



- [Top 100](#)
- [News](#)
- [Reviews](#)
- [Debates](#)
- [CEOs](#)
- [Forum](#)
- [Verticals](#)
- [Job Board](#)

[Subscribe to RSS](#) [Email Updates](#)

« [Gimme gimme gimme a search engine for Karaoke!](#)

[MultiSearch Engine Limmz](#) »

## An interview with Kevin Chang, co-founder of Cazoodle



An interview by Hope Leman.

Hi, Kevin. Before we start, just a bit of background to put this interview into context for our readers. I recently attended the Web 2.0 Expo in San Francisco. I wandered around the Expo Hall and came across [Cazoodle's](#) booth there. I was given a demonstration of both the apartment search and product search. I was quite impressed by the swiftness and incredible detail provided by both.

When I got home I watched the very useful screencast about Cazoodle here:

<http://www.cazoodle.com/about.php>

I recommend that anyone new to the term “vertical search” watch that screencast.

Now, on to our interview.

First of all, one reaction I had, is impressed as I was, I am not in the market for an apartment. But there is way more to Cazoodle than that, right? The apartment search engine is just the first of what will become a series of search products, correct?

Yes, apartment search is where we started; we plan to develop a series of vertical search products, with electronics shopping being the next.

Since Cazoodle came from our research work on “data-aware” search at the University of Illinois, we hope to bring these enabling technologies to build vertical search for the domains where deep search on structured data—such as finding apartments by addresses—is desired but is difficult to achieve.

How did you hit upon apartment search as your first product?

Cazoodle was started by a bunch of graduate students and me, so you could imagine what we got when I asked the students to brainstorm when we were looking for “driving applications.” They came up with two areas—apartments and used cars. And, you know why! We eventually picked apartments according to their ranking of

pain. Even in our small town, it was a major pain to have to look everywhere online for finding apartments to check out. There were many useful sources, from Craigslist, Rent.com, to local apartment owners' sites—just too many! In the old days, our mom and dad had to drive everywhere in town to look for rental signs; now we have the Internet but we had to surf everywhere online for property listings. We thought our technologies could do something useful for finding apartments.

As another major reason, I also had experience, in my faculty consulting work, with a major real estate service. They wanted to expand their good coverage from properties for sale to those for rent—which turned out to be so different and so difficult. In for-sale real estate, you get the MLS database from the National Association of Realtors. However, the for-rent market is rather fragmented and MLS does not cover very much. The company I did consulting work with was struggling to build an organic search for apartment properties, since there was no MLS-equivalent for rentals. So I knew since long ago that the industry did have this on their wish list.

How does apartment search showcase the potential for your products?

In hindsight, that was a very good choice, because apartment is a tough domain to deal with. For building structured search, the challenge really boils down to crawling, indexing, and integrating data from thousands of or even more sources. Apartment search very well represents these key barriers—the domain is very fragmented, with not only some large posting services but also numerous small landlord websites. To be both comprehensive and precise, we needed to tackle crawling and integration of data from numerous data sources at the order of hundreds of thousands. Large scale integration is what my research group has focused on, and we certainly have an edge here.

Furthermore, it was also good to pick a domain where the industry had encountered a bottleneck to breakthrough. While paid ads/posting services for apartments have been in existence since the dawn of the Web age, only until recently we started to see some organic crawl-based searches—but of rather small data coverage. As I said earlier, I witnessed stalled attempts of leading real estate companies. In such a domain, our potential would be clearly testified as we become the first and only comprehensive apartment search to cover data sources of all sizes.

Tell us about your plans for branding. Will you have, say, Cazoodle Drug Search, Cazoodle Car Search, and Cazoodle Book Search? You already have a Cazoodle for Electronics, for example.

Yes, that's right. We hope to deploy the same technology in many new domains where vertical search is hard to build. Our strength in data harvesting can quite generally extend to various application areas.

However, to leverage our data harvesting capability, we also hope to serve as a data provider to collaborate with service providers in each domain. With our limited manpower, we could not develop all these many verticals so quickly. So we are always talking with partners to power up their vertical applications using our data harvesting and integration technology.

One of the things I liked about the screencast was the obvious enthusiasm of the speaker and the pride he takes in Cazoodle. Is that because he is a student and excited to be creating a real-world tool—something so few students get the opportunity to do?

The best part of being a faculty at UIUC is that I have the privilege of working with very strong students. They are so full of youth and, of course, enthusiasm as you noticed. Some of my previous students are now leading Cazoodle, and I am witnessing their "fire" from the great results they are producing, including this screencast. Govind Kabra, our CTO, led the production, and Jason Hertenstein contributed his professional voice.



Can you tell us about EnterpriseWorks, the incubator facility University of Illinois at Urbana-Champaign?

After the little fight with Netscape, the university seemed to have realized the potential of its strong engineering school, and has since then built up quite complete supports for faculty startups: the three arms of incubator, technology transfer, and venture capital investment.

The EnterpriseWorks is the incubator arm and is truly the best gift I could imagine from the University for doing startup! It provides not only hard-walled office space but also much soft-minded mentoring of entrepreneurship. Since we started in the building in January 2007, as the only “scale intensive” internet company there, we have added much new challenge to the administration, such as accommodating our data center. They have always been open-minded to learn new stuff with us!

Are you unique in enabling students to develop commercial search products? Do you know of any other programs like yours?

I am certainly not the only one doing this. Many know that the story of search, if search equals Google, has a lot to do with academic research—and I was particularly inspired because when I was a graduate student at Stanford, I had the pleasure to watch Google growing around us. We were in the same Digital Libraries project lead by Professor Hector Garcia-Molina, who was my advisor. When I was doing circuit interviewing for academic positions, one of the most asked interview question was—could you imagine—“why don’t you join Google?”

And I think I had a good answer. I was so inspired by how technologies—through startup—can change the way people live in a good way. I thought being a professor in academic research, if done right, could be the best way towards this goal. Why? Everyone doing startup would tell you recruiting good guts to join your dream is every bit as hard as coming up the dream itself. But no worry—As a professor, your research will develop not only technologies but also students! That’s exactly how I started Cazoodle, with the students and technologies developed at the university.

What are the majors of the students involved? Computer science? Electrical engineering? Do you recruit kids from non-techie department like marketing?

I am in the computer science department so, in research and teaching, I mostly work with CS students, with occasional undergraduates from electrical engineering. However, at Cazoodle, we have quite a few kids of various backgrounds in our data engineering process—our best performer was a music major specializing in trombone. We just wanted to make our data harvesting something kids can do for fun (or for an allowance). In addition, we also work with IBC, a group of part-time business student consultants at Illinois, for helping our marketing.

How do you recruit your students and how do you all work together to design marketable products?

Students actually like to work on product-able ideas that may have a market—you could imagine that, given we are in the department where Netscape, PayPal, YouTube, and Yelp came from. Even when writing academic papers, you always have to explain why the stuff is useful.

To this end, recruiting is not too hard. We simply demo what we are working on and what we have worked out. My research group holds “open house” to give demos. My research talks would always include demos. I recruited many graduate and undergraduate students with a “making-real-things” mindset this way.

What I found harder is to get students to work on search in particular. They would tell you Google has done it. I had to convince a new graduate student that her idea can go much beyond what Google is. But as soon as students are motivated, they can do wonder!

Who came up with the idea of Cazoodle? Do students get a royalty for their contributions?

The idea of Cazoodle came from my research focus: integrating and search structured data all over the Web. We started the [MetaQuerier](#) project exploring the “Deep Web,” where we developed techniques for integrating numerous online databases. Then, we continued with the [WISDM](#) project exploring structured data embedded in the vast amount of “Surface Web” text. In all these efforts, I worked with many students and some of them graduated at the right time and we co-founded Cazoodle together.

All inventions at a university belong to the university and the co-inventors. So students do share whatever rights associated with the inventions.

As a woman, I am eager for women to get more into technology. Are any of your students women?

Sure. Although Computer Science is not a department with many women, I do work with a few female students at any time. In fact, two of our co-founding students, Dr. Zhen Zhang and Huiting Yang, are women. Of course, quite some of our key developers are female.

As someone who works in a center for health research and quality, I immediately began to wonder how Cazoodle’s technologies could be used in healthcare. Do you envision it being used to search for consumer health products such as non-prescription drugs and medical supplies? I find, for instance, that there are many fascinating, worthwhile assistive technologies for the disabled but there don’t seem to be easy ways to find them. I often read about such technologies on message boards at social network sites like Patients Like Me. Could you assign a brilliant student to render such data as easily findable as you have managed to make apartment listings?

Yes, we are hoping to expand to other domains beyond current products. The healthcare domain could certainly use data aware search to make data more findable. At this point, we are focusing on apartments, shopping (electronics), and local events.

I try to list as many grants, awards and scholarships in the health sciences as I can find on the site I help maintain, ScanGrants. I find such data on the Web via hours of searching and enter them into ScanGrants one by one. When I saw Cazoodle demonstrated, I was envious and wished could that with grants and scholarships! What is your personal dream application of Cazoodle’s powerful technology?

First, as a professor I have to look for grants all the time—I think you are certainly right in how we really can make grant finding a much less laborious process. In fact, we are helping the [IRIS](#) (Illinois Researcher Information Service) service, a grant search service from our university, to automate their data indexing process.

Second, dream application? As Cazoodle has the technologies to understand structured data everywhere, in a large scale, I am hoping we can bring search to everywhere users need it, rather than requiring users to drop their task at hands and come to Cazoodle to search. Can we develop search technologies that can, when users are looking at something in the browser, understand what users are looking at and bring back relevant information automatically, without explicit search? I think, when search can go invisible and ubiquitous, it will open up a whole range of new applications.

Also, you mention Deep Web searching. Are there plans to apply your technologies to such Deep Web materials as chemical compounds and genes?

Yes, we are very interested in making science more findable, but at this point we need to stay focused on what we can do with a small group of us. We do look for partnership to co-develop applications for novel domains like science!

Tell us about the Deep Web. How is Cazoodle different from the search products of Deep Web Technologies, for instance?

As mentioned earlier, we started with the MetaQuerier project, one of the earliest efforts in exploring and integrating the numerous databases online, or the so-called “Deep Web” because their contents could not be easily reached by a typical crawler that only follows static links.

What we focused on, and may have excelled, are the set of technologies to automatically discover such databases for a particular purpose, to comprehensively crawl, extract, and index their data, and to precisely integrate and rank them for search.

What we differ from most other companies is the capability to do this in a large scale, with an automatic approach. For instance, as far as I know, most state of the art deep-Web technologies focus on federated “meta-search” with built-in interoperability protocols among participants, which has been effective in close-knit communities. In contrast, we hope to deploy to open, consumer-oriented domains like apartments and shopping.

I am in library school and am struck by how little discussion there is about search technology. By contrast, social networking is all the rage. How do you get kids interested in search? And how do you foresee search engines dealing with the fact that so much information is now being generated in member-only sites like Sermo, in microcontent fora like Twitter and via Open Science?

Since I am in computer science it is a little easier since kids are more into technologies to begin with. Many of my colleagues in CS agree that search is just a perfect playground for various CS fields—databases, artificial intelligence, text retrieval, high-performance computing, to name a few. As I said earlier, it is not hard to get students interested in search, it is harder to establish their confidence that they can do something Google has not covered well.

The new trends such as Twitter simply highlights that search will never be a solved problem. People not only consume information but of course also produce information. Search will have to tackle with new ways information is produced and the forms it exists. Twitter for example shows that information could be generated by everyone at real time. Search must always progress along with the course that our civilization evolves.

Can you describe for us the design process for a search engine from inception to launch? How do you decide what data to search? Do you have to get permission from each database? Any headaches about copyright?

I will focus on vertical search as we have been doing. There appear to be a generic 3-step formula for the process: discover, crawl-and-index, integrate-and-rank, to finally deliver search to users. Think apartments. First, we must discover, among the billions out there, which websites provide apartment listings. Second, we must crawl and index every listing from each site. Finally, we must integrate repeated listings, rank, and show them on a map.

Our crawler does follow the common robot crawling protocol and we always work hard to be a good Internet citizen with polite behavior. We always respond to crawling issues quickly.

We did run into some permission issues—mostly from site owners who did not want to see their data listed in the way we showed them, or who syndicated and did not actually own data on their sites to further disseminate. We would simply explain and honor their requests. On the other hand, we more often get requests from websites asking how their data could be included.

Where might readers see Cazoodle demonstrated this year?

We are now promoting our Apartment Search product in the rental industry tradeshows, like the NAA (National Apartment Association) conference in June. In addition, we will also be at various technology expos such as Web 2.0 conferences later this year.

Thank you for your time.

Thanks for the opportunity to speak up!

## [Scholarship in Aust](#)

Get Your Guaranteed Place in Uni. Scholarship Offers.  
Enquire Now!  
[TaylorsCollege.edu.au](http://TaylorsCollege.edu.au)

## [Corporate Accommodations](#)

Furnished and Fully Equipped Short Term Apartments in  
Pittsburgh, PA.  
[www.corpaccom.com](http://www.corpaccom.com)

## [Apartments in Milan](#)

Fully Furnished Apartments in Milan For a Short or  
Medium Stay in Italy  
[www.Rentok.it/Apartment](http://www.Rentok.it/Apartment)


## [Landmark Asia Realty](#)

Online Visual Sales/Rental Property Listings with Photos.  
Search NOW!  
[www.landmarkasia.com.hk](http://www.landmarkasia.com.hk)

Ads by Google

This entry was posted on Thursday, June 11th, 2009 at 8:48 am and is filed under [CEO Views](#). You can follow any responses to this entry through the [RSS 2.0](#) feed. You can [leave a response](#), or [trackback](#) from your own site.

## One Response to “An interview with Kevin Chang, co-founder of Cazoodle”

-  [Nancy Leman](#) Says:  
[June 11th, 2009 at 11:50 am](#)

My reaction: Excellent questions, with just as excellent responses. An interviewer has to know something (at least) about the field of interest to get the best information. So what is the interviewer’s background in searching?

And how did she record the answers? Quite likely, she didn’t write or keyboard them down as the interviewee was speaking. Or did she?

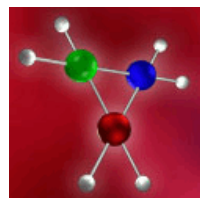
## Leave a Reply

Name (required)

Mail (will not be published) (required)

Website

## • Our Sponsors



[Sponsor this site](#)

## • Search AltSearchEngines