The Case of the Strangerationist: Re-interpreting Critical Technical Practice

Vera Khovanskaya
Information Science
Cornell University
vdk9@cornell.edu

Maria Bezaitis
Intel Corporation
Maria.Bezaitis@intel.com

Phoebe Sengers
Information Science and Science & Technology Studies, Cornell University
pjs54@cornell.edu

ABSTRACT
We describe a method for critically informed development of new technical systems by combining analysis of historical discourse with critical technical practice. We take the case of social recommender systems, a class of algorithms that calculate which people should be recommended to whom. We demonstrate similarities between limitations of "social network" rhetoric in contemporary social matching algorithms and discourse on planning in Artificial Intelligence. We develop an algorithm for social matching that recombines "lost" ideas from the history of AI, orienting around situated behavior and algorithmic transparency. By implementing this approach in a functioning prototype called "The Strangerationist", we examine directly how conceptual commitments inform low-level technical decisions, and how available technologies shape conceptual vision. Our goal is not to design a 'better' algorithm but to explore the challenges and opportunities of weaving together historical discourse and critical analysis of values embedded in technology with the experience of designing it.

Author Keywords
Critical technical practice, critical methods, HCI.

INTRODUCTION
In "The Relevance of Algorithms," Tarleton Gillespie describes algorithms as embedded, multi-dimensional entanglements between technical processes and the social tactics of the designers and audiences which gives rise to them. He calls for simultaneous sociological inquiry into the technical process by which algorithms evaluate and present data and the social contexts and processes which yield and legitimate algorithms [20]. Yet, as noted by Gillespie, along with many other social science scholars [31, 38, 40], algorithms continue to be fraught ground for sociological inquiry because algorithms, like many technical infrastructures are notoriously embedded, "squirreled away in semi-private settings or buried in inaccessible electronic code" [45, p 378] so much so that they are often only visible in malfunction [45] or at their seams [11]. Even if researchers can locate themselves in the presence of code and have the technical literacy necessary to analyze its contents, algorithms pose further unique difficulties. The sheer scale and technical complexity of the data manipulations of the algorithm is often impossible to recreate manually, leaving even technically trained persons unable to follow the many iterative transformations by algorithms that render inputs into "intelligent" outputs [38].

Beyond mapping and making visible the embedded representations of the social reality, the task of critical engagement with algorithms is motivated by an interest in experiencing the of decisions made on the part of technical practitioners—a mixture of "choice, necessity, pragmatism, and unquestioned 'home truth'"[40]—which in turn is shaped by the constraints of what is practically calculable and implementable. The Strangerationist project is one attempt to answer the call to social analysis of algorithms in light of these difficulties. It answers that call by engaging directly with algorithms, specifically taking the case of social matching algorithms—algorithms which recommend people to people—simultaneously on the level of social and historical discourse, and on the level of technical system design.

Our work is grounded in critically oriented and reflective design traditions in Human-Computer Interaction (HCI), and in a parallel earlier thread of inquiry conceived by Philip Agre called *critical technical practice* [1]. When he was a researcher in Artificial Intelligence (AI), Agre developed critical technical practice as a means to integrate philosophical, critical reflection and technical development. This method works by recognizing technical impasses as philosophical problems, applying philosophical and critical methodologies to find ways around these impasses, and thereby driving new technical innovation. We see system development here not as an end in itself, but as a means to reflectively explore underlying assumptions and attitudes about technology and humanity.

Broadly following Agre's development of critical technical practice, our methodology here starts with a critical analysis of contemporary technical discourse to identify recurring metaphors and assumptions on which the discourse

DIS 2016, June 04 - 08, 2016, Brisbane, QLD, Australia

ACM978-1-4503-4031-1/16/06 $15.00
DOI: http://dx.doi.org/10.1145/2901790.2901860

rests. Through this analysis, we pinpoint and explain systemic issues that follow from those assumptions and the points of connection between these limitations. To further explore the relationship between conceptual commitments and algorithms, we use inspiration from the history of technical discourses to identify, revive, and recombine "lost" ideas which embody different points of view from those currently prevalent in the design of social matching algorithms. We explore the potential impact of these conceptual commitments within algorithms by embodying those heuristics in a specific technical system. We use the practice of system building both as a means to identify and alter low-level decisions that would otherwise flow unthinkingly from the dominant technical discourse into systems design. Throughout the process, we track the ways in which technologies resist, constrain, or reroute those commitments. While the work we present here is focused specifically on social matching, more broadly we are concerned with questions about the role of critique and design in a world increasingly structured by algorithms. As we will describe next, in this paper we contribute a method for integrating historical analysis of technical discourses with critical technical practice as a means for exploring the mutual embedding of discursive concepts and algorithmic development.

THE RELEVANCE OF SOCIAL MATCHING

Recommender systems are part of a growing family of algorithms that curate one's experience of technology through the careful mining of specific types of data. Hand in hand with the proliferation of personal and social data, recommender systems are being explored through commercial systems and research prototypes as a way of profitably sifting data-overflow by suggesting which information may be of most value to you. While recommender systems generally seek to automate the everyday process of word-of-mouth sharing of opinions in various domains—e.g. music, movies, or books—social *matching* systems specifically automate the process of 'bringing people together" by recommending people to people based on their activities. Popular modern social media applications like Facebook and Google+ calculate matches between users of the system and other people that they are likely to know by constructing intricate friendship graphs, or "networks," using interaction data.

Network-based models for social matching algorithms have a clear practical utility. Nevertheless, given the reach of social matching applications and their potential to shape people's on-line experiences, it is worthwhile to ask for what reasons and with what consequences these algorithms rely so heavily on social networks. In this paper we critically examine the conceptualizations of human relations that are incorporated into these algorithms, using technical design coupled with analysis of historical discourse as a means for social study of algorithms.

We will do so by drawing on historical analysis of debates within AI. In analyzing social matching algorithms, we will identify similarities between the historical discourse in which social matching algorithms and the concept of the "social network" arose, and arguments challenging "top-down" modeling approaches in Artificial Intelligence, which similarly narrow the scope of experience to what can be effectively modeled. In particular, we will argue that networked-based matching algorithms are based on an intellectual legacy of adopting structural sociology into computer science which tends to see closeness between people as reasonably static and as quantifiable based on data traces. This legacy is tied to two well-known problems: first, that this understanding of closeness has clear limits in representing real human relationships given their dynamic and incompletely documented nature, and, second, that to the extent the algorithms effectively do capture human relationships, they tend to connect people who would already know or meet each other and otherwise create 'filter bubbles' of like-minded people [37].

EXPLORING ALGORITHMS I: HISTORICAL DISCOURSE

Most recommender systems arise from a desire to use computers to effectively manage data overload [50]. Traditionally, this problem has been tackled in computer science through AI, i.e. as a computer intelligence task: can computers be "taught" how to trace useful patterns through a milieu of data to show the optimal collection of music, books, etc. for another person. In the field of social matching in particular, where, instead of books, *people* are recommended to people, the task has similarly been approached as an intelligent modeling task. For example, in a seminal survey paper in AI, Terveen and McDonald describe this general methodology for developing social matching recommenders: "(1) modeling the set of users who can be matched, (2) matching users in response to an explicit request or implicit opportunity, (3) introducing matched users, enabling them to (4) interact with each other... note that the results of the process must feedback to the system, possibly causing it to update its models" [50, p 404]. Terveen and McDonald articulate two significantly different classes of techniques for modeling users.

The first set of approaches identified by Terveen and MacDonald calculate tie-strength (the distance between two people, represented as nodes in a graph of the social network) using data about people's' *social interactions* to compute who is likely to know each other [50]. Early forms of such social-modeling recommendations include the "friend-of-friend" algorithm -- "if many of my friends consider Alice a friend, perhaps Alice could be my friend." Canonical examples of this kind of system include ReferralWeb [25] and Expertise Recommender [32], which model professional social networks through co-authorship in hopes of recommending experts on a certain topic that the user of the system is likely to know. A more recent example is the "people you may know" feature of social networking websites like Facebook.

The second set of approaches identified by Terveen and McDonald rely on modeling *content characteristics* like users' preferences and behaviors, based on the assumption that matches can be recommended through similarity of interests or content produced. An early example of this approach was an early-2000s recommender I2I [12] which matched people who were on the same website at the same time, reasoning that "if we both post content on similar topics, we might be interested in getting to know each other". This approach to matching can be seen in several current, interest-specific social websites like Flickr (photography), Pandora (music), and dating websites such as OkCupid, which perform social matching based on user content and characteristics (e.g. [26]).

Although these approaches are distinct in their use of different kinds of data for determining a 'match' between two users – the social network ties of a user versus the content produced or consumed by the user – it is important to note that recommendation algorithms are not inherently limited to just one approach. For example, Chen et al. empirically evaluate a hybrid model that combines content matching with social network "links" [13].

At the time of Terveen and McDonald's survey, 2005, it did not seem clear that any one approach dominated the design space. But five years later, research showed that social network modeling algorithms appeared to produce better-received recommendations and found more known contacts for users [13]. In more recent work, trends show that for general-purpose social matching, social network modeling is the prevalent technique. While some domain-specific social matching recommender systems facilitate connections over content-based interest, the dominant algorithms for calculating social matches in mainstream social media rely on the construction of social networks.

The trend toward network models for social matching has a legacy in an area of sociology called structural analysis, which is commonly drawn on in this work. Historically, this field has conceptualized the social network as a graph-based representation of social interactions that can be used for various computations such as tracking the spreading of disease or the dissemination of news. In social network models, social matches are generally computed as the measurable distance between two human nodes on a graph. This conceptualization has been represented and analyzed using mathematical graph theory, notably in work by mathematician Manfred Kochen and political scientist Ithiel de Sola Poole [46].

In conjunction with advances in personal and interpersonal data gathering and reiterations of popular sociological experiments, particularly Stanley Milgram's "Six Degrees of Separation" [33] and Mark Granovetter's "Strength of Weak Ties" [21] these earlier works have motivated many computational efforts to model and quantify social networks [15]. Algorithms have been developed to calculate social closeness though social network modeling; for example, work by Gilbert and Karahalios found a series of features (e.g. number of mutual friends, number of words exchanged) that can, notionally, predict tie strength between friends in an online social network [17]. These algorithms have been applied and iterated by many researchers in social matching. Work in this field has provided an intellectual legacy that frames how researchers think about the strength of social ties and central and peripheral social roles [15,19].

Trouble on the Network

There is no denying the practical benefits that have accrued from matching people with other people based on social network modeling. At the same time, limitations of quantifying social closeness and modeling social networks have become markedly visible in recent HCI research. We here identify two key systemic issues.

The first issue is that, to the extent that social networks truly capture real-world ties, there is limited novelty to the predictions and connections that arise from network models. Empirical work suggests that algorithms based on social network information are able to produce better-received recommendations and find people that users already know, while algorithms using similarity of user-created content were stronger in discovering "new friends" [13]. The sense that social network algorithms may reproduce real-world connections rather than develop new ones has led to concerns in academic and popular-press writing about social matching leading to "echo chambers" and "filter bubbles" [37].

The second issue is that there are substantial and perhaps fundamental difficulties in capturing novel and useful data for the computation of social closeness. For example, in reflecting on close friendships as revealed in Facebook data, Sosik et al highlight that the data available to data scientists is limited because closest ties are often left out of computer mediated communication [47]. Other identified difficulties include persuading users to provide clean, semantically interoperable data [23] and recognizing the recursive malleability of social networks because the recommendation process itself can influence social ties [39**]**. The legacy of Granovetter's strength of weak ties research in the social network software space has even been directly challenged by Panovich et al.'s work which suggests that, in the context of online question and answer features embedded into social networks, users report getting the most useful answers from people that, when assessed by various accepted social closeness algorithms, including those by Karahalios, were computed as close - not weak - ties [36].

It is important to note that these impasses are neither new nor unanticipated by proponents of structural analysis. Work from the 1970's on modeling contact and influence in social networks observed the difficulty of quantifying and computing basic information about "man-to-man" contact, for example because of the near-impossibility of obtaining accurate values for many of the variables [46]. The authors present an influential model for estimating the scale and connectivity of a “human contact net” and argue that the model requires precise and necessarily reductive definitions of “knowing.” They describe significant practical caveats about friction between the model and the “lumpiness” of daily, lived experience. These concerns have been echoed in the context of social matching by Terveen and McDonald, who describe ways in which the setting of a relationship often influences how the relationship develops [50]. They argue that “whole network analyses” have consistently revealed that the social and information seeking behaviors of people evade formal structures, and that personal roles and interpersonal experiences are lurking variables in modeling networks.

These concerns suggest that underlying the practical difficulties associated with social-network-based models are deeper questions about what kind of relationship exists between the social networks discovered by algorithms based on available data and actual relationships between people in the world. One interpretation of 'found' social networks would be that they ideally correspond to the actual real-world relationships between people. Since it is easy to realize that (1) people have multi-faceted, dynamic relationships which cannot easily be reduced to simple, static links, (2) many aspects of those relationships take place outside of monitored data streams, and (3) even monitored data streams have human meanings outside of the plausible reach of contemporary algorithms, such a simplistic interpretation is unlikely to be adhered to long by even the most ardent proponent of social networks.

And yet there is a seduction to the term "network" itself that makes the equation of data structure and relationship a practical, quotidian accomplishment of algorithmic work despite its proponents’ critical reflection. In a similar vein, Agre has analyzed the extraordinary power of certain broad technical terms such as "planning" when they are used, like "networks," both to identify a family of technical specifications and as metaphorical descriptors of empirical phenomena [2]. The power of these terms derives from their under-recognized plasticity, in that they are flexibly and seamlessly deployed to refer sometimes to a specific technical representation (e.g. a graph data structure) and sometimes in a much more vague sense to real-world phenomena (e.g. a network of friends). It is precisely the ability of these terms to travel between these two kinds of meanings that provides them with the heady power to apparently calculate the world.

Within social matching, we see the consequences of this rhetorical plasticity in slippages between 'networks' understood as structures that can be calculated from conveniently captured data and 'networks' as actually existing, real-world relationships. One kind of slippage is when the networks that can be calculated are taken as necessarily revealing one’s real-world relationships. Another kind of slippage happens when network models are used to give you opportunities to connect with the match that has been calculated – in this case your actual social network is not ‘found’ but transformed, through the system, to more closely resemble the calculated structure. A third kind of slippage happens when the forms of networks that are calculated bear noticeable traces of the kind of data that happens to be available; we see this for example in the evolution of social recommender systems moving from co-authorship networks in early work to social media interactions, Facebook interactions, Twitter interactions, etc.

Certainly, all these forms of data bear relationships to the real world of human experience; equally certainly, it can be informative to consider real-world relationships in terms of social networks and to use calculation to help give a sense of the shape of those relationships. But the issues of limited novelty and of limited data caution against an unintentional equation between a model of reality and reality itself. They suggest an untapped potential for finding ways to suggest social matches that do not depend conceptually on an equation between real-world relationships and what computers can sense and model. Is it possible for algorithms to embody such a worldview? It is this opportunity we now turn to explore in depth.

EXPLORING ALGORITHMS II: ROADS UNTRAVELLED

Our analysis above highlights possibilities, constraints, and limits of the focus on social networks as a technically instantiated metaphor of human relationships. Here, we take another tack to understand social recommendation algorithms, by identifying other conceptual commitments that could also be embodied in algorithms for social matching. Our goal in this section is to demonstrate concretely that an equation between computational representations and human relationships or practices is not a requirement for a social matching algorithm by laying out how alternative conceptual commitments could be embodied in design. In the next section, we will explore how these conceptual commitments come under stress, shift and are further shaped in technical instantiation.

The first source for rethinking social matching comes from following different paths through the history of sociology than those normally drawn on in social-network-based matching, specifically ideas from sociology about how to frame how humans relate to each other that arose before present conceptions of networks had been cemented. The second source is from the history of technical discourses around modeling prevalent in AI in the 1980's and early 1990's, which questioned and developed technical and design alternatives to the idea that computing should be based on complex models that directly mirror real-world phenomena. Our goal here is to explore the social commitments embedded in algorithms for social matching by examining and recombining paths not taken by contemporary social matching.

What is a 'match'?

Previously, we discussed how earlier researchers in sociology had articulated difficulties arising in the modeling of human relationships through networks that later surfaced in HCI as concerns about network-based social matching models. One way, then, to potentially address these difficulties is to return to prior moments in sociology where alternative understandings of network models were formulated. Precisely because these ideas may not have been picked up on at the time, they may now provide us with a different lens for conceiving of what social matching algorithms could or should be.

One place to start is with Granovetter's discussion of the limitations of weak ties [21]. In the concluding sections of "The Strength of Weak Ties", Granovetter expresses concerns about reducing relationships to strong and weak ties. He encourages future social modeling practitioners to elaborate the social network model, specifically since tie-strength alone is a "very limited model for social networks" [p. 1380]. One direction Granovetter suggests for future work is to consider tie strength as a "continuous" [p. 1380] variable. He also points to the value of considering the developmental, rather than static, structure of networks.

Since Granovetter, this line of thinking has been tackled from a modeling perspective. This has led to "context sensitive" frameworks for social modeling that attempt to use probabilistic modeling to capture context sensitivities in the network [e.g. 12,50]. Yet we note, in the context of the concerns previously mentioned around data availability, that these approaches also largely depend on the assumption that real-world relationships can be effectively captured in the model, given sufficient data.

However, the distance between people does not necessarily need to be approached through a technical program that posits it as having a definite value, whether absolute or probabilistic. In this line of thinking we are inspired by early 20th century sociologist Georg Simmel's work, which tackles the distance between people by challenging classic distinctions between "the stranger" and the "native." In doing so, Simmel points to how even the most intimate relationships have degrees of distance and alienation. Importantly, he points to a degree of fundamental uncertainty: "we do not know how to designate the peculiar unity of this position other than by saying that it is composed of certain measures of nearness and distance" [44, pp 3].

Inspired by this conceptualization of an omnipresent, shifting closeness and distance to social matching, we suggest an alternative to trying to calculate (and/or iteratively re-calculate) the distances between people and interpreting these as "strong" or "weak" ties. Another, perhaps more effective way to think about tie strength is as a "continuous" variable in Granovetter's formulation, in terms of a closeness and distance, where at any given point in time, with any two people, a relationship is a to some degree unknowable and constantly shifting combination of the two. Under this conception of tie strength, it becomes feasible to drop some of the assumptions about the calculability of the network, and pursue a different model for social matching. For inspiration about how to do so, we turn to the history of Artificial Intelligence (AI).

Action beyond calculability

It may sound paradoxical to argue for building algorithms for social matching while seeing actual social matches as only partially calculable. We draw inspiration for how to do so from debates around the relationship between computational representations and the activity of computational systems that took place in AI in the 1980's and have more recently been revisited in HCI [29,30]. AI at the time was in crisis, as the ambitious attempts to build "artificial minds" in the '70's through the manipulation of complex symbolic representations appeared to be at a standstill. One strand of research termed "situated action" began to explore the idea that autonomous agents might be able to react to what is happening around them without necessarily directly representing it. As Brook's influential slogan, "The world is its own best representation" [9], suggests, the idea was to tie agent behavior directly to what could be sensed about the world. Rather than focusing effort on building complex models of the world which quickly became cumbersome and out-of-date, these researchers argued that directly tying sensing to action would allow agents to exhibit reactive, real-time, apparently intelligent behavior without necessarily "knowing" much in a representational sense. This approach led to substantial successes in robotics and AI. It also led to substantial controversy, pitting its ability to drive impressively complex, reactive behavior with challenges in framing such behavior as 'intelligent' and in generalizing the model beyond bodily action [9].

For our purposes here, the framing of technical systems as agents who do not know the world but who can react to its visible features in ways that can be narrated as intelligent suggests an alternative approach to social matching. What if we think about social matching systems as 'agents' reacting to a world of user behavior? In this sense, existing network-based matching can be thought of in the vein of 1980's AI as creating symbolic systems that construct complex representations of that world, with good matching achieved as a consequence of a correct representation. But we could also consider social matching as a form of situated action, emphasizing matching as continuously reacting to and embedded in human activity. This reframing removes our reliance on the model of the network and focus instead on optimizing the moment-to-moment interaction between the matching algorithm and what users do. In the next section, we describe potential design strategies for social matching based on this alternative model.

Embodying Alternatives in Algorithmic Design

We argued previously that the primary models for social matching based on networks rest on an implicit equation of computational representations and human activities. Here, we develop two design heuristics for social matching that are based on an alternative viewpoint. These heuristics embody the continuous, dynamic, and partially incalculable vision of social ties arising from our analysis of sociology. They do so by leveraging ideas from AI about how to design systems that can respond dynamically in real-time to human action without depending on long-term modeling. In particular, we develop heuristics based on two different, related strands of work termed "situated action" that drew on the previously discussed debates about AI in the 80's. Our first heuristic, *reactive social matching*, draws from work by technical researchers to improve the ability of agents to couple with external environments. Our second heuristic, *interpretable social matching*, draws from parallel work by social scientists who emphasized the role of human interpretation in human-agent interaction.

Reactive Social Matching

Technical researchers in situated action emphasized the use of simple rules to allow 'agents' or programs to immediately react to ongoing conditions in the local environments. A canonical example is Agre's Pengi, a quick-moving game in which agents are able to interact with their immediate context in an efficient and "intelligent" way by following situation-action rules, as opposed to calculating and maintaining models [3]. This is based on the idea that relevant 'knowledge' is stored not as representations in the mind but as a bodily readiness to respond to the solicitations of situations in the world. Similarly, we could think of situated recommender 'agents' as dynamically responding to behaviors in a world that is constantly in flux. Data from the host user could trigger immediate reactions instead of accreting in a model of the user's role in the broader social network. In other words, such an approach reframes matching from being about connecting me to others who are inherently like me on a long-term scale which a machine can eventually correctly model, to thinking of matching as being about connecting users who are behaving similarly in the current situation.

One implication of this refocusing is a potential return to a kind of content-based social matching, since information about data I am currently focused on could be a major component of what agents react to. But this design commitment has deeper implications for the processes used to make recommendations, as well as the ways we frame what these processes are doing. Instead of using data about users to represent their relationships and then use those models to make match recommendations, reactive social matching would use simple, "on the ground" techniques to present recommendations in response to user behaviors. Some early existing social matching recommenders already did this, such as I2I, which matches people who are on the same website; I2I reacts to users' immediate action (browsing the same webpage) simply by inviting the users to a shared chat room without pre-supposing any relationships [12].

It is important to note here that recommenders that use data associated with social networks can also follow this heuristic. For example, if Alice "friends" Bob, and is also "friends" with Jane, the system could respond to Alice's action by recommending Jane to also "friend" Bob. This is another framing of the "friend-of-friend" algorithms. The difference is that, conceptually, the system does not represent the Bob/Jane friendship, or claim to have any intelligence on Bob's relationship to Jane. As we will see later, such conceptual differences eventually can end up having significant implications for the resulting technical implementation.

Social matching without social modeling may have pragmatic benefits: we have already discussed how available data necessarily represents only a limited part of human dynamics, and how those dynamics themselves are constantly in flux. From a technical design point of view, relaxing the commitment to modeling in a highly nuanced problem with sparse data is not unreasonable and allows social matching researchers to focus less on the intelligence of the system, and more on user interaction with and interpretation of the system. This shows that technical designs are not limited to modeling only for highly nuanced problems, and that the 'problem' of sparse data only exists as such within a worldview that assumes that algorithm construction is always about ideal data fit, about creating the perfect alignment between representation and world. What this re-imagining suggests, is that it is possible to relax this commitment, that instead of focusing on the "intelligence" of the system or its degree of correspondence 'to the real world,' computational technics are also amenable to designs that center instead on user interaction with and interpretation of the system.

Interpretable Social Matching

In HCI, the most well-known proponent of "situated action" is Lucy Suchman, whose classic work *Plans and Situated Actions* explores how formal representations function as descriptions, not drivers, of action, and always require additional work to unpack and make relevant in the situation of action [49]. This social-science perspective of situated action highlights the work done by both technologies and humans in interaction to make sense of each other's behavior. One way in which these ideas have been built on in agent-based AI is to refocus agent construction on clearly communicating what agents are doing. For example, the Expressivator agent architecture supports comprehensibility of agent behavior through "narrative agent architecture", in which the agent continually entails its state to users [41].

A design heuristic following from this work is that social matching agents can and should communicate to users not only what matches have been calculated but also how.

Clearly, such a commitment places constraints on the complexity possible within a social matching process to those which are plausibly narratable to a human user.

Nevertheless, there are several potential benefits to algorithmic interpretability in social recommender systems. A body of work on user-evaluation of recommender systems shows that users prefer systems that explain their processes and favor those tools [7, 22, 39]. Empirical work with music recommenders, for example, suggests that explanation and interaction with visualizations of the recommendation results in "higher levels of user satisfaction and perceived relevance of predicted recommendations" [39]. Recent qualitative sociological analysis by Karakayali et al further suggests that users are *already* actively analyzing their presence on social media—including active classification and categorization of their relationships with other users [24]. Accounting for the recommendation processes could support engagement in a way that goes hand-in-hand with documented user practices in the field of social matching systems and has been shown to elicit positive reactions in closely related design spaces.

EXPLORING ALGORITHMS III: HOW IMPLEMENTATION RESHAPES CONCEPTUAL COMMITMENTS

In the previous section, we explored the conceptual assumptions embodied in social matching algorithms by tracing "lost" ideas from the history of sociology and AI and by demonstrating how these ideas could be drawn on to generate different design strategies for social matching than those currently in use. In this section, we explore how such conceptual commitments are embodied, altered and reworked through the process of embodying them in implementation.

Our decision to examine the ways that alternative design heuristics work out through implementation is grounded in critically-oriented and reflective design traditions in HCI, in which system development is not simply an end in itself, but also a means to reflectively explore underlying assumptions and attitudes about technology and humanity. Within HCI, critically oriented systems have explored design practice with reference to, and commenting on, technology's cultural and historical situation [e.g.27, 42, 43]. This means building technologies to change not only what people can do with but also the way they think about technology. By tracking our attempt to implement a simple, but functioning prototype that embodies alternative values, we shed light on the process by which algorithms come to both embody and rework their authors' conceptual commitments.

The Strangerationist

The prototype system we designed and implemented is called "the Strangerationist." This tool was designed to perform content-based semantic matching to engage "strangers" using the system in conversations about things they have (and do not have) in common. It embodies situated agent behavior by matching users based on analysis of their recent typing activity. It embodies interpretable agent behavior by exposing the process of calculating the match—along with some of the relevant assumptions and biases—as part of the recommendation itself.

The Strangerationist is a Firefox browser add-on. The tool, which is constantly active while Firefox is being used, appears as a sidebar on the browser. While the participant is using the browser, the Strangerationist gathers all text typed into HTML textbox areas (e.g. emails, web searches, blog posts) and sends it to a secure database for interpretation. The system continuously runs an open-source clustering algorithm—adapted from Carrot2— on the newest content by each participant [35]. The adapted Carrot2 algorithm automatically groups small collections of documents into thematic categories using a weighted Singular Value Decomposition. If clusters form, the Strangerationist arbitrarily runs an additional subset of simple language and grammar algorithms on both users' queues to find additional similarities and differences. It may compare, for example, the most frequent pronoun used, average sentence length, or rarest words used. The system provides a narrative account of its entire process in the extended output of the recommendation along with a means for the user to contact their matches.

An example extended output is the following: *"Several algorithms have been used to observe the ways in which you are similar (and not similar) to a stranger named Ken. To make the initial observation, we ran a clustering algorithm that uses documents from you and Ken every time that you contribute data to our service. We use a weighting scheme aimed at balancing the local and global term occurrences in the documents, down-weighting terms that are likely to appear very often. We favored an even contribution by both users to the cluster, preferring a cluster in which you and stranger Ken added similar number of documents. We also numerically rewarded the size of clusters, arbitrarily setting the ideal cluster size to 20 documents (clusters with more than 20 documents were not up-weighted.) After casting off lower-ranked cluster names (e.g. "spud", "growing" and "food") we found the largest value in the resulting vector to determine the best matching phrase: "POTATOES."*

To identify other writerly similarities and differences between you and Stranger Ken, we ran a set of other algorithms to look at aspects of the language in your documents: we have compared your recent vocabularies to Standard English and discovered that the most unusual word you and stranger Ken both use is "maleficent."

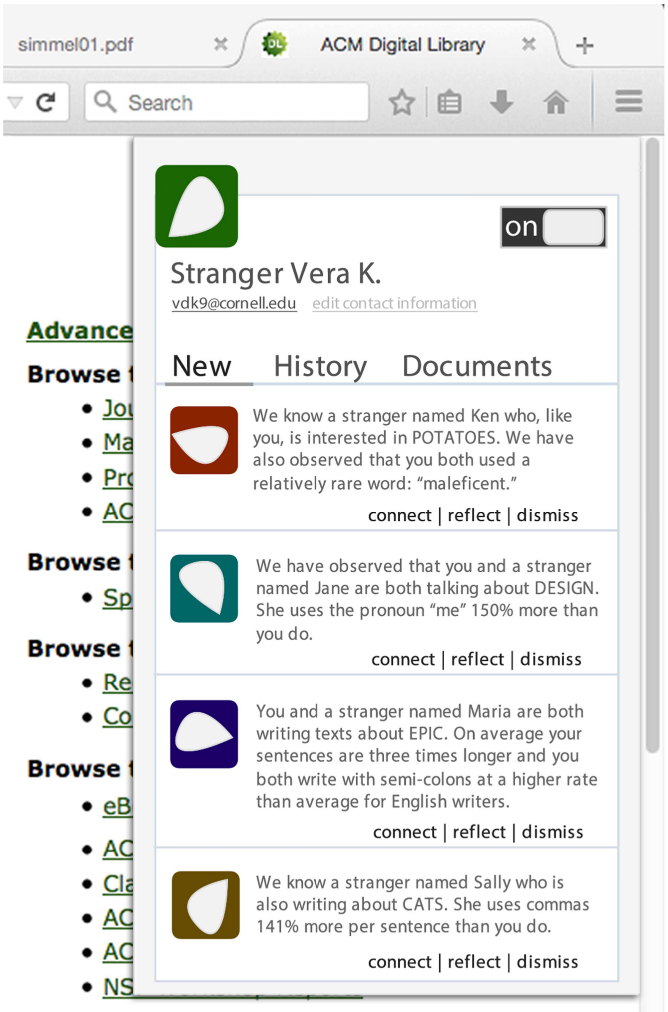

Figure 1. Screenshot of the Strangerationist.

Users of the system also have the option to contact their "matches" or dismiss incoming recommendations. Because the Strangerationist is meant to encourage interaction outside of one's traditional "social network" the design of the system explicitly reminds users of the various benefits of strangers when they select the "dismiss" function. These short quips, e.g. "relating to people who are not like you can lead to new experiences" or "many peripheral relationships are satisfying," were pulled from a variety of sources, including ones that initially motivated our work (e.g. [21]). The interactive design of the "dismiss" action, like the extended computational narrative, is a form of making some of the system's tacit commitments explicit.

In the design of the Strangerationist, we tried to follow our design heuristics, namely, we followed simple rules based on in-the-moment computations based on user-generated content, and we tried to express them, along with our implicit biases, through a narrative output of the system's processes—which was facilitated by the system's writerly and blatantly pro-stranger interaction personality. The emerging personality of the Strangerationist—obsessively wordy in its own account of users' words, aggressively pro-engagement with strangers, should one attempt to dismiss its observations—is reminiscent of other critical design work in which the technologies took on a personality based on the flavor of conceptual work they were doing to further underscore their critical commitments. For example, in a critical design piece that explores the potential social effects of disturbing, uncomfortable systems, a social agent that spreads false, strange gossip throughout an office space is named after a mythological Norse trickster and has an extensive backstory, including an affinity for purple shoes [16]. The output of a home health sensing system that promotes user interpretation through deliberately ambiguous output, expressed its sensor readings as automatically constructed, ambiguous "horoscopes" [18]. In both cases, the systems had personalities which affected their interactions with users and were in line with their conceptual design commitments.

Lessons Learned from Building

The process of designing and implementing the Strangerationist frequently led us to reconsider the conceptual analysis we had initially done. Specific technical decisions described below shed light on how decisions that appeared to be low-level or "technical" simultaneously engaged with, underscored, or altered our critical conceptual work.

Situated, not networked - Documenting the output of the clustering algorithm required us to interpret the results we had obtained algorithmically in light of our new design commitments. An initial temptation in implementing the clustering algorithm was to use word-matching as a way of calculating "strangeness" between people—and the Strangerationist could recommend most strange and least strange people for a user. We realized we had been more deeply influenced than we thought by the rhetorical plasticity of matching language, because even though we were implementing our agent architecture, we were still thinking of it as calculating (modeling) tie strength. In this way, we experienced first hand the "transfer" [31] that computational methods incur on knowledge practices outside of computer science.

To underscore the different idea we were aiming to approach here, we developed a new metaphor: rather than using clustering to measure "distance" between users, we used the clusters to provide an opportunity for people to connect. Rhetorically, it was like a shared "airplane seat"—an excuse for two people to relate to each other. To further this line of thinking, we removed language that directly referred to "matching" (calling the output "observations" instead) and followed the clustering with deliberately simplistic analysis of writerly qualities (e.g. "the most unusual word you and stranger Ken have in common is "maleficent") to make it clear that we were creating an opportunity for matching without making an argument for an objectively existing tie.

Asymmetrical observations – This design lesson emerged from a conversation about how to store the matches. Initially, we imagined that they would be stored in a lookup table such that, if person A is matched to person B, then person B also receives a notification of being matched to person A. We rejected this implementation for several reasons. First, since we were explicitly not calculating distance between people or building a top-down model of the social interactions, there was suddenly no reason to have symmetrical relations. We instead chose deliberately to "reward" contributors of data with observations and stored them in unlinked data structures, such that if person A contributes data that clusters with data given once by person B, person A receives the observation immediately, but person B will not receive any observations until they give the Strangerationist new data. Because the queue is regularly cleared, there is no reason that, at the time of person B's typing, the same clusters will still be calculable. This design move was in line with being reactive because it responded to individual behavior without assuming that relationships needed to be mutual, and it was in line with being interpretable because the asymmetrical recommendations provided an opportunity to clearly communicate that data contributions resulted in observations from the system.

Inherited aesthetics - When we were implementing our clustering algorithm, we directly inherited assumptions about the latent structural relationships between words from the algorithm we adapted. We tried to make these clear in the expanded text, for example, by explaining that the weighting scheme in Carrot2 has a bias toward novel word-phrases: "down-weighting terms that are likely to appear very often" and has a very specific ideal cluster, favoring "an even contribution by both users to the cluster... and setting the ideal cluster size to 20 documents" [34]. They were thresholds that we could change, but the metrics for assessment came as part of the code. The consequence of our decision to use semantic clustering was that, instead of looking for a latent structure between people, we had found a calculable proxy in modeling the latent structure of words in their documents.

As we struggled to make use of the available alternatives to social network analysis without absorbing the positivist motivations of the algorithms, we saw more clearly the broader motivations for doing social matching work. One thing we learned about social matching recommenders through the reframing of the "matching" metaphor in the Strangerationist is that social-network modeling—especially gathering feedback about user behavior to improve the model through iterative representation and re-representation—was not necessary for basic social matching. Through the parallels between historical discourse on social networks and the discourse around Agre's critical technical practice, we see another logical reason for conventional social recommenders to do the iterative modeling that they do. For research enterprises, it directly follows a long history of research in AI, where people are simultaneously users of an interactive system and also data points for improving and assessing the success of an artificial intelligence.

DISCUSSION

Reflecting on the legacy of critical technical practice in the form of situated action from the perspective of reforming practices in artificial intelligence, Agre wrote about the necessary "split identity" of critical technical practitioner between technical design and reflexive critique rooted in non-technical fields [3]. In this work, we ended up rhetorically taking on the role of a technically oriented researcher with expertise in social matching recommenders. We negotiated this stance so what we could explore how critical technical practice could be combined with historical discourse to shed light on values and assumptions undergirding algorithmic practice. It is important to note, however, that we do not identify as experts in the domain of recommender algorithms, and do not believe that the Strangerationist should be considered primarily as a contribution to the domain of recommender systems.

Our work is grounded in critically oriented and reflective design traditions in HCI, which is a home discipline for several of the authors. It is interesting to note that while Agre's methodological contributions to critical technical practice have been re-absorbed into what he considered to be the dominant practice within the field of AI itself [49], critical technical practice has strongly shaped the intellectual traditions HCI. We see system development as not simply an end in itself, but also a means to reflectively explore underlying assumptions and attitudes about technology and humanity. This work has also drawn from the strange, uncomfortable, and provocative "critical design" methods described by Dunne and Raby [14, pp 12] as a means to raise questions about the political implications of design practice. Critical design as a term has been elaborated and repositioned by Bardzell and Bardzell to recognize their ties to a longer critical theory legacy including the Frankfurt School of critical theory and post-structuralism [5]. Within HCI, critically oriented systems have explored design practice with reference to, and commenting on, technology's cultural and historical situation [27, 42, 43]. This means building technologies to change not only what people can do with but also the way they think about technology. For example, critically oriented design methods have been used to alter objective, informatics approaches to affective computing— which structure, formalize, and represent emotion as informational units—by centering technical decisions around the indefinable complexities of human affective experience [6].

Since doing the project, we have been able to make use of historical analysis, substantiated by our familiarity with the details of technical implementation, to discuss social network analysis and recommendation with practitioners in HCI. Because we set out to explore a methodology where the critique of values in design can be a guiding method in technical development, we count technically and historically grounded engagement with social recommender systems among our successes.

But is the Strangerationist successful as a social matching system? This is not possible for us to say. We did not engage in conventional evaluation of the Strangerationist in the form of a user study that assesses its usability and value in quotidian use. This was because our goal in the project was less to create a "better" social matching than to explore the values and commitments embodied in social matching algorithms. For this end-user evaluation is not a directly relevant measure. More deeply, we felt standard models of evaluation significantly misrepresent what Strangerationist was doing as a project.

Moreover, what the 'historical' aspect of the Strangerationist reveals, as an extreme case, are tacit temporalities in technical system design and evaluation in HCI. In order to evaluate a technical system, a prototype design is often deployed back 'in the wild'—so that it can be tested against the realities which inspired the design. As we contemplated the multiple privacy implications of deploying the system—keystroke logger and all—it became clear to us that the temporal frame of swinging time back around and putting the design into the world was not possible in the case of the Strangerationist. In response to over a decade of anxious intertwining of internet users' personal and online lives (e.g. e-commerce, personal data mining, and security panics) it seems that attitudes about gathering keystroke data have shifted in the passage of time. We imagine that the world of finding friends, for example by being on the website, may not exist in the way it served as inspiration for early social recommender algorithms. The Strangerationist itself may be an anachronism for which there is no longer any 'where' against which to test the algorithm on HCI's normative terms. What this method offers instead is an opportunity to reconfigure the relationship between social science and technical methods through the historical analysis and design of technical systems.

CONCLUSION

In this sense, our goal with the Strangerationist differs from what Agre set out to do with critical technical practice in AI: we did not set out to reform the field of social recommenders so much as to imagine and implement an alternative configuration between historical work on technical practice and technical practice itself. What we hope can be drawn from our work is an approach to borrowing ideas from critical technical practice and historical discourse on technical practices to inform sociological and STS understanding of algorithms. In some ways, by exploring how technical practices can be used in such a way, we articulated a method which resembles the conceptual inverse of critical technical practice: taking ideas from technical discourse to innovate on research practices in the social sciences. We hope that this combination of approaches can be used to explore the entangled facets of algorithms that are often left unexamined by researchers operating outside of niche technical research programs, and thought to be exclusively visible through (to borrow Agre's phrase) "the daily work of trying to get things built and working" [2].

ACKNOWLEDGEMENTS

This work was funded by NSF grant 1217685 and the Intel Science and Technology Center for Social Computing. The Strangerationist was developed at Intel Labs with Ken Anderson, Richard Beckwith, and Cian Montgomery. Thanks to Nick Seaver, Ellie Harmon, and all of our anonymous reviewers for their helpful feedback.

REFERENCES

1. Agre, P. (1997). *Computation and human experience*. Cambridge University Press.
2. Agre, P. (1997). Beyond the mirror world: Privacy and the representational practices of computing. *Technology and Privacy the New Landscape (PE Agre and M. Rotenberg, Eds.). Massachusetts, MIT Press*, 29-62.
3. Agre, P. "Toward a critical technical practice: Lessons learned in trying to reform AI." *Bridging the Great Divide: Social Science, Technical Systems, and Cooperative Work, Mahwah, NJ: Erlbaum* (1997): 131-157.
4. Agre P., and Chapman, D. 1987. Pengi: an implementation of a theory of activity. In Proceedings of the sixth National conference on Artificial intelligence - Volume 1 (AAAI'87), Vol. 1. AAAI Press 268-272.
5. Bardzell, J., and Bardzell, S. 2013. What is "critical" about critical design? *Proc* CHI '13, 3297-3306.
6. Boehner, K., Sengers, P., and Warner, S. 2008. Interfaces with the ineffable: Meeting aesthetic experience on its own terms. ACM Trans. Comput.-Hum. Interact. 15, 3, Article 12 (December 2008)
7. Bostandjiev, S., O'Donovan, J., and Höllerer, T. 2012. TasteWeights: a visual interactive hybrid recommender system. In *Proc* RecSys '12. 35-42.
8. Bowker, G., Star, S. L., Gasser, L., & Turner, W. (Eds.). (2014). *Social science, technical systems, and cooperative work: Beyond the great divide*. Psychology Press.
9. Brooks, Rodney A. "Elephants don't play chess." *Robotics and autonomous systems* 6.1 (1990): 3-15.
10. Brooks, R. A. (1989). A robot that walks; emergent behaviors from a carefully evolved network. *Neural computation*, *1*(2), 253-262.

11. Chalmers, Matthew. "Seamful design and ubicomp infrastructure." *Proceedings of Ubicomp 2003 Workshop at the Crossroads: The Interaction of HCI and Systems Issues in Ubicomp*. 2003.
12. Budzik J., Bradshaw, S., Fu, X., and Hammond, K. J. 2002. Clustering for opportunistic communication. In *Proc* WWW ’02. 726-735.
13. Chen, J., Geyer, W., Dugan, C., Muller, M., Guy, I. 2009. Make new friends, but keep the old: recommending people on social networking sites. In *Proc* CHI '09. 201-210.
14. Dunne, A., and Raby, F. *Design noir: The secret life of electronic objects*. Birkhäuser, Basel, Switzerland, 2001.
15. Ergin, E., and Lee, D. "On six degrees of separation in DBLP-DB and more." *ACM SIGMOD Record* 34.2 (2005). 33-40.
16. Foucault, B., Mentis, H. M., Sengers, P., and Welles, D. 2007. Provoking sociability. In *Proc* CHI '07. 1557-1560.
17. Freeman L. 1998. Computer programs in social network analysis. *Connections 11*, 26–31.
18. Gaver, W., Sengers, P., Kerridge, T., Kaye, J., and Bowers, J. Enhancing ubiquitous computing with user interpretation: field testing the home health horoscope. In *Proc* CHI '07. 537-546.
19. Gilbert, E., and Karahalios, K. Predicting tie strength with social media. In *Proc* CHI '09. 211-220.
20. Gillespie, T. The relevance of algorithms. In T. Gillespie, P. Boczkowski, and K. Foot, eds., *Media Technologies*. MIT Press, forthcoming.
21. Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 1360-1380.
22. Herlocker, J. L., Konstan, J. A., and Riedl, J. Explaining collaborative filtering recommendations. In *Proc. CSCW '00*. 241-250.
23. Kadima, H., & Malek, M. (2010, December). Toward ontology-based personalization of a recommender system in social network. In *Soft Computing and Pattern Recognition (SoCPaR), 2010 International Conference of* (pp. 119-122). IEEE.
24. Karakayali, N., & Kilic, A. (2013). More network conscious than ever? Challenges, strategies, and analytic labor of users in the Facebook environment. *Journal of Computer-Mediated Communication*, *18*(*2*).
25. Kautz, H., Selman, B., and Shah, M. 1997. ReferralWeb: Combining social networks and collaborative filtering. *Comm. ACM 30*, 3.
26. Krzywicki, A., Wobcke, W., Kim, Y. S., Cai, X., Bain, M., Mahidadia, A., & Compton, P. (2015). Collaborative Filtering for people-to-people recommendation in online dating: Data analysis and user trial. *International Journal of Human-Computer Studies*, *76*, 50-66.
27. Leahu, L., Schwenk, S., & Sengers, P. Subjective objectivity: negotiating emotional meaning. In *Proc* DIS ’08. 425-434.
28. Leahu, L., Thom-Santelli, J., Pederson, and C., Sengers, P. Taming the situationist beast. In *Proc* DIS '08. 203-211.
29. Leahu, L., P. Sengers, and M. Mateas. 2008. Interactionist AI and the promise of ubicomp, or, how to put your box in the world without putting the world in your box. In *Proceedings of UbiComp '08*, 134-143.
30. Leahu, L. Representation without representationalism. PhD thesis, Cornell Computer Science, 2012.
31. Machlup, F. and Mansfield, U. (1983), *The Study of Information: Interdisciplinary Messages* (New York: Wiley).
32. McDonald, D. W.and Ackerman, M. S. 2000. Expertise recommender: A flexible recommendation architecture. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work. (CSCW’00)*. 231–240.
33. Milgram, S. (1967). The small world problem. *Psychology today*, *2*(1), 60-67.
34. Osiński, S. *An algorithm for clustering of web search results*. Diss. Poznań University of Technology, Poland, 2003.
35. Osiński, S., & Weiss, D. (2004). Conceptual clustering using lingo algorithm: Evaluation on open directory project data. In *Intelligent Information Processing and Web Mining* (pp. 369-377). Springer Berlin Heidelberg.
36. Panovich, K., Miller, R., and Karger, D. 2012. Tie strength in question & answer on social network sites. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW '12).
37. Resnick, P., R. Garrett, R. K., Kriplean, T., Munson, S. A., and Stroud, J. 2013. Bursting your (filter) bubble: strategies for promoting diverse exposure. In Proceedings of the 2013 conference on Computer supported cooperative work companion (CSCW '13).
38. Röhle, T., and Rieder, B. (2012) ‘Digital Methods: Five Challenges,’ in D. Berry (ed.), *Understanding Digital Humanities: The Computational Turn and New Technology*. (London: Palgrave Macmillan), pp. 67–84.
39. Sharma, A. 2013. PopCore: a system for network-centric recommendation. In *Proceedings of the 2013*

conference on Computer supported cooperative work companion (CSCW '13).

40. Seaver, N. *(2015)* The nice thing about context is that everyone has it. *Media Culture Society,* vol. 37 no. 7 1101-1109
41. Sengers, P. (1998, May). Do the thing right: an architecture for action-expression. In *Proceedings of the second international conference on Autonomous agents* (pp. 24-31). ACM.
42. Sengers, P., Boehner, K., David, S. and Kaye, J. Reflective design. *Proc. Critical Computing*, ACM Press (2005), 49–58.
43. Sengers, S., McCarthy, J., and Dourish, P. 2006. Reflective HCI: articulating an agenda for critical practice. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '06).
44. Simmel, Georg. "The stranger." *The Sociology of Georg Simmel* (1950): 402-408.
45. Star, Susan Leigh. "The ethnography of infrastructure." *American behavioral scientist* 43.3 (1999): 377-391.
46. de Sola Pool, Ithiel, and Manfred Kochen. "Contacts and influence." *Social networks* 1.1 (1979): 5-51.
47. Sosik, V.S., Zhao, X. and Cosley, D. See friendship, sort of: How conversation and digital traces might support reflection on friendships. *Proc. CSCW '12*
48. Suchman, L. *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press, 2007
49. Suchman, Lucy A. *Plans and situated actions: the problem of human-machine communication*, Cambridge University Press, New York, NY, 1987.
50. Terveen L., and McDonald, D. W., 2005. Social matching: A framework and research agenda. *ACM Trans. Comput.-Hum. Interact.* (September 2005), 401-434.