

Joseph Konstan on Human-Computer Interaction

Recommender Systems, Collaboration, and Social Good

[Joseph Konstan is an associate professor of computer science at the University of Minnesota. His background includes a bachelor's degree from Harvard College and a Ph.D. from the University of California, Berkeley. His principal interests are human-computer interaction, recommender systems, multimedia systems, information visualization, internet applications and interfaces.]

UBIQUITY: How does your work fit you into the general field of human/computer interaction?

KONSTAN: My work as a researcher spans several different parts of human/computer interaction. The biggest project I have been working on, and one that I have been working on now for nine years, is the GroupLens project which is about recommender systems—systems that do real-time personalization. It is very much like you see on Amazon.com when you are recommended books or movies that they think you might like. I joined that project a decade ago, and it had already been going for a couple of years. We've been exploring both the technology for how you create those recommendations and, what I think is more important, the understanding of what designs and what properties lead users to find them useful. So a chunk of this work is understanding, given what a computer can do, what is better to present to a person to be helpful to them. I will give you one concrete example of that. We have, and this is work that dates back to '99 or so, studied explaining to users what the system was doing as a way of helping them understand whether they should trust the computer systems' recommendations and we found that most of the explanations that were intuitively appealing to a computer scientist, things that got into the statistics and the processing, completely turned off ordinary people. At the same time, really simple three point charts or analogies were much more compelling to the average user. Knowing this has been helpful. In recent years that work has also moved much more into understanding the whole idea of online communities and how people participate in them, and what's most interesting in my perspective, how you can design a community to elicit a level of participation. What do you set up in the design of a website, whether it's for conversation or in our case, for getting people to rate movies or rate other content, what do you do in that design to get people to contribute? Do you tell them about what other people are doing? Do you help them measure themselves against their peers? Do you show them how much other people are benefiting from their work, or how much they've benefited from the work of others? I think these are interesting questions that hit at the overlap between computer science and psychology, sociology, economics, other social sciences.

UBIQUITY: There are of course various social aspects to these issues, but all your degrees are in computer science, right?

KONSTAN: All my degrees are in computer science. I did my minors when I was a graduate student for experimental psychology and in education, and so I have gotten a little bit of training in that area, but I've had to learn this as I go along and collaborate with people that know it better than I do.

UBIQUITY: You mentioned Amazon; you haven't had anything to do with Amazon have you?

KONSTAN: The technology we created was at one point licensed by Amazon, though I doubt they are using it anymore. But the basic idea of a recommender at its core--the

system that learns your preferences either by watching you or just having you tell it, tell the system here's what I like and dislike, and comes back and helps you find other things you might like. The simplest system we run today is a system we call MovieLens (www.movielens.org). When you go into the system it shows you a bunch of movies, you can rate them if you have seen them. Over time it gets more and more accurate at recommending movies that you will like or at predicting for a given movie you might ask it about, how much you'll like it. The way it does that is by matching your ratings up with those of tens of thousands of other people.

UBIQUITY: And has your movie-reviewing system won good reviews for itself?

KONSTAN: You know, we have a lot of users who really love it, which is great, but more important we're learning things about how you build and design these systems, which is what we really care about. The whole intuitive idea is word of mouth. You find what other people say and you use that to help you form an opinion. A computer can do this better than a human can because it can consider thousands of opinions. If you were thinking what movie am I going to go see, you would go and talk to a half dozen people in your office and they suggest some movies and not others, and the problem is their tastes may not match yours at all. MovieLens has tens of thousands of people who have evaluated movies; we've got over 10 million movie ratings in our database. We can find opinions that line up pretty well with what your tastes have been so far, and so our challenge becomes how do we make sure we get enough information about you so that we can match you up with other people, and then deliver that information back to you in a way that's useful.

UBIQUITY: What problems have you found in the way people feel about it?

KONSTAN: One of the challenges that always comes up is that there is a set of rare films, and getting information on things that are obscure is always hard. You also get selection bias; most of the time when someone tells us about a movie they're telling us they liked it, but part of that is that people don't go to movies that they don't think they're going to like and they are usually right. But it also means that on average we're going to see almost all of our scores be much more positive than a random selection of movies would be. But in general I think what we have found is the stuff really works; we spun out a company which had an internet bubble meteoric rise and fall in the late '90s and early 2000's but it did get this technology out and adopted by a lot of e-commerce websites and other businesses that were interested in seeing ways to personalize these large catalog offerings that they would otherwise be making to their customers or members, and the technology has been a fairly unqualified success.

UBIQUITY: Can you mention any prominent companies where it's used?

KONSTAN: Some of the well-known customers included businesses such as Amazon.com, CDnow, Brylane, GUS, E!Online, and Ticketmaster.com. These companies used the software for a variety of personalization and marketing tasks including real-time web personalization and call-center sales.

UBIQUITY: So what would one find on first encountering a movie rating system?

KONSTAN: In our system, because we're building this for research, we force people to rate about ten movies before we'll do anything for you. We don't want to clutter our database with users who aren't at least that serious. Once you have rated those ten movies you'll get back a screen that recommends some new movies and new DVD releases that we think you will be interested in, and a search facility. You could come in and say, what I am in the

mood for today is a comedy, and it will come back and say, well here are what we think are the top comedies for you. Or you could say here is a movie I heard about, Million Dollar Baby. Would I like it? It will come back and say on a one to five star scale, oh, this one for you, three and half stars--sort of a toss up. You could connect with other people, we use the term buddies like many other people do. One you and your buddies are connected, you can ask MovieLens to find a movie that is in theatres that I would like and also these buddies would like. It will return a list of films sorted by how much you are going to like them collectively so you can find movies to go to as a group. You could also sort or search in a lot of other fields; you can search by movies by genre, by date, by whether they have been released on VHS or DVD or other details like that.

UBIQUITY: Is there much content in the database?

KONSTAN: We're not providing very much content. You can look up things like directors and actors on some of the movies but we send you off to other sites if you want deep content. We're there to answer the question how much would I like this, and to give you a list if you say here's what I am in the mood for, can you give me back a list. We'll give you a list to start your decision making with. And then we do all sorts of odd and interesting things once we have this user community as a researcher; what's most powerful to me is we've got hundreds of new users coming every week and thousands of people coming back every month and we can experiment with them. We can try out a new feature and see if they like it. We can try out different versions of the system and see if one version or another leads people to be happier, or to rate more movies, or to log in more often or to do other things like that, and we do that regularly.

UBIQUITY: Approximately how many movies listed are in your database?

KONSTAN: We are up around 10 thousand movies. Certainly nothing close to all the movies that are out there -- but we're trying to focus on the ones that people in the system want.

UBIQUITY: Talk about some of the other applications beyond movies and products for sale.

KONSTAN: One of the ones we have been looking at is in the area of digital libraries. I've got a student who's working with a couple of other people that built a prototype of a research paper recommender. You can tell it which papers you've already read and it will recommend papers that you should read next? He's actually now working with data from the ACM digital library to see what types of recommenders we can build that would help you discover that an article just published is something you should know about. You're doing research now in this new area? Here's a set of things to get you up to speed. I think that recommending information is an important area; other people have done research using news as a domain (e.g., Mike Pazanni at UC Irvine built news recommenders that would send you news based on a profile of what would interest you most). In general recommender systems work in domains where you can identify that there is something about human preference that's consistent and that's a better model of what people want than the next best alternative. I am not sure I would want to use this to pick a doctor.

UBIQUITY: Why not?

KONSTAN: Because I think there is some objective data that might be more important than the preference of other people who are like me.

UBIQUITY: But wouldn't it reinforce that objective?

KONSTAN: I might use a hybrid system that on the one hand went and said, well gee I want a doctor who is board-certified in the area and whose patients tend to survive and thrive, and who also is one that people like me like, because bedside manner and responsiveness do matter. But it is the same argument that I am not sure that you pick investments this way; you don't necessarily want to pick an investment because other people have picked it, you want to pick an investment before the other people have picked it, because they have probably bid the price up. In a competitive and dynamic environment you have to be a little more careful about this kind of technology because the fact that something was good yesterday doesn't mean it is still good today. Whereas most of the things we're talking about are things that have some persistence in your opinion of them. A movie that you thought was great yesterday didn't become bad today.

UBIQUITY: What about social kinds of things, or political kinds of things, any application?

KONSTAN: There are people who have been trying to look at this technology to help people identify everything from political candidates and issues to meetings and groups of other people. We haven't done that ourselves but I think it is an exciting area to look at. The whole idea of matchmaking is an exciting one. Not matchmaking for dating and romance, but instead for example putting together ten people who have interesting things to talk about. How do you put together a group of ten people? There are some interesting challenges with how you gather the data. And with what you tell people about why they were matched together. We've talked about using this in areas like computer games. When you go online to one of these game sites and want to play backgammon, there are 7,000 people playing; how can you figure out who would be fun to play with? It might be useful to be able to track who you enjoyed playing with, and who other people enjoyed, and then let the system suggest, hey there is someone sitting at table number 23 who would be a good match for you. It might know why, it might be that their attributes are understood; this is somebody whose chatty but always makes their moves on time. It might be that you don't know why this person is good, but you just trust that if other people think so then you believe it too. It's a huge issue with bridge because of the nature of the partnership. In online bridge finding a compatible partner is difficult. Some people don't care as much about the opponents, others probably do. But I think this is something we're going to run into if we want the online world to be a social world.

UBIQUITY: What are your thoughts about online communities?

KONSTAN: The topic of online communities is closely related to what we are doing here. We actually got funded about two years ago to start working with psychologists and economists to understand what leads people to contribute to these online communities. We've done a variety of experiments already that look at things like creating discussion groups and seeing what kinds of prompts lead people to discuss more or less. One example grounded in psychology theory says that if somebody believes that they have a unique contribution to make they are more likely to contribute than if they think that other people can make the same contribution and therefore they are redundant. So we did an experiment with discussion groups where for a particular topic of discussion we would find something in that person's history, again looking at their movie tastes in this case, and tell them, here's something that's true about you that nobody else in the group has that property. You're the only person in the group who has seen this movie, or you're the only person in the group who's liked this movie. Telling someone how they're unique leads them to discuss more. It teaches us something about how we can facilitate those discussions. We're interested in how big an online community can be before it feels like it's too large and you've lost a sense of people's identity.

UBIQUITY: What else have you been doing with on-line communities?

KONSTAN: We've been running some experiments around the question of what happens when you take the members of a site like MovieLens and give them control to maintain the database. Let them add movies, let them correct mistakes, and while our first results are a bit disastrous, they're quite interesting. The people go and do things, but they don't necessarily do things in the interest of the community. At times they put their favorites in, even if what people asked them to do was something else. We're also finding that you can start to design mechanisms where people are guided to act in the interest of the community. I have recently been working with some economists at the University of Michigan; we've built economic models of people's participation in one of these sites. How much happiness do they get just from using the site versus what they get by influencing others or from the quality of the recommendations they get out. And we're using those economic models to start testing theories that suggest we can give you an economic report card that says at the end of the month you've gotten this much value from the site but on average other people have gotten a lot less because you've taken more than you have given and maybe you want to find a way to give back.

UBIQUITY: Don't the same social principles apply both to online communities and to writing to newspapers and so forth?

KONSTAN: That's one of the questions we're trying to understand. There have been a lot of studies of face-to-face communities and we're not sure that all of the results that hold at a face to face community hold online. Now it is certainly true you can think of online as a case non-face-to-face, so if you had a club where people always communicated by mail, e.g., by playing mail chess, that might have the same property as online rather than a bridge club where people see each other. But there are a lot of issues of identity online that are different from the real world. You can disappear and come back five minutes later with another name and nobody knows it's you. There's a question of whether the social cues that exist in the real world that stop people from doing things that might be harmful to a community exist online, or whether there's that feeling of anonymity and isolation that lets people behave differently. So online is not purely incidental here, part of what we're trying to understand is how much of the psychological research that's been done in the lab really works in this kind of real-world online setting.

UBIQUITY: Let's pause here for a minute and ask you to tell us about your group.

KONSTAN: Oh, sure. Here at the University of Minnesota there are three faculty working in this area: me; John Riedl, who's been here longer than I have and we started working together about a decade ago; and Loren Terveen, who joined us about three years ago, after having the beginning of his career at AT&T research. Loren had also been doing work in related areas, so we're lucky we were able to grow the group. We have in our lab together, at any given time, somewhere around 20 students working with us, sometimes 25, a mixture of PhD students, master's students and undergraduates. We typically will have one or two PhD students a year graduate out of the program and another three to five master's students getting degrees with us. We have alumni who now are spreading out to a whole bunch of places including four universities and industry jobs as well. We were one of the first places where this kind of work was done.

UBIQUITY: How many places now exist that are prominent in the field?

KONSTAN: If you look in recommenders work specifically, there are probably between half a dozen and a dozen that are prominent that are regularly active in publishing, and maybe another dozen that do this related to some other area that's their real specialty. If you take this more broadly and say well what about the people doing work in online communities, people doing more general work in human/computer interaction it obviously, that number gets much larger.

UBIQUITY: How did you happen to get into this particular specialty?

KONSTAN: For me it was an accident. John Riedl my colleague and Paul Resnick, who at the time was a graduate student at MIT and is now a faculty member at Michigan, were talking at a conference where they saw an interesting keynote address about the information economy and ...

UBIQUITY: Who gave it?

KONSTAN: Shumpei Kumon. It was the CSCW 1992 conference, which I think was in Toronto, and at that conference they started coming up with ideas and the two of them and a small group of students at both places put together the first GroupLens system which was a recommender system for Usenet newsgroups (for people who aren't familiar with Usenet news, it is a set of public discussion groups). They did a proof of concept system with just the two sites, Minnesota and MIT, showing that you could have people rate articles that they had read and then display predictions for the articles to people before they choose which articles to read. That work was published in '94 and was the first published example of a working automated collaborative filtering system, which is the name of the technology (the term recommender systems came later in the field). They received very positive feedback from that work and John came back here very excited and convinced me to get involved with it in '95; we then went did another much larger study with a couple a hundred users, across multiple sites and distributing newsreaders out to different people. We ran that study in early '96, and by that time there was two or three other groups doing work in this area and we were frantically trying to keep somewhat of a lead. We formed a company with that technology in the summer of '96, that's the company Net Perceptions, and we actually got our first research funding from the National Science Foundation to do work in that area in '97 and that gave us the ability to start attracting more students to this work and it's consistently grown from there.

UBIQUITY: Thinking about the half a dozen or so prominent groups that are working in this field, is your approach significantly different from theirs in any way?

KONSTAN: Oh, very much so, I think everyone of them has their own specialization. Some of them specialize in a particular recommender technology and applying it to lots of problems. There are some people who try to specialize in a particular problem. There are some who are taking aspects like privacy or other personalization features and trying to go forward with them. We lately have been fairly broad but we've had certain areas where we've tried to build our expertise to be strongest. One of them is in this area of evaluation, understanding that what makes a recommender system better than another, or an algorithm better than another, in deep ways. We've been doing a lot of work in that area for several years really dating back to early work we were doing in '95. The important idea is to get beyond the simple measures of how many stars away from the user's true rating you are. The naive solution to understanding how good a recommender is to simply say how close is it to being right. So if you believe this is a four star movie, and I say it's four and a half stars then I am a half a star off; I can take the average error and see how good I am. What we have been arguing now and pretty much convincing the field for ten years, is that

there many more sophisticated ways of looking at value to the user. You probably don't care whether I say four stars or four and a half stars for a movie, that's high enough that you are going to want to go see it. If I make a mistake though and I take a movie that was at four stars and I give you three instead of five, you may see the three stars and decide not to bother ever considering it. Worse yet, it may move off the first screen of recommendations and the user may never see it.

UBIQUITY: How do people deal with these possibilities?

KONSTAN: People have been building ways of evaluating based on how good are the top ten items that I give you. How many times did I mislead you? Some of our recent research which is just coming out this coming May at the World Wide Web conference in Japan shows that other features that matter, specifically how diverse a set of recommendations I offer you. It's not that exciting if I tell you that the top ten movies that you should go see are Star Trek The Motion Picture, Star Trek 2, Star Trek 3, Star Trek 4, Star Trek 5. Once I told you one of those, the other ones might be obvious. But if I give you two science fiction movies and a couple of comedies and a romance and a documentary, most users will find that to be a more valuable tool. It gives them a chance to explore some diversity, which is just one of the many other measures that you can look at when you evaluate a recommender.

UBIQUITY: How might recommenders apply to something like, if I dare suggest it, evaluation of faculty at the University of Minnesota, for example?

KONSTAN: Sure, go ahead and dare suggest it. I think there are two questions. Are you doing this to evaluate faculty (in which case it wouldn't be a great tool) or are you doing it to evaluate the fit between students and faculty (where it might be just the right tool). Today we have students fill out numerical surveys about the faculty and the courses they took. Right now we just take those numbers and give you a statistical summary and you know I find out on a scale of seven you're a six as a teacher and the course is five for how much I learned from it. The classroom was a 5.5. We could instead personalize that data. The system could come back and say, hey, based on people who agreed with you in past years, for you this instructor is a 7 out of 7. For somebody else this, on average the instructor was a six but this is the kind of instructor that you like. Or this is the type of class that you like. At the same time, another student may hear, for the same class, that this isn't a good fit.

UBIQUITY: When is such a system like to be least useful and when is it likely to be most useful?

KONSTAN: I am not sure if that would be the most useful tool because we don't have a lot of data, it might work for very large classes, it's not going to work well when you have a class with six or ten students in it. Also, because there are so many other factors affecting student evaluations, it would probably work best in elective classes or in classes where you get to choose one of several instructors. After all, for the average student going into computer science, it doesn't matter if you tell them you don't think they're going to like calculus. It's a requirement, and they're going to take it anyway. You're not necessarily supposed to *like* calculus anyway; you're going to work your way through it to learn something useful. Still, it might be helpful in saying identifying the two or three teachers, from among the twelve choices, who seem like the best fit for a student. I will say that in any of these domains, if you have a better way to make a match, I normally would advocate that you use it. In learning and teaching it might be that using one of the learning style assessments would be better; if so, that would be great. It may be that you can do an

assessment of a teacher's teaching style and a student's learning style and find out that you have a good match. What our technology is really powerful for is where you don't have that kind of a match, and then you can come back and say, oh, I could just use the raw power of lots of data to get me something that's pretty good.

UBIQUITY: What about opinion research? Could you do Gallup-style research?

KONSTAN: In work on opinion research the biggest challenge is in getting the opinions from the right people, which usually means a random sample of a population that you want to focus on. And so for election research you may say we want a random sample of likely voters. But often our volunteer subjects show up and express a lot of strong personal opinions, and volunteers like that don't give us a great way of estimating what the population as a whole thinks. For example, with regard to movies, you could take almost any really bad movie and we'll find you a bunch of people who rated it highly. The reason they rated it is because they liked that movie and that doesn't mean it's a popular movie. There are interesting questions you can ask here. You can find out whose opinions are influential. Pedro Domingos at the University of Washington has done some work showing how to analyze a population to find out which people you should give a free sample to if your goal is to spread positive word of mouth. A combination of people who are influential to others and who are likely to like what you give them. I know there are some trade groups that sell access to influential populations, people with big social networks who are known to be communicative, as a way to try to create opinion. But when you start getting into assessing opinion you start getting into really some hard problems of sampling. Which are the kinds of problems that we were facing when we were trying to do a different project on HIV prevention, where we would like to be able to generalize to a larger population. One of the first questions we had to ask was whether we were actually reaching people who were representative of the larger population?

UBIQUITY: Do you involve undergraduates in your research?

KONSTAN: Yes, as much as possible. Some of our graduate students were undergraduates working with the project, others we involve as undergraduates sometimes for just for a semester or two; when we can we like to get them as juniors and give them two years to work on some interesting research before they go on to whatever they do next.

UBIQUITY: I assume that you see this as a growing field -- or is there a natural limit on how much work will be done in this?

KONSTAN: I think that the field broadens so that it can grow. When we were seeing the field as primarily people who were worried about algorithms for predicting opinion there was going to be a limit; there's a finite number of approaches that people already know and a limit to the number of people who can engage themselves in trying to create new approaches. As you start going broader in dealing with the human interface issues, as well as online community and social issues, I think the problem becomes much bigger and the solutions that come up and the research results have a bigger impact on other problems. The people who are working in this area come from very different backgrounds. Some come from human/computer interaction, some come from artificial intelligence and machine learning, some of them are coming out of business schools and marketing and economics and psychology, not to mention other parts of computer science, data mining and different areas. As you bring all those different people together its not just that they're solving this problem better but they're solving related problems and broader problems. I don't think you'll ever see a million people doing research in this area but I don't think we're at a point now where somebody considering it should stop because the field is saturated.

UBIQUITY: You're now in the middle of another big project now, which is about HIV infection and the Internet, right? Tell us about it.

KONSTAN: We're now in the second phase of a multiphase project. In the first phase our question simply was, can we assess whether there is a different and larger risk of a population (for the purpose of the study it was the U.S.-resident Latino men who've had sex with other men and who use the Internet) for HIV. The part I was most involved with was figuring out how you find this out. And we developed online survey techniques that we advertised online and then all sorts of security mechanisms because we had people who were trying to take the survey dozens of times to get paid extra. We wanted to try to get to a population that roughly modeled the Latino Gay population in the US, and in the end, after we got rid of all the duplicates and invalid responses, we reached that population. We also found that the men in that population who sought or met other men online for sexual liaisons, rather than first meeting them in person, were exposing themselves to more HIV risk. Some of that risk could be attributed to taking reduced precautions, but most of that risk was the fact that they were meeting more sexual partners online than they did in bars and conventional offline places.

UBIQUITY: And did you say that you're now in the second part of the study?

KONSTAN: Yes, the second part of this work, which we're in now, is repeating the study for a larger population. We're moving outside of just Latinos, which had been picked originally because they were a known high risk population, to look at all men in the US who engage in sex with other men. The more interesting and fun part, however, is that we're developing an online tool to reduce that risk. This is a team that is lead by medical school people who specialize in sexual behavior, but we've also got people in curriculum and online design to create a multi-media online experience that if it works successfully, when somebody has gone through that experience on average, they will have improved sexual health and take fewer risks over the course of the next year and the rest of their life. And it has the nice property of being something where you know that the result matters. It also is something that we do not know how to do successfully. We're building it based on the best in-person techniques that are known and trying to figure out will they work online and so it's been rather exciting.

UBIQUITY: So what are your tentative conclusions?

KONSTAN: The biggest conclusion so far was the fact that men in this group that we studied are exposing themselves to more risks meeting other men online than those same men do when they meet men in conventional settings. We've also learned that these surveys have tremendous technical hurdles to make them work. In public health and medicine they're used to trying to do real random sampling and you can't do that here. There's no way to find people online at random, so we now have a better understanding of who you reach when you go online and how you make sure that the people you're reaching are legitimate. We're pretty happy; we've managed to mirror in the census for the distribution of the people and found ways to do that pretty well. I think the big results are the ones that are going to come out, if everything goes well, in about four years when we've finished not only building this tool but trying it out on about 600 people to see if we can make a change in their behavior.

UBIQUITY: A very interesting project. Computer scientists don't always try to change people's social behavior.

KONSTAN: My takeaway message for the computer scientists here is there are some very interesting opportunities to collaborate with people solving big problems in the world, whether you're interested in AIDS and medical problems, or the kind of work that Negroponte was talking about with the hundred dollar computers for the developing world, or dozens of other things. There are a lot of opportunities there where you can make a difference.

UBIQUITY: How do you find those opportunities?

KONSTAN: I don't always know. In this case I was lucky. The person who was running this project came to me and said I need a computer scientist who understands the internet. If you're at a university it is usually not that hard because you can go to this immense collection of talks and seminars and meet people. If you're out in the rest of the world, the real world to many of us, probably it depends if you're trying to do it professionally or as a volunteer. If you want to do it as a volunteer almost any of the good volunteer agencies, if you call them up and say that you have technical skills and want to use them in a way to better people's lives, will find a way to use your skills. Professionally it's a much greater challenge because most professional jobs are not lined up with social good.

[END]

Joseph A. Konstan
<http://www.umn.edu/~konstan>
konstan@acm.org

Source: Ubiquity, Volume 6, Issue 10 (March 24 - April 4, 2005)
<http://www.acm.org/ubiquity>