

# TLIR 2008 – Introduction

London - October, 18th, 2008.

*Andrew MacFarlane, Juan M. Fernández-Luna,  
Juan F. Huete & Efthimis N. Efthimiadis*

# Introduction to the Workshop

- **Why TLIR?**
- Format of the workshop
- Review of Teaching and Learning in IR
- Summary/Q&A

# Why TLIR?

- IR is an important subject
- Many forms of dissemination: teaching
- Workshop Aim: provide common space for discussion
- 2nd Workshop in series
- First Held in January 2007
- Want to continue good work done here.

# Introduction to the Workshop

- Why TLIR?
- **Format of the workshop**
- Review of Teaching and Learning in IR
- Summary/Q&A

# Format of the Workshop

- 9:00am - 10:00am: Introduction, review of teaching and learning in IR.
- 10:00am - 10:30am: Coffee break.
- 10:30am - 12:30am: Papers.
- 12:30pm - 13:30pm: Discussions and close.
- 13:30pm - Lunch.

# Format of the Workshop

- Papers
- 10:30 - Clare Thornley. "Teaching information retrieval (IR) as a philosophical problem"
- 11:00 - Frances Johnson. "On the relation of search and engines"
- 11:30 - Ian Ruthven, David Elsweiler and Emma Nicol. "Designing for users: an holistic approach to teaching Information Retrieval"
- 12:00 - Fidel Cacheda, Diego Fernandez, Rafael Lopez. "Experiences on a Practical Course of Web Information Retrieval: Developing a Search Engine"
- Note: Review + papers inform breakout session

# Introduction to the Workshop

- Why TLIR?
- Format of the workshop
- **Review of Teaching and Learning in IR**
- Summary/Q&A

# Review of Teaching and Learning in IR

- Currently undertaking a review of literature in the area
- For Special Issue in Teaching and Learning for IR. Journal: Information Retrieval
- Consists of
  - Introduction + Review
  - 5 Papers revised and expanded from TLIR 2007 Workshop

# Review of Teaching and Learning in IR

- Categories/Taxonomy:
- Level 1
  - [A] Technical Level (Non-technical Mid-way to Technical continuum)
  - [B] Educational Goals
    - [b1] Library and information Science
    - [b2] Computer Science
    - [b3] MIS
    - [b4] Linguistics
- Level 2
  - [1] Teaching and Learning methods:
    - [1a] classroom;
    - [1b] elearning (distance learning);
  - [2] Assessment and feedback
  - [3] Curricula.

# Level 1: [A] Technical Level

- In terms of what students need to know from LIS to CS.
- More details very soon... 😊

## Level 2: [B] Educational Goals

- [b1] Library and Information Science
- Knowledge/skills: search intermediary
- Role of professional and accrediting bodies in defining goals - STRONG
- Range from abstract (analysis) to specific (skills). CILIP more abstract.
- 3 levels: HIB, seeking and retrieval
- General agreement on skills needed

## Level 2: [B] Educational Goals

- [b2] Computer Science.
- Precedent: Mooers (1950) introduces for the first time the term IR in Library documentation.
- Students' skills: formulating questions based on a user's information need, identifying potential sources of information, developing successful search strategies or evaluating the results of a search.

# Level 2: [B] Educational Goals

- Focusing on CS, IR is not a core part in the curricula.
- It is difficult to find courses in a CS degree. More common in Masters and doctoral programs.
- The focus is more on the technical and implementation aspects:
  - how information is stored in computer systems,
  - retrieval models,
  - efficiency issues,
  - evaluation methods, and
  - how to develop implicit techniques to facilitate search.

## Level 2: [B] Educational Goals

- (CS Curricula, 2001) expected that computer scientists devise new approaches for searching, develop the products and design complex, high-performance networks to cope with immense quantities of data.
- Not in one IR course but through the whole degree.

# Level 2: [B] Educational Goals

- The learning objectives of the elective unit “Information storage and retrieval (IM11)” from the Information Management (IM) area (CS Curricula, 2001), are:
  - Explain basic information storage and retrieval concepts.
  - Describe what issues are specific to efficient information retrieval.
  - Give applications of alternative search strategies and explain why the particular search strategy is appropriate for the application.
  - Perform Internet-based research.
  - Design and implement a small to medium size information storage and retrieval system.

# Level 2: [B] Educational Goals

- That's theory, but practice?
- To show basic concepts of IR in a first course.
- Defence of the acquisition of knowledge about basic technical concepts for search engines.
- To get a general idea of what IR is, its areas and interactions.
- In other works, not only IR basics, but to learn advanced techniques of IR is found as a goal.

# Level 2: [B] Educational Goals

- With respect to search strategies...
- To develop the learners' practical capability to search and understand the heuristic nature of IR techniques.
- Another educational goal is to improve the search skills of the students.
- Also to retrieve information using a variety of systems, learning their main features and how some are good for some tasks and other appropriate for other tasks.

# Level 2: [B] Educational Goals

- With respect to IR systems in more advanced courses...
- To acquire the necessary skills in order to develop new IR methods using software modules already developed.
- The understanding of the whole IR process.
- To train the students to think about relevant issues when analyzing problems and offer viable solutions by means of IR tools.

# Level 2: [B] Educational Goals

- Therefore, in the literature we can find three main educational goals:
  1. Knowledge on IR foundations (1 and 2).
  2. Training in search strategies (3).
  3. Knowledge on Information retrieval systems: processes and components (4 and 5).

# Level 2: [B] Educational Goals

- Some announcements in the field of IR, obtained from several mailing lists as SIG-IRList, IR, Doceng, WebIR, UM, UAI, among others, we have found the following requirements in most cases:
  - Strong knowledge of the Information Retrieval (IR) field.
  - Deep familiarity and hands-on experience in IR techniques.
  - Ability to conduct experiments involving massive data sources.
  - Understand those features relevant to Search Engine Optimization.
- Most of these requirements can be placed in at least one of the three educational goals that we have extracted from the literature and in the five learning objectives from the CS Curricula.

# Level 2: [B] Educational Goals

- [b3] Management Information Systems

Very soon... 😊

# Level 2: [B] Educational Goals

- [b4] Linguistics

Also very soon... 😊

# Level 3: [1] Teaching and Learning Methods

- [1a] Classroom
- Objective: delivery techniques and students' outcomes.
- Before starting teaching, acquisition of knowledge about the students in the class and their context.
- An experience showed several stages in the query formulation: initiation, topic selection, pre-focus exploration, focus formulation, information collection, and search closure.

# Level 3: [1] Teaching and Learning Methods

- Five strategies were developed for coaching students in search:
  - collaborating (to work with other students),
  - continuing (refinement of the information need),
  - conversing (to talk about what they already know about the problem and what they might be interested in finding out),
  - charting (organizing ideas graphically), and
  - composing (writing all the information gathered).

## Level 3: [1] Teaching and Learning Methods

- Cohen (2001): better to help understand the information need, analyse queries and redirect to the corresponding information sources than to exhaustively show the operation with the different search engines.

# Level 3: [1] Teaching and Learning Methods

- “Theory-based” vs. “practice-based”:
- The idea of teaching the underlying theoretical concepts vs. teaching how to use a specific system without understanding why the system does what it does.
- The argument is that when students understand the underlying principles then they can transfer the skills from one system to another easier.

# Level 3: [1] Teaching and Learning Methods

- Strategies are a key-concept: students must have a wide range of strategies to face the problem of searching, a skill in which they are often deficient.
- Bhavnani et al (2001):
  1. Learn the existence of specific strategies to execute frequent tasks;
  2. Learn when to use a particular strategy;
  3. Know how to execute a strategy;
  4. Learn to use the strategies across applications.

# Level 3: [1] Teaching and Learning Methods

- An example of a method to teach search strategies is presented in Walker and Engel (2006).
- Students are asked to carry out a first research exercise.
- The answers are collected and processed, allowing the instructors to analyse the search strategies used and show new ones based on examples.
- This two step process is repeated twice.
- The results showed an increasing use of sophisticated search strategies, which lead to a better quality of answers.

# Level 3: [1] Teaching and Learning Methods

- Analogously to search strategies, interesting to know the students' conception of how search engines works.
- “Conceptual approaches, metaphors, representations, and misconceptions”.
- Efthimiadis et al (2005) elicited this information by getting the students to draw a sketch about how they think a search engine works.

# Level 3: [1] Teaching and Learning Methods

- Summing up, most of the methodological approaches found are based on offering tools for searching than searching tools themselves.
- Instructors usually offer general knowledge about the IR field, and ask the students to apply them to specific contexts, problems or tools. But this is not a direct application = elaboration and study of the context.
- If this application is successful, then the objectives of the teaching and learning process will be achieved in terms of delivery and learning outcomes.

# Level 3: [1] Teaching and Learning Methods

- The use of IRSs in teaching and learning.
- Two perspectives:
  - For CS and specialized LIS students: how search engine works internally.
  - For non-specialized students: IRSs are tool to support learning.

# Level 3: [1] Teaching and Learning Methods

- Example of the second approach: Dialog system.
- Command language for interacting with the text databases.
- It is argued that this method is the best as it isolates the students from the graphical interfaces thereby giving the student an understanding of how an underlying search engine works,
- all the commands that they can find in this online searching application are available in other types of search systems.

## Level 3: [1] Teaching and Learning Methods

- Several papers about a computer-supported learning environment, called IR Game - Query Performance Analyser (QPA).
- Students perform searches against a test collection and are given specific feedback on the quality of their searches.

# Level 3: [1] Teaching and Learning Methods

- With respect to the first approach (how IRSs work)...
- IR Toolbox (Efthimiadis and Freier, 2007), “an experiential teaching tool for learning about information retrieval systems”. The student can learn the whole IR process (document analysis, indexing, searching and evaluation), without having to program, with different levels of complexity, and containing individual and group exercises.

## Level 3: [1] Teaching and Learning Methods

- IR Base (Calado et al, 2007). This is a toolkit, object oriented-designed, whose aim is the “integration of components, documentation and services, focused on the rapid development of prototypes for research and teaching”.
- A set of classes for IR utilities and models.

## Level 3: [1] Teaching and Learning Methods

- Learning outcomes in this last example:
- to be able to build a search engine using a set of classes giving basic functionalities,
- and to develop new modules to include new requirements (for example, new retrieval methods, indexing techniques, and so on).

# Level 3: [1] Teaching and Learning Methods

- Another approach: medium or large scale projects.
- students get a holistic view of the IR process, in opposition to the assessments approach, where the students deal with a specific problem each time, in the context of the subject, although they usually tackle them separately.

## Level 2: [1] Teaching and Learning Methods

- In this example, the students would also “acquire experience in a variety of aspects, including new technologies, system integration, database administration, and project management”, and not only knowledge in IR.
- In order to work in IR related jobs it is also very important this training.

## Level 3: [1] Teaching and Learning Methods

- [1b] e-learning (distance learning)
- Distance learning with a digital base is also a common technique used in the instruction of IR, in its different facets.
- The development of course for e-Learning in the IR field does not change very much from other fields, sharing common elements and basic e-Learning methodologies-

# Level 3: [1] Teaching and Learning Methods

- Two examples:
- (Henrich and Morgenroth, 2007): a wide range of distance learning courses on IR, in different levels of education.
  - Based on the Moodle platform as well as on DocBook standard for documentation.
  - Typical elements of such a learning style: e-Learning material in different formats (mainly, HTML, PDF and MS PowerPoint presentations), communication tools (e-mail, chat, and forums), evaluation modules, etc.
  - They also integrated interactive elements, like Java applets, with which the student can evaluate and learn methods and algorithms.

## Level 3: [1] Teaching and Learning Methods

- (Sacchanand and Jaroenpuntaruk, 2006): a “web-based self-training package for information retrieval using the distance education approach”.
  - Study modules (10 instructional multimedia modules on IR) and
  - self-assessments through pre and post-test) and
  - references and further readings.

# Level 2: [2] Assessment and feedback

Again... very soon... 😊

## Level 2: [3] Curricula

- What is a curricula?
- Who decides to put IR in curricula? Why does it go in?
- Professional & accrediting bodies
- Input: varying degrees of specificity
- Source is outcome focused
- Researchers own work
- Source is delivery focused

# Level 2: [3] Curricula

- What to put IR in curricula?
- Driven by educational goals
- Computer Science
  - Design and build IR systems
- Library and Information Science
  - Search intermediary
- Core subject in LIS
- LIS had broader view, but CS moves much quicker: implications?

# Level 2: [3] Curricula

- When does it go in?
- Pre-University students: information literacy
- Undergraduate: ideas in computing which make up IR e.g. data structures, interfaces etc
- Postgraduate: research and new developments.
- Specialist or generalist?

# Introduction to the Workshop

- Why TLIR?
- Format of the workshop
- Review of Teaching and Learning in IR
- **Summary/Q&A**

# Summary/Q&A

- Categories/Taxonomy:
- Level 1
  - [A] Technical Level (Non-technical Mid-way to Technical continuum)
  - [B] Educational Goals
    - [b1] Library and information Science
    - [b2] Computer Science
    - [b3] MIS
    - [b4] Linguistics
- Level 2
  - [1] Teaching and Learning methods:
    - [1a] classroom;
    - [1b] elearning (distance learning);
  - [2] Assessment and feedback
  - [3] Curricula.

# Discussion Session

- Textbooks
- Theory and Practice
- Specialist or Generalist? Or Teaching - Tailor made?
- Assumptions about prior knowledge
- Going under the hood
- Assessment: pool
- Separating concerns (indexer/crawler)
- TLIR Network: wiki, online, purpose? IRSG hosting?
- <http://www.bcs.org/server.php?show=nav.7927>

# Bibliography

- S. Bhavnani, K. Drabenstott, D. Radev (2001). Towards a Unified Framework of IR Tasks and Strategies. ASIST Annual Meeting, ASIST Annual Meeting.
- P. Calado, A. Cardoso-Cahopo, A. Oliveira (2007). IR-BASE: An Integrated Framework for the Research and Teaching of Information Retrieval Technologies. Proc. Of the 1<sup>st</sup> TLIR workshop.

# Bibliography

- L.B. Cohen (2001). 10 tips for teaching how to search the Web. American Libraries. November 2001.
- E.N. Efthimiadis, N.G. Freier (2007). IR-Toolbox: an experiential learning tool for teaching IR. SIGIR'07 Procs., 914. ASIST Annual Meeting,
- E.N. Efthimiadis, D.G. Hendry (2005). Search Engines and How Students Think They Work. SIGIR'05 Proceedings, 596 - 595.

# Bibliography

- A. Henrich, K. Morgenroth (2007).  
Information Retrieval as eLearning  
Course in German – Lessons Learned  
after 5 Years of Experience. Proc. Of  
the 1<sup>st</sup> TLIR Workshop.

# Bibliography

- C. Sacchanand, S. Thammathirat (2006). Web-based Development of a web- based self-training self-training package for package information retrieval using the distance education approach. The Electronic Library, Vol. 24 No. 4, 501-516.

# Bibliography

- H.M. Walker, K. Engel (2006)  
RESEARCH EXERCISES:  
IMMERSION EXPERIENCES TO  
PROMOTE INFORMATION LITERACY.  
Journal of the Consortium for  
Computing Sciences in Colleges 21, 4,  
61-68.