

# Navigating the (Systems) Literature

The Way Things  
Are

Publishers are  
Definitely Stupid

David Cock

8 May 2014



Australian Government  
Department of Broadband,  
Communications and the Digital Economy  
Australian Research Council

NICTA Funding and Supporting Members and Partners



Australian  
National  
University



UNSW  
THE UNIVERSITY OF NEW SOUTH WALES



Trade &  
Investment



THE UNIVERSITY OF  
SYDNEY



Queensland  
Government

# Why Read Papers?

---



NICTA

At least 3 reasons:

The Way Things  
Are

Publishers are  
Definitely Stupid

# Why Read Papers?

---



NICTA

At least 3 reasons:

- Novelty — To convince yourself that people care about the problem, and that it hasn't already been solved.

The Way Things  
Are

Publishers are  
Definitely Stupid

# Why Read Papers?

---



The Way Things  
Are

Publishers are  
Definitely Stupid

At least 3 reasons:

- Novelty — To convince yourself that people care about the problem, and that it hasn't already been solved.
- Expertise — Aim to become the local (or even global) expert in your topic. By 3rd year, expect other group members (including researchers) to ask **you** whether X has been done before.

# Why Read Papers?

---



The Way Things  
Are

Publishers are  
Definitely Stupid

At least 3 reasons:

- Novelty — To convince yourself that people care about the problem, and that it hasn't already been solved.
- Expertise — Aim to become the local (or even global) expert in your topic. By 3rd year, expect other group members (including researchers) to ask **you** whether X has been done before.
- Communication — In your papers, and your thesis, you need to direct the reader to your sources. Aim to extend and improve the literature.



## The Way Things Are

Publishers are Definitely Stupid

- The Way Things Are
- Publishers are Definitely Stupid

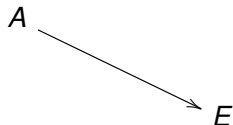


A

The Way Things  
Are

Publishers are  
Definitely Stupid

It's 2025. You've found a great new paper on verifying Angry Birds XII, A.

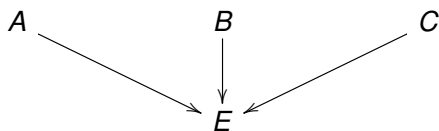


The Way Things  
Are

Publishers are  
Definitely Stupid

It cites *E*, saying “It is well-recognised[1], that all commercially-successful kernels are formally verified.”.

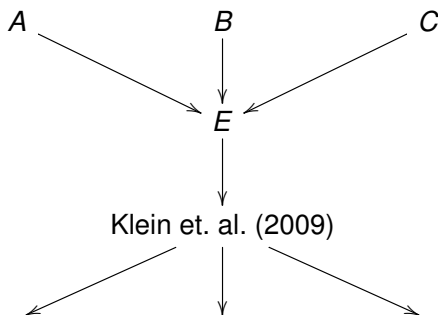




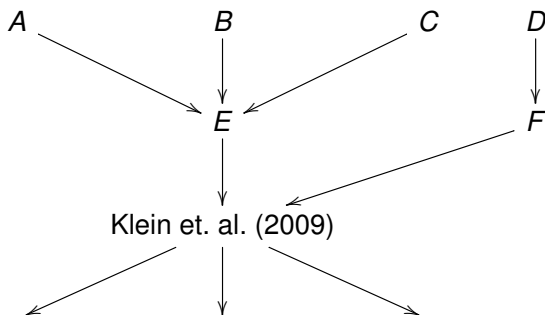
The Way Things  
Are

Publishers are  
Definitely Stupid

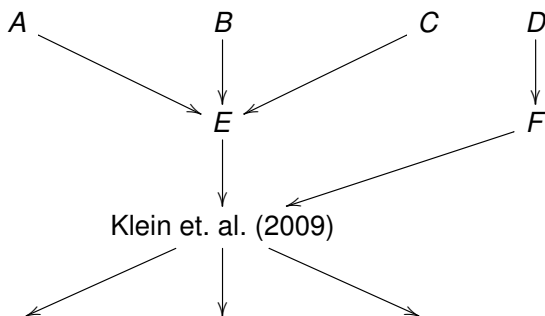
You put *E* into Google Scholar, and click on “cited by”, and find *B* and *C*. *B* is a paper on verified GC, while *C* is a paper on timewasting. You decide that *C* is directly relevant, while *B* less so.



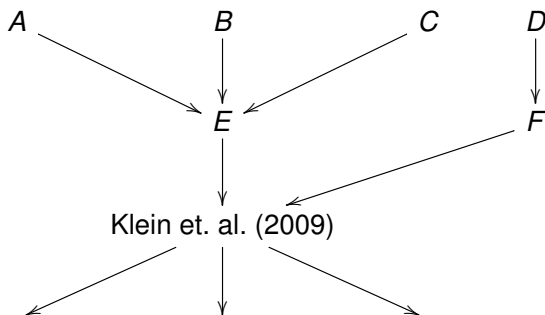
You see that *E* cites another paper, as the first to show that large-scale verification works. This paper is clearly pivotal, but is now mostly of historical interest.



Chasing citations of this paper is pretty futile. *D* is a paper in a naval engineering journal on the use of seL4 in the control systems of on-board fusion reactors.



The literature is huge DAG, and you need to walk it efficiently. Identifying 'sink' articles, like *E*, and 'landmarks', like Klein et. al. is crucial, as they either cite or are cited by everything worth reading.



You can't read everything, so you need to learn to prune effectively.

# Finding Things

---



NICTA

The Way Things  
Are

Publishers are  
Definitely Stupid

You need to know where the work in your field is published, and how to find it. Generally, this comes down to knowing three things:

You need to know where the work in your field is published, and how to find it. Generally, this comes down to knowing three things:

- Venues

You need to know where the work in your field is published, and how to find it. Generally, this comes down to knowing three things:

- Venues
- Publishers



You need to know where the work in your field is published, and how to find it. Generally, this comes down to knowing three things:

- Venues
- Publishers
- Aggregators

# Venues (Systems)

---



NICTA

The Way Things  
Are

Publishers are  
Definitely Stupid

# Venues (Systems)

---



NICTA

Prestige venues: SOSP, OSDI, ...

The big-story results will be here.

The Way Things  
Are

Publishers are  
Definitely Stupid

# Venues (Systems)

---



NICTA

Prestige venues: SOSP, OSDI, ...

The big-story results will be here.

The second tier: HotOS, Eurosys, ...

Higher volume, not necessarily lower quality. Often more focused.

The Way Things  
Are

Publishers are  
Definitely Stupid

# Venues (Systems)

---



**Prestige venues: SOSP, OSDI, ...**

The big-story results will be here.

**The second tier: HotOS, Eurosys, ...**

Higher volume, not necessarily lower quality. Often more focused.

**Workshops**

Incremental or early work, highly focused.

The Way Things  
Are

Publishers are  
Definitely Stupid

# Venues (Systems)

---



Prestige venues: SOSP, OSDI, ...

The big-story results will be here.

The second tier: HotOS, Eurosys, ...

Higher volume, not necessarily lower quality. Often more focused.

Workshops

Incremental or early work, highly focused.

Journals

Longer papers, not a thing in systems.

The Way Things  
Are

Publishers are  
Definitely Stupid



The Way Things  
Are

Publishers are  
Definitely Stupid

- In systems, mostly ACM, IEEE & Usenix.
- Springer and Elsevier for some journals (sometimes proceedings).
- Some open-access or independents.
- Sometimes (e.g. ACM DL) index each other's pubs — don't trust citation details. Go to the source.

The publisher's database is the canonical source for citation information:

- ACM Digital Library (DL)<sup>1</sup>
- IEEE Xplore<sup>2</sup>
- Usenix<sup>3</sup>
- SpringerLink<sup>4</sup>
- Elsevier ScienceDirect<sup>5</sup>

---

<sup>1</sup><http://dl.acm.org>

<sup>2</sup><http://ieeexplore.ieee.org/>

<sup>3</sup><https://www.usenix.org/publications/proceedings>

<sup>4</sup><http://link.springer.com>

<sup>5</sup><http://www.sciencedirect.com>

The Way Things  
Are

Publishers are  
Definitely Stupid



Cite  
Seer  
X BETA

Google  
scholar

ACM DL DIGITAL LIBRARY

The Way Things  
Are

Publishers are  
Definitely Stupid

Index articles across publishers.

---

<sup>6</sup><http://citeseerx.ist.psu.edu/>

<sup>7</sup><http://scholar.google.com/>

Cite  
Seer  
X BETA

Google  
scholar

ACM DL DIGITAL LIBRARY

The Way Things  
Are

Publishers are  
Definitely Stupid

Index articles across publishers.

- Citeseer<sup>6</sup> — Slow, crappy bibtex, don't bother.

---

<sup>6</sup><http://citeseerx.ist.psu.edu/>

<sup>7</sup><http://scholar.google.com/>



The Way Things  
Are

Publishers are  
Definitely Stupid

Index articles across publishers.

- Citeseer<sup>6</sup> — Slow, crappy bibtex, don't bother.
- Google Scholar<sup>7</sup> — Better, basically a narrow google search, noisy. Use for cited-by.

---

<sup>6</sup><http://citeseerx.ist.psu.edu/>

<sup>7</sup><http://scholar.google.com/>



The Way Things  
Are

Publishers are  
Definitely Stupid

Index articles across publishers.

- Citeseer<sup>6</sup> — Slow, crappy bibtex, don't bother.
- Google Scholar<sup>7</sup> — Better, basically a narrow google search, noisy. Use for cited-by.
- ACM DL — Indexes most of CS, pretty well cultivated. Don't trust for citation details other than ACM.

---

<sup>6</sup><http://citeseerx.ist.psu.edu/>

<sup>7</sup><http://scholar.google.com/>



- The Way Things Are
- Publishers are Definitely Stupid

Many publications are still paywalled, although some publishers have got the message (e.g. Usenix).

- Your university library will usually have a subscription.
- At UNSW, use the library proxy by appending to the URL e.g.

`http://dl.acm.org.wwwproxy0.library.unsw.edu.au/`

- This can be automated, see next slide.
- Some very old papers aren't digitized, but that's rare in CS.
- Most authors put a prepress version on their own website. At NICTA this is automatic.

# zotero

There are  $n + 1$  of these, I used Zotero<sup>8</sup>, and found it useful.

- Integrated with Firefox.
- Detects ACM DL, IEEExplore, ...
- Saves reference details with a click.
- Free cloud sync for references and PDFs.
- Detects library proxy and automatically redirects.

---

<sup>8</sup><https://www.zotero.org/>





# Questions?