooperation for Decision-Support

Due to the dynamic and inherently complex nature of PSM, the matching decisions cannot merely be automated or offloaded to an algorithm. For example, machine learning generally produces relevant results only if initialized with good-quality training data and well-defined goals whereas human reasoning is suited for multi-faceted challenges where the desired pattern is unknown a priori. The different limitations and strengths in the human and computational analytical capabilities call for effective collaboration between computational intelligence (*deep yet narrow*) and human intelligence (*broad yet shallow*). This relates to the general notions of Augmented Intelligence and human-in-the-loop thinking. As this approach has already shown its power in, for example, classification problems, the complexity of PSM offers an even more opportune application area.

The human-in-the-loop approach has been envisioned[6] as useful, for example, when (a) the cases that need to be identified are rare (class imbalance, e.g., rare type of collaboration), (b) the cost of error is high (e.g., time spent on browsing irrelevant matching options), (c) human annotations are already used (e.g., recruiting processes), and (d) generic pre-trained models exist but need to be customized. The holistic thinking and contextual adaptability of the user are needed, for example, to steer the deep yet narrow algorithms (e.g., refinement of what an ideal match is, or prioritizing the sought features for each matching case) and to make sense of and choose between the resulting recommendations. Particularly when considering non-experts as users of recommender systems, we need user interface solutions that

---

[6] https://blogs.technet.microsoft.com/machinelearning/2016/10/17/the-power-of-human-in-the-loop-combine-human-intelligence-with-machine-learning/

support decision-making with alternative options, a multi-dimensional systemic viewpoint, and ways to communicate and deal with the algorithmic uncertainty that this application area entails.

Figure 5 outlines the potential collaboration points along the computational analytics process. First, we need methods that guide the user in selecting appropriate training data and analytical goals to enable accurate profiling and, eventually, predictions. Semi-supervised approaches could allow training with a very small amount of labeled training data, minimizing both the *cold start problem* and need for manual work by the user. Second, we need feedback from the user regarding which factors are of top priority in the current matching case. Third, we need to support user exploration for enacted sensemaking, which has been found to be important both in visual analytics in general [6] and visual network analytics in particular [17].

User interactions and inputs

Defining the matching case
to initialize the process:
- Parameter tuning
- Adding critical missing data

Data enrichment:
- Labeling clusters
- Refining the user model

Interactive sensemaking & decision-making:
- Tagging & bookmarking the results
- Filtering & prioritizing
- Recalibrating visualizations

| Goals & data | Labels & refinement | | Interactions & reasoning |

| **Data curation** | **Profiling** | **Parallel methods for matching logic** | **Presentation** |

Unclear goals
Conflicting ideals
Cold start problem

Choice of relevant factors

Opacity and trust issues

Communicating uncertainty
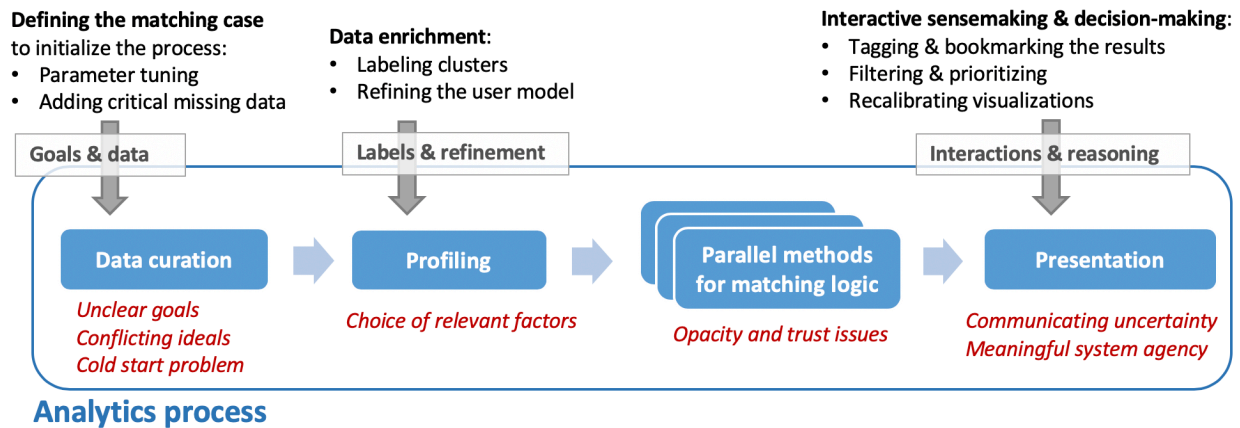*Meaningful system agency*

Analytics process

Figure 5. Outlining the user interactions in the human-in-the-loop analytics process. Typical research and design challenges in red italics.

Due to the relatively high risks and decision making complexity involved in PSM recommendations, as discussed earlier, future PSMs should be able to better explain the reasoning behind a recommendation. For example, the uncertainty of the prediction should be translated into human-comprehensible forms so that the user can understand and trust them, is able to assess the factual relevance of the recommendation, and can adjust preferences accordingly. AI-based human-like reasoning and the related capabilities for a human-like explanation of recommendations should be developed in PSM, just like the recent focus in the development of AI-based knowledge-based systems and expert systems that deal with complex human expert decisions, rather analogously to PSM.

Throughout the process, the cost in terms of burdening the user must be in line with the benefits of using the system. This raises the question of what user input and feedback is sufficient to hone the algorithms while causing minimal burden with tedious tasks for the user. In other words, it introduces a trade-off between the certainty of the matching decision and invested user effort.

An example of interactive visualization that can support exploration of relevant other people is the Conference Navigator [38]. The recommended conference talks are presented as sets, each of which are

identified by a recommendation engine. The active user is able to explore different recommendation strategies and their combinations through an interactive interface. This kind of hybrid recommender system with several visualization approaches supports user-driven exploration.

In the long term, following Leifer's hu-mimesis epiphany,[7] we envision future PSM systems taking the form of collaborative robots, that is, *cobots* or Socially Interactive Agents working alongside the user. The user could delegate some of the decision making to a personal agent that, over time, builds an increasingly accurate model of the user and their preferences and typical cognitive biases. This could allow maximally automated tasks (e.g., continuous pre-selection of relevant candidates amongst the whole population and recommending them at an opportune moment), implementing the idea of prescriptive analytics [9]. Such automatic modeling of other people and their suitability naturally poses challenges considering not only data availability but also data protection regulations and other ethical issues, which can become greater hindrances than the engineering of such mechanisms.

Proactive Nudging for Behavioral Change

While the previous section revolved around how the user steers the system, user-system cooperation also touches the system's influence on user behavior. To address the goals regarding serendipity, optimal diversity, and systemic perspective, we expect that typical end users need to be nudged away from their habitual and individually oriented preferences in networking. The notion of *diversity-sensitive design* can be considered to mitigate trust issues in exposing the user to diversity [16]. To this end, the concepts of persuasive computing [12] and gamification have been long studied in human-computer interaction as mechanisms for supporting behavioral change. While the general approaches have been successfully applied in application areas like exercising and education, it remains an open question how user interfaces could meaningfully support and affect the rather intimate decisions around professional collaboration.

Matching people and interests is already enabled by various digital platforms and social media. However, to date, there are few solutions that would take a more explicit and proactive role in social matching (e.g., recommending a person to meet someone when the situation is opportune). In content recommender systems, the recommendations are typically provided while the user is actively using the service. However, PSM calls for technology that takes a more active role and has its own judgement of what are optimal matches—i.e., a moral stance—as well as when and where to present the recommendations.

Proactive systems for PSM pose new challenges for user interface design as they manifest design ethics beyond the will of a user, the traditional fundament in human-computer interaction. The computational design ethics would need to balance between an individual's preferences (i.e., their habitual behavior) and the greater good for an organization, business ecosystem, or other social entity (i.e., systemic perspective). This requires careful encoding of the ideals and mechanisms for the user to adjust them when necessary. This also requires new interface design solutions that sufficiently preserve the user's agency to maintain trust and acceptance.

The conventional subtle ways of enabling serendipity in content recommenders, such as adding randomness to the results or removing top-ranked results, might not be enough for PSM systems. Rather, this application area calls for solutions that avoid bias and prejudice. For example, relevant approaches

---

[7] Hu-mimesis: Design Requirements for Personal Relevance, https://www.youtube.com/watch?v=pg0xU-6PQII

could include hiding a recommended person's profile picture or name (as strong indicators of gender, ethnicity, and socio-economical stance) or hiding details about professional background and education. Revealing some commonalities is important for building interpersonal trust, whereas the relevance of the complementary qualities needs to be considered with respect to the collaboration need at hand. Rankings and ratings tend to encode human bias [29], which means that the presentation of the matches should focus on qualitative descriptions rather than numeric ones.

Persuasion can also manifest itself via facilitation of building the connection. For example, in an event context with low-risk matching, the matched actors can be given a playful challenge to explore the commonalities and complementariness by not revealing the exact reason behind the recommendation, such as in [8]. Here, the human capabilities in identifying possible matches could also be harnessed beyond the primary end user, for example, encouraging a third party to serve as a matchmaker between people they know, thus facilitating newly recommended connections and validating the ill-reasoned recommendations that an algorithm might provide.

Conclusion

Social matching is ubiquitous in professional life, yet suboptimally supported by computational systems. Based on a multi-disciplinary review of the literature, we contribute an overview of a complex problem space, new design targets for computer-supported PSM, and requirements and research problems for developing next-generation systems.

Our review implies that many conventional approaches in recommender systems and social link prediction, when applied to PSM, could have detrimental long-term implications for organizations' performance. Conventional mechanisms, such as optimizing for similarity and triadic closure, involve risks of strengthening the human biases of homophily and echo chambering. We call for diversity-enhancing yet contextually sensitive designs of future PSM systems, tapping on multi-dimensional analytics of not only the potential matched actors but also the intended type of collaboration and organizational context. Further, we call for a paradigmatic shift from automated recommender systems towards decision-support systems that are based on meaningful user-system cooperation.

All in all, analysis of this application area underlines the importance and urgency of rethinking some paradigmatic algorithmic approaches. We argue that in such a complex and inter-disciplinary area as PSM, reconsidering the conventions is not only beneficial; it is a necessity.

References

1. Adomavicius, Gediminas, and Alexander Tuzhilin. "Context-aware recommender systems." Recommender systems handbook. Springer, Boston, MA, 2011. 217-253.
2. Aggarwal, Ishani and Anita Williams Woolley (2013) Do you see what I see? The effect of members' cognitive styles on team processes and errors in task execution. Organizational Behavior and Human Decision Processes, Volume 122, Issue 1, 2013, Pages 92-99, https://doi.org/10.1016/j.obhdp.2013.04.003.
3. Alkhatib, A., Bernstein, M. S., & Levi, M. (2017). Examining Crowd Work and Gig Work Through The Historical Lens of Piecework. In Proceedings of the 2017 CHI Conference on Human Factors in

Computing Systems, CHI '17 (pp. 4599–4616). New York, New York, USA: ACM Press. http://doi.org/10.1145/3025453.3025974

4. Aral, S., & Van Alstyne, M. (2011). The Diversity-Bandwidth Trade-off. American Journal of Sociology, 117(1), 90–171. http://doi.org/10.1086/661238

5. Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509–512. http://doi.org/10.1126/SCIENCE.286.5439.509

6. Bendoly, E. (2016). Fit, Bias and Enacted Sensemaking in Data visualization: Frameworks for Continuous Development in Operations and Supply Chain Management Analytics. Journal of Business Logistics, 37(1), 6–17.

7. Burt, R. S. (2004). Structural Holes and Good Ideas. American Journal of Sociology, 110(2), 349–399. http://doi.org/10.1086/421787

8. Chen, Jay and Azza Abouzied (2016). One LED Is Enough: Catalyzing Face-to-Face Interactions at Conferences with a Gentle Nudge. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, ACM Press, pp. 172–183.

9. Davenport, T.H. (2013). Analytics 3.0 Harvard Business Review, December 2013.

10. Easley, D., & Kleinberg, J. (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World. New York, NY, USA: Cambridge University Press.

11. Faliagka, Evanthia; Kostas Ramantas; Athanasios Tsakalidis; and Giannis Tzimas (2012) Application of Machine Learning Algorithms to an online Recruitment System. Proceedings of ICIW 2012, The Seventh International Conference on Internet and Web Applications and Services

12. Fogg, B. J. (2002). Persuasive technology: using computers to change what we think and do. Ubiquity 2002, December 2002.

13. Gonzaga GC, Campos B, Bradbury T. (2007) Similarity, convergence, and relationship satisfaction in dating and married couples. J Pers Soc Psychol. 2007 Jul; 93(1):34-48. https://www.ncbi.nlm.nih.gov/pubmed/17605587

14. Granovetter, Mark (1973). The Strength of Weak Ties. American Journal of Sociology, Vol. 78, Issue 6, May 1973, pp. 1360-80.

15. Guy, Ido. (2015). Social Recommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), Recommender Systems Handbook (pp. 511–543). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_15

16. Helberger, N., K. Karppinen, and L. D'Acunto (2018) Exposure diversity as a design principle for recommender systems. Information, Communication & Society, vol. 21, pp. 191–207, 2018.

17. Huhtamäki, J., Russell, M. G., Rubens, N., & Still, K. (2015). Ostinato: The exploration-automation cycle of user-centric, process-automated data-driven visual network analytics (pp. 197–222). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-18552-1_11

18. Hupa, K. Rzadca, A. Wierzbicki, and A. Datta (2010) Interdisciplinary Matchmaking: Choosing Collaborators by Skill, Acquaintance and Trust. Comput. Commun. Netw., pp. 319–347, 2010.

19. Kahneman, Daniel; and Amos Tversky (1973) On the psychology of prediction. Psychological Review, 80(4): 237-251. http://dx.doi.org/10.1037/h0034747

20. Koren Y., Bell R. (2011) Advances in Collaborative Filtering. In: Ricci F., Rokach L., Shapira B., Kantor P. (eds) Recommender Systems Handbook. Springer, Boston, MA

21. Kossinets, G., & Watts, D. J. (2009). Origins of Homophily in an Evolving Social Network. American Journal of Sociology, 115(2), 405–450. http://doi.org/10.1086/599247

22. Kotkov, Denis, Shuaiqiang Wang, and Jari Veijalainen. "A survey of serendipity in recommender systems." Knowledge-Based Systems 111 (2016): 180-192.

23. Kunaver, Matevž and Tomaž Požrl. (2017) Diversity in recommender systems – A survey. Knowledge-Based Systems, Vol. 123, 2017, pp. 154-162. https://doi.org/10.1016/j.knosys.2017.02.009.

24. Li, Z. (Lionel), Fang, X., & Sheng, O. R. L. (2017). A Survey of Link Recommendation for Social Networks. ACM Transactions on Management Information Systems, 9(1), 1–26. http://doi.org/10.1145/3131782

25. McCay-Peet, Lori and Elaine G Toms. 2010. The process of serendipity in knowledge work. In Proceedings of the third symposium on Information interaction in context. ACM, 377–382.

26. McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being accurate is not enough. In CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06 (p. 1097). New York, New York, USA: ACM Press. http://doi.org/10.1145/1125451.1125659

27. Mitchell, Rebecca and Stephen Nicholas (2006) Knowledge Creation in Groups: The Value of Cognitive Diversity, Transactive Memory, and Open-mindedness Norms. The Electronic Journal of Knowledge Management, Volume 4 Issue 1, pp 67-74, available online at www.ejkm.com

28. Obstfeld, D. (2005). Social networks, the tertius iungens orientation, and involvement in innovation. Administrative Science Quarterly, 50(1), 100–130. Retrieved from http://asq.sagepub.com/content/50/1/100.short

29. O'Neill, Cathy (2016) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown.

30. Pham, M. C., D. Kovachev, Yiwei Cao, G. M. Mbogos, and R. Klamma (2012) Enhancing Academic Event Participation with Context-aware and Social Recommendations. In 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 464–471, IEEE, 8 2012.

31. Power, D. J. (2008). Understanding Data-Driven Decision Support Systems. Information Systems Management, 25(2), 149–154. http://doi.org/10.1080/10580530801941124

32. Rodan, Simon; and Charles Galunic (2004). More than network structure: how knowledge heterogeneity influences managerial performance and innovativeness. Strategic Management journal, Vol 25, Issue 6, June 2004. Pp 541-562, John Wiley & Sons, Ltd.

33. Shilton, Katie (2018). Values and Ethics in Human-Computer Interaction. Foundations and Trends in Human- Computer Interaction, Vol. 12, No. 2, pp 107–171.

34. Simon, H.A. (1957) Models of Man. New York, Wiley & Sons.

35. Tanghe, Jacqueline, Wisse, B., & van der Flier, H. (2010). The formation of group affect and team effectiveness: The moderating role of identification. Br. J. Manag. 21, 2 (July 2010), 340–358.

36. Loren Terveen and David W. McDonald. (2005). Social matching: A framework and research agenda. ACM Trans. Comput.-Hum. Interact. 12, 3 (September 2005), 401-434. DOI: https://doi.org/10.1145/1096737.1096740

37. Tsai, C.-H., & Brusilovsky, P. (2018). Beyond the Ranked List: User-Driven Exploration and Diversification of Social Recommendation. In Proceedings of the 2018 Conference on Human Information Interact

38. Verbert, K., Parra, D., & Brusilovsky, P. (2016). Agents Vs. Users: Visual Recommendation of Research Talks with Multiple Dimension of Relevance. ACM Transactions on Interactive Intelligent Systems, 6(2), 11:1–11:42. http://doi.org/10.1145/2946794

39. Wallach, H. (2018). Computational social science ≠ computer science + social data. Communications of the ACM, 61(3), 42–44. http://doi.org/10.1145/3132698

40. Wang, G. A., Jiao, J., Abrahams, A. S., Fan, W., & Zhang, Z. (2013). ExpertRank: A topic-aware expert finding algorithm for online knowledge communities. Decision Support Systems, 54(3), 1442–1451. http://doi.org/10.1016/j.dss.2012.12.020

41. Wu, Wen, Li Chen and Yu Zhao (2018) Personalizing recommendation diversity based on user personality. User Modeling and User-Adapted Interaction, July 2018. https://doi.org/10.1007/s11257-018-9205-x

42. Sharon Zhou, Melissa Valentine, and Michael S. Bernstein. 2018. In Search of the Dream Team: Temporally Constrained Multi-Armed Bandits for Identifying Effective Team Structures. In Proceedings of