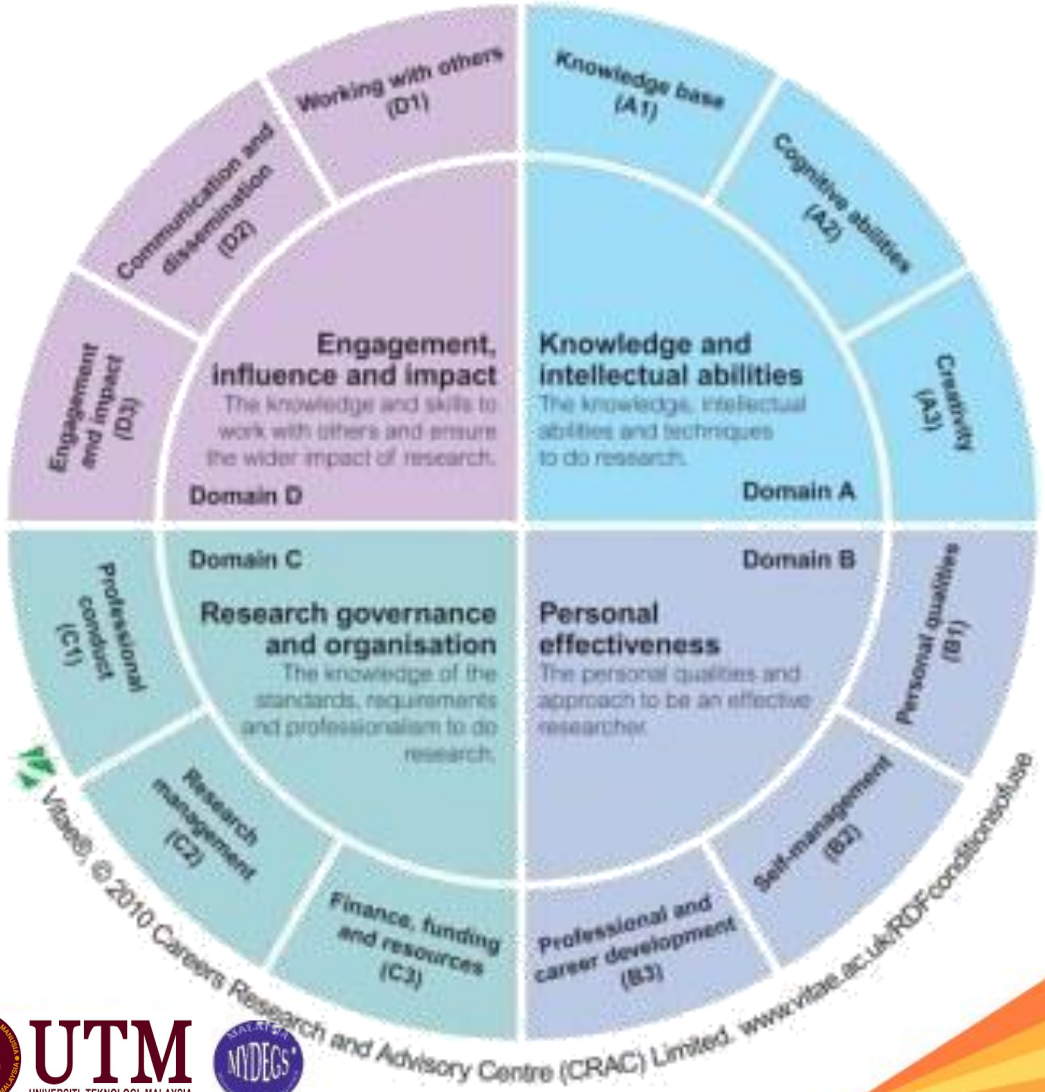
MENTORING FOR "GRADUATE-ON-TIME (GOT)"

Professor Dr. Naomie Salim
Faculty of Computing
Universiti Teknologi Malaysia

Outline of Talk

- The challenges to GOT
- Mentoring for GOT throughout the PhD research life-cycle
  - Preparation phase
  - Problem definition phase
  - Literature review phase
  - Methodology design phase
  - Experiment and data collection phase
  - Analysis of results phase
  - Writing-up phase

PhD - Training & Developing Competent Researchers



Problem Statement

High percentage of non-completion

Long process towards completion

Candidates regard PhD period as lonely and stressful episode, at high personal cost, and without 'mentor' to discuss problems;

Meetings have inadequate frequency and depth, often no regular performance and progress interviews;

Completed PhDs have low scientific and social impact, wasteful research time and money.

Sources of Problems

- Poor research design, no focus, no adequate research question;
- Lack of realistic expectations (‘targeting the sky’);
- Inadequate research background; lack of training in methodological and writing skills (inadequate Bachelors and Masters training);
- Problematic research facilities (time, office, computer, assistance, money, flexibility in rules, under-funding of essential tasks);
- Many competing tasks (teaching, consultancies, family life);
- Bad planning, bad phasing, bad time management;

Sources of Problems (cont.)

- Major problems with writing academic English;
- Negligent or inadequate supervision; often unclear, strained relationship;
- Inadequate networking: not aware of others working in the same field of studies, no contact with peers, parochial local research cultures;
- Breakdown of motivation, psychological stress due to isolation, feelings of uselessness;
- Lack of possibilities to participate in a vibrant research culture.

Different Student Needs for Supervision

- **Immature, unconfident student - supervisor as "big daddy."**
  - student may have requisite tools and skills to do research, but ability to work independently not well developed.
- **Somewhat mature, somewhat confident student – supervisor as "mentoring" colleague**
  - Preparatory work helped student be somewhat confident, but the student still needs moderately detailed direction to get going
- **Very mature, confident student - supervisor as a "senior" colleague**
  - Student recognizes need for guidance and supervision, but the need is at a fairly general level.

Supervision Styles- Strong Master/apprentice style

- Supervisor is the master, student works as an apprentice on problems selected by master
- Assuming a competent master, the advantages:
  - **significant reduction in the time** to formulate a problem,
  - strong guidance and direction in doing the work,
  - development of **specific skills for the type of problem** with the master.
- Severe disadvantages:
  - may **not develop an ability to formulate research** and conduct it independently.
  - research methods may be limited.
  - focus on the master's problem may becloud the apprentice student's understanding.
- Fits best the condition of
  - a relatively immature, inexperienced student who needs strong direction.
  - well-defined, funded streams of research.

Supervision Styles- Collegial master/apprentice style

- Limited domain advising, not restricted by advisor's current research activity but fit within the general domain of expertise.
- Supervisor willing to advise on problems that are within the scope of his research and methods within his skill set.
- Puts more responsibility on student than master/apprentice style
- Work well when
  - both supervisor and student were interested in a problem
  - supervisor had sufficient expertise to provide good guidance.

Supervision Styles – Collegial development style

- Extended domain advising, not in domain of supervisor's current or past research but is extended to areas in which the supervisor has an interest and willing to invest in becoming reasonably proficient.
- There is a joint learning experience
  - Supervisor starts with more experience, but both are learning details of research area.
- Fits when
  - Supervisor willing to expand his or her research competence
  - Student willing to engage in a joint learning experience.
- Fits very well for a dissertation that opens up a new or fairly new area of research.
- Fail when supervisor was not willing to make the investment to be competent.

Supervision Styles – Guidance and Suggestion style

- General advising over a range of problem domains.
- Some supervisors have good skills at problem identification and problem formulation over a range of problems and research methods.
  - They conceptualize well and are good at “sense making.”
- Works best with students who are willing and able to take initiative and take responsibility for learning the research domain and the appropriate research methods.
- Student gets good general guidance and good evaluation of the dissertation but usually does not get detailed feedback and detailed mentoring of methods.
- It is not very good for immature students who need more detailed guidance.
- Good with mature students who took initiative.

Supervision Styles – Passive hands-off style

- Laissez faire style, the supervisor takes the role of a general quality control reader.
- Student must take the initiative to define a problem, decide on a research method, develop a research plan, and so forth.
- Supervisor responds to student plans and initiatives with some suggestions, but the responsibility is almost entirely with student.
- Given a competent supervisor who gives good suggestions in response to student initiatives and plans, the advantages are that the student develops independent skills at formulating problems and planning research.
- The disadvantages are that the student may meander from problem to problem and take too long to do a dissertation. Under these conditions, a student may not develop good skills and may drop out of the program.
- For fairly mature students with an ability to take initiative, this style may work well.
- Significant danger with project for which the student does not have the necessary background for doing a good dissertation or the supervisor is unable to do reasonable quality review.
- For immature students, it is likely to be a disaster.

Supervision Style

Style	Advisor Role and Behavior	Student Role and Behavior
Strong master/apprentice style	Advisor is master. Advisor has a well specified domain of expertise and set of problems within it.	Student is an apprentice working for the advisor. Student works on advisor's problems.
Collegial master/apprentice style	Advisor is expert who limits advising to problems that are within scope of his or her research skill set but will work on student's problem.	Student develops a problem within advisor's domain and skills and works under the advisor to develop the research plan and procedures.
Collegial development style	Advisor is senior colleague who will respond to student research problem and extend his or her advising domain to include new problems and new skills.	Student takes initiative to introduce new problem that requires new skill set and works as a junior colleague with advisor in joint development of new domain.
Guidance and suggestion style	Advisor is a senior colleague who gives good general guidance over a wide range of problems and methods but does not have personal skill in all of them.	Student is an independent, junior colleague who takes initiative for presenting problems and research plans for discussion and guidance. Student develops required skills.
Passive hands-off style	Advisor has quality control role and responds only to requests or documents and performs only general quality control review	Student is an independent researcher who takes initiative for developing problem, developing skills, and presenting research plans for general review and approval.

Advantages & Disadvantages of Styles

Advising Style	Advantages	Disadvantages
Strong master/apprentice style	Advisor is heavily involved and gives expert direction for research activities. Student learns how to do research within advisor's domain.	Student works on advisor's problems and within advisor expertise and may not develop independence.
Collegial master/apprentice style	Advisor knows the research domain and research methods and can give expert advice on them. Student can take initiative in formulating problem and working with advisor.	Student is limited to advisor research domain and advisor research skills. Student may do work that is not within his or her long term research plan.
Collegial development style	Advisor and student develop together to explore new domain and new research methods. Student develops independence within relationship.	Risk of exploring new research area that does not work. Risk that necessary development of both advisor and student does not occur or occurs unevenly.
Guidance and suggestion style	Student is able to develop independent research and research management skills while receiving guidance and suggestions. The student may research a broad range of topics and employ broad range of methods.	Student may not get expert advice from advisor on many issues, so student must search for expert advice. Student has significant responsibility for research quality and management of process.
Passive hands-off style	Student is able to act independently with little interference from advisor. Student can work on problems of his or her choosing.	Student may make serious mistakes because of lack of advice and suggestions. Student may flounder and not complete on timely basis.

Supervision Style and Fit

Advising Style	Likely Fit with Student Archetypes
Strong master/apprentice style	Good fit with immature, unconfident student. "Do what I do" provides clear direction. Possible good fit with somewhat mature, somewhat confident student, but the style creates some stress between heavy direction of advisor and initiative and independence of somewhat mature student. Likely stress and conflict between this style and the independence of a very mature, confident student
Collegial master/apprentice style	Some stress but a workable fit with immature, unconfident student because some initiative is required. Fairly good fit with somewhat mature, somewhat confident student because it provides boundaries for what is expected and allows some independence. Some stress from this style for mature, confident student because of constraints on what can be done by student.
Collegial development style	Stressful for immature, unconfident student because of need for significant student initiative. Good fit with somewhat mature, somewhat confident student because it builds confidence through development interactions. Reasonable fit and reasonably low stress for very mature, confident student if advisor and student are compatible relative to problems and methods.
Guidance and suggestion style	Very stressful for immature, unconfident student because of vagueness of process and need for initiative that may exceed capacity of novice. Stressful but workable relationship for somewhat mature, somewhat confident student because of high initiative required from student. Good fit and reasonably low stress fit for very mature, confident student who is given much freedom and good feedback.
Passive hands-off style	Likely disaster for advising relationship with immature, unconfident student because not sufficient guidance. Stressful relationship between advisor with this style and somewhat mature, somewhat confident student because of insufficient feedback. This style may work for a very mature, confident student but introduces risks because of lack of clarity in expectations.

Helping Students Prepare for a GOT PhD Journey

PROGRAM GOT & U

Successful, Timely Completion of PhD

- Help in Analysis and Writing
- Coach for Mastery in Research Methods
- Help Plan for GOT
- Coach for Deep & Critical Comprehension of Subject Matter
- Help Identify Interesting & Suitable Research Topic
- Prepare Researcher Mindset
- Set-up Supportive Research Environment

SOARING UPWARDS

UTM UNIVERSITI TEKNOLOGI MALAYSIA

MYDEGS

Setting-Up of a Supportive Research Environment

Don’t Slow Your Student because of Unconducive Research Environment

Setting Up A Supportive Research Environment

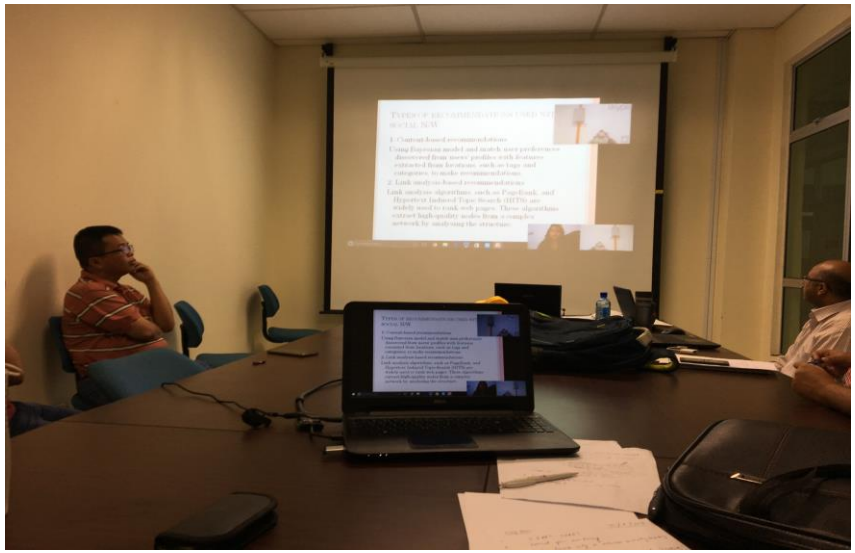
- Build systematic datasets/gold standards
- Have organized, well-documented standard programs (pre-processing, benchmarks)
- Have well-administered servers & tools
- Provide non-research support
- Develop mentoring system
- Build a strong, cohesive and family-like research group

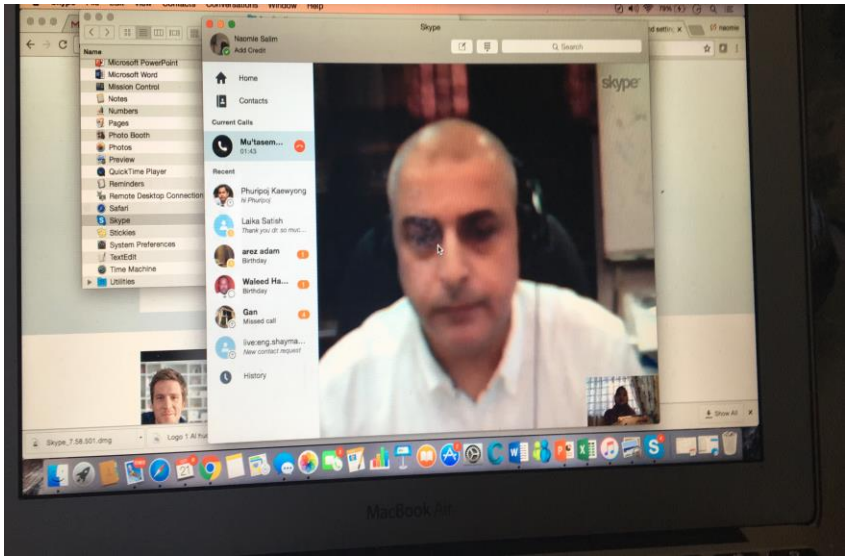
How to Help Student to Publish (cont.)

- Schedule regular presentations to group – to be developed into publication
- Monitor progress (eg. every 2 months) through paper
- Set appointment by draft paper
- Give fast feedback
- Suggest and ready to pay for proof-reading of papers
- Have list of journals and reviewers for papers
- Suggest collaborator(s)

Monitoring Attendance in WA

Remote PhD




Skype
Naomie Salim
Add Credit
Home
Contacts
Current Calls
Recent
History
skype

Mindset Preparation

Wrong mindset, and it takes forever to finish PhD

Mindset Preparation

- Matured
- Ability to focus and concentrate
- Disciplined
- Independent
- Hardworking
- Innovative & Creative
- Critical Thinking
- Available (time)

Help Them Understand Characteristics of Research

- Systematic
  - Work in a sequence of steps which were in order and thus systematic
  - Order: Observation, Problem Definition, Hypothesis, Testing, Conclusion
- Follows a scientific method of enquiry
  - Researcher should not just jump at the conclusions,
  - But used a scientific method of enquiry in reaching conclusion

Emphasize What Research is NOT

- Research is not information gathering
  - Gathering information from resources such as books or magazines.
- Research is not the transformation of facts
  - No contribution to new knowledge although this might make knowledge more accessible
- Research is not about having successfully developed something
  - What can others learn from it?

Helping Student Find Good Research Topic

Never underestimate the importance of choosing the right topic

Finding A Good Research Problem

- Novelty of the Idea.

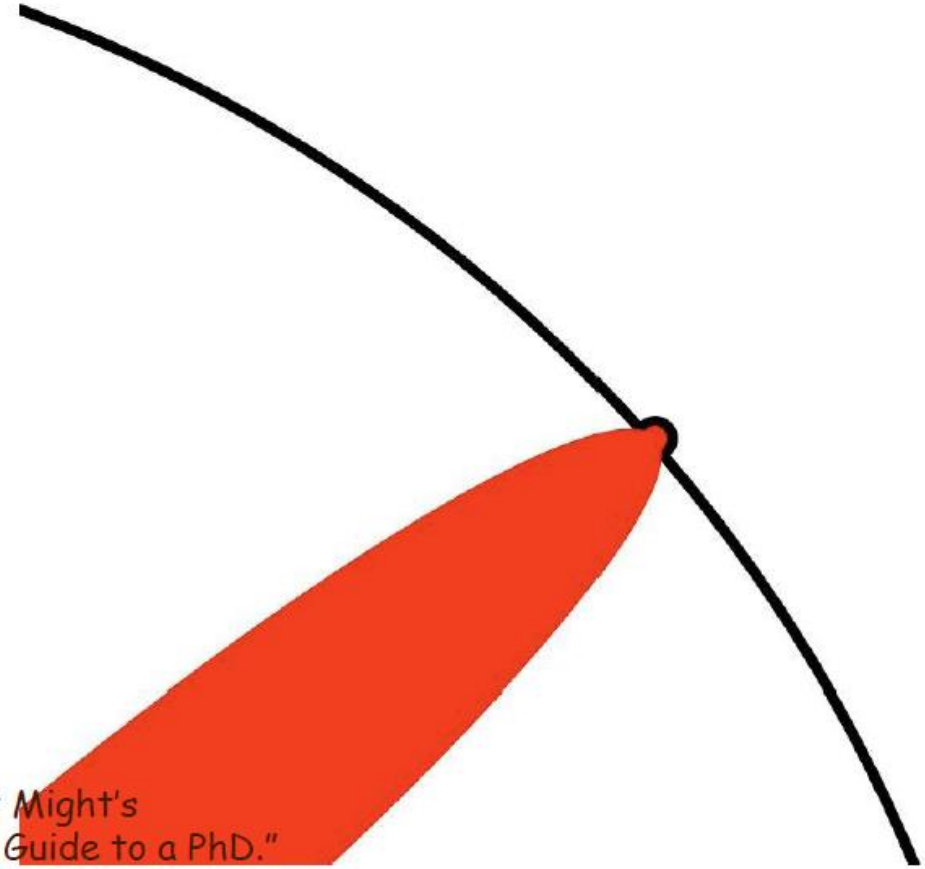
  Research is a study of new ideas in the field
- Significance for the Community.

  What idea is actually needed for the community “today”.
- Contribution from the Researcher.

  An amount of efforts made by a researcher to study the idea.

Novelty:

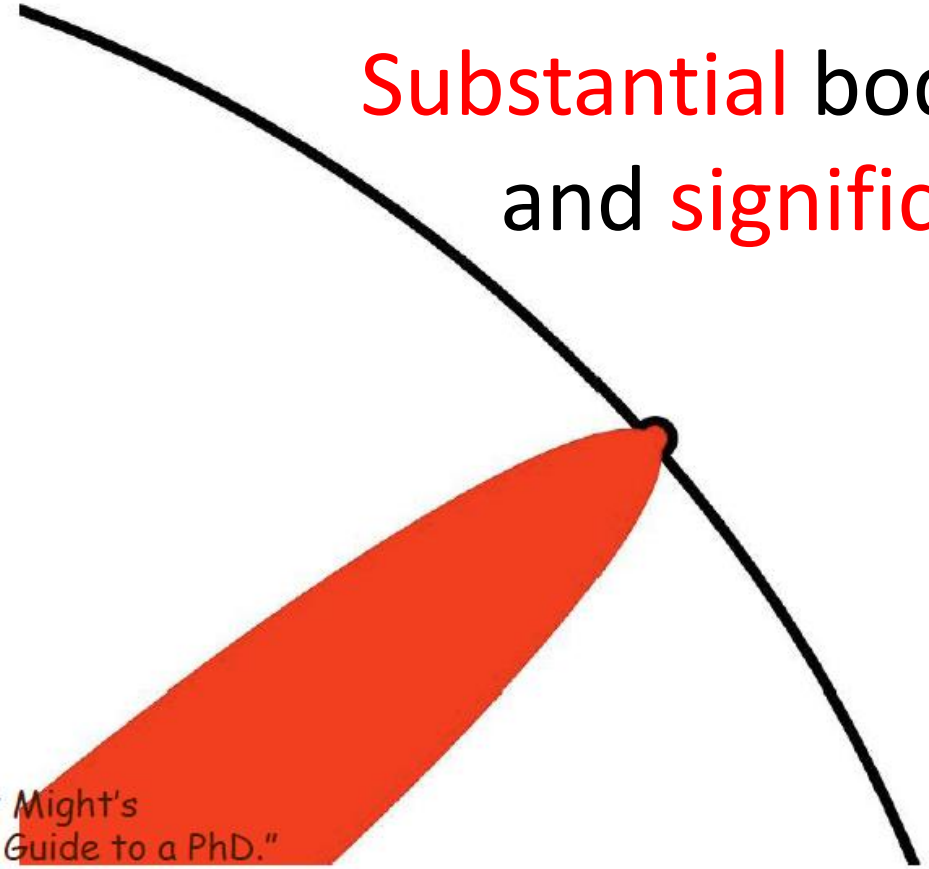
What to Expect of a PhD Research



From "Matt Might's
Illustrated Guide to a PhD."

What to Expect of a PhD Research

Substantial body of original and significant work



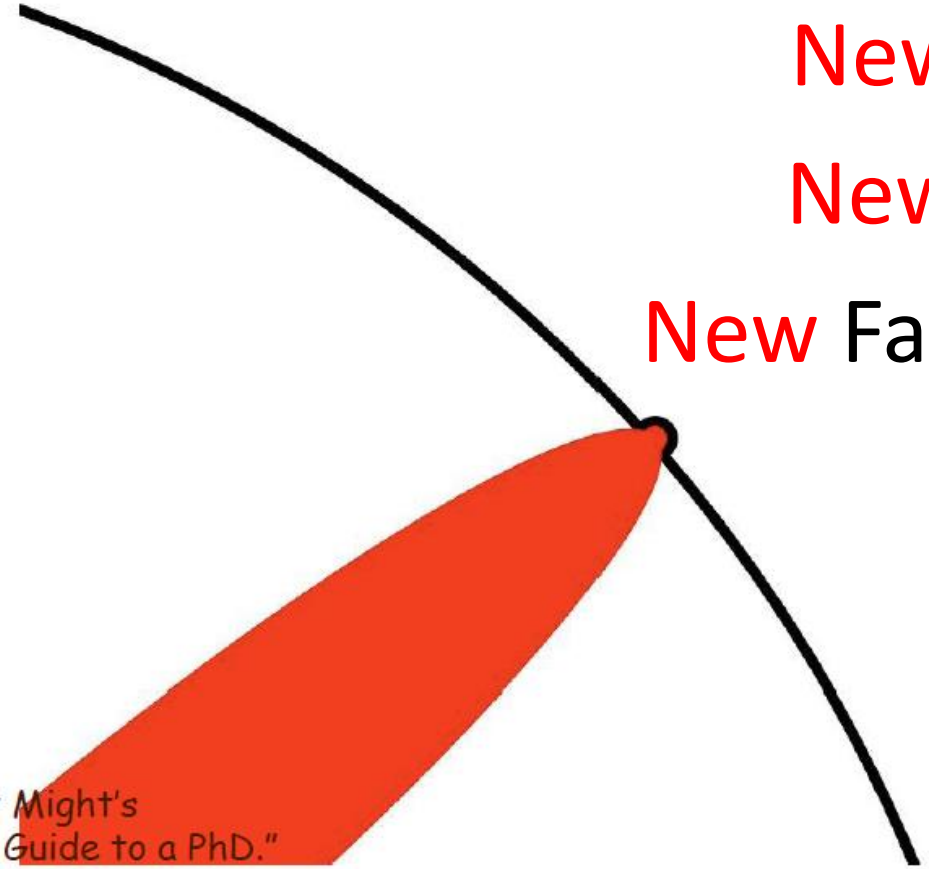
From "Matt Might's Illustrated Guide to a PhD."

What to Expect of a PhD Research

From "Matt Might's Illustrated Guide to a PhD."

Winning a Nobel Price through PhD Research is not Neccessary

Expectation of a PhD Research



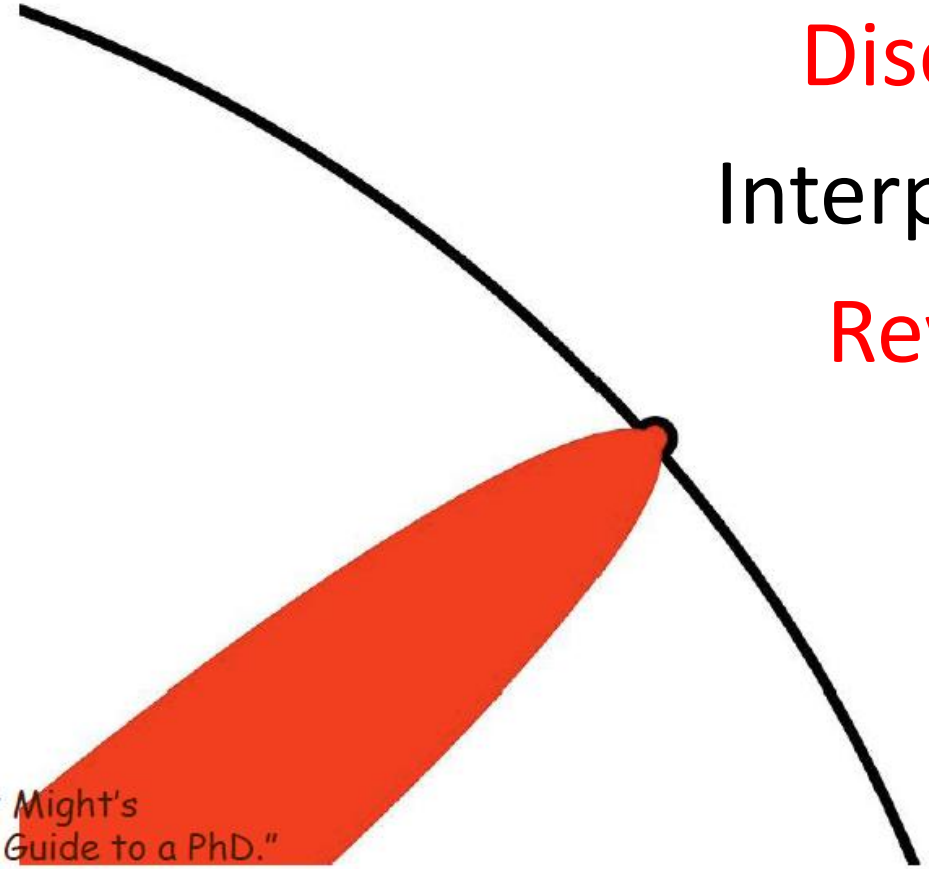
New Facts

New Ideas

New Facts + Ideas

From "Matt Might's Illustrated Guide to a PhD."

Expectation of a PhD Research



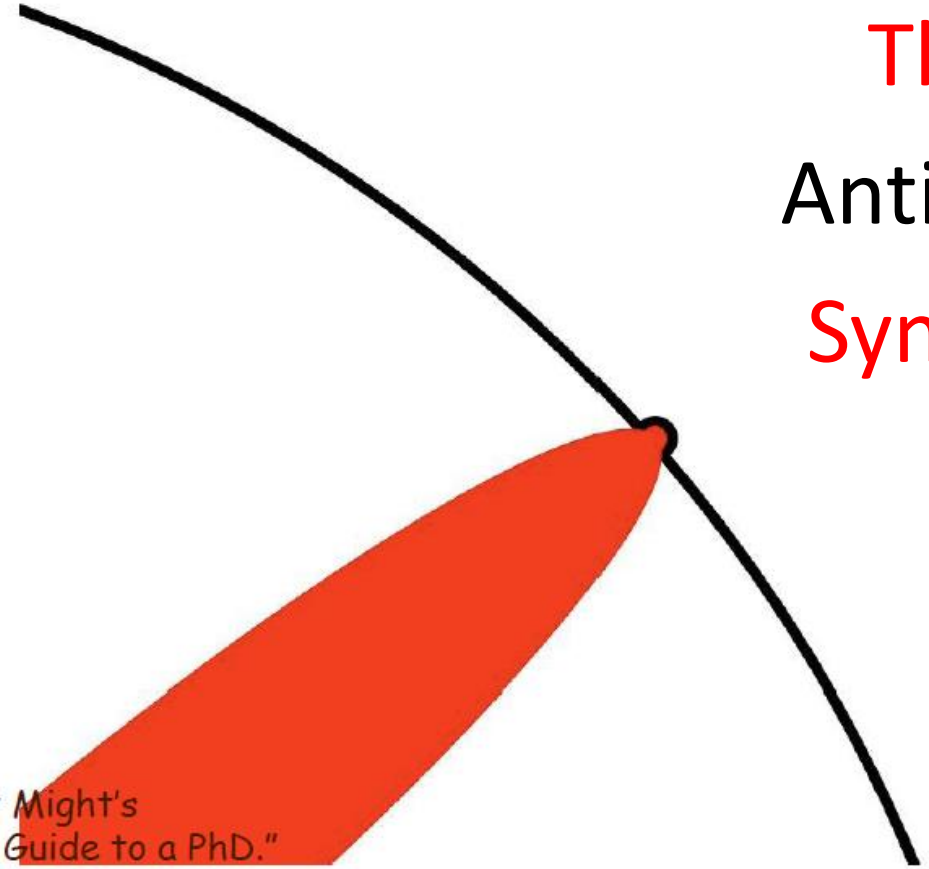
From "Matt Might's Illustrated Guide to a PhD."

Discovery

Interpretation

Revision

Expectation of a PhD Research



Thesis

Anti-thesis

Synthesis

From "Matt Might's Illustrated Guide to a PhD."

Problem Formulation

Problem

A situation, person or things that needs attention and needs to be dealt with or solved

Problem Formulation

Formulation

To develop all the details of a plan for solving problem

Problem Identification Process

- Identify Domain Area (Broad Topic)
- Make a conceptual map by identifying related topics (Narrow Topic)
  - Divide the domain area into progressively smaller sub-area until one reaches a subject of interest
  - Tools: mindmap, fishbone
- Identify topic of interests student wish to explore
- Identify "real" problem area in the sub-area
  - What is the most important and possible relative to scope and scale of study?

Problem Solution Process

- Problem analysis/Root-cause analysis
- Duration analysis/Activity-based costing
- Outcome analysis
- Breaking assumptions
- Technology analysis
- Proxy benchmarking
- Activity elimination/process simplification

Example - Identifying Gaps in Information Retrieval Research

- Other data that can be used to enhance?
  - Eg. Explicit vs implicit, multimedia
- Other ways to represent data?
  - Eg. Graph? Passages?
- Challenge assumptions and rules.
  - Eg.: crisp vs fuzzy? Deterministic vs probabilistics?
- Other external knowledge bases or sources to enhance?
- Combination? Optimization? Weightage?
- Adapting ideas from other fields? Eg. Diversity analysis, Cross-structural theory, Game theory, etc.

Significance to Community

- What is the Need?
- What is your Approach to address the need?
- What is the Benefit of your approach over cost?
- In what ways are your research better than the Competitors?

Problem tree analysis

- Analysis of researchable issue to gain insights about its possible cause-effect relationships.
- Helps in identifying the critical areas where an intervention would provide a solution to the problem of concern

Impacts No Protection of IP Questionable Integrity

Effect Intelligent plagiarism cannot be detected

Problem **No way to detect beyond cut-and paste plagiarism**

Causes Plagiarist change words Change order of words

Change Problem Tree to Objective Tree

Impacts Protection of IP Preserve Integrity

↑

Effect Avoid Intelligent plagiarism

↑

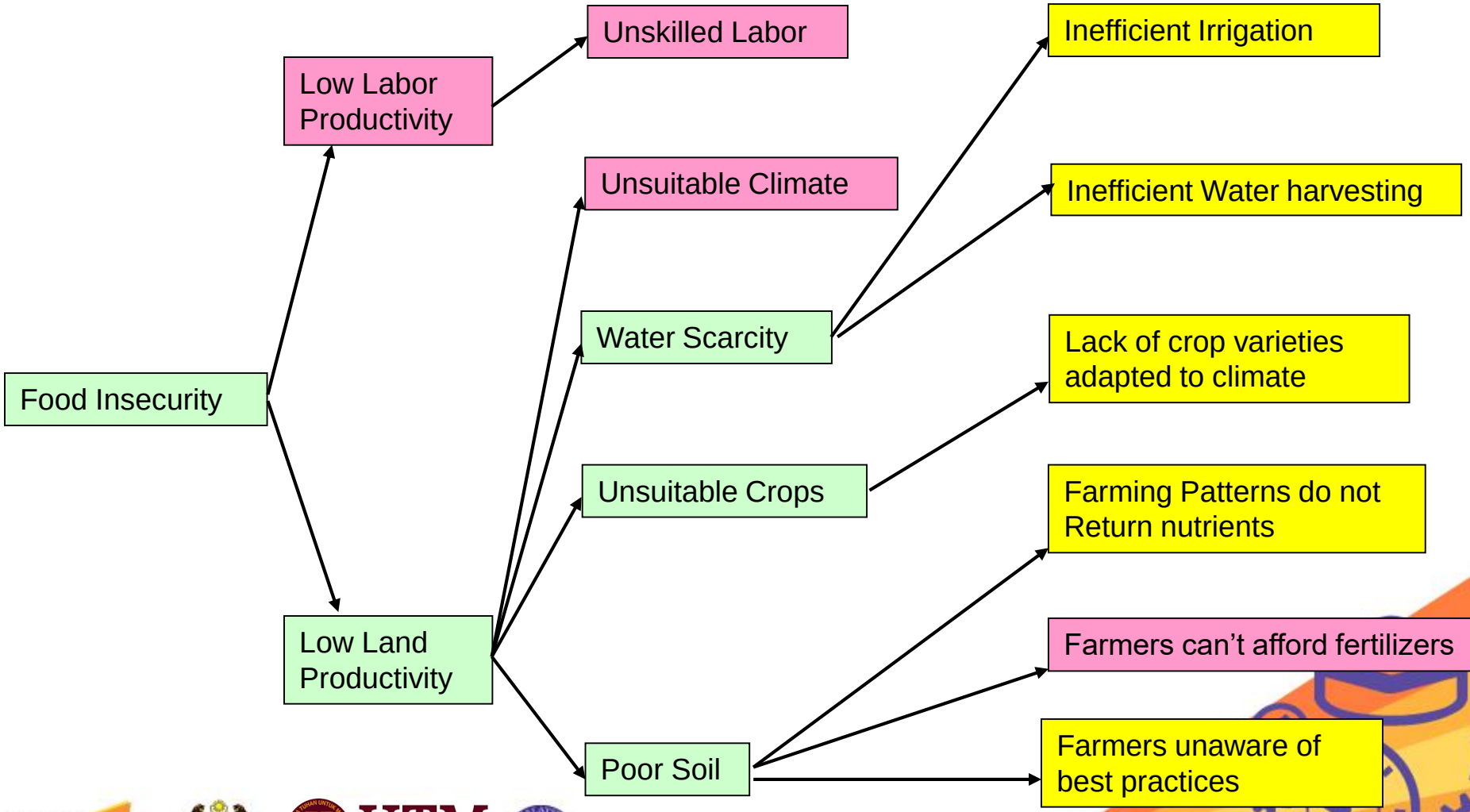
Aims **Develop ways to detect beyond cut-and paste plagiarism**

↑

Objectives Semantic-based detection Predicate-based detection

Root Cause Analysis in Problem Tree – Keep asking Why?

(SA Prathapar, Research Methodology Slide, 2012)



Horizontal Analysis

Output:

Decisions?

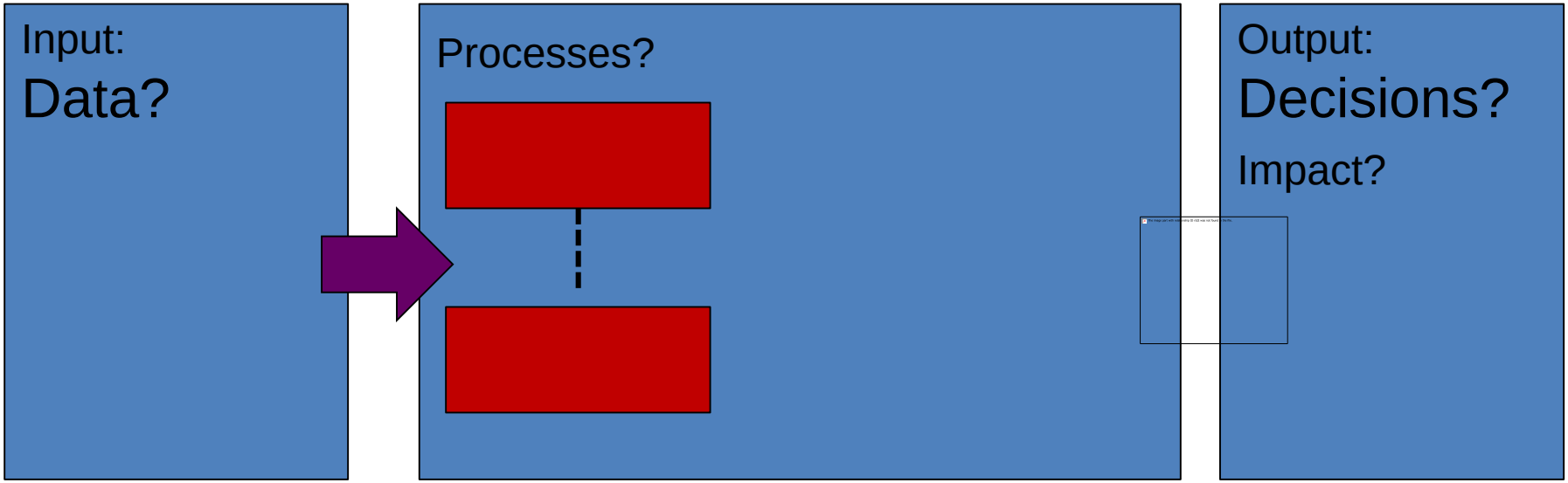
Impact?

Horizontal Analysis

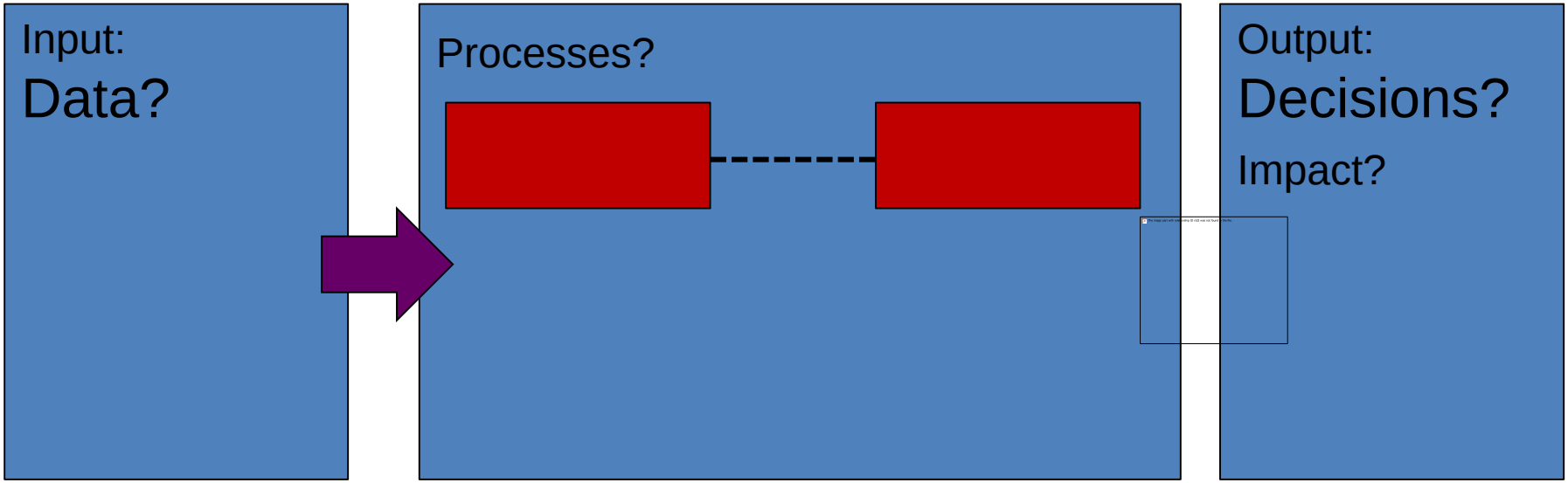
Input:
Data?

Output:
Decisions?
Impact?

Horizontal Analysis



Horizontal Analysis



Example of Horizontal Analysis

Input:

- Historical Stock Values
- Other Commodities
- Weather
- News

Processes?

Crude Palm Oil Stock Market Prediction: Turning Points? Values?

How many research objectives?

3 objectives

- Objective 1 — 6 months
- Objective 2 — 6 months
- Objective 3 — 6 months

Make Your RQ Specific Enough

- Very weak (in an empirical sense)
  - *Is the new technique any good?*
- Weak
  - *Is the new technique better than multi-tap?*
- Better
  - *Is the new technique faster than multi-tap?*
- Better still
  - *Is the new technique faster than multi-tap within one hour of use?*
- Even better
  - *If error rates are kept under 2%, is the new technique faster than multi-tap within one hour of use?*

Topic Identification

– Can he finish it on time?

Does It Suit Him?

- Student's Background
- Student's Interests
- Student's Preparation
- Student's Capabilities

How Fast Can He Start?

- Equipment/Hardware
- Data Availability
- Software/Simulator

Supportive Environment?

- Supervisor's Knowledge
- Supervisor's Research Experience
- Senior Students
- Available Networks

Literature Support?

- Is it well-published? In reputable venues?
- Is it a growing research area?
- Can the data and methodology be obtained from literature ? Understood by student?

In case he is not prepared at all…

Where should I look for ideas? What are the top journals in my area? Hot topics?

Who are the active researchers in the area?

Guide him to read …

Have a feel of good research and find his interest in good, reputable journals

https://jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my/JCRJournalHomeAction.action?SID=B1-x2BLr2zm3x2BRxx4pxx8

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

Web of Science™ | InCites™ | Journal Citation Reports® | Essential Science Indicators℠ | EndNote™ | Sign In | Help | English

InCites™ Journal Citation Reports®

THOMSON REUTERS™

Home

Go to Journal Profile

Compare Journals

View Title Changes

Select Journals

Select Categories

Journals By Rank | Categories By Rank

Journal Titles Ranked by Impact Factor

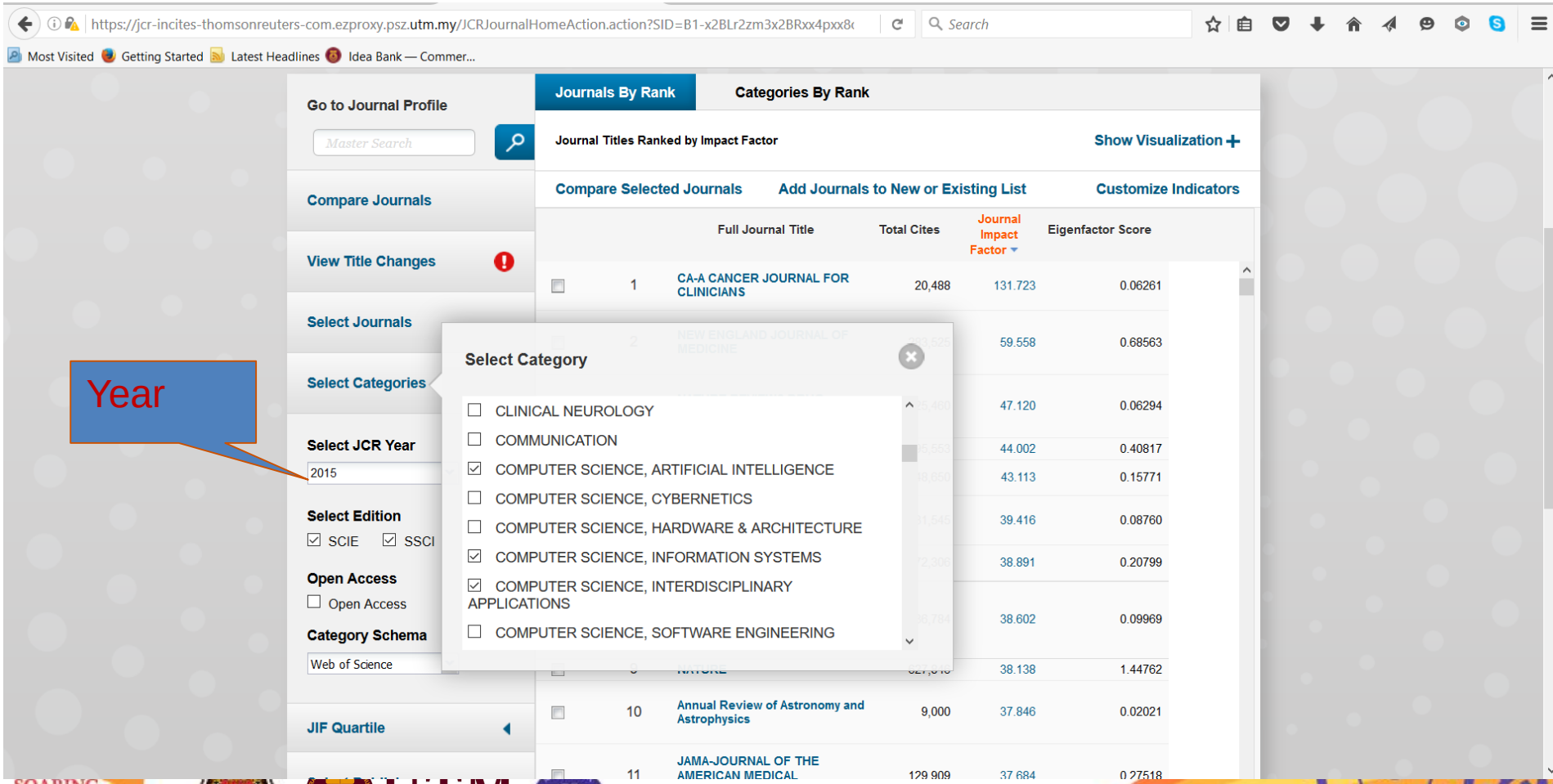
Show Visualization +

Compare Selected Journals | Add Journals to New or Existing List | Customize Indicators

	Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
1	CA-A CANCER JOURNAL FOR CLINICIANS	20,488	131.723	0.06261
2	NEW ENGLAND JOURNAL OF MEDICINE	283,525	59.558	0.68563
3	NATURE REVIEWS DRUG DISCOVERY	25,460	47.120	0.06294

Looking up jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my...

Select Categories based on Your Area



Choose Quartile, IF Range or Percentile

https://jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my/JCRJournalHomeAction.action?SID=B1-x2BLr2zm3x2BRxx4pxx8

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

Open Access

☐ Open Access

Category Schema

Web of Science

JIF Quartile

☑ Q1 ☐ Q3
☐ Q2 ☐ Q4

Select Publisher

Select Country/Territory

Impact Factor Range

to

Average JIF Percentile Range

to

Clear | Submit

7	NATURE MATERIALS	72,306	38.891	0.20799
8	NATURE REVIEWS MOLECULAR CELL BIOLOGY	36,784	38.602	0.09969
9	NATURE	627,846	38.138	1.44762
10	Annual Review of Astronomy and Astrophysics	9,000	37.846	0.02021
11	JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	129,909	37.684	0.27518
12	CHEMICAL REVIEWS	148,154	37.369	0.24548
13	NATURE REVIEWS GENETICS	30,286	35.898	0.10725
14	Annual Review of Immunology	17,023	35.543	0.03670
15	Nature Nanotechnology	40,881	35.267	0.16769
16	SCIENCE	568,210	34.661	1.15726
17	NATURE REVIEWS CANCER	41,846	34.244	0.08819
18	CHEMICAL SOCIETY REVIEWS	99,930	34.090	0.27041
19	REVIEWS OF MODERN PHYSICS	41,133	33.177	0.09735

Click Submit

https://jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my/JCRJournalHomeAction.action?SID=B1-x2BLr2zm3x2BRxx4pxx8

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

Open Access

☐ Open Access

Category Schema

Web of Science

JIF Quartile

☑ Q1 ☐ Q3

☐ Q2 ☐ Q4

Select Publisher

Select Country/Territory

Impact Factor Range

to

Average JIF Percentile Range

to

Clear | Submit

☐	7	NATURE MATERIALS	72,306	38.891	0.20799
☐	8	NATURE REVIEWS MOLECULAR CELL BIOLOGY	36,784	38.602	0.09969
☐	9	NATURE	627,846	38.138	1.44762
☐	10	Annual Review of Astronomy and Astrophysics	9,000	37.846	0.02021
☐	11	JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	129,909	37.684	0.27518
☐	12	CHEMICAL REVIEWS	148,154	37.369	0.24548
☐	13	NATURE REVIEWS GENETICS	30,286	35.898	0.10725
☐	14	Annual Review of Immunology	17,023	35.543	0.03670
☐	15	Nature Nanotechnology	40,881	35.267	0.16769
☐	16	SCIENCE	568,210	34.661	1.15726
☐	17	NATURE REVIEWS CANCER	41,846	34.244	0.08819
☐	18	CHEMICAL SOCIETY REVIEWS	99,930	34.090	0.27041
☐	19	REVIEWS OF MODERN PHYSICS	41,133	33.177	0.09735

javascript:catSubmit()

List of Journals Sorted by IF

ers-com.ezproxy.psz.utm.my/JCRJournalHomeAction.action?SID=B1-5u9otq9kX3sD5NOx2Bb

Search

adlines Idea Bank — Commer...

Compare Journals

View Title Changes

Select Journals

Select Categories

Select JCR Year

2015

Select Edition

☑ SCIE ☑ SSCI

Open Access

☐ Open Access

Category Schema

Web of Science

JIF Quartile

Select Publisher

Select Country/Territory

Compare Selected Journals | Add Journals to New or Existing List | Customize Indicators

Select All		Full Journal Title	Total Cites	Journal Impact Factor	Eigenfactor Score
☐	1	IEEE Communications Surveys and Tutorials	4,172	9.220	0.02037
☐	2	IEEE TRANSACTIONS ON FUZZY SYSTEMS	9,220	6.701	0.01354
☐	3	International Journal of Neural Systems	1,235	6.085	0.00247
☐	4	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	31,757	6.077	0.06131
☐	5	IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION	7,999	5.908	0.00910
☐	6	MIS QUARTERLY	11,320	5.384	0.01137
☐	7	COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING	1,935	5.288	0.00367
☐	8	INTEGRATED COMPUTER-AIDED ENGINEERING	533	4.981	0.00124
☐	9	IEEE Transactions on Cybernetics	2,246	4.943	0.01017
☐	10	IEEE Transactions on Neural Networks and Learning Systems	12,919	4.854	0.02248
☐	11	JOURNAL OF INFORMATION TECHNOLOGY	1,695	4.775	0.00268

Click Rank to Get Quartile of Journal

File Edit View History Bookmarks Tools Help

ugent - naomi... | Inbox (2,010) ... | Online Database v... | Web of Scienc... | InCites™ | Model of onli... | Comparing str... | PDLK: Plagiari... | InCites™ | utm library - ...

https://jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my/JCRJournalProfileAction.action?pg=JRNLPROF&journalImpactFa

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

2011	770	4.284	1.837	2.574	0.333	36	3.9	6.6	0.00142	0.396	100.00	Not A...	96.847
2010	816	4.237	1.710	2.581	0.421	38	3.6	6.9	0.00122	0.318	100.00	Not A...	94.907
2009	669	2.988	1.682	Not A...	0.361	36	5.0	7.1	0.00164	Not A...	94.44	Not A...	88.835
2008	493	0.901	0.888	Not A...	0.075	40	8.7	8.3	0.00132	Not A...	100.00	Not A...	30.319
2007	347	0.488	0.450	Not A...	0.095	42	8.9	6.7	0.00089	Not A...	100.00	Not A...	20.968
1998	144	0.295	0.295	Not A...	Not A...	Not A...	5.1	4.9	Not A...	Not A...	Not A...	Not A...	20.545

Source Data

Rank

Cited Journal Data

Citing Journal Data

Box Plot

Journal Relationships

JCR Impact Factor

JCR Year	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE			NEUROSCIENCES		
				Rank	Quartile	JIF Percentile
2015	2/130	Q1	98.846	NA	undefined	
2014	2/123	Q1	98.780	NA	undefined	
2013	3/121	Q1	97.934	NA	undefined	
2012	2/115	Q1	98.696	NA	undefined	
2011	4/111	Q1	96.847	NA	undefined	
2010	6/108	Q1	94.907	NA	undefined	
2009	12/103	Q1	88.835	NA	undefined	
2008	66/94	Q3	30.319	NA	undefined	
2007	74/93	Q4	20.968	NA	undefined	
1998	NA	undefined		161/202	Q4	20.545

https://jcr-incites-thomsonreuters-com.ezproxy.psz.utm.my/JCRJournalProfileAction.action?pg=JRNLPROF&journalImpactFactor=6.085&year=2015&journalTitle=International Journal of Neural Systems&edition=SCIE&journal=INT J NEURAL SYST

Find Most Recent, Most Read and Most Cited Articles

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

• Current Contents®/Engineering, Computing, and Technology

Add to Favorites | Citation Alert | View Abstracts | Send to Citation Mgr

Select All

Most Recent | Most Read | Most Cited

SPIKING NEURAL NETWORKS
SAMANWOY GHOSH-DASTIDAR, HOJJAT ADELI
International Journal of Neural Systems Vol. 19, No. 04, pp. 295-308 (2009)
Abstract | PDF (252 KB) | PDF Plus (371 KB)

LOW-RANK APPROXIMATION BASED NON-NEGATIVE MULTI-WAY ARRAY DECOMPOSITION ON EVENT-RELATED POTENTIALS
FENGYU CONG, GUOXU ZHOU, PIIA ASTIKAINEN, QIBIN ZHAO, QIANG WU, ASOKE K NANDI, JARI K. HIETANEN, TAPANI RISTANIEMI, ANDRZEJ CICHOCKI
International Journal of Neural Systems Vol. 24, No. 08, 1440005 (2014)
Abstract | PDF (441 KB) | PDF Plus (675 KB)

INTRINSIC CURVATURE: A MARKER OF MILLIMETER-SCALE TANGENTIAL CORTICO-CORTICAL CONNECTIVITY?
LISA RONAN, RUDOLPH PIENAAR, GUY WILLIAMS, ED BULLMORE, TIM J. CROW, NEIL ROBERTS, PETER B. JONES, JOHN SUCKLING, PAUL C. FLETCHER
International Journal of Neural Systems Vol. 21, No. 05, pp. 351-366 (2011)
Abstract | PDF (464 KB) | PDF Plus (497 KB)

TRANSCRANIAL MAGNETIC STIMULATION (TMS) MODULATES EPILEPTIFORM DISCHARGES IN PATIENTS WITH FRONTAL LOBE EPILEPSY: A PRELIMINARY EEG-TMS STUDY
VASILIOS K. KIMISKIDIS, DIMITRIS KUGIUMTZIS, SOTIRIOS PAPAGIANNOPOULOS, NIKOLAOS VLAIKIDIS
International Journal of Neural Systems Vol. 23, No. 01, 1250035 (2013)
Abstract | PDF (1416 KB) | PDF Plus (1302 KB)

APPLICATION OF EMPIRICAL MODE DECOMPOSITION (EMD) FOR AUTOMATED DETECTION OF EPILEPSY USING EEG SIGNALS
ROSHAN JOY MARTIS, U. RAJENDRA ACHARYA, JEN HONG TAN, ANDREA PETZNICK, RATNA YANTI, CHUA KUANG CHUA, E. Y. K. NG, LOUIS TONG
International Journal of Neural Systems Vol. 22, No. 06, 1250027 (2012)
Abstract | PDF (798 KB) | PDF Plus (849 KB)

CORTICAL EXCITABILITY AND REFRACTORY EPILEPSY: A THREE-YEAR LONGITUDINAL TRANSCRANIAL MAGNETIC STIMULATION STUDY
RADWA A. B. BADAWY, GRAEME D. JACKSON, SAMUEL F. BERKOVIC, RICHARD A. L. MACDONELL
International Journal of Neural Systems Vol. 23, No. 01, 1250030 (2013)
Abstract | PDF (1517 KB) | PDF Plus (581 KB)

Search by Topic in WOS

File Edit View History Bookmarks Tools Help

ugent - naomi... | Inbox (2,010) ... | Online Database v... | Web of Scienc... | Web of Scienc... | Model of onli... | Comparing str... | PDLK: Plagiari... | InCites™ | Impact factor ...

apps.webofknowledge.com.ezproxy.psz.utm.my/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearc

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

Web of Science™ | InCites™ | Journal Citation Reports® | Essential Science Indicators℠ | EndNote™ | Sign In | Help | English

WEB OF SCIENCE™ THOMSON REUTERS

Search | Web of Science™ Core Collection | My Tools | Search History | Marked List

Welcome to the new Web of Science! View a brief tutorial.

Basic Search

plagiarism detection system | Topic | Search

Click here for tips to improve your search.

+ Add Another Field | Reset Form

TIMESPAN

All years

From 1990 to 2016

MORE SETTINGS

Customer Feedback & Support | Additional Resources | What's New in Web of Science? | Customize your Experience

Result List by Topic in WOS

Conference Titles
Publication Years
Organizations-Enhanced
Funding Agencies
Languages
Countries/Territories
ESI Top Papers
Open Access

For advanced refine options, use

Analyze Results

Full Text from Publisher | View Abstract

7. **High time for a common plagiarism detection system**
By: Jaric, Ivan
SCIENTOMETRICS Volume: 106 Issue: 1 Pages: 457-459 Published: JAN 2016
Full Text from Publisher | View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

8. **DOCODE 3.0 (DOcument COpy DEtector): A system for plagiarism detection by applying an information fusion process from multiple documental data sources**
By: Velasquez, Juan D.; Covacevich, Yerko; Molina, Francisco; et al.
INFORMATION FUSION Volume: 27 Pages: 64-75 Published: JAN 2016
Full Text from Publisher | View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

9. **Software Plagiarism Detection with Birthmarks Based on Dynamic Key Instruction Sequences**
By: Tian, Zhenzhou; Zheng, Qinghua; Liu, Ting; et al.
IEEE TRANSACTIONS ON SOFTWARE ENGINEERING Volume: 41 Issue: 12 Pages: 1217-1235 Published: DEC 2015
Full Text from Publisher | View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

10. **PDLK: Plagiarism detection using linguistic knowledge**
By: Abdi, Asad; Idris, Norisma; Alguliyev, Rasim M.; et al.
EXPERT SYSTEMS WITH APPLICATIONS Volume: 42 Issue: 22 Pages: 8936-8946 Published: DEC 1 2015
Full Text from Publisher | View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

Select Page | Save to EndNote online | Add to Marked List

Sort by: Publication Date -- newest to oldest | Show: 10 per page | Page 1 of 3

29 records matched your query of the 3,458,326 in the data limits you selected.

Click on Journal Title to Find Out Its Quartile and Rank

File Edit View History Bookmarks Tools Help

Online Database via Ezproxy | Web of Science [v.5.22.2] ... | InCites™ | Terlalu Cinta by Rossa ... | (7) Naomie Salim | IJNS Aims and Scope (Wo...

apps.webofknowledge.com.ezproxy.psz.utm.my/Search.do?product=WOS&SID=W23JvWCEYAQWaLUapgd&search_mode=Ger

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

Conference Titles

Publication Years

Organizations-Enhanced

Funding Agencies

Languages

Countries/Territories

ESI Top Papers

Open Access

For advanced refine options, use

Analyze Results

Full Text from Publisher | View Abstract

7. **High time for a common plagiarism detection system**
By: Jaric, Ivan
SCIENTOMETRICS Volume: 106 Issue: 1 Pages: 457-459 Published: JAN 2016
Full Text from Publisher | View Abstract
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

8. **DOCODE 3.0 (DOcument COpy DEtector): A system for plagiarism detection by applying an information fusion process from multiple documental data sources**
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

9. n Sequences
-1235 Published: DEC
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

10. hed: DEC 1 2015
Times Cited: 0 *(from Web of Science Core Collection)*
Usage Count

Select F

Sort by:

Page 1 of 3

29 records m

INFORMATION FUSION

Impact Factor
4.353 *2015* **4.049** *5 year*

JCR® Category	*Rank in Category*	*Quartile in Category*
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	9 of 130	Q1
COMPUTER SCIENCE, THEORY & METHODS	4 of 105	Q1

Data from the 2015 edition of Journal Citation Reports®

Publisher
ELSEVIER SCIENCE BV, PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS

ISSN: 1566-2535
eISSN: 1872-6305

Research Domain
Computer Science

Close Window

Create Citation Alert

File Edit View History Bookmarks Tools Help

ugent - naomi... | Inbox (2,010) ... | Online Database v... | Web of Scienc... | Web of Scienc... | Model of onli... | Comparing str... | PDLK: Plagiari... | InCites™ | Impact factor ...

apps.webofknowledge.com.ezproxy.psz.utm.my/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=

Most Visited | Getting Started | Latest Headlines | Idea Bank — Commer...

WEB OF SCIENCE™ THOMSON REUTERS

Search | Return to Search Results | My Tools | Search History | Marked List

Full Text from Publisher | Look Up Full Text | Save to EndNote online | Add to Marked List

2 of 166

Efficient top-k similarity document search utilizing distributed file systems and cosine similarity

By: Alewiwi, M (Alewiwi, Mahmoud)[1

CLUSTER COMPUTING-THE JOURN
Volume: 19 Issue: 1 Pages: 109-1
DOI: 10.1007/s10586-015-0506-0
Published: MAR 2016
View Journal Information

Abstract

Document similarity has important real
k-similarity algorithms have been long
similar documents in reasonably short amount of time. In particular, pairwise comparison of documents' features, a key operation in calculating document similarity, necessitates prohibitively high storage and computation power. In this paper, we propose a new filtering technique that decreases the number of comparisons between the query set and the search set to find highly similar documents. The proposed filtering technique utilizes Z-order prefix, based on the cosine similarity measure, in which only the most important features are used first to find highly similar documents. We propose a three-phase approach, where the phases are near duplicate detection, common important terms and join phase. We utilize the Hadoop distributed file system and the MapReduce parallel programming model to scale our techniques to big data setting. Our experimental results on real data show that the proposed method performs better than the previous work in the literature in terms of the number of joins, and therefore, speed.

Keywords

Author Keywords: Z-order; Document similarity; MapReduce; Hadoop; Cosine similarity

Sign In

E-mail Address:
Password:
Sign In | Cancel
Keep me signed in
Forgot Password
Register

To create or access your Citation Alerts, please sign in or register.

Creating Citation Alerts lets you:

- Add an article to your alerts from its full record page
- Access your selected articles from any computer at your institution

Citation Network

0 Times Cited
22 Cited References
View Related Records
View Citation Map
Create Citation Alert
(data from Web of Science ™ Core Collection)

All Times Cited Counts

0 in All Databases
0 in Web of Science Core Collection
0 in BIOSIS Citation Index
0 in Chinese Science Citation Database
0 in Data Citation Index
0 in Russian Science Citation Index
0 in SciELO Citation Index

Scopus - evaluated by SciMago Journal Rank (SJR)*

www.scimagojr.com/journalrank.php?category=1706&area=1700&year=2012&country=&order=sjr&min=0&min_type=cd

SJR SCImago Journal & Country Rank

EST MODUS IN REBUS
Horatio (Satire 1,1,106)

- Home
- Journal Rankings
- Journal Search
- Country Rankings
- Country Search
- Compare
- Map Generator
- Help
- About Us

Journal Rankings

Ranking Parameters

Subject Area: Computer Science
Subject Category: Computer Science Applications
Country: All Year: 2012
Order By: SJR
Display journals with at least: 0 Citable Docs. (3 years) Refresh

Subject Area: **Computer Science.**
Subject Category: **Computer Science Applications.**
Year: **2012.**

Download data in MS Excel format (50 Kb)

How to cite this website?

1 - 50 of 230 << First | < Previous | Next > | Last >>

	Title	SJR	H index	Total Docs. (2012)	Total Docs. (3years)	Total Refs.	Total Cites (3years)	Citable Docs. (3years)	Cites / Doc. (2years)	Ref. / Doc.	Country
1	Bioinformatics	Q1 4,223	204	747	2.251	17.874	14.933	2.160	5,38	23,93	
2	Argument and Computation	Q1 3,173	5	8	17	377	68	17	4,00	47,13	
3	Journal of Field Robotics	Q1 3,161	46	48	148	1.610	492	134	3,39	33,54	
4	Journal of Informetrics	Q1 3,075	26	78	172	2.242	787	158	4,80	28,74	

*http://www.scimagojr.com/journalrank.php
(developed from the Google PageRank algorithm)

SJR is developed by:

Analyze trend

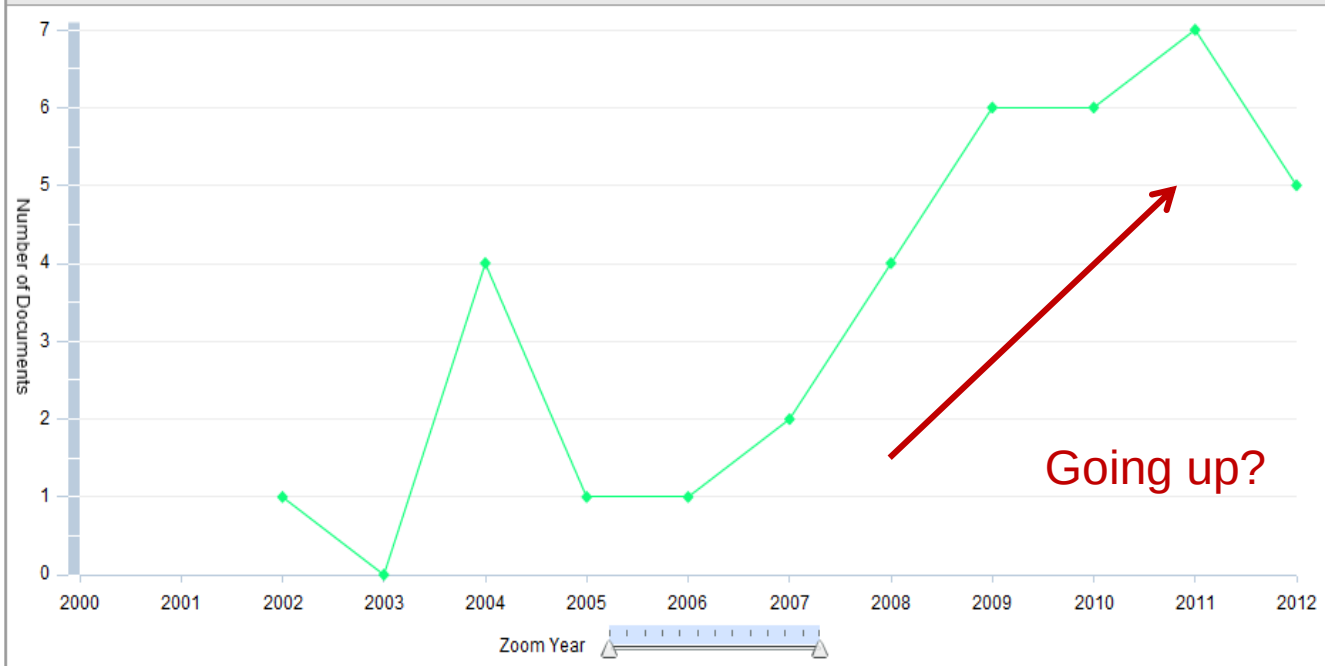
Analyze results | Back to results

Date range 2000 to 2012 Analyze Document results 37

Year | Source title | Author name | Affiliation name | Country | Document type | Subject area

Export | Print | E-mail

Years This chart shows the total number of documents for this query by Year.



Year	Documents
2012	5
2011	7
2010	6
2009	6
2008	4
2007	2
2006	1
2005	1
2004	4
2002	1

Conference & Journal for Area

Analyze results | Back to results

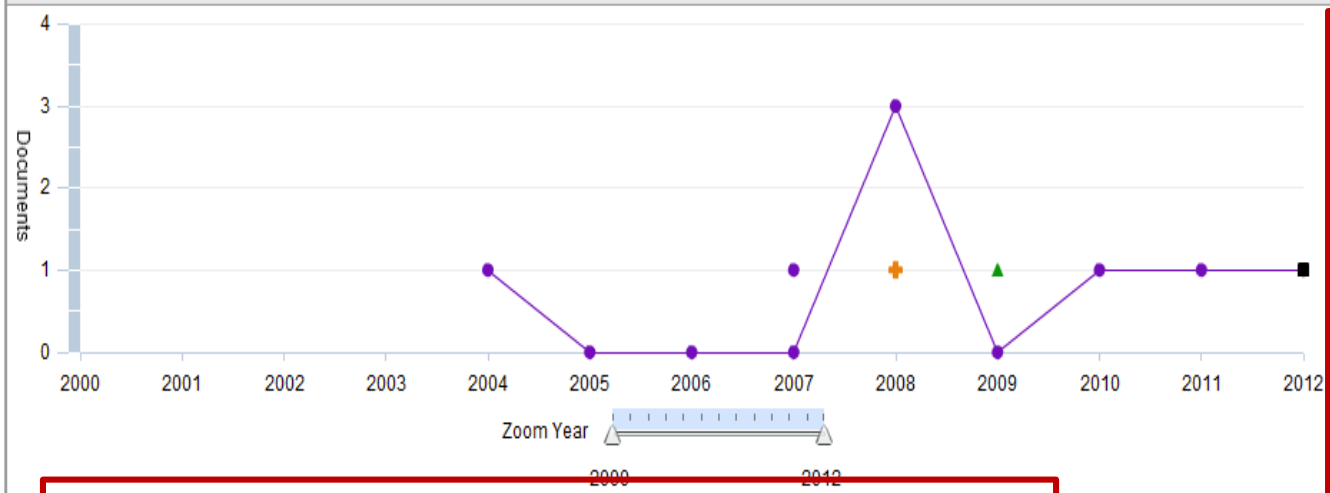
Date range 2000 to 2012 Analyze Document results 37

Year | Source title | Author name | Affiliation name | Country | Document type | Subject area

Export | Print | E-mail

Source Title This chart shows the total number of documents per year for this query by Source Title.

Compare journals in Journal Analyzer



	Source	Documents
☑	Lecture Notes in Computer Science	7
☑	2007 Asia Pacific Conference on C	1
☑	3rd International Conference on Ge	1
☑	3rd International Conference on Inr	1
☑	Applied Soft Computing Journal	1
☐	Beijing Hangkong Hangtian Daxue	1
☐	Dr Dobb S Journal	1
☐	IEEE Transactions on Computers	1
☐	IEEE Transactions on Systems Mar	1
☐	Journal of Information Science and	1
☐	Journal of Multimedia	1
☐	Journal of Theoretical and Applied I	1
☐	Kese 2009 2009 Pacific Asia Confe	1
☐	Lecture Notes in Artificial Intelligenc	1

- Lecture Notes in Computer Science Including Subseries Lecture Notes in Artifici...
- 2007 Asia Pacific Conference on Communications Apcc
- 3rd International Conference on Genetic and Evolutionary Computing Wgec 2009
- 3rd International Conference on Innovative Computing Information and Control Ici...

To add more source titles to the graph use the checkboxes in the right hand table.

People who work in same area

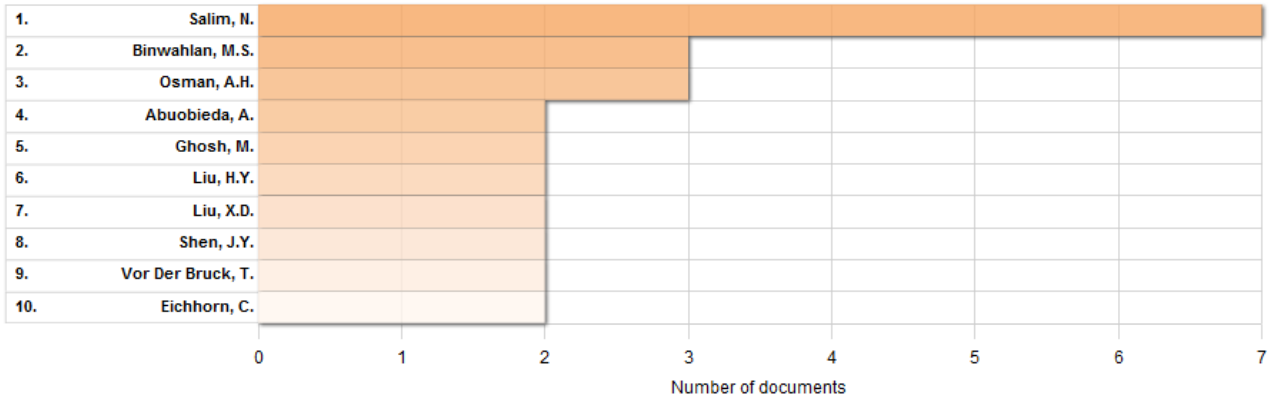
Analyze results | Back to results

Date range 2000 to 2012 Analyze Document results 37

Year | Source title | Author name | Affiliation name | Country | Document type | Subject area

Export | Print | E-m

Author Name This chart shows the total number of documents for this query by Author.



	Author	Documents
☑	Salim, N.	7
☑	Binwahlan, M.S.	3
☑	Osman, A.H.	3
☑	Abuobieda, A.	2
☑	Ghosh, M.	2
☑	Liu, H.Y.	2
☑	Liu, X.D.	2
☑	Shen, J.Y.	2
☑	Vor Der Bruck, T.	2
☑	Eichhorn, C.	2
☐	Bao, J.P.	2
☐	Abraham, A.	2
☐	Zhang, X.D.	2
☐	Hartrumpf, S.	2

To add more authors to the graph use the checkboxes in the list on the right.

Institution Doing Active Research in Area

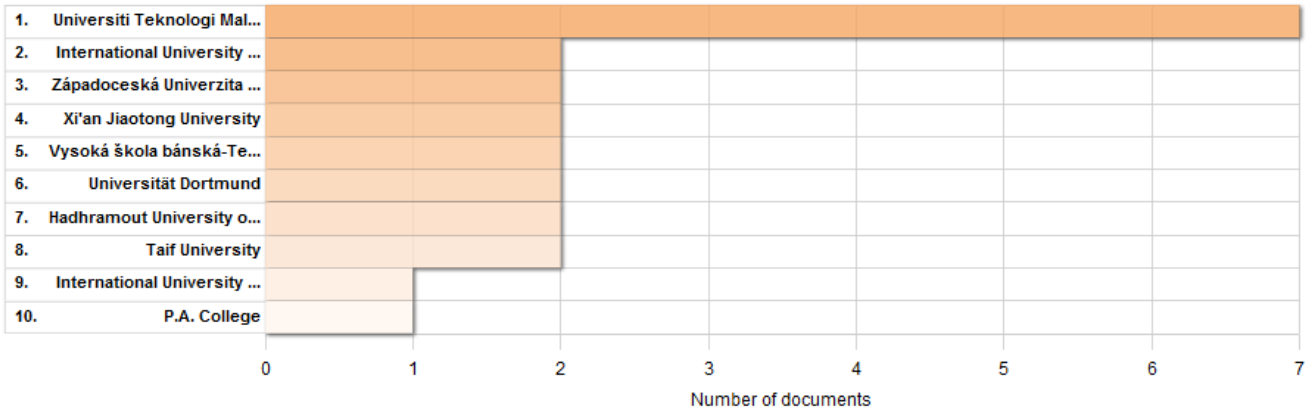
Analyze results | Back to results

Date range 2000 to 2012 Analyze Document results 37

Year | Source title | Author name | Affiliation name | Country | Document type | Subject area

Export | Print

Affiliation Name This chart shows the total number of documents for this query by Affiliation Name.



	Affiliation	Documents
☑	Universiti Teknologi Malaysia	7
☑	International University of Africa	2
☑	Západoceská Univerzita v Plzni	2
☑	Xi'an Jiaotong University	2
☑	Vysoká škola bánská-Technická unive	2
☑	Universität Dortmund	2
☑	Hadhramout University of Science and	2
☑	Taif University	2
☑	International University of Africa	1
☑	P.A. College	1
☐	School of Management	1
☐	Zeidman Consulting	1

Coaching For Deep & Critical Apprehension of Literature

Guding Students to Do LR

- Conventional Review
- System Literature Review
- Experimental Review

After initial topic/area has been determined, ask student to find a focus

- A review is usually organized around ideas
  - not the sources themselves as an annotated bibliography would be organized.
  - not just simply list your sources and go into detail about each one of them, one at a time.
- Construct thesis statement
  - Justify the need for research through LR

Example of Thesis Statement for Review

- How effective is semantic-based database integration?
- Can current plagiarism detection techniques identify intelligent plagiarism?
- What are the fuzzy aspects of plagiarism? Can current and fuzzy logic approaches detect these fuzzy aspects in plagiarism?

Parts of a Good Review

- Introduction
  - To the context & importance/significance of the work
- Analysis
  - Main framework used to review a particular topic
- Synthesis
  - Reorganization/re-shuffling of main parts inside the analytical framework
- Evaluation
  - Comparison – similarities, difference
  - Critical discussion of strength, weaknesses
  - Gap analysis
- Suggestion
  - Of further work that can or need to be done

Writing the Introduction Section

1. What is the problem? Define.
2. How has it been solved? Show the general ways. Broad, to specific (to your focus).
3. Why it need solving? Significance, impact.
4. Unique viewpoint of your review.

Example of Introduction

I. INTRODUCTION

The problem of plagiarism has recently increased because of the digital era of resources available on the World Wide Web. Plagiarism detection in natural languages by statistical or computerized methods has started since the 1990s, which is pioneered by the studies of copy detection mechanisms in digital documents [42], [43]. Earlier than plagiarism detection in natural languages, code clones and software misuse detection has started since the 1970s by the studies to detect programming code plagiarism in Pascal and C [28], [44]–[47]. Algorithms of plagiarism detection in natural languages and programming languages have noticeable differences. The first one tackles different textual features and diverse methods of detection, while the latter mainly focuses on keeping track of metrics, such as number of lines, variables, statements, subprograms, calls to subprograms, and other parameters. During the last decade, research on automated plagiarism detection in natural languages has actively evolved, which takes the advantage of recent developments in related fields like information retrieval (IR), cross-language information retrieval (CLIR), natural language processing, computational linguistics, artificial intelligence, and soft computing. In this paper, a survey of recent advances in the area of automated plagiarism detection in text documents is presented, which started roughly in 2005, unless it is noteworthy to state a research prior than that. Earlier study was excellently reviewed by [48] and [52]–[55].

This paper brings patterns of plagiarism together with textual features for characterization of each pattern and computerized methods for detection. The contributions of this paper can be summarized as follows: First, different kinds of plagiarism are organized into a taxonomy that is derived from a qualitative study and recent literatures about the plagiarism concept. The taxonomy is supported by various *plagiarism patterns* (i.e., examples) from available corpora for plagiarism [60]. Second, different *textual features* are illustrated to represent text docu-

What is the problem? Define

Why it need solving? Significance, impact.

How has it been solved? Show the general ways. Broad, to specific (to the review focus).

Unique viewpoint of review.

Organizing Your Review: The Analysis

- Set out your thinking on paper through **maps** and **trees**.
  - Build conceptual map/theoretical framework
  - Build taxonomy/trees of area

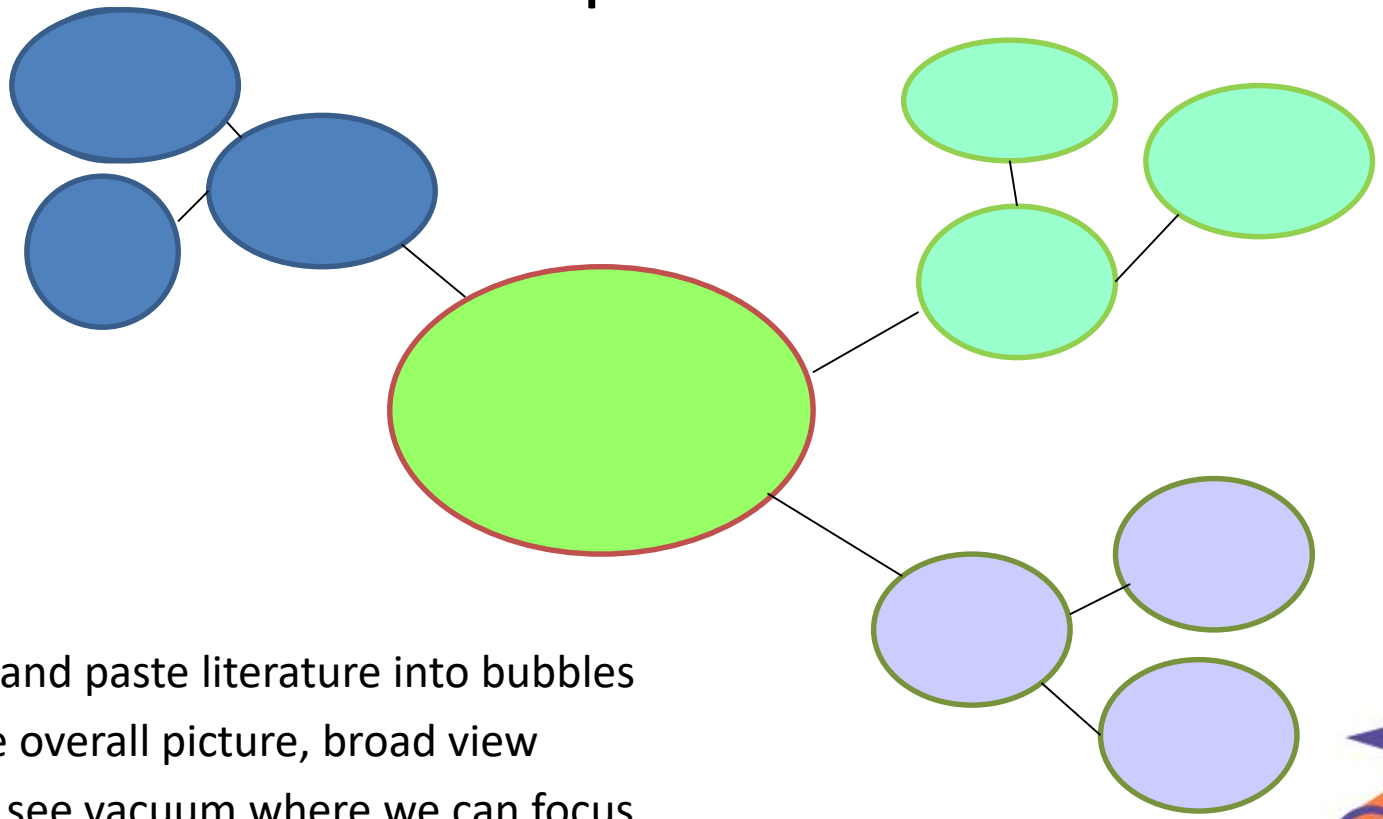
Feature map	Classifies and categorises your thought in tabular form
Concept map	Links between concepts and processes, or shows relationship between ideas and practice
Tree construction	Shows how topic branches out into subthemes and related questions or represents stages in the development of a topic.

Example of Analysis Framework

- What data used
- Different representation of data
- Different ways of comparing data or calculating values of similarity
- Different paradigms used for comparisons
- External knowledge-bases used to enhance. Eg.
  - Wordnet, Social networks

Conceptual framework

Start with a mind-map



- Cut and paste literature into bubbles
- Give overall picture, broad view
- Can see vacuum where we can focus
- Know where to put boundaries, scope, limitations

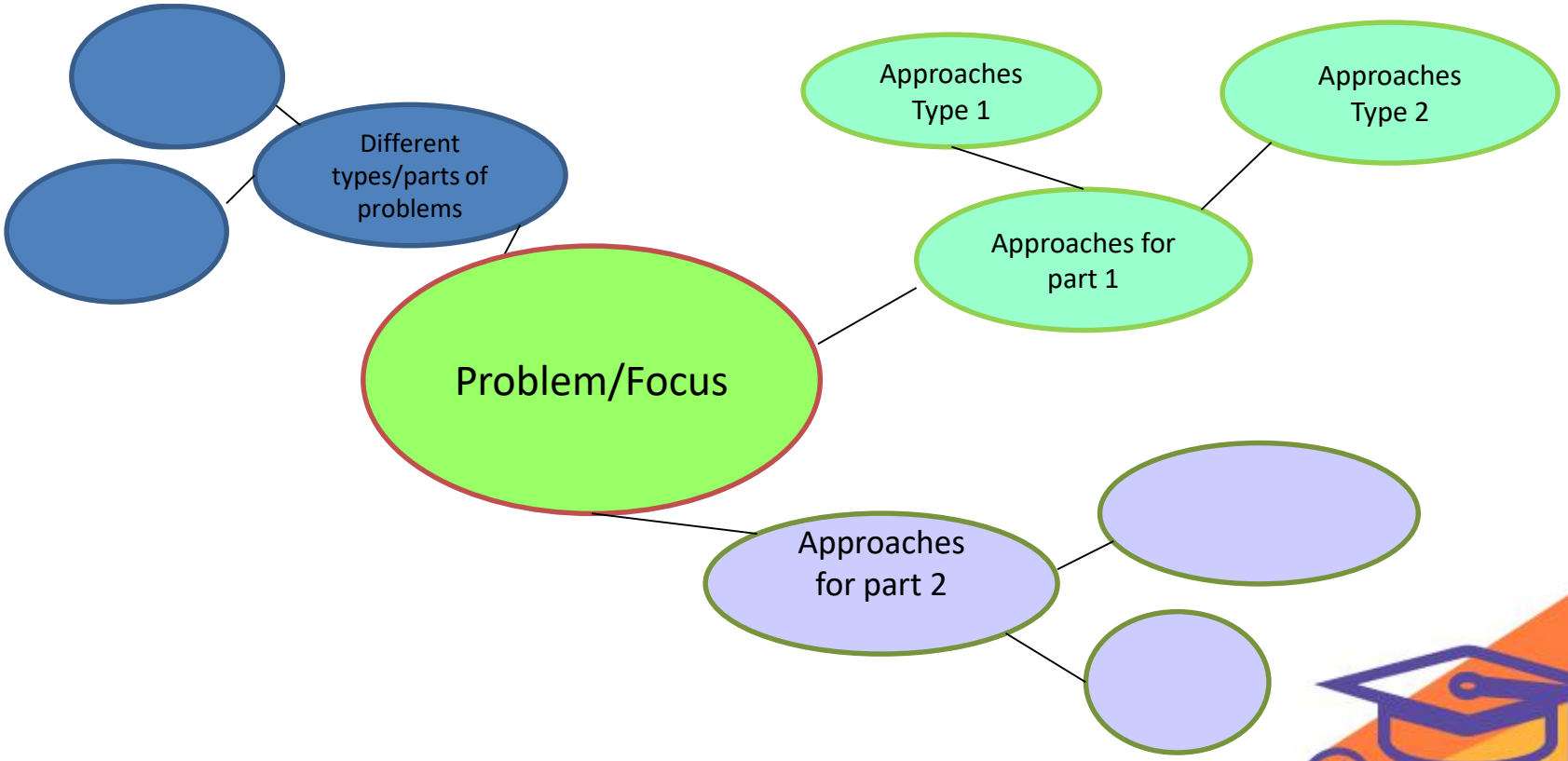
Conceptual Framework

- Building up of concept of work through literature

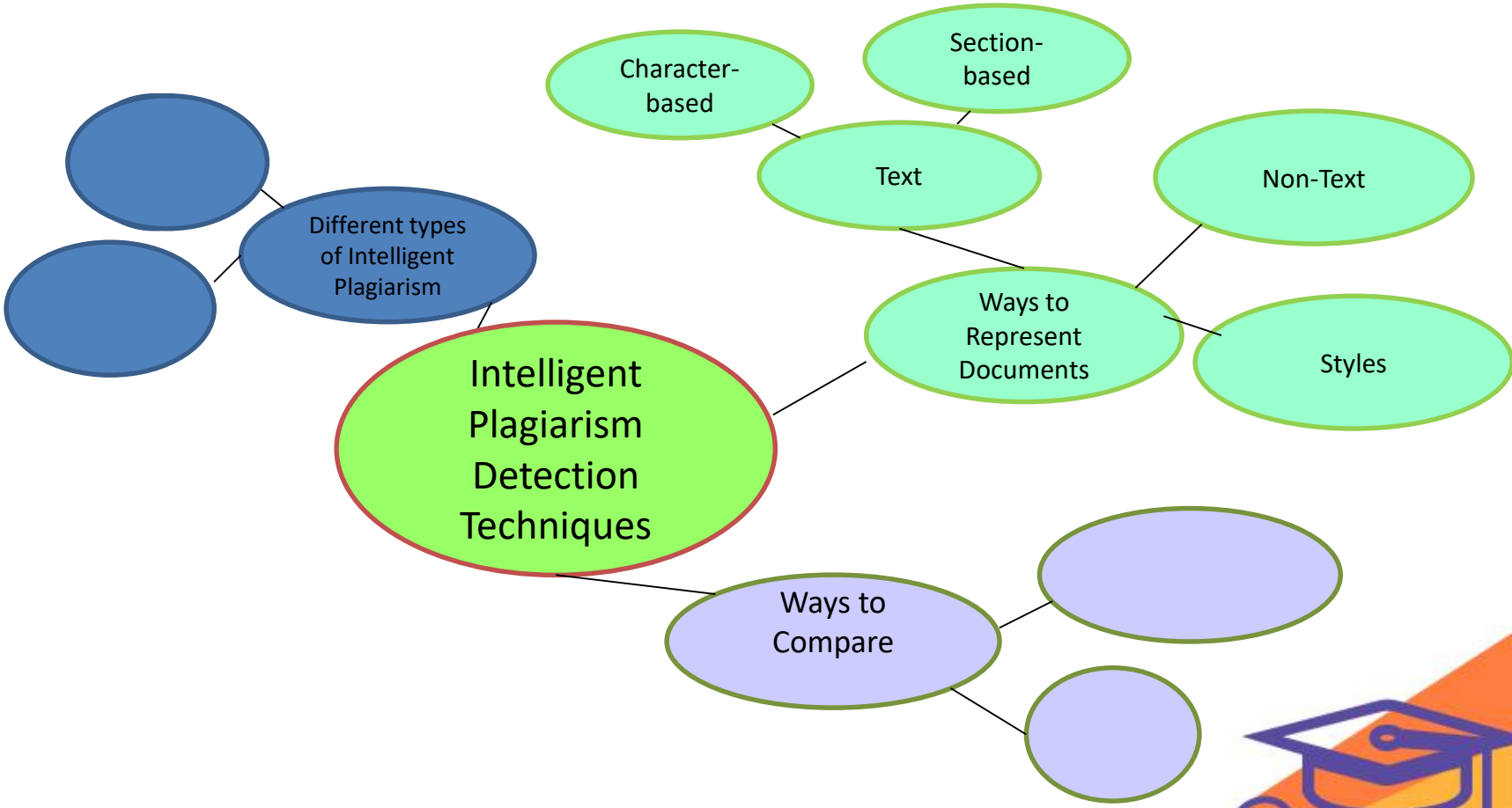
Theoretical Framework

- What theory support each component of the mind-map

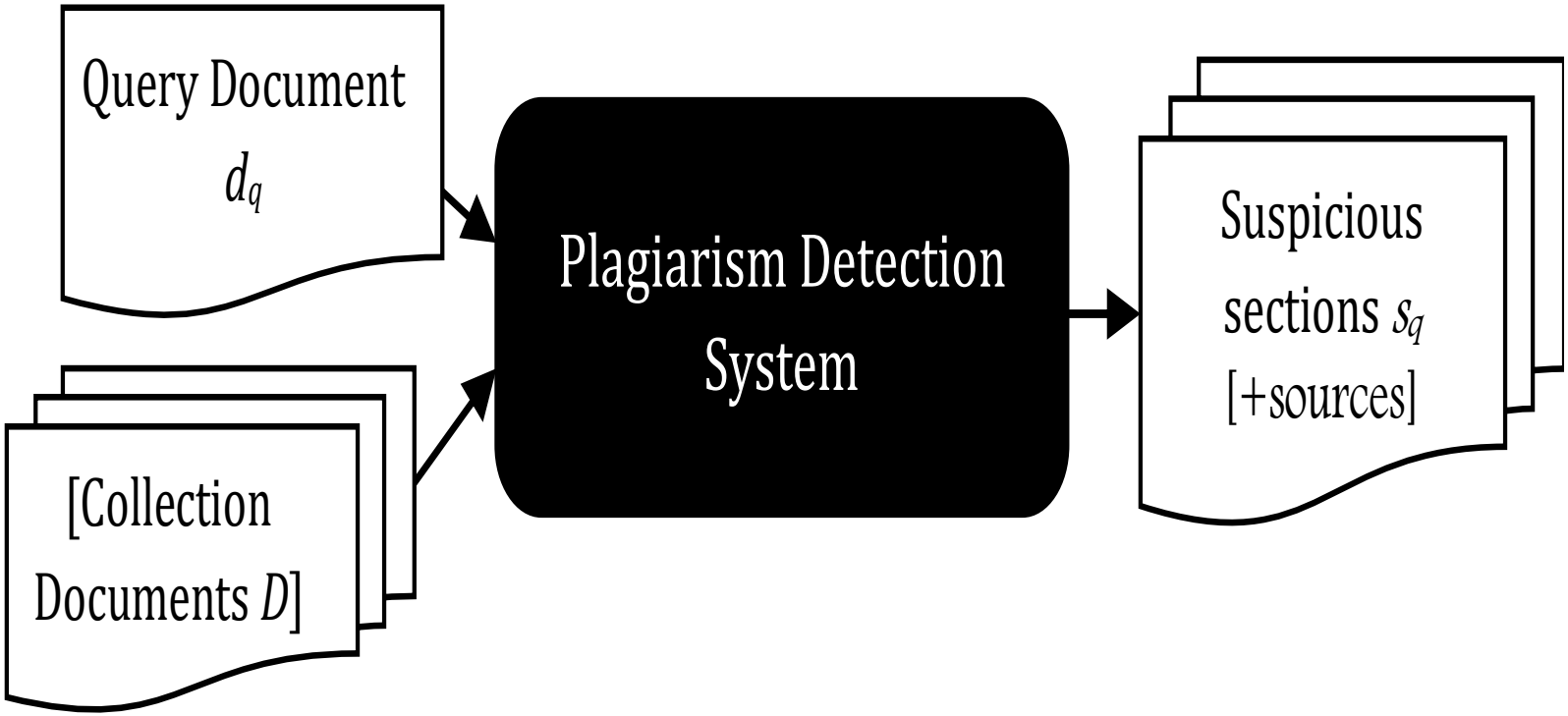
Conceptual framework



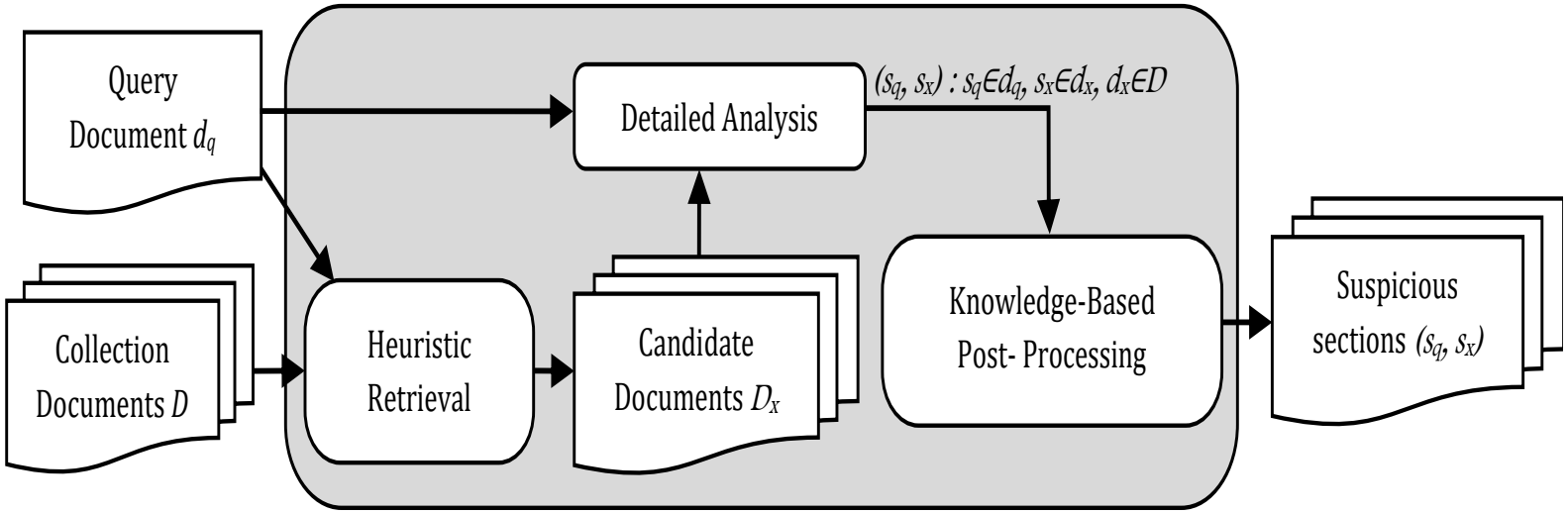
Conceptual framework



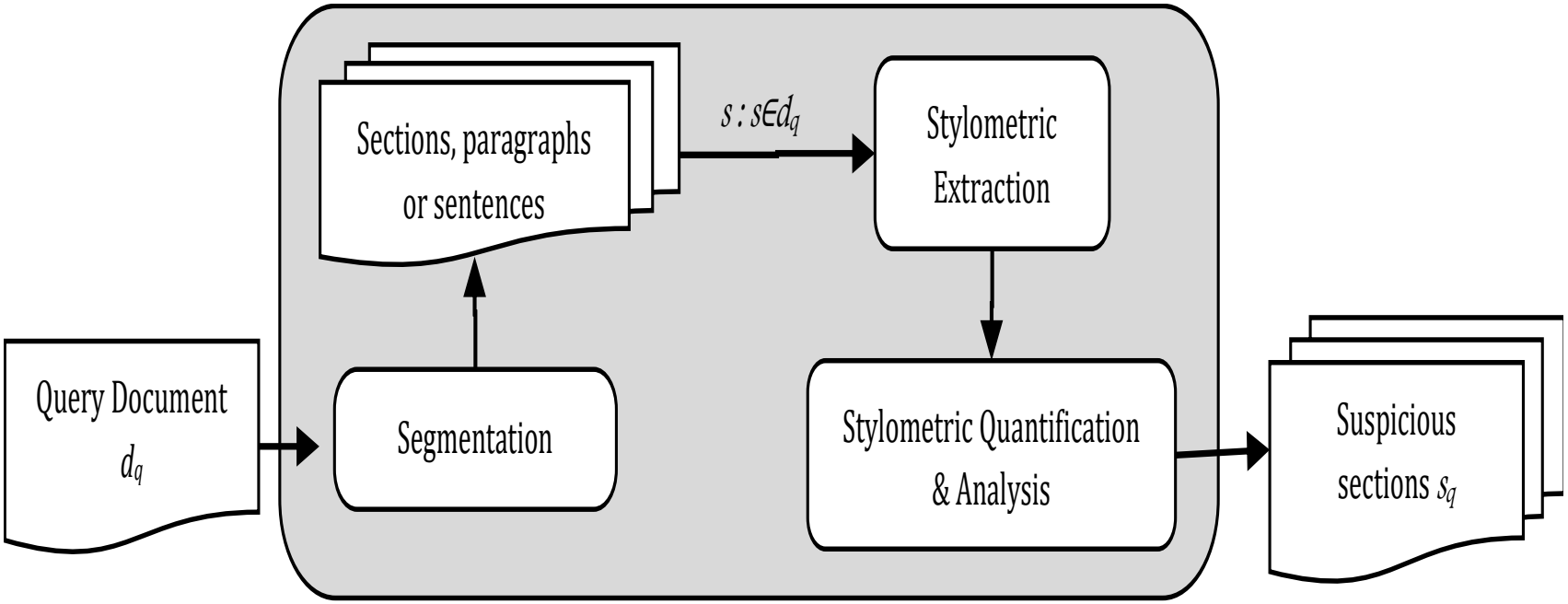
Eg. Framework of Problem Area



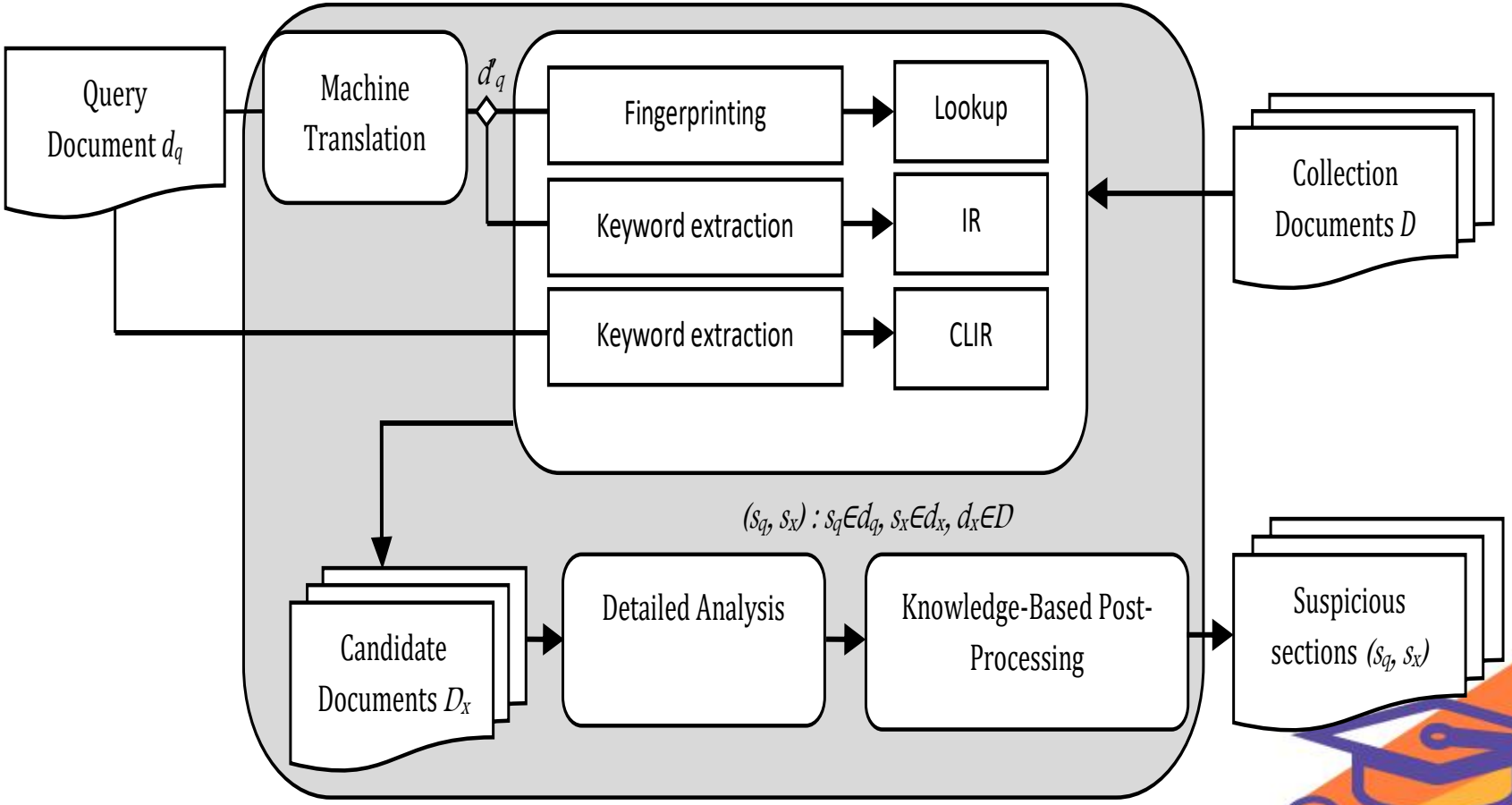
White-box design for extrinsic plagiarism detection system



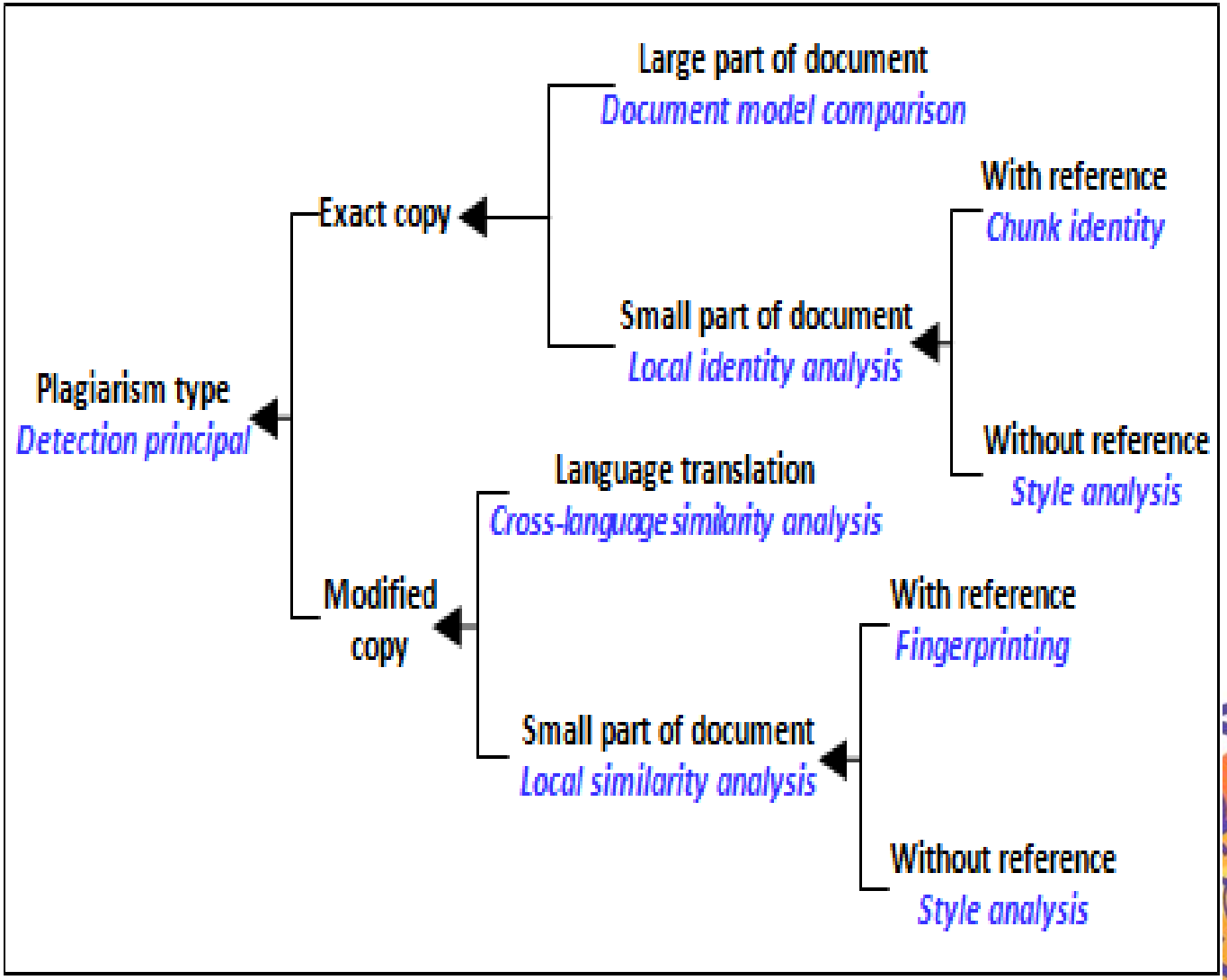
White-box design for intrinsic plagiarism detection system



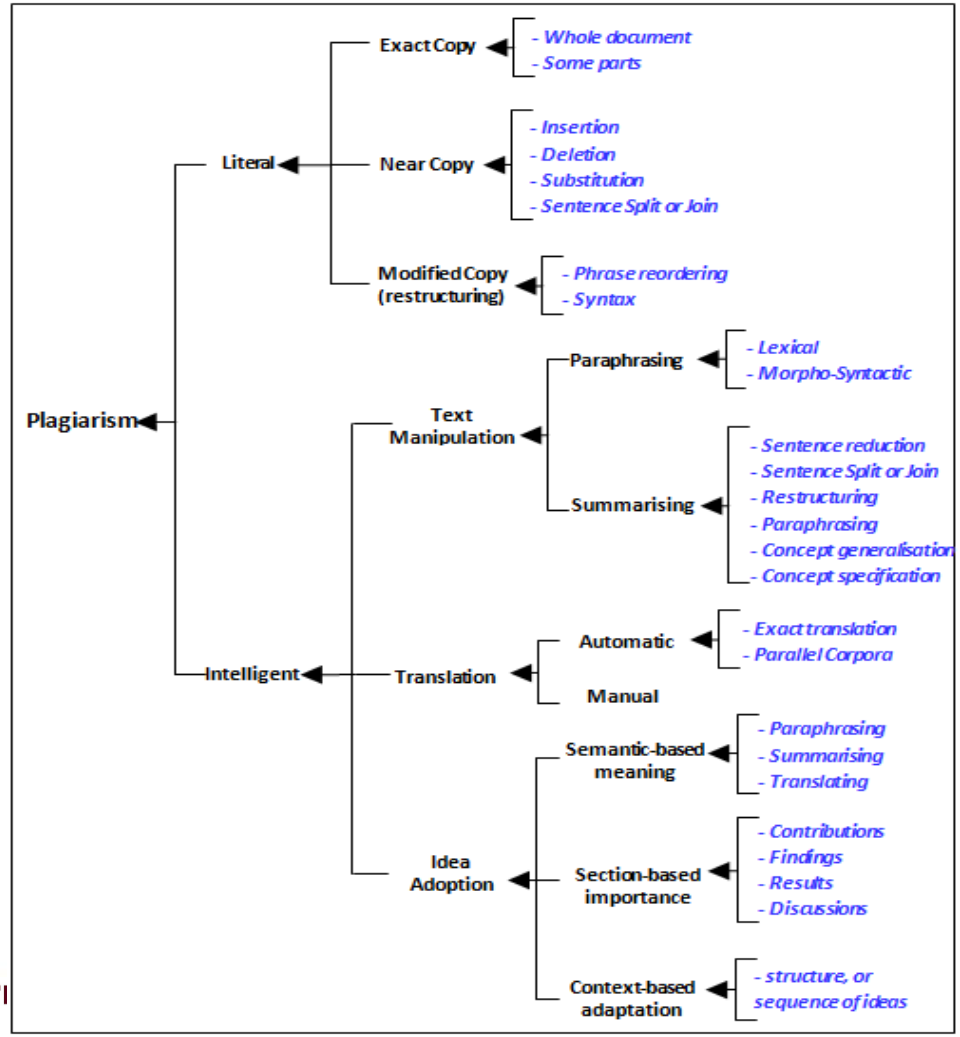
White-box design for cross-lingual plagiarism detection system



Eg. Current Taxanomy

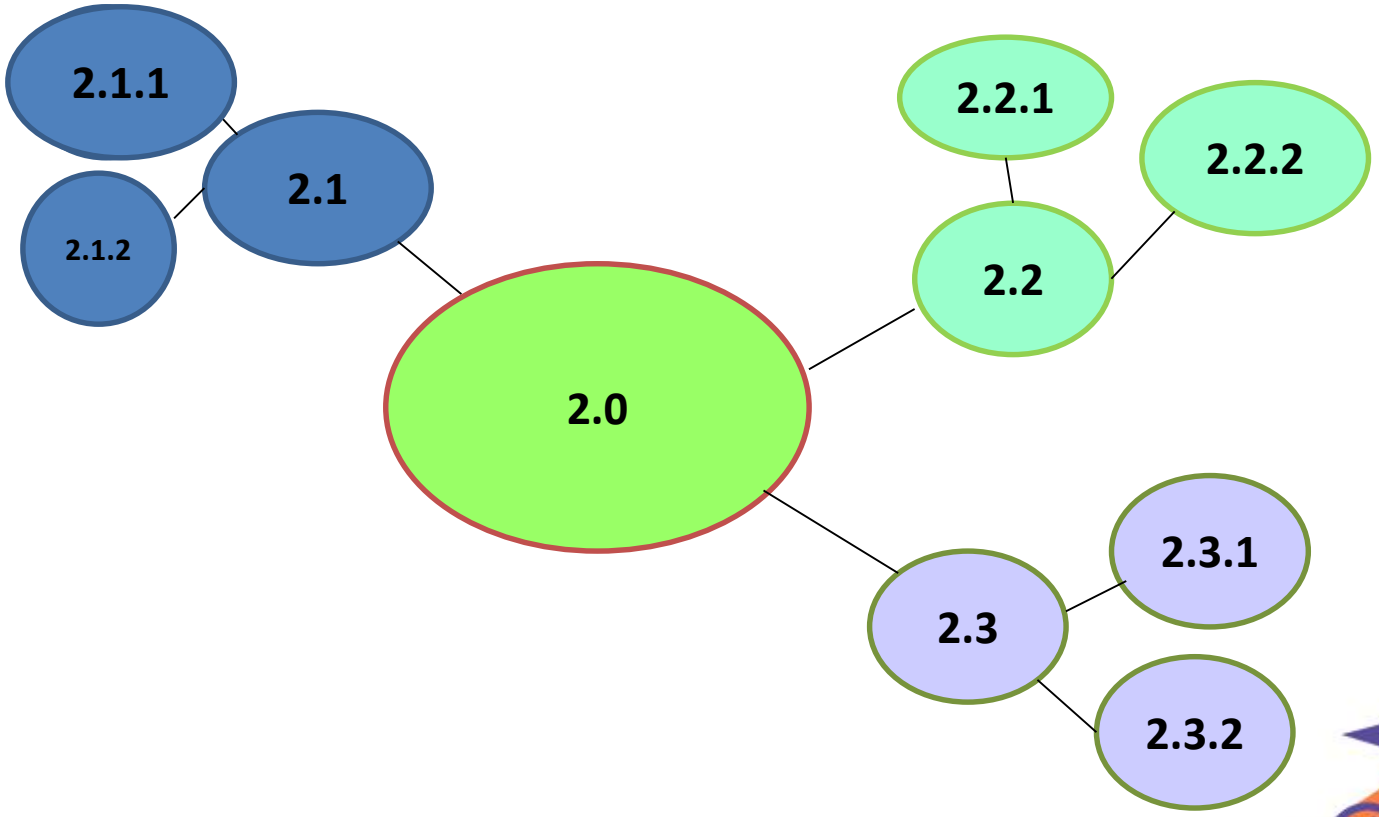


Modified Taxanomy of Concepts

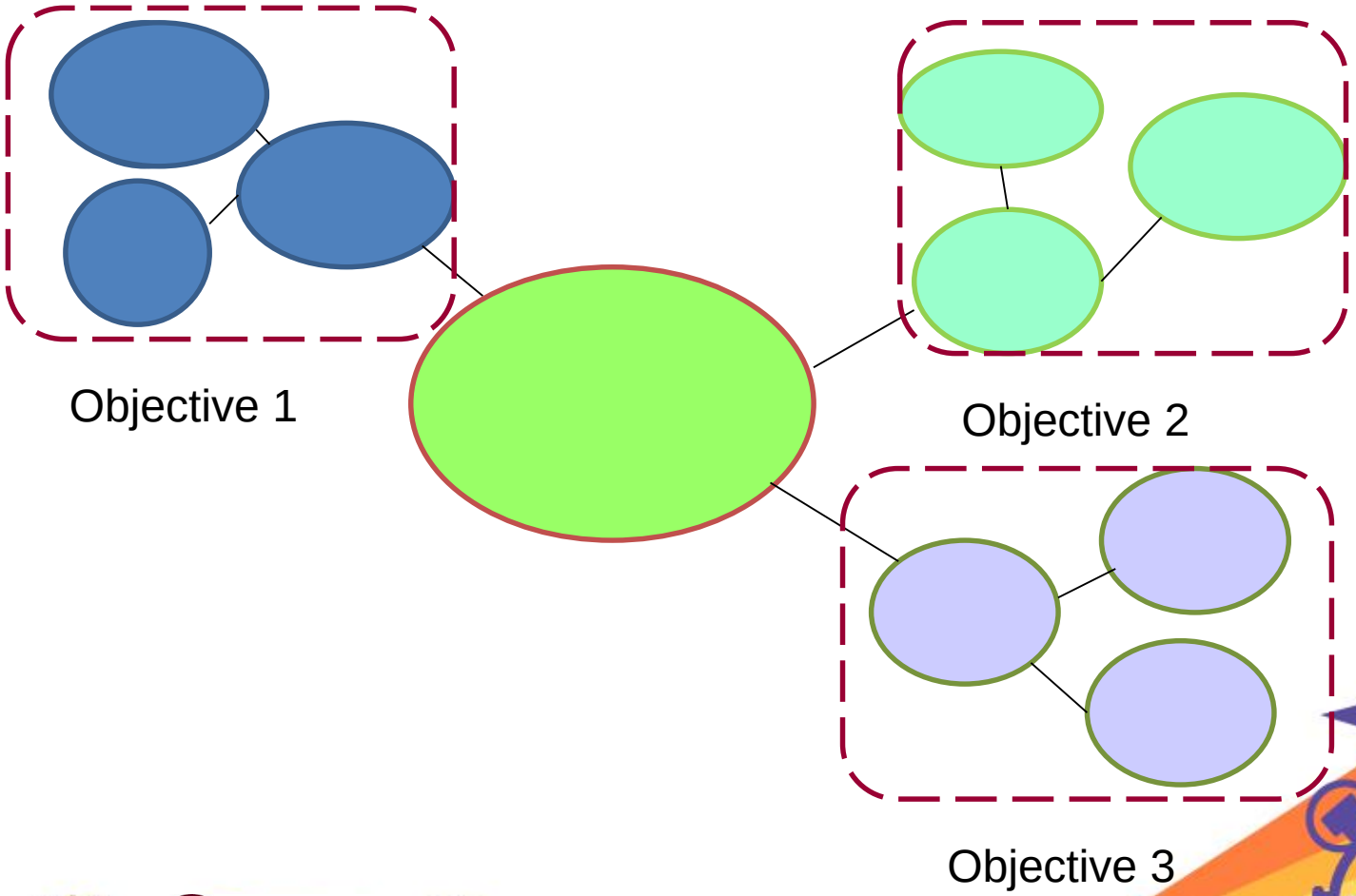


Conceptual framework

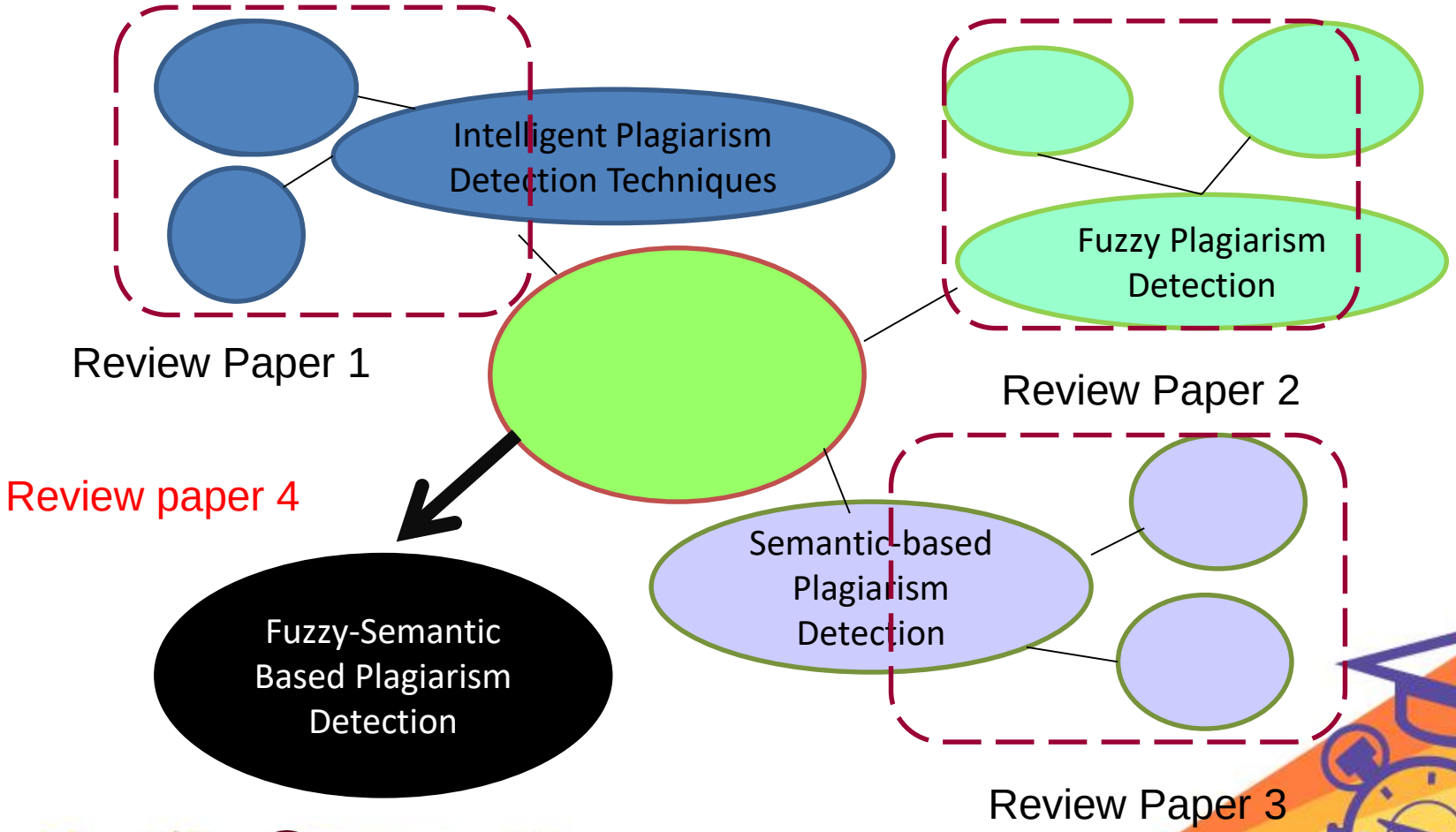
Organize LR chapter from mind-map



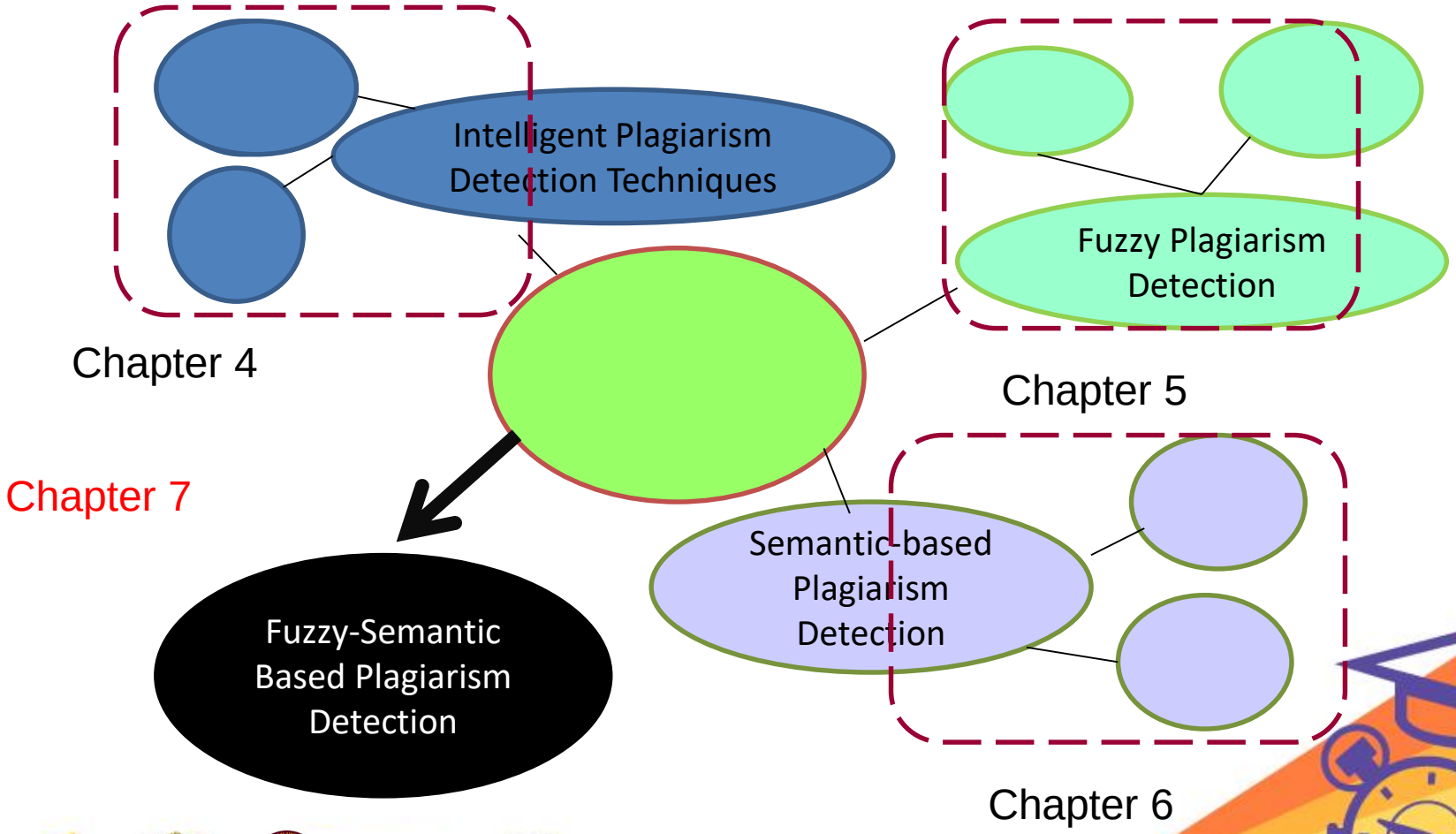
Conceptual framework – can result in research objectives



Conceptual framework into Review papers



Conceptual framework into Chapters



Tree Based Organization

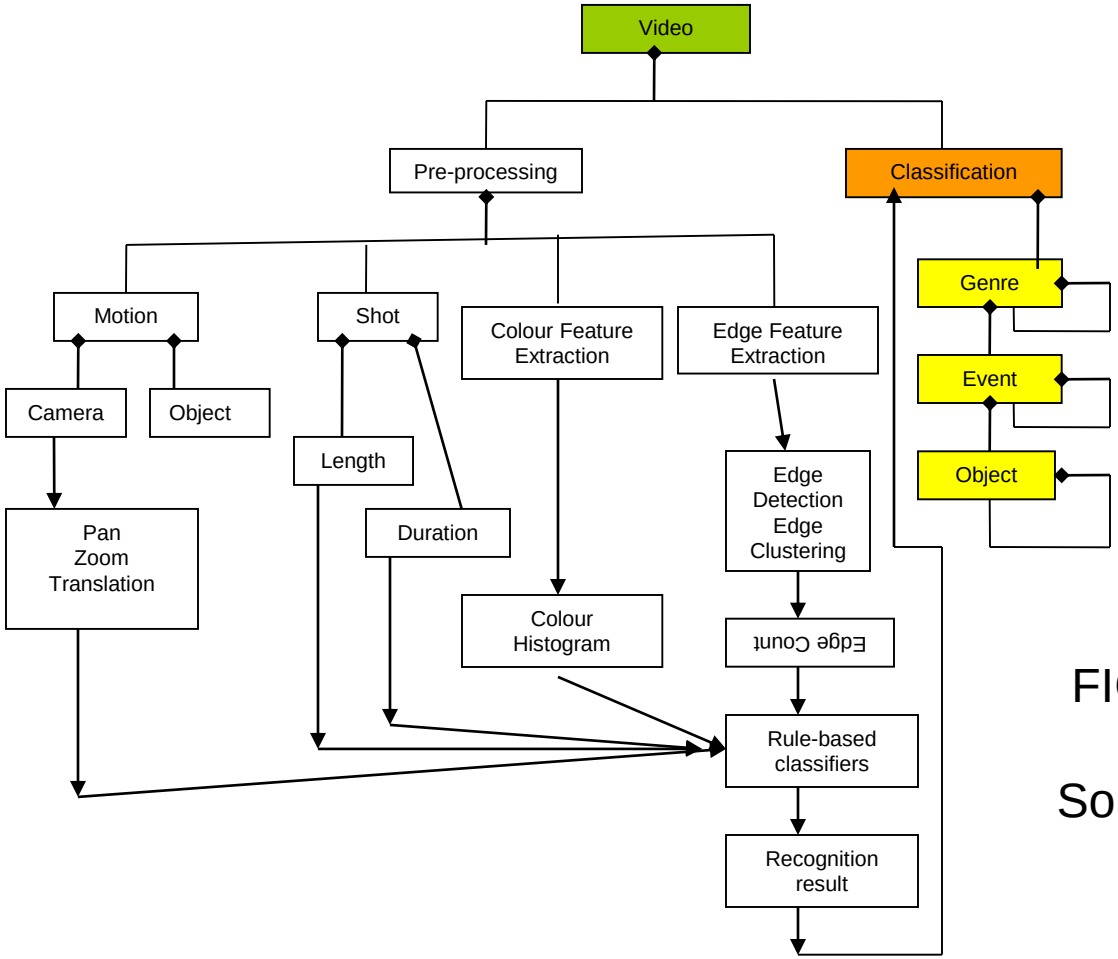


FIGURE 2.3 Schematic diagram for video classification
Source: L.N. Abdullah et al. 2005.

Synthesizing Sources in Each Section of the Review Paper

- Based on groupings in tables
- Inverted pyramid
- Thematic
- By trend
- Questions for Further Research
- Chronological

Representation

-	Examples	Required Tools and Resources	Ref.
Lexical features	Character n-grams (fixed-length)	-	[1]
	Character n-grams (variable-length)	Feature selector (e.g. n-gram weights)	[16]
	Word n-grams	Tokenizer, [Stemmer, Lemmatizer]	[2, 3, 17, 26] [30]
Syntactic features	Chunks	Tokenizer, POS tagger, Text chunker (Windowing)	[4]
	Part-of-speech and phrase structure	Tokenizer, Sentence splitter, POS tagger	[6, 12, 48]
	Word position/order	Tokenizer, Sentence splitter, Compressor (e.g. Lempel-Zif)	[13, 14]
	Sentence	Tokenizer, Sentence splitter, POS tagger, Text chunker, Partial parser	[16, 58]
Semantic features	Synonyms, hyponyms, hypernyms, etc.	Tokenizer, [POS tagger], Thesaurus	[14, 16, 18, 58] [30]
	Semantic dependencies	Tokenizer, Sentence splitter, POS tagger, Text chunker, Partial parser, Semantic parser	[14, 61]
Structural features	Block-specific	HTML parser, Specialised parsers	[21, 29]
	Content-specific	Tokenizer, [Stemmer, Lemmatizer], Specialised dictionaries	-

Similarity Evaluation

Vector Similarity Metric	Description & Equation	Equation	Range	Example	Ref.
Matching coefficient	-similar to Hamming distance but between vectors of equal length.	$M(x,y) = \lvert x \rvert - \lvert x \cap y \rvert$	0 to |x| Where |x|=|y|	x=[0.1, 0.2, 0.3, 0.4] y=[0.1, 0.2, 0.3, 0.5] M(x,y)= 1	[11]
Jaccard (or Tanimoto) coefficient	-defines number of shared terms against total number of terms. This measure is computed to one if vectors are identical.	$J(x,y) = \frac{\lvert x \cap y \rvert}{\lvert x \cup y \rvert}$	0 to 1	J(x,y)=3/5=0.6	[3, 7, 8, 21]
Dice's coefficient	-similar to Jaccard but reduces the effect of shared terms between vectors. This measure is computed to two if vectors are identical.	$D(x,y) = \frac{2 \lvert x \cap y \rvert}{\lvert x \cup y \rvert}$	0 to 2	D(x,y)=6/5=1.2	-
Overlap (or containment) coefficient	-if v_1 is subset of v_2 or the converse, then the similarity coefficient is a full match.	$O(x,y) = \frac{\lvert x \cap y \rvert}{\min(\lvert x \rvert, \lvert y \rvert)}$	0 to 1	O(x,y)=3/4=0.75 (or 75%)	[10]
Cosine coefficient	-finds the cosine angle between two vectors.	$Cos(x,y) = \frac{\sum_i (x_i, y_i)}{\sqrt{\sum_i (x_i)^2} \sqrt{\sum_i (y_i)^2}}$	0 to 1	Cos(x,y)=0.34/0.3421 =0.9939 ≈1	[9, 21, 26, 28]
Euclidean distance	-measures the geometric distance between two vectors.	$Ec(x,y) = \sqrt{\sum_i \lvert x_i - y_i \rvert^2}$	0 to ∞	Ec(x,y)=0.1	-
Squared Euclidean Distance	-places progressively greater weight on vectors that are further apart	$SEc(x,y) = \sum_i (x_i - y_i)^2$	0 to ∞	SEc(x,y)=0.01	-
Manhattan Distance	-measures the average difference across dimensions and yields results similar to the simple Euclidean distance	$Manh(x,y) = \sum_i \lvert x_i - y_i \rvert^2$	0 to ∞	Manh(x,y)=0.1	-

METHODS AND THEIR EFFICIENCY IN DETECTING DIFFERENT PLAGIARISM TYPES

Technique	Tasks		IR		Language(s)	Plagiarism Type(s)								Reference
						Literal			Intelligent					
	extrinsic	intrinsic	mono-lingual	cross-lingual		copy	near copy	restructuring	paraphrasing	summarising	translating	idea (section)	idea (context)	
Char-Based (CNG)	☑		☑		any	☑	☑							[1-6]
Vector-Based (VEC)	☑		☑		any	☑	☑	☑						[7-11]
Syntax-Based (SYN)	☑		☑		specific	☑	☑	☑						[6, 12, 13]
Semantic-Based (SEM)	☑		☑		specific	☑	☑	☑	☑	☐				[14, 15]
Fuzzy-Based (FUZZY)	☑		☑		specific	☑	☑	☑	☑	☐				[16-19]
Structural-Based (STRUC)	☑		☑		specific	☑	☑	☑	☐	☐		☐	☐	[21, 29]
Stylometric-Based (STYLE)		☑	☑		specific	☑	☑	☑						[22, 23, 32-35]
Cross-Lingual (CROSS)	☑			☑	cross						☑			[31, 36-38]

Suggested words to discuss table

- There are a number of main approaches in….
- One of the most popular/used/oldest technique is …., which has been used by ….(give refs)
- Another technique is …. (give refs)
- The two techniques are similar in terms of …… However, the first technique …(highlight difference)

For each (group) of technique, describe

- What is it?
- How is it done/performed?
- Why is it introduced/proposed? What advantages it offered over other techniques?

Discussion can be …

- Embedded in each (group of) technique described
- At the end of each section after the techniques are introduced
- In a separate “Discussion” section

Discussion and Evaluation

- Review should be evaluative and not merely descriptive.
  - For example possible reasons for similarities or differences between studies are considered rather than a mere identification of them.

Examples …

- Similarities
  - As can be observed, all the techniques discussed above used …
- Weaknesses
  - The problems that could arise with the use of such techniques …. Similar problems have been observed in ….
- Link to research questions
  - A possible ways to improve is ….

Critical Framework

- **Regardless of the method of one's research—subjective, textual, historical, empirical, etc.—an analytical lens must be used to interpret literature and data.**
- **For quantitative research**
  - **this framework is the logical or mathematical method by which the data is analyzed**
- **When analyzing or interpreting qualitative or textual research**
  - **choose an individual or, more likely, interdependent approaches or lenses through which that data or material is interpreted**

Eg. Of Discussion

Discussion

When choosing between clustering methods, a few factors need to be taken into account. These factors are discussed in the following sections.

Computational efficiency

Table 2.4 summarises the computational complexity of some of the clustering method discussed. Basically, non-hierarchical methods are usually more computationally efficient than hierarchical methods. The Jarvis-Patrick method is very computationally efficient because ……..

Ability to recover natural clusters in dataset

A study by Blashfield [1976] revealed that single linkage has the lowest agreement between cluster solutions and actual structure, whilst Ward's method has the highest. The superiority of Ward's method in producing meaningful clusters is confirmed by Adamson and Bawden [1981] ……

Effectiveness for intended application

Empirical results of tests that use evaluation criteria specific to the problem being studied can be used to get an idea of the most suitable clustering method. ……

Example of Comparisons

Selection method	Time-complexity	Space-complexity	Applicability
Hierarchic agglomerative (stored matrix algorithm)	$O(N^3)$	$O(N^2)$	Small files only
Reciprocal nearest neighbour (Ward's)	$O(N^2)$	$O(N)$	Up to a quarter of million molecules
Reciprocal nearest neighbour (Jarvis-Patrick)	$O(N^2)$	$O(N)$	Up to more than a million molecules, due to its lower constant proportionality in the time-complexity
Maximum-dissimilarity	$O(N^3)$	$O(N^2)$	General algorithm implies that it is applicable only to small files. However, $O(N^2)$ time complexity has been described for the MaxMin and MaxSum versions [Holliday et al., 1995; Higgs et al., 1997]. These versions can thus be applied to a million molecules [Higgs et al., 1997].

Review + Experimental Paper

Can Use Empirical Comparison After Dry Review

- Evaluation criteria
- Gold Standard, Benchmark datasets or Development of unique datasets based on criteria
- First, can do Baselines identification
  - Dry comparison based on criteria
  - Selection
- Evaluate alternatives empirically
  - Discuss based on performance criteria (efficiency, effectiveness, ease of use, etc.)
  - Justification, reasoning – look at specific formulation or nature of algorithms, mathematical proving, relate to current findings in the area or other areas
  - Identify weaknesses, gaps that lead to novel technique or fusion or hybrid proposed

Planning for GOT

PhD Research Schedule

No.	ACTIVITIES	Year 1				Year 2				Year 3			
		1	2	3	4	1	2	3	4	1	2	3	4
1	Literature Review												
2	Problem Formulation												
3	Initial Results												
4	Proposal Writing (Chapters 1, 2, 3, 4)												
5	Objective 1												
6	Objective 2												
7	Objective 3												
8	Thesis Writing												

No.	MILESTONE	Year 1				Year 2				Year 3			
		1	2	3	4	1	2	3	4	1	2	3	4
1	Review Paper												
2	Problem Formulation												
3	Completion of Proposal Writing & First Assessment												
4	Paper with some results												
5	Completion of Objective 1												
6	Completion of Objective 2												
7	Completion of Objective 3												
8	Thesis Writing Completion/Submission												

Read and Write Throughout PhD

- Ask student to write at least one chapter every semester
- Set-up publication agenda
  - A detailed, realistic, time-bound, publication plan for the research degree, including significant milestones and maintain progress towards its achievement

Publication Agenda

- Concept paper (s) -> Merged concept paper
- LR -> Experimental paper (s) -> Merged experimental paper
- Presentation to group -> paper
- 1 paper every three month
- Appointment by paper

Sample Publication Agenda

Time	What	Where	Sample Inclusion
1st semester	Critical Analysis of Literature	• Conference • Journal	• Framework of analysis • New Taxonomy • Specific review • Hypothesis dy
2nd semester	• Concept Paper • Empirical Comparison of Techniques • Assumption testing • Corpus design	• Journal	• Choose a number of performance/ selection criteria • Select a number of best techniques omparison ation of ments

Al Zahrani S., and Salim, N. (2011). "Understanding Plagiarism Linguistic Patterns, Textual Features and Detection Methods". IEEE Transactions on Systems, Man and Cybernetics. Vol 41(1): pp. 1-17. (WOS & Scopus indexed, Impact Factor: 2.06)

Salim,N, Whittle, M.W., Holliday, J.D. and Willett, P., (2003) " Analysis and Display of the Size Dependence of Chemical Similarity Coefficients" , Journal of Chemical Information and Computer Sciences, vol. 43(3): pp. 819-828. (WOS & Scopus indexed, Impact Factor: 3.643)

Salim,N, Holliday, J.D. and Willett, P. (2003), Combination of Fingerprint-based Similarity Coefficients using Data Fusion, Journal of Chemical Information and Computer Sciences, vol. 43(2): pp. 435-442. (WOS & Scopus indexed, Impact Factor: 3.643)

Sample Publication Agenda (cont.)

Time	What	Where	Sample Inclusion
3rd semester	• LR + Suggested framework • 1st Objective/ Experimental Paper	• Conference (framework) • Journal	• Introduction • Experimental Design • Results • Discussion • Conclusion
4th semester	• 2nd Objective/ Experimental Paper	• Conference • Journal (extended dataset)	• Introduction • Experimental Design • Results • Discussion • Conclusion

A. Abdo, Salim,N (2009), " Similarity-Based Virtual Screening with Bayesian Inference Network", ChemMedChem 4(2): pp. 210-218 (WOS & Scopus indexed, Impact Factor: 3.51)

A. Abdo and Salim,N. (2009) "Bayesian Inference Network Significantly Improves the Effectiveness of Similarity Searching Using Multiple 2D Fingerprints and Multiple Reference Structures", QSAR &Combinatorial Science, 28(11-12):1537-1545, (WOS & Scopus indexed, Impact Factor: 2.594)

Sample Publication Agenda (cont.)

Time	What	Where	Sample Inclusion
5th semester	• 3rd Objective/ Experimental Paper	• Conference • Journal (extended dataset)	• Introduction • Experimental Design • Results • Discussion • ion
	Paper	journal	• ction • Experimental Design • Results • Discussion • Conclusion

A. Abdo, Salim,N, C. Mueller, and P. Willett. "Similarity-Based Virtual Screening Using Belief Network", Journal of Chemical Information and Modeling, 50(6): pp. 1012-1020 (WOS & Scopus indexed, Impact Factor: 3.631)

Binwahlan, M. S., Salim, N. and Suanmali, L. (2010). "Fuzzy Swarm Diversity Hybrid Model for Text Summarization". Information Processing & Management. Vol 46(5). Pp. 571-588 (WOS & Scopus indexed, Impact Factor: 2.106)

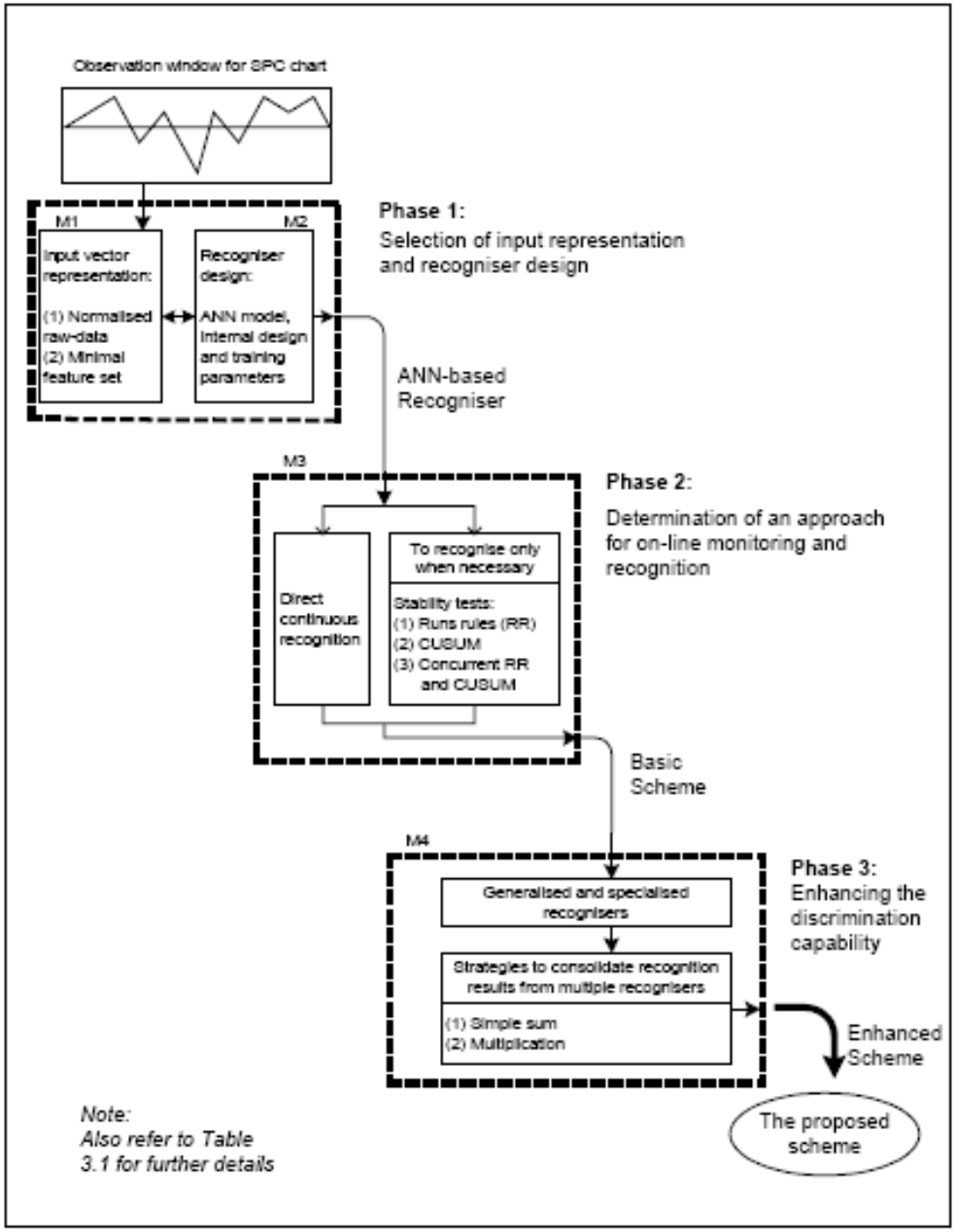
Mentoring in Experimental Design & Data Collection Phase

Do standard things

- Get idea from literature or graduated student. Why?
  - Someone has thought method out carefully
  - Saves time
- Learn what those standard things are (add only to test new ideas)
  - Datasets
  - Methods
  - Evaluation
- Statement must be supported by a reference to the scientific literature or by original work.

Draw Up a Research Framework

Phase	Objective
1	To develop an on-line recognition scheme that can perform timely and accurate recognition of CCPs even as they are developing
2	To develop improved recognisers that can perform accurate classification of partially developed CCPs. In particular, this research focuses on improving input representation and design of the ANN-based recognisers.



Draw Up a Research Framework

Phase	Activities	Resources Needed	Benchmark Data	Baselines for Comparison	Perfor-mance Evaluation	Objectives Addressed

- Diagram
- Table
- Description
- Gantt Chart

Performance Measurement

- **Analytical analysis**
  - will not give the final answers but help understand the concept
  - Eg: proof of validity of the major idea, rough estimation of the performance, rough estimation of the complexity,calculation of initial values for simulation analysis to follow,
- **Simulational analysis**
  - Use simulation
- **Implementational analysis**
  - Actual implementation

Performance Measures

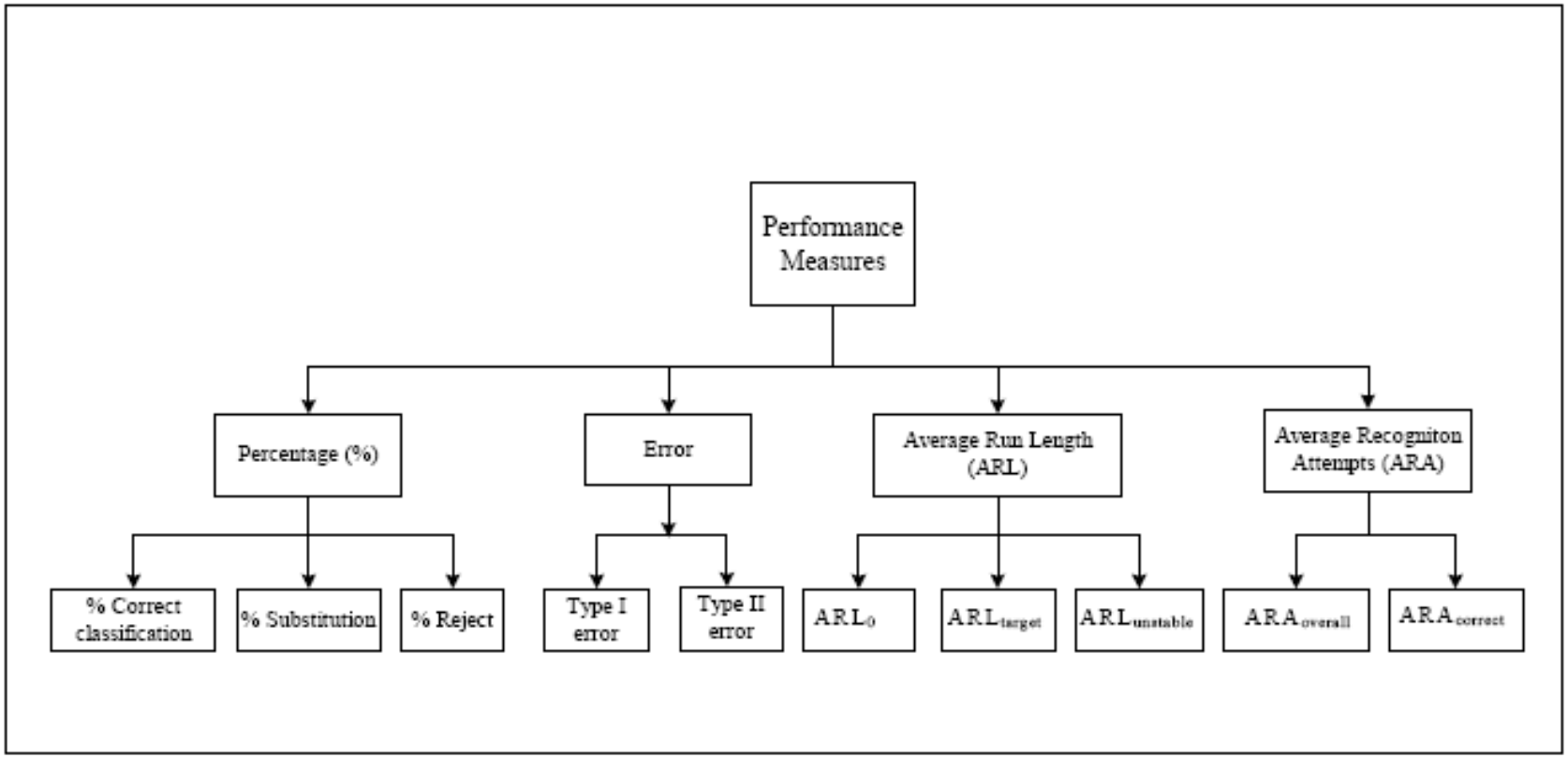


Figure 3.1: Performance measures used in the investigations

Coaching in Analysis of Results and Writing Up

Make an outline

- Think about the plan of chapters and decide what is best to report work.
- Then make a list, in point form, of what will go in each chapter.
- Try to make this rather detailed, ending up with a list of points that corresponds to subsections or even to the paragraphs of thesis.
- If planning each chapter and section before sitting down to write, the result will probably be clearer and easier to read. It will also be easier to write.

Order of Writing

- Draft outline and write literature review section
- Write up a preliminary version of the background section
- Draw up research questions and objectives
- Identify scope of study and (expected contribution)
- Write up the methods section.
  - include a description of the research equipment and relevant calibration plots.
- When you have some data, start making plots and tables of the data.
  - help you to visualize the data and to see gaps in your data collection
  - If time permits, you should go back and fill in the gaps.
  - You are finished when you have a set of plots that show a definite trend (or lack of a trend).
  - Be sure to make adequate statistical tests of your results.

Order of Writing (cont.)

- Once you have a complete set of plots and statistical tests, arrange the plots and tables in a logical order.
  - Write figure captions for the plots and tables.
- Once your plots and tables are complete, write the results section.
  - describe your results, but you must NOT interpret them. Be factual and orderly.
- Once you have written the results section, you can move on to the discussion section
  - talk about your ideas about the data

Order of Writing (cont.)

- In writing the discussion session:
  - explain your results, find reasons, derive conclusions
  - discuss the work of other authors who collected data on the same or related scientific questions.
  - discuss how their work is relevant to your work. If there were flaws in their methodology, this is the place to discuss it.
- After you have discussed the data, you can write the conclusions section.
  - take the ideas that were mentioned in the discussion section and try to come to some closure.
  - If some hypothesis can be ruled out as a result of your work, say so.
  - If more work is needed for a definitive answer, say that.

Order of Writing (cont.)

- Finally, write your recommendation
  - Make recommendations for further research or policy actions
  - If you can make predictions about what will be found if X is true, then do so.
- After you have finished the recommendation section, look back at your original introduction.
  - Your introduction should set the stage for the conclusions by laying out the ideas that you will test
- Write your abstract last.

Results Chapter

- Ask student to draft figures/tables first
- Make captions for every figure and table
- Explain figures and tables
- Discuss and interpret results
- Compare results with previous works

Results vs. Discussion Sections

- Quarantine observations from interpretations.
  - physically separate statements about observations from statements about the meaning or significance of those observations.

Discussion

Start with a few sentences that summarize the most important results. The discussion section should be a brief essay in itself, answering the following questions and caveats:

- What are the major **patterns** in the observations? (Refer to spatial and temporal variations.)
- What are the **relationships, trends and generalizations** among the results?
- What are the **exceptions** to these patterns or generalizations?
- What are the likely **cause**s (mechanisms) underlying these patterns resulting predictions?
- Is there **agreement or disagreement with previous work**?
- Interpret results in terms of background laid out in the introduction - what is the **relationship of the present results to the original question**?
- What is the **implication of the present results for other unanswered questions** ?

Summary

- Prepare before start of PhD
- Plan for GOT early in the course of study
- Ask student to commit for GOT
- Set up conducive environment for GOT
- Choose a topic that suits student best
- Write early, from the beginning and polish for coherent, smooth flow at the end
- Maintain effective, regular supervision

Thank you