

# Recommendations 2.0

John Riedl, Ph.D.

**Discovery Expert Series** is a series of interviews by John Furier, the founder & former CEO of PodTech Network, Inc. John has a history with Internet companies serving as Vice President of the Product Group and Strategic Development for RealNames, an Internet search and naming company, and founder and CEO of Labrador Software, a paid keyword search and navigation company. John also has experience at large technology companies, having spent nine years at Hewlett-Packard in various technical, marketing, and sales roles.

*John Riedl, author of **Word of Mouse: The Marketing Power of Collaborative Filtering**, discusses the evolution of recommendation systems and lessons learned from his experience at Net Perceptions. He outlines how the technology has evolved from difficult and expensive to deploy, to simple and effective. Now, virtually any online businesses can have an efficient, low-risk way to integrate discovery into their marketing and merchandising decision-making. Riedl explains why just employing a keyword search on a site isn't enough in the Web 2.0 world.*

*John Riedl was co-founder and Chief Scientist for Net Perceptions, an early leader in online personalization technology. Riedl is currently a professor in the computer science department at the University of Minnesota where his research includes the GroupLens Project one of the most famous collaborative filtering and recommendation research groups in the world. In 1999, Riedl and other Net Perceptions co-founders shared the MIT Sloan School's award for E-Commerce Technology.*

---

[Edited from original transcript]

**We're here with John Riedl, professor at The University of Minnesota, Department of Computer Science. John, welcome to the podcast on the Discovery Series. Tell us about who you are, and what you're doing at the university and some of the projects that you're working on.**

Sure, I'm a professor in the computer science department and I work on the GroupLens project, which is a sort of overarching program of research that explores the applications of computer technology, broadly, to help people find the information, products, and services that they're most interested in, especially on the Internet.

**What are some of the things that you're seeing that you're focusing your research on, some of the big trends?**

Well, the biggest thing I see in the Web 2.0 age is that it's all about letting people contribute content to the Web and then letting other people comment, discuss, rate, review that content so that the stuff that we all see is really put together by other people like us. I mean, in a way of thinking about it, I see Web 2.0 as a democratization of the editorial process, in a sense.

The original web let anybody who wanted be able to write things, and the problem was that there were still a very limited number of people who got to choose which of the things that were written or produced (for instance, as an audio podcast) which of those things were viewed by other people, and in the 2.0 era, we use technologies like rating, reviewing, tagging to help other people find the stuff that they'll be most interested in from all of the things that are available on the Web.

**You're one of the founders of GroupLens, you mentioned, perhaps one of the most famous collaborative filtering and recommendation research groups in the world. Tell me how the group got started.**

Well, sure, I can remember exactly because it was a very exciting moment for me. In 1992, I was sitting with a friend, Paul Resnick, in a talk at the Computer Supportive Cooperative Work conference, and as we were sitting listening to that talk, we realized that the researcher was envisioning a world in which the most important things that everybody would produce and consume for the economy of the world were going to be information items.

And it was really kind of a cool vision, but there were two things missing from the vision that he described. The first one is where were we going to actually get enough food to eat in this world where all anyone ever did was information. And I don't know a solution to that one; so I won't address it further.

But the second one is that we realized that this was really the way the world went. There was going to be a terrible challenge for people to pick out, of all this available information, the stuff that they individually were most interested in. We saw that the technology he imagined, that he was hoping would solve that problem, was basically an artificial intelligence agent that would read the newspaper for you every morning and then clip out the articles that you would be most interested in.

Well, we realized that that technology was not going to be ready nearly in time to enable this new world of information exchange. And so we started thinking well, what technologies could make it possible? And we came up with the idea that one really powerful concept would be to use computers to aggregate the ideas, the thoughts, the values, the evaluations of humans - rather than try to use computers to directly make these decisions, we would use computers to leverage human decisions in making - in helping make the information decisions that they wanted.

There's a cute little story from Artificial Intelligence that I think really puts that in perspective. Some people say that Artificial Intelligence is the idea of having computers do badly what humans do well, and what we wanted to do was turn that on its head. Paul and I wanted to say well, what if we let humans make these value judgments that we humans are so good at and we just use computers to do the statistical analysis, an aggregation of all of those opinions, to then add value to other users? In a sense we were having computers do well what computers do well, and humans do well what humans do well.

**Talk about how you see the AI software environment evolving. I mean, is that where software ultimately is going to get to?**

Well, you know, over the long-term I have a very open mind about where AI is going and I just finished reading Ray Kurzweil's book, *The Singularity Is Near* about his vision of a world where the AIs are all smarter than we humans, and I'll tell you, I find that vision compelling. I think that really, human brains are limited by their biology and by the processes of evolution and ultimately we're going to get brains and silicon that are stronger, more powerful, bigger, better, faster than human brains.

The one question I have is when that ultimately is going to happen. And I'll say that in my view there's this terrible danger, which you can see throughout the history of AI, of always expecting it to be 10 years away.

And ever since I can remember in my life, 20 years as a computer scientist, is people have always been saying well, in just 10 years we're going to have that. Or some people say in just 20 years we're going to have computers smarter than humans, and as far as I can tell, we're not really making rapid progress towards that goal.

But there is another goal, in some ways technically less ambitious goal, but I think equally ambitious in terms of social impact, which is to say what we're going to try to do is build computers that are going to amplify the ability of humans or computer programs that are going to amplify the abilities of humans.

So, for instance, when people come to Amazon nowadays and they get all those cool recommendations for stuff to buy, they are getting exposed to a computer experience that is very much amplified over what they could do individually. But the way it's amplified is Amazon is collecting lots of information from people all over the world about what products they like to buy and is leveraging that information with some very clever computer algorithms to make suggestions about things that you might want to buy. That's a great example of a computer program that amplifies human abilities.

It's taking what we've got now, which is computers that can deal with terabytes of data reasonably, rapidly, can present us user interfaces that we can understand and take advantage of, but that certainly don't have the ability to do human value judgments or human understanding of documents, pictures, audio/video.

And yet humans are great at that stuff. We find it really easy to look at a movie and say whether we like it or not, and so collaborative filtering is the idea of taking all of that information and using some algorithm that some people would call artificial intelligence algorithms, some people would not, you know, there's this other danger in artificial intelligence, which is the field has been around for many decades now and it has done some tremendous advances in human understanding, but some people, every time we understand something they say, "Well, that can't be artificial intelligence, that's just a computer program."

And so I think there's this real danger of saying that anything we understand is not anything important and I reject that. I think that these contributions are - from artificial intelligence as a discipline - enormously valuable; they're just wonderful. And we should just accept them for what they are, which are great ways of making humans even more effective at the things that we try to do.

**You've been involved with recommendations as CTO of Net Perceptions, one of the first recommendation of companies. What are the issues around some of these new navigation techniques?**

Well, the thing that I think is really cool is that recommenders have gone beyond a technology that, when we founded Net Perceptions, we thought was mostly going to help people find information that they would find valuable. Frankly that was a failure for our company. We would go on a sales call and we'd say "Hey, I've got a technology that can double the number of times that a user will come back to your site because they're just going to love the information they find." And the guys on the site would say, "I absolutely believe you can do that, but I can't afford to double the volume of traffic on my site because all it'll help me do is lose money faster, because I'm not making any money yet from those page views."

And now we're in an era of the Internet where people have finally figured out how to monetize page views, I think led by Google; I mean Google has just been enormously successful at that.

Oh, by the way, to put in a plug for a different way of seeing Google, remember that Google is a company that has been one of the leaders of this idea of leveraging human abilities through computer algorithms. I mean PageRank, the heart of the Google search algorithm is fundamentally a relatively simple piece of computer science algorithm applied over all of the decisions of the millions of people who've contributed to making up the Web. Right? What Google looks for is which pages have the most links to them. Well, that's, in a sense, an expression by those people of their confidence in the site that they linked to. And that's why PageRank is such a wonderful algorithm.

**Well, they extrapolated human behavior and put it into an algorithm that could scale and provide in essence a metric for users.**

Exactly, and it's just an explosive breakthrough; I mean, it has really changed the way the whole world works. You think about the kind of information that you can just sit down and type into a Google search and, boom, you get an answer to something where 20 years ago it would have been literally hours in the library. When we look at the types of productivity improvements that we're seeing in information workers, I think one of the reasons is because of their amplification by these early AI technologies that have now been applied to the masses of data that are available on the Internet.

**You recently became an advisor to Aggregate Knowledge. What about their approach was interesting to you?**

Well, there are two things that Aggregate Knowledge has done that I think are really exciting. One of them is they've taken the basic recommender technology and they've built sort of a 2.0 version of that technology. They use all the state of the art stuff and just get it right, and they provide a way to deploy it that is much easier than the Web 1.0 approach. You just take a piece of Javascript and insert it in your Web site in the places where you want users to be able to see recommendations, and nearly instantly you're providing recommendations to your users. This recommender 2.0 approach is a complete turnaround from the old Web 1.0 days where it could take months to install an engine and integrate it with whatever Web technology you had on the site. It's very exciting to see the industry move in this direction.

I think the thing that's going on there is that Aggregate Knowledge is agnostic about how people use the recommendations; they see themselves as providing a state of the art, world class recommendation engine and then providing very simple APIs to let their customers leverage those engines in any way that they want.

**Users want navigation but they also want good search too. They're not mutually exclusive but they are separate in theory. So how does social networking and collaborative filtering fit into this in their approach that's different?**

That's a fascinating question. I mean, how do search and browse relate? I would say that in general we're in an era in which search has just dominated. Google has just turned search into the dominant paradigm in the information world. I do think that we're going to see a rebalancing of that over time. I don't know that I'm right about that; that's speculation, but I think one of the things – whether or not I'm right about the rebalancing towards browse or not – one of the things that I'm certain of is that we're going to see increasingly that search that is only information-based, that's only based on things like the keywords that are in the documents, is going to be a failure in the Web 2.0 world.

And the reason for that is that people need searches that also have some concept of equality or the fit of the documents that are returned. And that's the cool thing about the Aggregate Knowledge technology is that you can implement that in a website where somebody can do a search and get back a set of results. Then by just plugging this little JavaScript widget into the result list the website can get a set of scores presented to the user in innovative interfaces that let users look at search results not just by which searches were relevant to the query but which searches are personally relevant to them as an individual user of the website.

**But at the end of the day, for the people who have these big sites who want to take advantage of these kind of technologies, it's a deployment issue. I mean didn't that cause a lot of the early firms to kind of not scale?**

Yeah, that's exactly right. I mean one of the challenges in that perception in the early Net Perceptions days was that we could go out to a company and make the case that we could literally make them millions of dollars with our recommendations over the technologies that they were using in-house. And yet we still couldn't convince them in some cases to do a deployment because it would take months of work from their IT department to do their deployment and the thing that's really exciting – you know, Web 2.0 has sort of two characters to it and one of the characters we've talked about already – socialization of the Web.

And the other character is this deep use of simple technologies, JavaScript and Ajax, to make it really easy to build interfaces that are much more interactive than the original Web 1.0 interfaces. And that's the thing that I think is very exciting about the recommendation 2.0 world is that you've gotten to a place where in order to make recommendations into a website; all you have to do is build a dictionary of the products that you have in your catalog, and then export that dictionary, and then build a simple JavaScript widget into the pages where you want the recommendations to appear, and voila you're done; you've got everything you needed.

**It pushes the complexity really into the cloud or into the network where any technology you put big iron and clusters of servers out there and Google is doing this now with this big kind of massive computing cloud and Amazon has their utility model, S3, EC and now payment so the trend is that the heavy lifting gets done on the network, not on the sites...**

Exactly, it's like in the early days if you wanted to run an Internet business you had to figure out how you were going to have a dozen servers all around the country and then Akamai showed up and they just commoditized that problem. All you had to do was pop up with your server and your service and sign up with Akamai and off you went. That's what Aggregate Knowledge is trying to do now for recommendations. They're saying hey, you want to do recommendations, great, that is our competency. Why don't you outsource that and put your effort into building the best products and services and user experience that you possibly can?

**So final question: what are some of the current projects that you're working on at GroupLens that have you most excited?**

Well, one of the things I'm really interested in is some explorations that one of my Ph.D. students, Shilad Sen has been leading where he's been exploring the impacts of tagging on how users behave on a website. And one of the things he's looking at particularly is how recommenders interact with tagging systems.

So the question is if you have a typical user tagging system, people just start applying tags to the site kind of randomly then over time you'll see the vocabulary that's being used on the site diverge and become less and less relevant to the users of the site.

Well, one of the ideas is that maybe a recommender could detect that, it could sort of watch the vocabulary as it's emerging and it could make sure that the tagging system itself applied some pressure to tagging users to try to encourage them to use tags that were the most valuable tags for those users.

**So is it building taxonomy?**

Exactly. What it's doing is it's watching the users tag and it's trying to understand the emerging taxonomy of the tags and then sort of encourage other users by the recommendations that it makes to fit their tags into that taxonomy. Now it never requires it - like all tagging systems, any user can apply any tag they want. I remember I saw a tag a little while ago, my favorite tag ever, it was "great music to listen to while drinking tequilas and driving in a convertible to Mexico."

What a tag! I mean how much more precise do you get? So you could always add that tag if you want, but what the recommender does is it sort of encourages you to use tags that are a better fit for the community.

# Recommendations 2.0

John Riedl, Ph.D.

Discovery  
Expert  
Series

**So what it really is doing is in essence setting a linguistic architecture so that when people come in, they can be more specific by having different diverse tag sets out there around cluster and content.**

Exactly.

**That's amazing. And then that scales with the machine learning; essentially it's the machine-learning environment, right?**

Exactly, it scales with the people and with the machine learning algorithms that you use.

**Wow that's exciting. Anything coming onto the network that we can play with at all? Or people out there who were interested?**

That's right, yeah; we'd love to have you come visit our research site. This is totally not for profit, just a National Science Foundation supported research site; it's called [www.movielens.org](http://www.movielens.org) and please come. You can see our tagging features, you can see a bunch of other community features that we're exploring. And if any of your listeners would like to see the published papers that underlie our work, they are all available on our website at [www.grouplens.org](http://www.grouplens.org).

**John thanks so much for the chat. I've always loved talking about computer science and some of the innovations out there and Web 2.0, modern web and recommendations, and group theory, group research, and group algorithms - great stuff. Thanks so much for taking the time.**

Well, thank you. I couldn't agree more. It's just a wonderful time to be alive.

aggregate  
knowledge

1510 Fashion Island Blvd., Suite 201  
San Mateo, CA 94404  
Tel 650.293.2920 | Fax 650.293.2919  
[www.aggregateknowledge.com](http://www.aggregateknowledge.com)

Copyright ©2007 Aggregate Knowledge, Inc. All rights reserved. Aggregate Knowledge, the Aggregate Knowledge logo, Pique, the Pique logo, Pique Discovery, Pique Onsite, Pique Email, Pique Affiliate, Pique Multi-Site, Pique Discovery Window, and Pique Discovery Igniter are trademarks of Aggregate Knowledge Corporation in the United States, other countries, or both. All other company, product, or service names may be trademarks or service marks of their respective owners. All statements regarding plans, directions, and intent of Aggregate Knowledge are subject to change without notice. DESC-1107-01