

# *Emerging Trends*

## *The next generation*

KENNETH WARD CHURCH

*IBM, Yorktown Heights, NY, USA*

*e-mail: kwchurch@us.ibm.com*

*(Accepted 19 August 2016)*

---

### **Abstract**

I'm sure you want me to tell you about the next new emerging trend, but I'm not going to do that. It is much easier to suggest where trends come from (the next generation), and how to distinguish passing fads (bubbles) from emerging trends. Young people are often the early adopters, the first to see what is about to happen, but most people don't see what's coming until well after the fact. Those with the most to lose (the establishment) tend to be the most resistant to change.

---

### **1 Who are the early adopters?**

How do we distinguish bubbles from emerging trends? Eventually, time will tell. Bubbles are followed by crashes, whereas emerging trends continue to take off. But that's not very helpful. We want to distinguish the two, while it still matters, before the take-off/crash. How do we do that?

In my experience, there is an important difference with early adopters. If the early adopters are unqualified to have opinions, then you have a bubble. It is not a good thing when conferences are flooded with unqualified participants. CACM recently ran an article describing how some conferences are attempting to cope with this reality (Katz-Basset 2016), but the article failed to mention that it generally ends badly when a field becomes so hot that everyone wants in on the game, especially people that are completely clueless about what the game is. I am reminded of IJCAI-1985, which was a complete zoo. There were 6,000 participants at a time when there were only a few hundred PhDs in the field. It was pretty obvious, even at the time, that an 'AI Winter' was just around the corner.

At Coling-1988, the early adopter was the mayor of Budapest. It is common for politicians to give short opening speeches, but this mayor went on much longer than the norm, boring much of the audience (including yours truly), but not the locals. When it was safe, I asked what I missed. What was so exciting about the mayor's speech? Apparently, the mayor had just declared that there was nothing to fear from the Soviet Union. The mayor talked at considerable length about many things,

but failed to mention the Soviet Union. The locals recognized this for what it was, a clear signal that great changes were about to happen.

Many paradigm shifts start with younger researchers, often students and recent graduates. This was my experience with the revival of empiricism in the 1990s. EACL-1993 had three invited talks, two by senior leaders in established areas (rationalism), plus one on empiricism by an upstart (yours truly). While there was still plenty of resistance to empiricism, mainly from established circles, I was pleasantly surprised by positive responses from the folks that matter, the next generation. Young people have little invested in the *status quo*, and much to gain by rocking the boat. But they also tend to know what they are doing (unlike unqualified participants in conference bubbles).

One might think that directions are set by established leaders of the field, but ironically, these decisions are actually made by students. Established professors may raise money, but students spend it. The next round of PhD theses are written by students. Students get to pick the next emerging trend.

It is not surprising that young people have so much influence with emerging trends in computational linguistics given how much influence they have elsewhere. Consider historical linguistics and language change. Linguistic innovations often start out as youthful slang in disenfranchised neighborhoods in inner cities, and slowly migrate into more mainstream settings. Received pronunciation and the ‘King’s English’ (and formal academic discourse) are most resistant to change, though eventually, they too will fall into line.

Popular music works much the same way. I don’t care for today’s teenage music and teenagers don’t care for the music that I listen to (which I started listening to when I was a teenager). It has always been that way, and it will always be that way. Bottom line: trends in music are determined by teens, and trends in computational linguistics are determined by people in their twenties.

## 2 Late adopters and ‘old fogies’

The scientific establishment is all too similar to the King’s English and received pronunciation, all too resistant to progress. Coling used to have a panel at the end of the meeting which we used to call ‘the old foggy session’. Senior citizens would get up on stage and pontificate about what happened at the meeting. One particular old foggy took a pot shot at empiricism when it was just beginning to take off with students. He was so clueless about what was happening that he actually lectured students on what they should be working on, namely what was hot when he was their age.

It can be a lot of fun to pick on old fogies. I remember how much my mother enjoyed a silly Beatles song, ‘Will you still need me... when I’m 64’, but that was before she was 64. Some slogans were less light-hearted and more down-right divisive such as, ‘don’t trust anyone over 30’.

In any case, now that I am old enough to join AARP, I hope I don’t come across like what I used to object to. We need to understand our station in life. We are

counting on the next generation to rock the boat, and the last generation to keep it from flipping over.

It is easy to poke fun at the ‘old foggy session’, but more seriously, I am concerned about how Coling is run. The organizational structure of Coling gives senior citizens too much of a voice, with little chance for other positions to be heard. When we set up EMNLP (formally known as the Workshop on Very Large Corpora), we intentionally selected chairs to be up-and-coming young researchers who could benefit from the exposure. These people did all kinds of wild and crazy (mostly good) things, like change the name of the meeting from a workshop to a conference, which made so much sense in retrospect that I wish I had thought of it. To maintain a little sanity among all the chaos, we paired the chair with a more seasoned co-chair who knew how things had been done in the past, and could offer constructive advice when necessary (if ever). It might seem upside down for the chair to be more junior than the co-chair, but we found the meetings often benefited when the chair was full of energy, and viewed the task more as an opportunity than a thankless chore.

### 3 Kuhn’s structure of scientific revolutions

It is ironic to look to historians for insight into the next big thing, but (Kuhn 2012)<sup>1</sup> has an amazing number of citations, presumably because so many people find his work so useful for predicting the future.

Unfortunately, Kuhn is difficult to read. He uses a long-winded flowery academic writing style that doesn’t work so well today where elevator pitches have to be compressed down to tweets. Yogi Berra understood how to tweet (before tweets were a thing). Here’s what Yogi had to say about new trends (in computational linguistics): ‘It’s tough to make predictions, especially about the future’.

Many of the points mentioned above were inspired by Kuhn. In particular, Kuhn criticizes the old fogies (Priestley) for doing what the establishment does (attempt to maintain the *status quo*):

[T]he fact that a major paradigm revision was needed to see what Lavoisier saw must be the principal reason why Priestley was, to the end of his long life, unable to see it. (Kuhn 2012, p. 56)

The establishment (e.g. Priestley) are typically the last to see change. But often, it also takes a long time for the mainstream (silent majority) to see what’s happening. History is easier to appreciate in retrospect:

How, then, are scientists brought to make this transposition? Part of the answer is that they are very often not. Copernican-ism made few converts for almost a century after Copernicus death. Newtons work was not generally accepted... for more than half a century after the Principia appeared. Priestley never accepted the oxygen theory, nor Lord Kelvin the electromagnetic theory, and so on. (Kuhn 2012, pp. 150–151)

<sup>1</sup> A pdf version became freely available in 2012 with the fiftieth anniversary of the 1962 original.

With 20/20 hindsight, the process of change is relatively straightforward. First, it starts small, and then it gets bigger, and eventually, it succeeds. But of course, bubbles start out the same way as successes. The difference is that bubbles don't keep growing (for long):

At the start a new candidate for paradigm may have few supporters, and on occasions the supporters motives may be suspect. Nevertheless, if they are competent, they will improve it, explore its possibilities, and show what it would be like to belong to the community guided by it. And as that goes on, if the paradigm is one destined to win its fight, the number and strength of the persuasive arguments in its favor will increase. More scientists will then be converted, and the exploration of the new paradigm will go on. Gradually the number of ... will multiply. Still more men, convinced of the new views fruitfulness, will ... until at last only a few elderly hold-outs remain. And even they, we cannot say, are wrong. Though the historian can always find men – Priestley, for instance – who were unreasonable to resist for as long as they did... (Kuhn 2012, p. 159)

We like to believe there was a strong case for the new paradigm, even at the very beginning, but actually, that's rarely the case. Early work tends to be more promising than convincing:

But paradigm debates are not really about relative problem-solving ability, though for good reasons they are usually couched in those terms... that decision must be based less on past achievement than on future promise... The man who embraces a new paradigm at an early stage must often do so in defiance of the evidence provided by problem-solving... A decision of that kind can only be made on faith. (Kuhn 2012, pp. 157–158)

Kuhn taught me that an emerging trend has to do two things. First, it helps to have a few (promising, if not convincing) initial successes that excite a target audience of early adopters (students). That's not surprising.

But what I find completely counter-intuitive is that those initial successes shouldn't be too successful. If the first paper is too definitive, the field will be still born. It is important to leave plenty of room for the next generation to contribute. As mentioned above, students get to set directions by writing the next round of PhD theses. Show them a way forward so they can contribute and join in on the fun, but don't do all the work for them by writing the definitive last word on the subject.

In my next column, I will discuss how Word2vec meets both of these desiderata: (1) a few initial successes that motivate early adopters to do more, as well as (2) leaving plenty of room for early adopters to contribute and benefit by doing so. The fact that Google has so much to say on 'How does word2vec work', makes it clear that the definitive answer to that question has yet to be written. This a great formula for racking up citations, as we can learn from the Word2vec experience. It also helps citation counts to distribute code and data to make it that much easier for the next generation to take advantage of the opportunities (and cite your work in the process).

## References

- Katz-Bassett, E., Sherry, J., Huang, T.-Y., Kazandjieva, M., Partridge, C., and Dogar, F. 2016. Helping Conference Attendees Better Understand Research Presentations. *CACM* **59**(9): 32–34.
- Kuhn, T. 2012. *The Structure of Scientific Revolutions*. Chicago, USA: University of Chicago Press.