


How can we spark the movement of research out of the Ivory Tower and into production?

Thursday, July 22, 2010 at 7:22AM

Todd Hoff

Over the years I've read a lot of research papers looking for better ways of doing things. Sometimes I find ideas I can use, but more often than not I come up empty. The problem is there are very few good papers. And by good I mean: can a reasonably intelligent person read a paper and turn it into something useful? 

Now, clearly I'm not an academic and clearly I'm no genius, I'm just an everyday programmer searching for leverage, and as a common specimen of the species I've often thought how much better our industry would be if we could simply move research from academia into production with some sort of self-conscious professionalism. Currently the process is horribly hit or miss. And this problem extends equally to companies with research divisions that often do very little to help front-line developers succeed.

How many ideas break out of academia into industry in computer science? We have many brilliant examples: encryption, microprocessors, compression, transactions, distributed file systems, vector clocks, gossip protocols, MapReduce, search, algorithms, networking, communication, and on ad infinitum. For every Google that breaks out there must be thousands of other potential ideas that go nowhere, even in this hyper-VC aware age.

We need to do a better job of using research. There's a lot out there in the literature that we could be making use of right now, but it's closed off from the people, i.e., developers, who can turn this research into gold.

And it's largely closed off because researchers don't consider developers as an audience and they don't write their papers with the intention of being applied. Change the publication process and we can save the cheerleader and save the world.

I'm bringing this up now because:

It's been an issue I've thought about for a long while, but I never thought there was much that could be done about it.

I attended a talk by [Daria Mochly-Rosen](#), PhD at Stanford titled: [How Are Drugs Developed and Why Is It So Expensive?](#) Her talk made me think maybe there's something we in computer science could do about this problem because they are already doing something about a similar problem in the drug discovery space.

The Gap Between Research and Development is More Like an Ocean

Dr. Mochly-Rosen is Senior Associate Dean for Research and George D. Smith Professor in Translational Medicine, School of Medicine. Her talk was fascinating. She walked us through the incredible gauntlet every drug must run before being approved for human use. She also talked about a project called [SPARK](#):

SPARK was created to advance promising discoveries from the research laboratory into medical practice. SPARK's mission is to facilitate collaboration between academia and industry by providing Stanford researchers with the support, knowledge, and partnerships that can bridge the gap between discovery and its application into important new medical therapies. SPARK was created to advance

promising discoveries from the research laboratory into medical practice. SPARK's mission is to facilitate collaboration between academia and industry Stanford researchers with the support, knowledge, that can bridge the gap between discovery and its application into important new medical therapies.

I thought the idea of SPARK might also work for the computer industry.

My thoughts on the state of research versus industry were reinforced after listening to Barbara Liskov during her [Turing Award Speech](#). In her speech she recounted what to her was an amusing little anecdote of how she discovered that one of the key rules of OO modeling, the [Liskov Substitution Principle](#) (LSP), came to be named after her. If you haven't read a design patterns book in a while, LSP says: *objects of subtypes should behave like those of supertypes if used via supertype methods*. This is a very famous design rule in OO.

I was absolutely floored while listening to learn that she had no idea that people in industry had taken some of her ideas and elevated them to a principle. She says "what happened is that 5 or 10 years later she discovered that this idea got picked up by a community on the internet and they were all discussing what they referred to as the Liskov Substitution Principle. They call it LSP. That's just sort of amazing."

While amusing to her, it signaled a very serious problem to me. It made me realize how distant academia is from industry. How could she possibly not know? How could she possibly think so little of this practical little gem of an idea? It's hard for me to imagine, but it illustrates that vast gulf between industry and academia, the same sort of divide seen in the drug industry.

During her talk Dr. Mochly-Rosen shared how difficult it was to get research acted on by industry. As an example she used one of her own discoveries, [a drug to reduce the damage from a heart attack](#) by up to 70%. Her discovery, like so many others, was published in a prestigious journal, was impeccably well researched, and showed a lot of potential in a sadly growing market, yet drug companies were not beating down her door. She eventually ended up starting [her own company](#) to get the ball rolling. But this experience taught her something was broken in the process. And it wasn't what you might think, it wasn't simply that drug companies are evil.

Only about 20 new drugs are approved for human use every year. Only 5 of those drugs are new compounds, the other 15 are called "me too" drugs, they are slightly different versions of existing drugs. Take Viagra, for example (insert joke here). The discovery of Viagra would be one of the five new drugs. The other five versions of Viagra-like drugs we see on the market would be me-too drugs. They don't exactly advance the state-of-the-art, but they are safer to develop.

It can take 10+ years and a billion+ dollars to develop a new drug. It takes numerous trials to determine if a drug is stable, safe for humans, and to then determine the effective dosage. You know when on a label it says take 2 pills every 24 hours? How do they know the dosage? Someone has to figure that out. And the only way to figure that out is with a very wide range of trials on a very wide variety of humans. This is extremely expensive and risky. Very few drugs make it through the entire process.

The Gap isn't Caused by Evilness, it's More of an Awareness Problem on Both Sides

When the drug companies weren't beating down their door it wasn't necessarily because they were evil, it's because developing a drug is a

risky and expensive business. Why should they take a risk on novel compounds they are unsure about when a miss is so costly? When drug companies market the hell out of an approved drug it's because they finally got an approved drug, it makes sense to make the most money they can from it given the investment costs. I know, I think this is a screwy situation too, but given market forces it's rational.

Dr. Mochly-Rosen's insight was to realize there was a disconnect between researchers like herself and industry, a situation very similar to that of computer science. To illustrate this point she told a story about a talk she gave at a Cardiology Convention about a discovery she made on how to control the rate at which a heart beats. An amazing and novel discovery. But the cardiologists didn't care and she was puzzled as to why. A friend told her that cardiologists simply don't have a heart rate problem. What they care about are heart attacks. So she figured out a way to make her research be something cardiologists would be interested in. And once you have a market you can make a case for developing a drug.

That's what the SPARK program is about, **figuring out how researchers can reframe and present their research in a way that is more easily consumed by potential productizers**. Typically the goals and motivations of researchers and productizers are diametrically apposed. Researchers only get rewarded for publishing something new and moving on to the next new thing. Productizers have to stick with something for long periods of time and ride out every vicissitude. These are not the same people. Dr. Mochly-Rosen thinks researchers can be a great at generating new drug ideas, companies can develop them, the problem is how to get them to speak the same language and agree on some basic goals and protocols?

In the drug development world there's a clear **process that can be mastered**: Trial design, Intellectual property law, Consent

forms Regulatory submissions, Regulatory documents and amendments, Case Report Forms, Coordinating with other departments and specialties, Confidentiality, Protocol development and deviations, Preparing adverse event documentation, Data Safety Monitoring Boards.

As a researcher once you understand the process it will be more straightforward for you to position your research in a way that's easier for industry to understand, see the value of, and to take the risk of financing. SPARK tries to *fill the void between laboratory work and the delivery of products, increasing the value and readiness of commercial interventions.*

Do we Bridge the Gap with Bureaucracy or Something Else?

What would a similar process look like for computer science? I don't know, I was hoping others might have some ideas. I don't think it needs to be an official agency. That would just end up a big bloated bureaucracy with little to show for it.

This is the internet age. What we need is to make all the pieces of the puzzle easier to find, easier to understand, and easier to put together. If we can make the pieces more linkable the system will self-catalyze.

If I am, for example, looking for a sharded, replicated, secure storage system and I find a research project to that effect, it would help me immensely if those researchers thought it a priority to make it easy for developers like me to use it. Research is so often just to publish a paper, not to use, which means the systems are dead once the paper is published. This isn't a sustainable model. How do we make a more sustainable research model?

A Few Simple Ideas

Here are a few simple ideas from my own experience. Hopefully you have some better ideas.

Tear down that paywall! How many times have you researched a topic only to find the object of your desire hidden behind a ridiculously priced paywall? It happens to me all the time. How ironic that an industry built on openness and sharing is the most secretive of them all. The first step to making research more useful is to make it so people can actually read the research. This is the 21st century. Journals. Really?

Solve problems people in the industry are interested in, will be interested, or should be interested in. I know, research is about the next big problems and isn't about being practical today.

Understood. But is there a middle ground where you can figure out a way to help people today while you work on tomorrow? Working on distributed systems in the abstract is cool, working on a MapReduce that works in the field is also pretty cool.

Write papers with future developers and practical implementations in mind. Try to explain ideas thoroughly. Use examples. Examples are a gift. Explore implications. Think about potential gotchas. Try to go beyond just formulas. Use plain language. A dense writing style may seem to make a paper more significant, but they are far less useful in practice. Amazon's [Dynamo paper](#) is a good example, but even that could be clearer.

Make buildable software that's available on a service like GitHub. I know it's just research, so there's no requirement to be formal, but please take a little extra time so other people can use the code. Add some documentation, build instructions, tests, examples. Make it modular. Turn it into something like a real open source project. If you start out with the idea of making a project from the

start then the chances are far less that you will be too embarrassed to show the code in the future.

Record talks on your research. Take a look at the talks Google produces on their research or the talks Apple makes on their frameworks. This can go a long way to make something useful. Everyone could do a lot better at this. Of the thousand conferences being run, how many are on-line? It's almost criminal that this information is being lost.

Teach Better basics. Schools need to teach better. There are a quite a few classes on-line now, but often they are way too high level. Often classes start out with the equivalent of teaching calculus by writing a big integral sign on the chalkboard. It wouldn't hurt to go slower and actually explain things better. A Taching Company for computer science might be a good idea.

Be available via discussion lists. It would help immensely if developers could ask questions and learn more about the ideas behind the research. Often the hardest task for a developer is to build up the expertise needed to grok a new line of research. Some help in building up that background would be priceless.

Change the game mechanics. The game mechanics for research are all weighted towards publishing. Is there some way to rework the reward mechanism so the ease of use is also rewarded?

Develop business savvy. There should be something about how to take research and turn it into a startup. That might make the practical applications aspect of research more attractive as it makes the link between how research is written to potential rewards. The easier it is for VCs and potential founders to clearly see how an idea can be applied, the easier it is for research to be turned into a product.

Anything else that might help?

Related Articles

[Data Abstraction and Hierarchy](#) by Barbara Liskov

Article originally appeared on High Scalability (<http://highscalability.com/>).

See website for complete article licensing information.