Literature Review

Dr.Muzammil Hussain

Département of CSCT, GCET, Oman

muzammil.h@gcet.edu.om

What is literature review?

- A summary of all papers in the area?
- A summary of papers you've read?
- A document to show how bad other research is?
- An argument for your research?
- Something you write after you've done your research?

What is literature review?

You make the *case* for your project

What is literature review

"… a systematic…method for identifying, evaluating and interpreting the …work produced by researchers, scholars and practitioners."

FINK, A., 1998. Conducting literature research reviews: from paper to the internet. Thousand Oaks, CA: Sage

" A summary of a subject field that supports the identification of specific research questions."

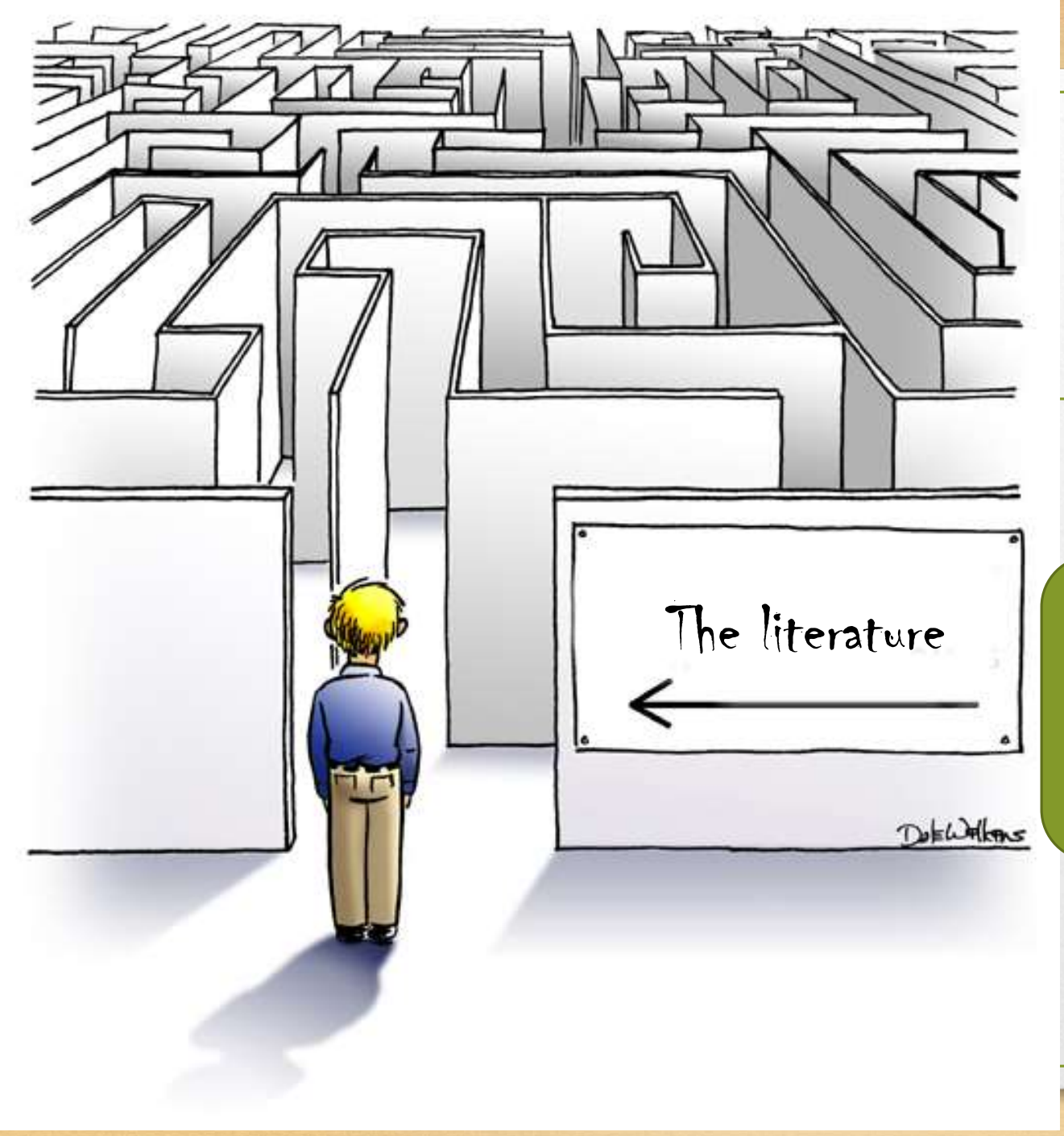
What is literature review

Novice researchers tend to approach the literature review as nothing more than a collection of summaries of papers or an elaborated bibliography of multiple research manuscripts (Webster & Watson, 2002).

A meaningful literature review is much more.

Goals

- To demonstrate to readers and examiners that we are familiar with the field
- To provide an overview of current knowledge in a particular area of application
- To review other studies critically
- To highlight a gap in areas of application
- To provide a context for our current study and to locate it within a specific field
  - *How the proposed research contributes something new to the overall Bank of Knowledge or advances the research field's knowledge.*



Before entering the maze

Be sure of what you are looking for

Sources of literature

- Books
- Patents
- Journals
- Conference Papers
- Theses
- Abstracts
- Electronic Databases
- Government publications
- Interviews and other unpublished research
- …..

Text-Independent Writer Identification and Verification Using Textural and Allographic Features

Marius Bulacu, *Student Member*, *IEEE*, and Lambert Schomaker, *Member*, *IEEE*

Abstract—The identification of a person on the basis of scanned images of handwriting is a useful biometric modality with application in forensic and historic document analysis and constitutes an exemplary study area within the research field of behavioral biometrics. We developed new and very effective techniques for automatic writer identification and verification that use probability distribution functions (PDFs) extracted from the handwriting images to characterize writer individuality. A defining property of our methods is that they are designed to be independent of the textual content of the handwritten samples. Our methods operate at two levels of analysis: the *texture level* and the *character-shape (allograph) level*. At the *texture level*, we use contour-based joint directional PDFs that encode orientation and curvature information to give an intimate characterization of individual handwriting style. In our analysis at the *allograph level*, the writer is considered to be characterized by a stochastic pattern generator of ink-trace fragments, or graphemes. The PDF of these simple shapes in a given handwriting sample is characteristic for the writer and is computed using a common shape codebook obtained by grapheme clustering. Combining multiple features (directional, grapheme, and run-length PDFs) yields increased writer identification and verification performance. The proposed methods are applicable to free-style handwriting (both cursive and isolated) and have practical feasibility, under the assumption that a few text lines of handwritten material are available in order to obtain reliable probability estimates.

Index Terms—Handwriting analysis, writer identification and verification, behavioral biometrics, joint directional probability distributions, grapheme-emission probability distribution.

An extract from PAMI

The writer identification experiments reported in this paper always involved two samples per writer: One was used as the query, while the other one represented the correct hit that the system was supposed to find in the database. Having more samples per writer enrolled in the database increases the chance of finding in the top positions of the hit list the correct author for a given query. We have run writer identification tests on the original IAM database that included at least three samples per writer for about a quarter of the total of 650 writers incorporated in the set. For the best performing feature combination, *f2* & *f4* & *f5*, we obtained writer identification rates of Top-1 92 percent and Top-10 98 percent. These values exceed the identification rates obtained on our modified IAM set that always contained only two samples per writer (see Table 4).

In another study performed on a subset comprised of 100 writers from the Firemaker data set, our methods largely outperformed two actual systems used in current forensic practice [9]. The use of automatic and computation-intensive approaches in this application domain will allow for massive search in large databases, with less human intervention than is current practice. By reducing the size of a target set of writers, detailed manual and microscopic forensic analysis becomes feasible. In the foreseeable future, the toolbox of the forensic expert will have been thoroughly modernized and extended. Part of our directional texture-level features have already been included in real-life applications.

REFERENCES

[1] R. Plamondon and S.N. Srihari, "Online and Offline Handwriting Recognition: A Comprehensive Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63-84, Jan. 2000.

[2] A. Vinciarelli, "A Survey on Offline Cursive Word Recognition," *Pattern Recognition*, vol. 35, pp. 1433-1446, 2002.

[3] R.A. Huber and A. Headrick, *Handwriting Identification: Facts and Fundamentals*. CRC Press, 1999.

[4] R. Morris, *Forensic Handwriting Identification: Fundamental Concepts and Principles*. Academic Press, 2000.

[5] F.J. Maarse, "The Study of Handwriting Movement: Peripheral Models and Signal Processing Techniques," PhD dissertation, Dept. of Experimental Psychology, Univ. of Nijmegen, The Netherlands, 1987.

[6] F. Maarse and A. Thomassen, "Produced and Perceived Writing Slant: Differences between Up and Down Strokes," *Acta Psychologica*, vol. 54, nos. 1-3, pp. 131-147, 1983.

[7] L. Schomaker, "Simulation and Recognition of Handwriting Movements: A Vertical Approach to Modeling Human Motor Behavior," PhD dissertation, Univ. of Nijmegen, The Netherlands, 1991.

[8] R. Schmidt, "A Schema Theory of Discrete Motor Skill Learning," *Psychological Rev.*, vol. 82, pp. 225-260, 1975.

[9] L. Schomaker and M. Bulacu, "Automatic Writer Identification Using Connected-Component Contours and Edge-Based Features of Uppercase Western Script," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 787-798, June 2004.

[10] R. Plamondon and G. Lorette, "Automatic Signature Verification and Writer Identification—The State of the Art," *Pattern Recognition*, vol. 22, no. 2, pp. 107-131, 1989.

[11] H. Said, T. Tan, and K. Baker, "Personal Identification Based on Handwriting," *Pattern Recognition*, vol. 33, no. 1, pp. 149-160, 2000.

A full English sentence database for off-line handwriting recognition

U.-V. Marti and H. Bunke
Institute of Informatics and Applied Mathematics
University of Bern, Neubrückstrasse 10, CH-3012 Bern, Switzerland
E-Mail: {marti,bunke}@iam.unibe.ch

Abstract

In this paper we present a new database for off-line handwriting recognition, together with a few preprocessing and text segmentation procedures. The database is based on the Lancaster-Oslo/Bergen(LOB) corpus. This corpus is a collection of texts that were used to generate forms, which subsequently were filled out by persons with their handwriting. Up to now (December 1998) the database includes 556 forms produced by approximately 250 different writers. The database consists of full English sentences. It can serve as a basis for a variety of handwriting recognition tasks. The main focus, however, is on recognition techniques that use

linguistic knowledge has been successfully applied [6, 5]. However, in handwriting recognition many systems have been developed for isolated characters, where no contextual knowledge is available at all [14, 15]. In other applications, such as bank check or postal address reading, contextual constraints can be applied, but only to a limited degree. Only few works in handwriting recognition address applications where broad linguistic knowledge is applicable. An example is the work by Kim, Govindaraju and Srihari [10] or of Oh, Ha and Kim [12] on unconstrained English sentence recognition. Our own research is directed towards similar goals [11]. The problem under study is the application of linguistic knowledge beyond the word level in the

An extract from a conference paper

Literature search

- Literature search is the process of querying the scholarly literature databases in order to gather research manuscripts related to the phenomenon under investigation.

- The major contributions are likely to be in the leading journals – It makes sense, therefore, to start with them

Where to look for….

- Google Scholar
  - http://www.scholar.google.com
- ACM Digital Library
  - http://www.acm.org/dl
- IEEE Digital Library
  - http://www.computer.org/publications/dlib/
- NEC CiteSeer
  - http://citeseerx.ist.psu.edu
- Online Libararies

Google Scholar

- Google Scholar provides a simple way to search for scholarly literature.

- Search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations.

Google Scholar

Web Images Vidéos Maps Actualités Gmail plus

Google scholar optical character recognition Rechercher Recherche avancée Scholar

Rechercher sur le Web Rechercher les pages en Français

Scholar date indifférente inclure les citations Créer une alerte par e-mail

Conseil : Recherchez des résultats uniquement en **français**. Vous pouvez indiquer votre langue de recherche sur la page Préférences Scholar..

[LIVRE] **Optical character recognition**
S Mori, H Nishida... - 1999 - en.scientificcommons.org
Publikationsansicht. 6945782. **Optical Character Recognition** (1999). ...
Cité 105 fois - Autres articles - En cache - Les 3 versions

Optical character recognition
WD Barber, TM Cipolla... - US Patent 4,339,745, 1982 - Google Patents
United States Patent [19] [ii] 4,339,745 Barber et al. [45] Jul. 13, 1982 [54] **OPTICAL CHARACTER RECOGNITION** [75] Inventors: William D. Barber, Jonesville; Thomas M. Cipolla, Ballston Lake; Joseph L. Mundy, Schenectady, all of NY 4,119,947 10/1978 4,175,236 11/1979 ...
Cité 14 fois - Autres articles - Les 2 versions

Optical character recognition
F Ghazizadeh - US Patent 5,077,809, 1991 - Google Patents
United States Patent [19] [n] Patent Number: 5,077,809 Ghazizadeh [45] Date of Patent: Dec. 31,1991 [54] **OPTICAL CHARACTER RECOGNITION** [76] Inventor: Farshad Ghazizadeh, 2267 Los Padres Blvd., No. B, Santa Clara, Calif. 95050 [21] Appl. No.: 358,955 [22] Filed: May ...
Cité 11 fois - Autres articles - Les 2 versions

OPTICAL CHARACTER RECOGNITION
US Patent 3,437,824, 1969 - Google Patents
United States Patent Office 3,437,824 Patented Apr. 8, 1969 1 3,437,824 **OPTICAL CHARACTER RECOGNITION** Adolf W. Lohmann, San Jose, Calif., assignor to Inter- national Business Machines Corporation, New York, NY, a corporation of New York Filed Oct. 16,1964, Ser. No. ...
Cité 9 fois - Autres articles - Les 2 versions

Google face recognition

Most of the time

Recherche avancée

Recherche

Environ 48 800 000 résultats (0,08 secondes)

Tout
Images
Maps
Vidéos
Actualités
Shopping
Livres
Plus

Recherche sur le Web
Rechercher les pages en Français
Pages en langue étrangère traduites

Date indifférente
Moins d'une heure
Moins de 24 heures
Moins d'une semaine
Moins d'un mois
Moins d'un an
Période personnalisée

Conseil : Recherchez des résultats uniquement en **français**. Vous pouvez indiquer votre langue de recherche sur la page Préférences.

MyHeritage.com **face recognition** - Find the Celebrity in You ... - [Traduire cette page]
www.myheritage.com/celebrity-**face-recognition** - En cache
Enjoy a free family site on MyHeritage.com where you can build a family tree and share photos. Get birthday alerts from your family site and never miss any ...

Face Recognition Homepage - [Traduire cette page]
www.**face**-rec.org/ - En cache
11 Aug 2011 – **Face Recognition** Homepage, relevant information in the the area of **face recognition**, information pool for the **face recognition** community, entry ...
Databases - Source Codes - Algorithms - General Info

Facial recognition system - Wikipedia, the free encyclopedia - [Traduire cette page]
en.wikipedia.org/wiki/**Facial_recognition**_system - En cache
A **facial recognition** system is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. ...

Images correspondant à **face recognition** - Signaler des images inappropriées

Literature search

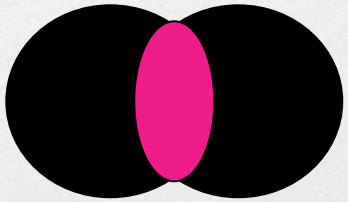
- Search Strategies
  - Keyword searching
  - Backward searching
  - Forward searching

Keyword search

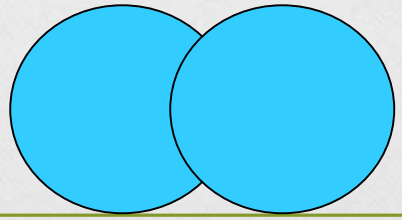
- Querying of scholarly databases by the use of a specific word or phrase (i.e. "keyword") when attempting to find relevant literature
- Keyword search should be just the initial, not the main step for a literature search
- Keyword search does not yield all that is available from the literature

Keyword search

- Not too broad
- Not too narrow
- Think of all possibly synonyms
- Make sure to follow the search engines 'rules' in combining the key words
- AND = Narrow



- OR = Expand



Backward search

- Using an effective keyword search will produce some initial insight

- *Backward literature search*
  - *Backward references search*
  - *Backward authors search*
  - *Previously used keywords.*

Backward search

- Backward References Search
  - Reviewing the references of the articles yielded from the keyword search
  - Example
    - You enter a keyword and download the first three articles
    - Download and review the references listed in each of these articles
  - A *second level backward references search* can also be done by pulling the '*references of the references*'.

Backward search

- Backward Authors Search
  - Reviewing what the authors have published prior to the article.
  - May search through author name
  - Visit authors home page etc.

Backward search

- Previously used keywords
  - Reviewing the keywords noted in the articles yielded from the keyword search
  - For example after reviewing the following article:

  - You may then look for Text-Independent or Text-Dependent writer identification

Forward search

- Can be divided into two specific sub-steps:
  - *Forward references search*
  - *Forward authors search*

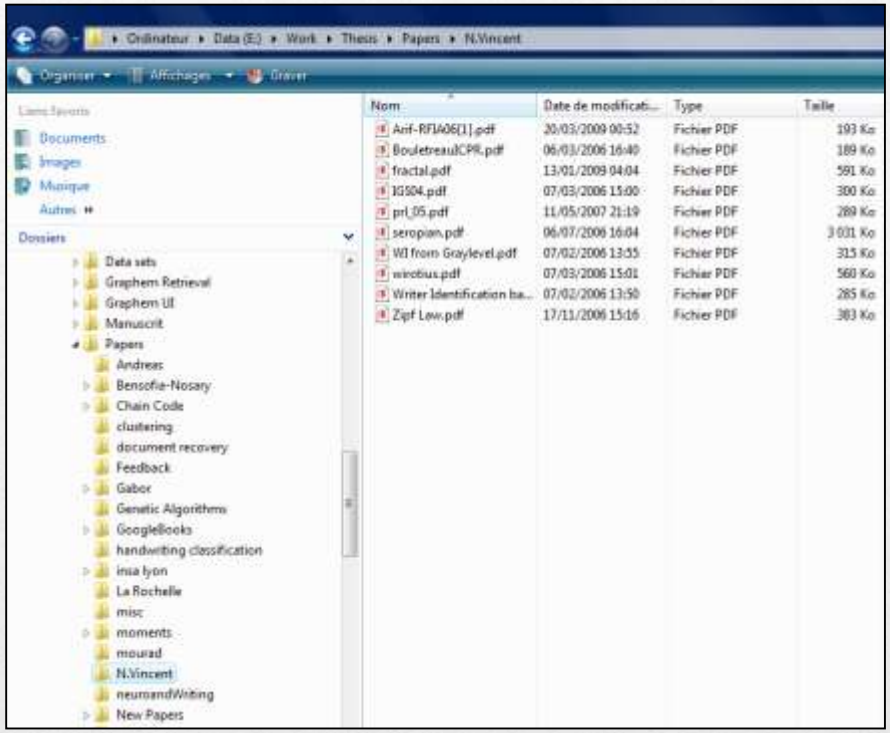
- Forward references search
  - Reviewing additional articles that have cited the article
- Forward authors search
  - Reviewing what the authors have published following the article

Information management

How will you manage the material you download?

Information Management

- Select a strategy for organizing information
- Folders?

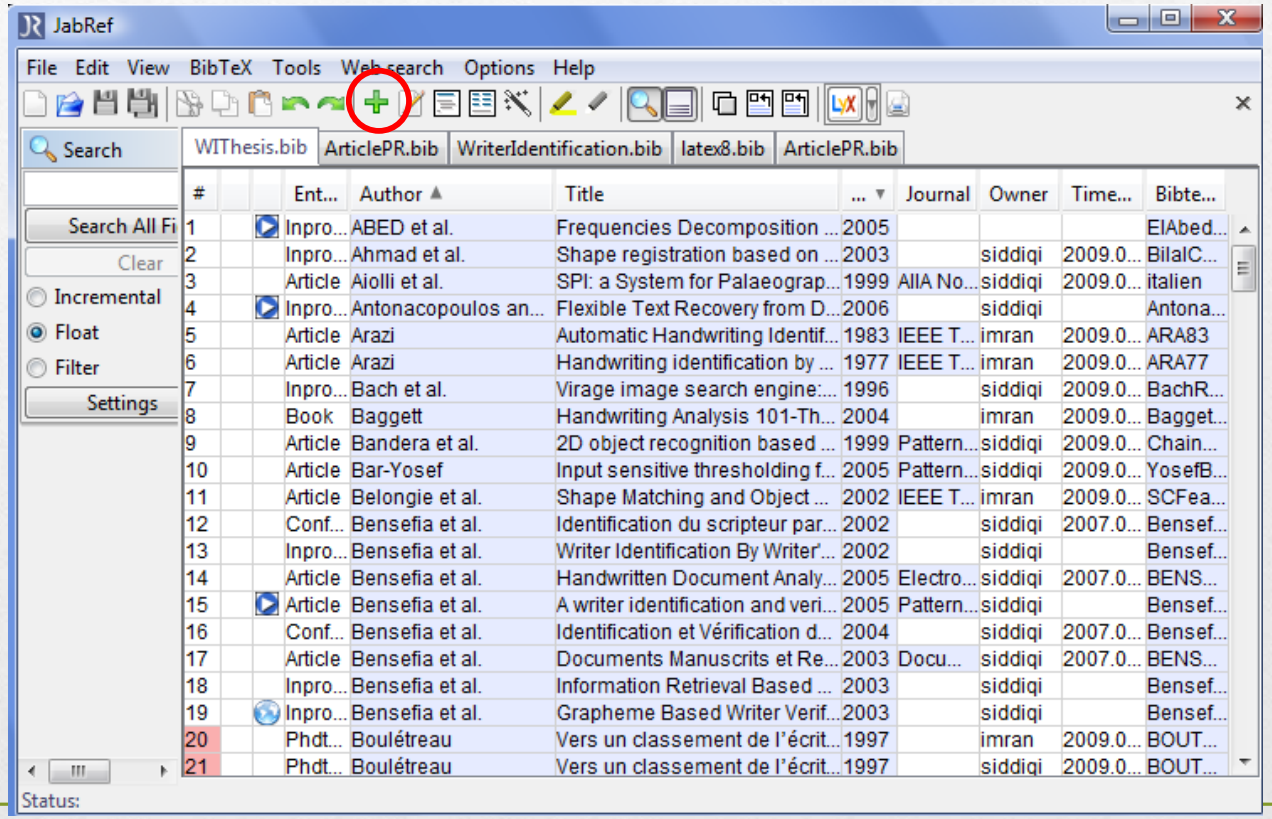


Digital

Physical

Information management

- Bibliography Managers
  - JabRef – Free



Reviewing the literature

Types of reviews

- **Systematic reviews:**
  - Use explicit and rigorous methods to identify, critically appraise, and synthesize relevant studies
- **Narrative reviews**:
  - Summaries of research
- **Meta-analysis:**
  - Quantitative review using statistical analyses(involves the quantifying of data with the help of some form of statistical analysis.)

Caution

Do NOT collect 100 papers before you start reading

Reviewing the literature

- Three phases to reading a paper:
- **First:** See if there's anything of interest– Scan the title, then the abstract, then -- if you haven't completely lost interest already -- glance at the introduction and conclusions.
- **Second:** Find the part that has the good stuff– Most fifteen page papers could profitably be rewritten as one-page papers; you need to look for the page that has the exciting stuff.
- **Third:** Go back and read the whole paper through if it seems worthwhile

Reviewing the literature

- Read with a question in mind
  - How can I use this? Does this really do what the author claims? What if...?
  - Most of the understanding is in figuring out the motivations, the choices the authors made
  - Whether the assumptions and formalizations are realistic
  - What directions the work suggests, the problems lying just over the horizon, the patterns of difficulty that keep coming up in the author's research program etc.

Reviewing the literature

- When you read a paper that excites you, make few copies and give them to people you think will be interested in it.
- Whenever you write something yourself, distribute copies of a draft to interested people.
- Conferences – How to meet people? Walk up to someone whose paper you've liked, say "I really liked your paper", and ask a question.

Reviewing the literature – levels of mastery

- Levels of mastery
  - Know the literature
  - Comprehend the literature
  - Analyze the literature
  - Synthesize the literature
  - Evaluate the literature

Know the literature

- Knowledge level is commonly demonstrated by activities such as *listing, defining, describing*

- At the very least, the researcher must demonstrate that he or she has read the article and extracted meaningful information from it

Comprehend the literature

- Comprehension is demonstrated by activities such as *summarizing, differentiating and interpreting,.*

- Researcher demonstrates that not only can he or she repeat what was included in the article but also knows the meaning and significance of the information being reported.

Comprehend the literature

Han and Kamber (2001) suggest an evolution that moves from data collection and database creation, towards data management, and ultimately, data analysis and understanding.

Pre-comprehension level

Han and Kamber (2001) suggest an evolution that moves from data collection and database creation, towards data management, and ultimately, data analysis and understanding. For example, *data processing* is a base function enabling manipulation and aggregation of data, thus facilitating searching and retrieval.

Comprehension level

Analyze the literature

- Analysis is demonstrated by activities such as *relating, and classifying and explaining* (why the information presented is important)
- Example1
  - Identify the major concepts relevant to the study
  - Place the citation in the correct category

	Concept 1	Concept 2	…	Concept n
Article 1	X			X
Article 2		X		
…			X	X
Article n		X	X	

Analyzing the literature

- Example 2

Data mining is the analyzing and interpretation of large amounts of information. Through analyzing vast amounts of data it is possible to find patterns, relationships and from these discoveries it is possible to make correlations (Chen & Liu, 2005).

Knowledge without analysis

Why it would be of any interest or value to find patterns and relationships in order to make correlations.

Data mining is a process of discovering new knowledge by using statistical analysis to identify previously unsuspected patterns and clustering in large data sets (Chen & Liu, 2005).

Analysis level mastery

Synthesizing the literature

- Synthesis entails activities such as *combining, integrating, rearranging and generalizing*

- The essence of synthesis is to assemble the literature being reviewed for a given concept into a whole that exceeds the sum of its parts.

Synthesizing the literature

The *Digital Object Identifier (DOI)* is an Internet-based system for global identification and reuse of digital content (Paskin, 2003). It provides a tracking mechanism to identify digital assets (Dalziel, 2004). The DOI is not widely employed across LOR and databases and is not universally adapted by content owners (Nair & Jeevan, 2004). The DOI does not provide provision for assets to be tagged with copyright information (Genoni, 2004).

Lack of synthesis – facts are presented as almost a series of isolated "bullet points"

One current DRM initiative, the *Digital Object Identifier (DOI)*, is an Internet-based system for global identification and reuse of digital content, and provides a tracking mechanism to identify digital assets (Paskin, 2003; Dalziel, 2004). However, despite being integrated in learning object technologies, this DOI is not widely employed across LOR and databases, nor is it universally adapted by content owners (Nair & Jeevan, 2004). Similarly, while most metadata schema enables assets to be tagged with copyright information, this method lacks technological enforcement (Genoni, 2004).

Synthases level mastery

Evaluating the literature

- Evaluation includes activities such as *assessing, deciding, recommending, selecting, judging, supporting, and concluding*.

Writing the literature Review

Writing the literature review

- Write as you read…….
  - Write a brief summary of each paper you read
  - List your findings, critics etc.
  - This will help you in writing the literature review

Writing the literature review

- **Find Models** – Look for other literature reviews in your area of interest and read them to get a sense of the types of themes you might want to employ to organize your final review.

- You can simply put the word "review" in your search engine along with your other topic terms to find articles of this type on the Internet or in an electronic database.

- The bibliography or reference section of sources you've already read may mention such review articles as well.

Organization

- Develop an organization – Global as well as local levels
- Atleast three basic elements in a review
  - **Introduction** – Central theme or organizational pattern
  - **Body** - Discussion of sources that is organized either chronologically, thematically or methodologically
  - **Conclusions/Recommendations** – What you have drawn from reviewing literature.

Organizing the sources

- Chronological
  - Arranging events in their order of occurrence in time
  - By Publication
    - Order your sources by publication chronology – i.e. the order in which they are published
  - By Trend
    - Subsections according to eras/trends
    - For example – If you are doing research on handwriting you may distinguish the following periods
      - *Proto-writing 4000-3000 BC*
      - *Bronze Age writing 3000 – 1500 BC*
      - *Iron Age writing 1000 BC – 500 AD*
      - *Medieval writing – 500 – 1500 AD*
      - *Modern writing*

Organizing the sources

- Thematic
  - Organized around a topic or issue, rather than the progression of time
  - Progression of time may still be an important factor in a thematic review
  - Example
    - Handwriting recognition
      - Online or Offline
      - Arabic or English

Organizing the sources

- Methodological
  - Focuses on the proposed methodologoy
  - Example – Handwriting recognition
    - Segmentation based approaches – first segment into characters and then reconize
    - Segmentation free approaches – recognize the complete words

Writing the literature review

- Bringing other writers into your writing – Keep your own voice
  - Ideas are of other people – start and end each paragraph with your own words
- Rephrasing in our own words & adding an acknowledgement
  - *Summary, Paraphrasing – Beware of plagarism*
  - *Common norm in science and engineering*
- Direct quotation and acknowledgement
  - *Less common in science and engineering*

Writing the literature review

- Integrate other studies into your own text, rather than read as a series of disconnected voices of other researchers
- Avoid narrative accounts – Find a way of grouping studies
- For each study discussed
  - *General idea of study*
  - *Strengths and weaknesses*
  - *If relevant, relationship with to the present study*

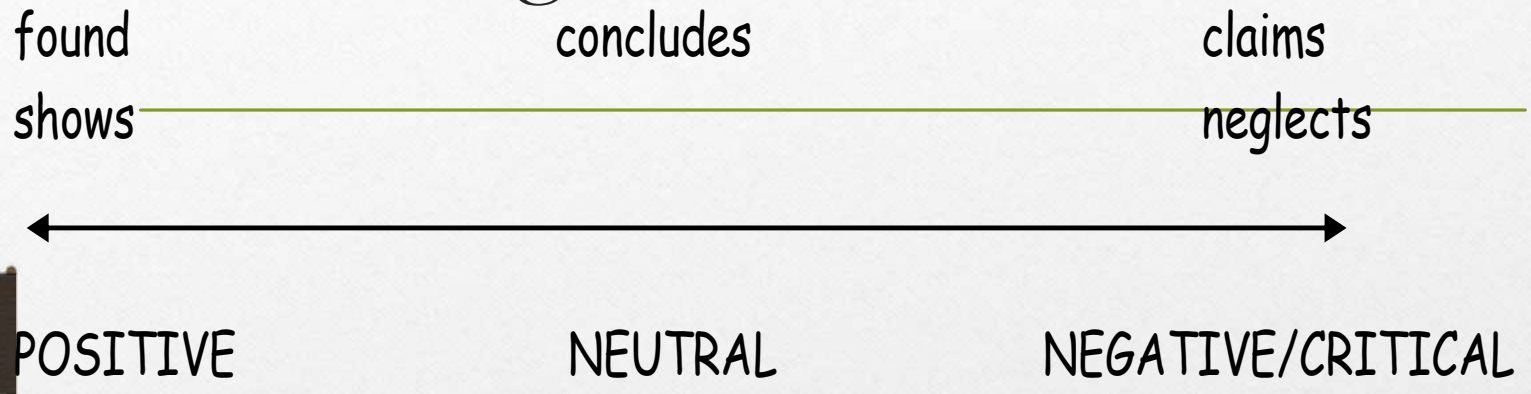
acknowledges	criticises	explores	observes	shows
admits	deals with	expresses	points out /to	signals
agrees	decides	finds	predicts	states
alleges	defines	focuses	presumes	studies
argues	demonstrates	highlights	proves	suggests
assumes	denies	hypothesises	proposes	tries to identify
believes	depicts	identifies	provides evidence for	sums up
challenges	determines	implies	questions	underlines
claims	diagnoses	indicates	recognises	views
classifies	discovers	infers	reports	wonders
comments	doubts	interprets	reveals	
concentrates on	emphasises	makes the point	says	
concludes	establishes	maintains	seeks to explain	
considers	explains	notes	seeks to identify	

Verbs for reporting other scientists' findings

Writing the literature review

- Position towards the sources
  - Reporting verb indicates your position towards the cited source
  - Selecting a particular verb involves taking a particular position in relation to other scientists' ideas
  - Strong level of agreement – strongly negative stance

Writing the literature review



- **Show** can be seen as positive as it reports an observation or finding as a proven fact

- **Claims** disassociates the writer from the position of the author cited. This allows the writer to establish a critical perspective and follow with a counterargument.

Listing references

[1] Journal, [2] Conference, [3] Thesis, [4] Book, [5] Web

[1] Hirschorn, R.M. and Miller, G. Control of nonlinear systems with friction, IEEE Trans. on Control System Technology, Vol.7, No.5, Sept. 1999.

[2] Whitfield, A. and Wallace, F.J. Study of incidence loss models in radial and mixed-flow turbomachinery, ***Proc. Cong. Heat Fluid Flow in Steam and Gas Turbine Plant***, Univ. Warwick, Coventry, UK, April 1973, pp 122-32.

[3] Murray, F. ***Time Series Forecasting Methodologies for Electricity Supply Systems***, PhD Thesis, Dublin City University, 1997.

[4] Kreyszig, E. ***Advanced Engineering Mathematics (7th Ed.)***, Wiley, 1993.

[5] Ringwood, J. and Galvin, G. ***Artificial Neural Networks - An Introduction***, http://www.eeng.dcu.ie/~annet/ , Accessed 5th November 2010.

Better Use Latex

Open Question

The program of an upcoming conference has just been announced, and there is a paper that will be presented that has a title which bears a remarkable resemblance to your ongoing thesis work. Now what ?

The material in these slides is based on the following resources.

References

- Research method in Computer Science, Jiaheng Lu, Renmin University of China
- Literature Search in Computer Science, Michiel van de Panne
- Literature Reviews, University of North Carolina.
- A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research, Yair Levy and Timothy J. Ellis. Information Science Journal, Vol 9, 2006
- How to do a literature review: An overview, Frederic Murray.