

SEARCHER

The Magazine for Database Professionals

Is the Price of
Cheap Chicken
Bird Flu?
page 8

Social
Bookmarking
page 26

A Failure to Communicate

Politics, Scams, and Information Flow
During Hurricane Katrina

page 40



by Jennie Starr
Consultant

Pandora: Finding New Music When You Have No Time to Hit the Clubs

I love to learn about new music, but, with two small children, my lifestyle doesn't permit much time to hit the clubs. The CD players in both car and home are filled with "Baby Baluga" and variations on that theme. I have to concentrate while I work, so I don't listen to music while writing. I don't have the time or patience to download music, rip CDs, create playlists, etc. But, I admit to owning and using an iPod. I use it only because my husband bought it for me, added the songs I like, and created the playlists for me. (And, nothing makes running more fun in sunny San Diego than some great music to listen to.) So it was a pleasant surprise to find myself, while reviewing Pandora, not only working for an hour-and-a-half on my PC listening to music, but checking back to the Pandora player on several occasions to see who was singing. What a great way to learn about new songs and artists. My favorites list on Pandora has grown over time and I'm thrilled with all the new music in my collection. This was a first for me.

Many companies have built products designed to help the music lover find new music. The products break down into two camps. Collaborative filtering applications have at their core social recommendations (what others around you like), relying on an artificial intelligence application to mine the music listening data in or-

der to make the connections. Most of these products leverage either what people listen to or what people buy to recommend music to others. The theory being, if you like an artist that a lot of other people also like, then you might like the other music those same people listen to or buy.

Alternatively, other applications have at the core the classification of the music itself, its inherent structure. These applications use humans to classify the traits of the music and enter them into a database. That data is then used to make recommendations that match the specific traits of artists or songs you enter to others in the database.

Collaborative Filtering Versus Classification

I should disclose that I've had some experience with digital music applications beyond casual use. In 2000, I was a product manager with MusicMatch [<http://www.musicmatch.com>], now a Yahoo! company that designed a digital music player and a music recommendation engine built on collaborative filtering. MusicMatch introduced its users to new artists and songs through a Music Guide and later radio stations. MusicMatch listens (if you let them — opt in or opt out) to the music you listen to in its digital music player or "Jukebox" and uploads a file that records your history

into its database. MusicMatch then leverages that information to recommend new music to you. We're not talking about a small sample size of listeners, either. According to its Web site, the engine is "fueled by the music tastes of more than 30 million MusicMatch Jukebox users who voluntarily share their listening habits with us. These music fans play more than 1 billion songs each month and we use this information to understand the relationships between thousands of artists around the world as well as the daily popularity level of hundreds of thousands of songs in the music universe. Think of it as the library of all the world's music organized by 30 million librarians."

The MusicMatch method always impressed me. Relying on the actions of music lovers to fuel recommendations always seemed to me an honest way to try and share music — albeit a little bit Big Brotherish. It seemed an infinitely better method than recommending what other people have purchased. But I do recall the founders from Savage Beast (now Pandora) coming by to show us their product during my tenure at MusicMatch. At the heart of their efforts was music classification. I have to admit that approach appealed to the librarian in me. Instead of trying to track behavior and serve up recommendations, Pandora's approach tried to really understand the nature of the music

and to use those attributes to recommend new songs and artists.

Pandora at a Glance

It was a thrill to see what Pandora has done since its Savage Beast days. Using Pandora didn't pose any difficulties. Within a few minutes, I dis-

A new song and artist I discovered on this station is "Sun" by Lisa Hayes & the Violets. Looking at the attributes for the song, you see that Pandora chose it because these artists appeared similar in the following ways: a subtle use of vocal harmony, mixed acoustic and electric instrumentation, a vocal-centric aesthetic, major

degrees, leveraging different types of data to drive the engine. Back when we were "Savage Beast," we spent a lot of time explaining how our product was not a collaborative filter. It was a challenge to get people to realize that what we did was very different.

So, people do want something better and there's a big need for it. People are not very satisfied by how recommendation engines work now. There's a problem with the industry: Collaborative filtering will always have a hard time helping people discover music that's not well-known, because the filtering is essentially a popularity contest. A core element for enduring value is music discovery. Most people tire of services that don't introduce them to new music. That's one of the reasons commercial radio has a bad rap. Now the [Pandora] service is out; we launched in the beginning of November 2005.

We also fit into the music services space. Until now, most existing services have had digital warehouses. Customers can get any of the music on demand. Most products have focused on access offering features and administrative services where you can create and manage your playlist, listen to whole albums, etc. It took a while to make digital music warehouses available, as companies had to work on both infrastructure and licensing. Now it's kind of a commodity. It was an arms race for a while — who has more songs. Now the companies have collaborative filtering engines. They have licenses to do on-demand stuff. But these services were designed for people who have time to use them. The average person with jobs and other responsibilities could not make the time for the detailed, time-consuming process required to analyze new music. It's that work that allows Pandora to make it simple for the end user. The other services chased the blogger, the guy who wanted lots of

Pandora will let you listen to five songs at no charge and then ask you to register. Registration requires an e-mail address but not much more.

covered some new artists I liked. The interface is extremely easy-to-use. Yet, the work behind why it works will knock your socks off. Musicians at Pandora have analyzed the songs and applied over 400 attributes to each one. When you enter an artist or song, Pandora creates a station that feeds you new music with similar attributes. Go figure. It works.

Try the site out for free. Pandora will let you listen to five songs at no charge and then ask you to register. Registration requires an e-mail address but not much more.

Create a new station by entering an artist or song. Use the "Guide Us" button to let Pandora know if you hate a song and don't want it on the station or like the song and want more. The "Guide Us" button also permits you to see why the song is on the station, i.e., the logic behind the comparison to the song you used to create the station.

The radio stations function as most do, permitting you to skip ahead to the next song, buy the song from iTunes or Amazon, add the song to a favorites page, and, of course, give Pandora feedback telling them if you like or dislike it.

The station I'm listening to the most is based on Natalie Merchant.

key tonality, and electric rhythm guitars. I have to say, while I would have had a hard time articulating the types of songs I enjoy, these attributes feel right.

The free version will be supported by ads, so enjoy this short window before Pandora starts inserting the ads. If you want to avoid the ads, a subscription will run you \$36/year (\$3/month) or \$12 for 3 months (\$4/month).

Talking to Pandora's Founder

I interviewed Tim Westergren, the founder of Pandora, about the company, the state of the music industry today regarding music recommendations, and where the company hopes to go.

Pandora offers music recommendations. But, if you were to define where it falls in the space of music products, where does it fit and how do you distinguish it from other similar products?

Pandora fits into the recommendation space, which until now has been dominated by collaborative filtering products in one flavor or another. That's what everyone does. Companies have refined it to differ-

features. Pandora makes it simple for the customer, but offers some features for the music lovers with more time who want more control.

Tell us about the feedback you've received since the service launched.

We're getting hundreds of e-mails giving us feedback. The most common piece of positive feedback is, "I'm finding new stuff I've never heard before and it has been years since that's happened." For these customers, the experience is qualitatively different.

What kind of feedback are you getting regarding your work with the attributes? Are you getting requests for the ability to have more control than you currently provide?

We do get a lot of feedback about the attributes. It really varies. There are some customers who use it that don't understand how it works. Other customers are very interested in it and they have faith in it. Those customers find it educational and want to know more. One thing we'll be doing soon is to allow people to tweak the playlist based on attributes. I like this "voice" rather than this "song." People do have the interest if its presented in a digestible way. We believe we can add this kind of functionality and keep it simple for both types of users; those that want to keep it simple and the power user as well.

Can you tell us more about your content choices? How do you determine what music to include? What's next in terms of music content?

With regard to the development of our database, there's work associated with obtaining the catalogs for the music from the four majors and satellite indie labels and then going directly to some other labels and artists to get licenses. We went through the Billboard charts back in the '50s and got everything since then for the database. Various independent charts went back to their beginnings as well. We go through

Other Music Recommendation Engines

The major music sales sites offer their own recommendation engines: iTunes calls its "Just For You" and it bases its recommendations on past purchases. Amazon offers recommendations too based on items you own.

There is no shortage of independent recommendation engines with their own twist. The difference between them in most cases is how or where they gather their data. I've listed several additional engines below:

Upto11

<http://www.upto11.net/>

Leverages hundreds of thousands of P2P network user song file collections. When you type in the name of a band or bands and conduct a search at its site, Upto11 searches a subset of those collections, the ones that have the bands that you searched for, and delivers back the names of other artists most likely to appear in those user collections.

Goombah

<http://www.goombah.com/>

Based on your iTunes music and listening behavior, Goombah displays like-minded members' collections and recommends the favorite tracks that you don't have. "We believe that the music is about the people that listen to it, not the science of the song."

Last.fm

<http://www.last.fm>

Downloads a software application which uploads the records of music you play in whatever player you use. They use those records then to compare across the database of listening behavior and make recommendations.

ChoiceStream

<http://www.choicestream.com>

ChoiceStream powers such services as eMusic, Yahoo! Movies, and AOL to make recommendations for customers. This service is the most like Pandora in its methods. It has a software application that classifies music by one of 25 attributes such as "macho" or "romantic" and has eight editors who review or QA the process.

And Name That Song

In researching this column, I ran across some unusual sites that I consider worth sharing. In no apparent order, and without any particular relevance to this column I give you:

Wildbits

<http://www.wildbits.com/tunatic/>

This site will help you identify a song's name. Simply download a small software application from the site, use a microphone to sing the song for the software, and it will display the title for you.

SongTapper

<http://www.songtapper.com/s/tappingmain.bin>

This one is great fun. Simply use your space bar to tap in the rhythm of the lyrics and it will display the song for you.

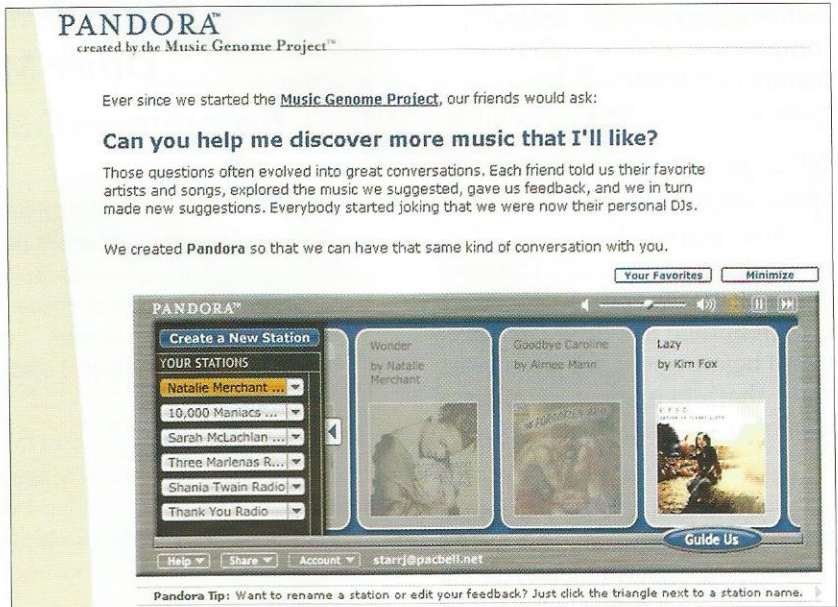
independent labels, anyone with a reputation for good music. More and more, we are also getting music sent to us. We've had 150,000 requests for new music, submissions, sent to us since November 2005.

So how did you build the product? As you get music, how does the process work to assign the attributes and add them to the database?

We built the product from scratch. We began by building the genome (or template), which has approximately 400 musical attributes. These attributes are the basic building blocks of a song. They include details from melody, harmony, rhythm, instrumentation, and lyrical content. One way to think of it is primary colors. There are over 30 attributes describing sound of voice. As much as a person has height, weight, etc., every song has these kinds of shapes and sizes. And all these elements are related. For 6 years now, we've been analyzing these songs. It takes up to 20–30 minutes to analyze a song. We have 38 musician analysts and two people who work solely on acquiring music. The analysts are trained professionals who generally have a 4-year degree in music and they get 40 hours of training before they start. The analysts do 10 percent of the songs twice, to get consistent data. We also have a steady Q&A process. Occasionally, there are debates among the analysts about attribute assignment and these are considered a good thing, as we want to obtain a consensus.

Let's talk about how consumers search for music. Do you have any statistics on what they are doing? How they are using the product?

There's a preference to use Artist to launch stations. But I think that's changing. Artists are usually what they can think of first. But an artist station gets more diffused in terms of results. If you type in an artist, it is as if you've entered all of his songs and then the station is built on all of that. The music will traverse his



Home page of Pandora

repertoire. You can thumbs up or thumbs down based on your reaction to any song. People are learning that using a song to launch a station produces a better result.

About 10 percent of the customers click on the “Buy” button. Anecdotally, a lot of musicians, independent ones, are experiencing an effect from having their music in Pandora. They've noticed that people come to their Web site and, until now, there's been no mechanism for these people [to be heard]. If I was a betting man, something like 70 percent [of the music we have] is what you'd call obscure. If I was to distill this company down into a couple of kernels, the feeling is what you find when you discover a new piece of music. For most people, it's a profound experience. People crave that. As soon as they do it, they want to buy it, own it, and share it. Right now, we lose a lot of sales because no one — not iTunes nor Amazon — has the music. And we're not linking to the individual Web sites at this time, either.

Tell us about any new services or features you have planned.

In terms of partnerships, last week we came out with a slim server to stream the music through your home system. Generally speaking, for us, two big areas are mobility and international [expansion]. We want it [Pandora] to be a global product, not just an English-speaking product. We already have a world genome put together. So now we have to deal with agreements with music labels. The labels have been really supportive of us. They're sending us music and it's working well. We are certainly interested in more partnerships. We are entertaining the idea of working with destination companies that have big user bases, for example, for whom Pandora would be an interesting addition.

Are you interested in selling the data that you are gathering from your customers about how they listen to, buy, and select new music?

I don't really believe in this idea. There are a lot of companies that

gather data that might be of interest to labels, but it doesn't translate into a good business.

Tell us about your background, why you founded Pandora, and how long it took from idea to launch.

We'd been chasing the business-to-business model for a few years and it's a hard business. There is something of a glass ceiling when you license technology. There aren't that many companies that will pay you a lot of money for it. So, we had really been in survival mode for a lot of years. We managed to finally raise a round of financing that allowed us to figure out where to go with the company. In the end, a bunch of things conspired: Broadband was much more ubiquitous, online radio was more popular, and radio was the best use of what we do. Because

what we do is at a song level, playlist generation plays into what we do. Our research told us that music is about listening. He who owns the listener, owns the business.

How are you funded?

We raised about \$8 million in March. We brought on Joe Kennedy, a new CEO with a strong consumer marketing background. It's been almost 2 years now. A year after that we raised another 12 million from our own investors and Venture Capital. Our second round of financing happened with our launch, so now people are very bullish about the company.

Have you looked at your competition? How is your company different?

There is no one who is doing this as far as we know. Eventually, there

will be all sorts of people doing this, but I think one of the really distinctive characteristics is you can't do this quickly. It's not a technology thing; the technology part is a commodity. There is no way to do that quickly. Machines can't do it. Collaborative filtering can't do it. When we signed a deal with AOL, a while ago, the first thing they did was put pen to paper to see how much it would cost to do it. Defensibility is the single most important thing for a startup. You can't just throw money at it. ♦

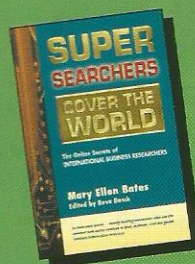
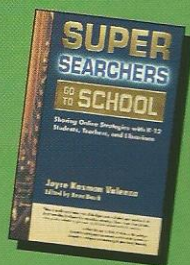
Jennie Starr [jennie@jenniestarr.com] is a consultant and lawyer specializing in software product marketing and a freelance writer living in San Diego, Calif. She also edits KMTool [http://www.kmtool.net], a knowledge management resource she launched in 1999.

SEARCH ONLINE LIKE THE PROS!

Super Searchers Go to School

Sharing Online Strategies
with K-12 Students, Teachers,
and Librarians

by Joyce Kasman Valenza
272 pages/ISBN 0-910965-70-6
\$24.95



Super Searchers Cover the World

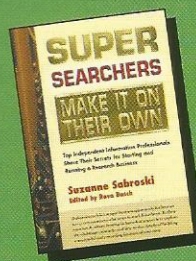
The Online Secrets of International
Business Researchers

by Mary Ellen Bates
290 pp/ISBN 0-910965-54-4
\$24.95

Super Searchers on Competitive Intelligence

The Online and Offline Secrets
of Top CI Researchers

by Margaret Metcalf Carr
352 pages/ISBN 0-910965-64-1
\$24.95



Super Searchers Make It on Their Own

Top Independent Information
Professionals Share Their Secrets
for Starting and Running
a Research Business

by Suzanne Sabroski
336 pages/ISBN: 0-910965-59-5
\$24.95

VISIT YOUR LOCAL BOOKSTORE OR
ORDER DIRECT FROM THE PUBLISHER

www.infotoday.com

For more information, call (800) 300-9868; outside the U.S. call (609) 654-6266; Write to: Information Today, Inc., 143 Old Marilton Pike, Medford, NJ 08055