Secondary Studies & Systematic Literature Reviews

Technical Reviews
The ‘review paper’ is a well-established model in many domains, where there are some journals that only publish reviews, with such papers being written by experts in the field who review the current state of the literature on a specific topic and draw appropriate conclusions. In computing we have ACM Computing Reviews and in the Information Systems area, MIS Quarterly. However, because these are ‘expert reviews’, it is possible that two different experts might select different papers and draw different conclusions!

The Evidence-Based Paradigm
Originated in clinical medicine over the past two decades, spurred on by Archie Cochrane’s concern about the quality and variability of research evidence being used to inform practice/teaching. Evidence-Based Medicine (EBM) seeks to employ secondary studies to find, judge and synthesise the outcomes of all relevant empirical studies to draw conclusions about particular treatments. While clinical medicine has pioneered the use of such studies (helped by the widespread use of Randomised Controlled Trials (RCTs) that permit statistical meta-analysis) other domains that are employing an evidence-based approach include Education (where such reviews often inform policy); other branches of healthcare such as Nursing & Midwifery; and various branches of the Social Sciences. Software Engineers are relative late-comers!

Secondary studies
A secondary study sets out to aggregate and synthesise the outcomes of many primary studies in an objective and unbiased manner using either qualitative or quantitative forms of synthesis. Primary studies are usually empirical (often some form of experiment, but other forms can be included), and because many factors may influence an individual primary study, sometimes without the experimenters being aware of this, the use of a secondary study should reduce the effects of individual bias in the results. A major form of secondary study is the systematic literature review or systematic review (both terms are used). This is motivated by a research protocol (plan) that identifies such elements as the research question, how it can be categorised in keywords, and how and where the search for source material will be conducted.

The first stages of a systematic review are sometimes used separately as a mapping study (or scoping review) that seeks to identify the scope and volume of relevant material, but is used to help formulate a specific research question rather than to answer one.

While motivated by the aim of improving practice in a discipline by providing practitioners with the best available evidence about ‘interventions’, secondary studies are also important to researchers because they:

• encourage improvement in the reporting standards used for primary studies, since for effective aggregation they need to have some common basis for extracting data
• help identify where more primary studies are needed, since they ‘map the research terrain’

Secondary studies in Software Engineering: examples
The (less formal, but systematic) series of studies conducted by Glass, Ramesh and Vessey (2004) involved sampling papers from a number of major computing journals and classified papers from these as being:

• descriptive (reporting, reviewing,...) – 28% for SE
• evaluative (deductive, interpretive, critical,...) – 14% for SE
• formulative (concepts, frameworks, guidelines,...) – 55% for SE (overwhelming for CS)

Another example was the study by Jørgensen and Moløkken-Østvold (2006) who looked at the widely-cited CHAOS report on project failure rates and found its conclusions were out of line with their own systematic study of available data (the Standish Group who conducted the original study would not identify their methodological approach used in this).

1 The Cochrane Collaboration (www.cochrane.org) was founded in 1993 to support this work.
Jørgensen also looked at studies that compared algorithmic cost models such as COCOMO with estimate made by experts: a third found the algorithmic estimates best; another third found expert estimation best; the remainder found no difference!

**Evidence-Based Practice**

The basic process is:
1. Convert need for information into an answerable questions
2. Find the best evidence with which to answer the question
3. Critically appraise the evidence for its validity (closeness to the truth), it impact (size of the effect), and its applicability (usefulness)
4. Integrate the critical appraisal with SE expertise and with stakeholders’ values
5. Evaluate the effectiveness and efficiency in steps 1-4 and seek ways to improve them

The first three of these steps are essentially the process of system**matic literature review**.

**Systematic Literature Review**

**Phase 1: Plan Review** consists of three steps:
- specify research question—a non-trivial task as used to construct search strings and may go through several stages of refinement, possibly helped by a wide scope mapping study
- develop review protocol—a key document, especially where a team is involved in conducting the review
- validate review protocol—among other things, may need some form of ‘dry run’

**Phase 2: Conduct Review** involves five activities:
- identify relevant research—consists of executing the search strategy defined in the protocol in terms of search strings, sources to search, bounding dates, etc., possibly leading to a large number of source documents
- select primary studies—sift through the candidate papers’ titles, abstracts and (if necessary) bodies using the inclusion/exclusion criteria specified in the protocol
- assess study quality—mainly concerned with assessing how a given primary study was conducted, and how well-structured the data reporting is
- extract required data—usually involves ‘data extraction forms’ to help structure this process, and ideally uses two analysts working independently to ensure consistency
- synthesise the data—depends on the form, may be qualitative, or quantitative using statistical forms

**Phase 3: Document Review** is essential the writing and validating of the resulting report.

**Experiences from Software Engineering**

No standard abstracting services (unlike Medicine). So finding primary studies is mainly performed using search engines such as CiteSeer, Web of Science, Google Scholar, etc. which search different sets of sources (mainly digital libraries), use different interfaces and do not always produce consistent outcomes. Even collectively they are unlikely to find all papers, and experience suggests that around 5-10% more papers (estimated) can still be found through snowballing (following up references in the included papers).

Also, reporting standards are poor. Abstracts for primary studies tend to be written badly and often omit important information. Also, the papers are apt to report only that part of the data that is relevant to their own research question, with no sense of adding to the overall set of knowledge about a topic.

**Mapping Studies**

These are often used as a preliminary study to identify whether there is scope for a systematic review. They have a fairly broad research question (‘what empirical studies have been conducted into…’) and use the same early steps as a systematic literature review, but rather than aggregating and analysing the papers, they usually categorise these into topics, and identify where ‘gaps’ and ‘clusters’ are, with the idea that the clusters could form the basis for a full review (if large enough).