Writing a literature review

CSW 2009
Steve King

based on a 2007 lecture by Helen Petrie

What's a lit review ...

"A critical analysis of a segment of a published body of knowledge through summary, classification and comparison of prior research studies …[and previous] reviews …" (University of Wisconsin-Madison, 2006)

A lit review for a project is also focussed on providing the necessary background to your work, not a general survey of the area

We stand on the shoulders of giants …

- First reason it's very important is that no-one is expecting you to make up the whole project, you should use as much as you can from previous work: ideas, methods, presentation style, critiques …

- For that reason, there should be a small number of highly relevant papers that you read over and over…

- When you think, "how should I do X?", look at how these papers tackled the problem

We stand on the shoulders of giants ...

- Second reason it's very important, is that it provides the rationale for what you do in your project
- Start from general considerations, present your lines of argument and end up with your particular problem/topic

More motivation

- Understand your project area and how to work in it
- Don't reinvent the wheel
- Don't reinvent the square wheel

[PhD Lit Review talk: Jim Woodcock, 2008]

A (real) example: lost in hyperspace

Say I'm interested in doing a project on why people get lost and confused on the Web

My first move would probably be to do a search in the ACM Digital Library and/or IEEE Xplore

Don't rely just on "open" web searches for research – much of the best material is copyright protected, so you won't necessarily find it in the open web

Start thinking about terminology – what keywords do people use, what will you need to define

Online sources: not just Google

- http://liinwww.ira.uka.de/bibliography/ : collection of CS bibliographies
- Bibliographies:
  - CiteSeer
  - DBLP
  - Google Scholar

See next week's lecture on CS Information Sources

What is 'literature'?

- Peer-reviewed publications
- Non-reviewed publications
- Software
- ….
- Anything necessary to understand the context of your project
- Make absolutely clear what's yours and what's not
- Non-standard projects?

A (real) example: lost in hyperspace

Turns out the key paper on the problem is called "Towards a practical measure of hypertext usability" (really old-fashioned title!) and the keywords were: human-computer interaction, hypertext, usability

So took us a long time to get to this key paper

However, we found a paper called "Lost in hyperspace: metrics and mental models" and that led us to other important papers

Following leads

Once you do find a really key paper you can work forward by finding out who cites that paper

Google Scholar is very good for this (and CiteSeer)
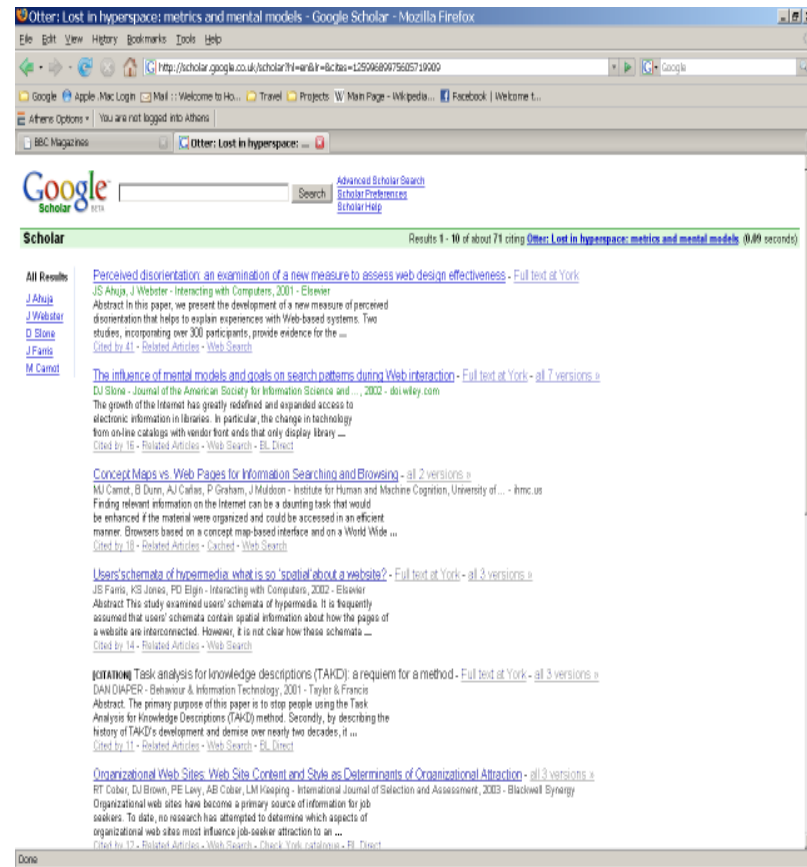
Enter the details of your key paper

Then follow the "Cited by" link

Following leads

71 papers cited the Otter and Johnson paper

Can obviously also trace back from recent papers by looking at what they cite

Developing an argument

- Need to consider what each paper contributes to your argument

- In reading the papers you might learn all kinds of things, but in the literature review, you are setting out the steps of the rationale underlying your work

- Just because you are discussing a particular paper, you don't need to discuss all of it

Example of building the argument

Problem: how am I going to measure "lostness"?

Smith (1996) argued that previous studies ("most existing usability studies …" – very poor, no specific citations) had used time taken to complete a task on the Web and errors made as the measure of success, but this does not tell you whether the person felt lost when doing the task on the web

Smith started a line of work about different measures of lostness, objective and subjective

Summarize key points from papers to build your argument

One of the most recent papers in this set is Gwizdka and Spence (2007)

Gwizdka and Spence (2007) start their paper with this sentence:

"Navigating large, complex websites is often difficult."

Excellent start - this is the problem they are interested in, stated clearly, we know what we are going to be reading about

This is the top of the funnel, the general statement of the purpose

Summarize key points from papers to build your argument

Later in the paper G &S (2007) introduce and summarize the Smith (1996) paper like this:

"*Getting lost,* or *disoriented,* is known to be one of the most important problems is hypertext navigation, yet there have been but a few attempts to assess and quantify lostness. Smith (1996) proposed an objective measure of lostness based on the ratios of visited and optimal node counts as shown in Equation 2."

Introduce terminology

"Getting lost, or *disoriented,* is …

You can either explain terminology **explicitly** (good for very complex terms):

Cognitive dissonance refers/means ….

Can also do it **implicitly**, as G&S did

Sometimes different authors use different terms for the same thing

Best to mention this and then choose one term you prefer and stick to that

Introducing terminology II

"Computer anxiety (Brosnan, 2001), computerphobia (Smith, 1989) and technostress (Rosen, 1998) are all terms that have been used to refer to a generalized fear of working with computers and other modern technology. In this paper, I will use the term computer anxiety to cover these terms."

You might also need to name your own terms - it might be next year's catchphrase

Helen's proudest moment – "the Tardis Effect"

Back to building your argument

"Getting lost, or *disoriented,* is known to be one of the most important problems is hypertext navigation, yet there have been but a few attempts to assess and quantify lostness."

Expression is a bit flowery!

"… is known to be one of the most important problems is hypertext navigation, but there have been few attempts to assess and quantify lostness"

Building your argument

"Getting lost, or *disoriented,* is known to be one of the most important problems is hypertext navigation, yet there have been but a few attempts to assess and quantify lostness."

BUT, this is the argument the authors are developing - there's not been much work on assessing and quantifying lostness - but we'll tell you what has been done and what we think of it …

Building your argument

"Smith (1996) proposed an objective measure of lostness based on the ratios of visited and optimal node counts as shown in Equation 2. Larson and Czerwinski (1998) compared user performance on information search in three different hypertext hierarchies: [details of the three hierarchies]. Using Smith's measure L&C showed that users were more lost on a hypertext with the 8x8x8 hierarchy than on either the 16x32 or the 32x16 hierarchies."

Building your argument

So all that's taken from Smith is the measure, whereas what's taken from L&C is the use of this measure and what it showed about using an objective measure of lostness - that it worked

I do have a problem with that bit of text - not clear why L&C is building your argument until you get to the second sentence

Always try and lay the argument out to the reader very clearly - you are doing the detective work on this research, not the reader

Lead your reader by the hand through your argument! If it can't be understood, is that the reader's fault or the writer's?

A re-write!

Smith (1996) proposed an objective measure of lostness based on the ratios of visited and optimal node counts as shown in Equation 2. **Using this measure**, Larson and Czerwinski (1998) compared user performance on information search in three different hypertext hierarchies: [details of the three hierarchies]. **They found** that users were more lost on a hypertext with the 8x8x8 hierarchy than on either the 16x32 or the 32x16 hierarchies. **This provided useful validation of the measure**."

Summarize -> classify

Remember the definition said a lit review classifies - what does this mean?

I'd say, presents a meta-level of discussion - what are the research papers trying to do, or what do they do for your argument, what is their nature/contribution

"This provided useful validation of the measure."

This is now discussing the L&C paper in relation to the objectives of a lit review - we are classifying this information

Very important to make this explicit to the reader

Summarize -> classify -> critique

Finally, thinking critically about research

This is hard and takes practice!

Are there important things that you think papers have not considered
- is the logic correct?
- was the method good?
- Many other factors, depending on the area

Critique can be positive as well - something good in one paper that could be applied to another paper

Critiques between papers

Often the critique can come in the gaps between the papers, not the problems in particular papers (any particular paper can only consider so many factors)

Also inconsistencies between papers are often very interesting - why are they inconsistent?

It's like doing a jigsaw puzzle - what bits have you got (particular bits of information from particular papers), what is lacking (that you could contribute)

Continuing the review from G&S

Otter and Johnson (2000) described two measures designed to assess user lostness. The first of their measures combines previous work by Smith (1996) with the effects of different types of links. Their second measure is concerned with the accuracy of users' mental models of websites. The authors suggested that to capture lostness in hypertext, a battery of measures was needed; Herder's (2003) work supported this viewpoint. Ahuja and Webster (2001) conducted an experiment demonstrating that user perceived disorientation in web navigation

Continuing the review from G&S

(assessed by a questionnaire developed by the authors) is only weakly related to user behaviour, and that perceived disorientation is a better predictor of performance …. In a study that examined perceived user disorientation in hypermedia (Herder, 2003) found that perceived disorientation (measured using the Ahuja and Webster instrument) was correlated with a combined page return rate (…) with median page view times, but not with the page revisitation ratio."

Several problems here - switching terms hypertext - hypermedia

Different between page return rate and page revisitation ratio not explained

Finally, the contribution from G&S:

"Thus, in contrast to Ahuja and Webster, Herder's work demonstrated that user lostness was correlated with diverse measures of user behaviour. *These findings suggest that lostness is not a simple unidimensional construct.*"

So, they have been working up to showing slight inconsistencies between results from several of these papers (although they don't really capitalize on that point) and that user lostness is not a single construct

Notice - they are not saying anyone is wrong, bad …

So, in conclusion

A lit review is really telling an interesting story about a particular issue/problem

It's mainly about good, precise and focussed summarization of previous work

Classification/high level thinking about the previous work is good

A dash of critique (negative or positive) is great

How to get started ...

Carefully study a set of related papers

All the lost in hyperspace papers will be linked from the course web page

Bear in mind that in published papers, the lit review is very compressed - typically they might have 500 - 1000 words, you might have 2000 - 3000 words

That doesn't mean you should waffle - give more detail about the papers, discuss the links between them more

Lost in hyperspace

Here's the set of papers that we found on this topic, with the main links between them