

# Publication Culture in Computing Research

## ***Design for impact: Rethinking academic institutions from the ground up***

Jon Crowcroft

University of Cambridge, Computer Laboratory, 7.10.2012

Why do we pretend that a publication is an event, rather than a part of an ongoing process?

Computer Science is a Soft Subject. We create artificial systems/artefacts, and explore their behaviours. We then report on this by talking about the behaviours at workshops and conferences, and writing about the systems in papers for web pages, online archives or even traditional print journals.

People assume that the artificial dichotomy between social events (workshops, conferences) and archival repositories (journals and the like) is right. And some of the debate about CS publication culture is oriented around trying to get people to use these two modalities more like other disciplines.

I think this is fundamentally wrong, and flies in the face of real scientific method. Science does not deliver truth. It delivers things that work, and explanations that are the best, current, simplest ones (c.f. Popper on Objective Knowledge, and of course Occam's Razor).

This means that a work is not the final word. It is just the current word. A goal of this proposal is to reduce the "slice and dice" culture present today due to various perverse incentives.

So the notion that an "archival paper" has been thoroughly checked and is infinitely more "correct" than a "rapidly" reviewed conference submission is not tenable. There is every chance that during the necessarily longer process to create an archival version of a work, subsequent work has improved over the results. Hence much archived material is actually less accurate because it is less timely.

The solution, for me, is to remove the notion of immutable publications, and admit that we should update work continuously

This can apply to the entire process of socialising our work, hence a dialogue (or multilogue) between authors, reviewers and readers, continually adds accuracy or timeliness (or invalidates a work). The same can apply to citations (which should, by the way, have a "sign bit" to indicate whether the citation is building on from a work, or citing it as the thing the new work invalidates).

Recognising this mutable publication model, would allow work to be presented at any point along the "production line", perhaps merely by "acclaim" - some work has reached a point where it is mature enough and timely and interesting enough to merit presentation at a social event (workshop or conference) - this could happen before or after some notional point when it is recognized that an archival version is the current best knowledge we have (a rare event).

Along side this continual process, I think one would have to abandon ideas of anonymity in both authorship of work, and reviews/critiques (viz, the “dialogues” mentioned above could only work in that open way). It goes without saying that code and data associated with a systems’ behaviour should also be openly available as part of this ongoing process (after all, since when did we declare code “bug free” correctly? Why, therefore do we declare journal papers “correct”?).

Finally, this isn’t exclusive to Computer Science, but we built the tools that would make the new approach viable, so we should use them first.

In fact we also have the next generation tools for this – we just need to combine Arxiv with Github (versioning repositories)<sup>1</sup>.

### ***Causes of paper count inflation.***

CS is notable (in most branches at least) for submitted to conferences more than journals. There are two pressures to do this

1. Urgency
2. Promotion

CS is a young disciple, and the young are noted for being impatient and impetuous - our slogan might even be said to be “Publish Early and Publish Often”<sup>2</sup>.

### **Urgency**

We live in a nanosecond world. More than other disciplines, partly because we built it.

We supplied the tools and tool chains (the net, e-mail, the web, PDF, bibtex/latex, databases, HotCRP/EDAS, etc) that let us cooperate to develop ideas, systems, results, and write papers faster, and deliver them for review, editing, and presentation more quickly than any previous generation. Surely, other disciplines use the tools, but we live and breath them.

As a result, there’s a feedback loop between publication of hot new work, This instant gratification leads to an increase in the rate of submission.

Our profession has also a tendency (at least anecdotally) to attract a share of people with OCD/Attention Deficit problems, who maybe (amateur psychologist’s hand waving here) seek instant rather than deferred gratification.

---

<sup>1</sup> Github because we want distributed repositories to avoid re-concentrating power in one place all over again.

<sup>2</sup> I could speculate here about whether these factors also contribute to the gender imbalance in Computer Science as a profession and academic career (whether directly, or simply as proxies for a root cause).

## **Promotion**

Our academic research culture is funded largely by tax payers money (NSF, DARPA, EU), and the tax payers seek metrics to see their money is well spent, and they seek such feedback on an annual basis. Paper counts (and to a lesser extent, citation counts) serve this. The same problem (inflation) has hit the industry research and development world, where patents are a proxy for real work, and are rewarded.

The amount rather than significance of work is measured - hence, the aforesaid dice and slice approach to work, producing minimal publishable units, and multiplying the number of venues and publishable units year on year.

Because CS is young and vigorous, we have in the past been able to keep up with this inflation. We are close to the limits though.

In the UK, we have a national Research Excellence Framework, for which researchers in universities do not return all their work. Instead, every 5 years, up to 4 “outputs” (e.g. papers) are returned. Secondly, and in addition, impact stories (pieces of work 10-20 years old, that have had a long term effect on the world, economically, socially, or in terms of further developments in a discipline) are employed.

It will be interesting to see the outcome of this process, but for me, it is probably a better basis for looking at some one person, or groups progress, so if we were to use these sorts of indicators for tenure or similar, this would remove the aforesaid perverse incentive to maximise the number of publications.

## **Acknowledgements**

Thanks to Richard Clegg and Ioannis Avramopoulos for comments on this draft.