Learning Ecosystems for Dealing with the Copy-Paste Syndrome

Narayanan Kulathuramaiyer and Hermann Maurer

Abstract

The fact that people of all walks of life are becoming more and more reliant on a wide-range of easily-available digital content is often called the Copy-Paste Syndrome. It implies the indiscriminate usage of material, i.e. without checking for reliability or a concern for violations of intellectual property rights or plagiarism, a kind of cheating that has become uncomfortably widespread. A holistic approach is required to address this universal problem combining an institutional approach together with the application of viable technologies, rather than a-posteriory checks with software of doubtful reliability. This paper describes a learning ecosystem, ICARE, that addresses the Copy-Paste Syndrome by minimizing the possibility for unwanted copy-and-paste situations.

Key Words

Holistic E-Learning, learner-centered technology support, copy-paste avoidance, guided learning environment.

Introduction

The Web is experiencing a phenomenal growth with the explosion of user-generated content. As tools get easier to use, and access becomes more widespread, it also becomes easier for networked learners to misuse the possibilities for plagiarism and IPR violation (Pannepacker, 2007). It also will continue to become much simpler to acquire information from the Web community as opposed to meeting up with co-learners and experts in the real world (Alexander, 2006). The openness of the Web environment thus poses a number of challenges in monitoring and keeping track of the explorative expressions of learners.

The term Copy-Paste is used in this paper to refer to an emerging practice of fast and easy publication by millions of people. The "Google Copy-Paste Syndrome" (GCPS) (Weber, 2006), describes a common

activity of performing a fast, easy and usually "not diligently researched" copying of passages of text by people of all walks of life including scientists, journalists, academics and students. The GCPS has resulted in a proliferation of infringements such as plagiarism and IPR violations. Acquiring insights is performed by "conveniently searching" the Web as opposed to a rigorous process of learning through scientific discovery. Information from Web sources such as Google and Wikipedia are often used without even considering the validity of the source. According to Weber, GCPS and Web mining can actually impede the inquiry-driven scientific process, as answers conveniently pop up, with minimal effort. This syndrome thus endangers original writing and thinking by de-emphasizing the need for deliberate and insightful reasoning (Weber, 2006). This emerging phenomenon in turn encourages mediocrity in published works due to the lack of careful thought and understanding.

Due to the potential danger of the Copy-Paste Syndrome and intellectual property violations, it is vital to explore innovative means of addressing these issues.

We will concentrate our attention on the phenomenon of plagiarism and the Copy-Paste Syndrome (CPS). Current learning environments are often more concerned about identifying problem situations after they actually happen. The detection of plagiarism or Copy-Paste activities after some work is finished is neither reliable nor a good approach. Rather, there is a need to explore preventive ways of making sure that unwanted versions of copy-and-paste just cannot happen. A learning ecosystem coupled strongly with pedagogical aspects and techniques that control copy-and-paste situations throughout will best serve the emerging needs of educational institutions.

Dealing with plagiarism (and the Copy-Paste Syndrome)

Students are often expected to read the policy in the handbook and thereafter comply with a non-plagiarizing attitude. This approach is likely to be unsuccessful as the core problem lies in the student's lack of understanding of the concept of plagiarism and, most of all, their inability to deal with it (Kennedy, 2004). Students are also generally not aware of the full implication of the acts of Copy-Paste. They also do not value the importance of intellectual property or take pride in their ability to produce creative works (Kennedy, 2004). As pointed out by Duff et al. (2006), there is the lack of appreciation of the Western

system of scholarship, especially among new students and foreign students. There is thus a need to teach the skills required for paraphrasing, summarizing and referencing accurately (Kennedy, 2004).

There is a need to instill moral and ethical values in students regarding their education. Students will begin to understand the need to respect other people's copyright when they themselves are actively engaged in creating their own intellectual property (Midolo & Scott, 2003). Best practices in teaching and learning and academic integrity can be further achieved if students are aware that their input has been valuable and considered carefully by instructors (Kennedy, 2004).

Another proposed approach to address Copy-Paste Syndrome is through the employment of well-structured and clearly articulated assessment tasks. Course designers will have to carefully design courses and course content to ensure that they do not indirectly encourage plagiarism. Factors that encourage plagiarism include the same questions being set repeatedly year to year, questions that cannot be understood clearly or when clear criteria are not specified (Kennedy, 2004).

There are a number of approaches that can be employed to reduce plagiarism as suggested by works in Harris (2004). Instructors are encouraged to enforce the use of one or more sources not written within the past year. This approach effectively invalidates results of paper mills (Harris, 2004). By enforcing the use of one or more specific articles or specific information, students can be encouraged to formulate their own thoughts. Another effective technique describe by Harris is to enforce the production of assignments as a series of process steps as they lead to the final completion of projects. Student learning can then be continuously checked and assessed at each stage. The administration of personalized student tracking and assessment tends to overwhelm instructors. A careful selection of viable technologies is required to minimize the effort required. A technological platform also can be applied to guide students in using material from various sources in a constructive way and promote critical thinking.

Typical Approach for Dealing with Plagiarism (and also Copy-Paste)

A typical approach used in dealing with plagiarism in educational institutions is to employ tools for plagiarism detection such as Turnitin or Mydropbox. However, a single tool by itself is not adequate for Copy-Paste detection. A suite of tools is required to detect plagiarism or Copy-Paste effectively to establish and substantiate the detection of plagiarism with as much evidence as possible. An overview of a broad range of tools required for fighting Plagiarism and IPR violation is presented in Maurer et al. (2006). A layered application of plagiarism detected has been further proposed by (Kulathuramaiyer & Maurer, 2007) to systematically perform elaborate mining by focusing on relevant subsets of documents. Table 1 describes the availability of multiple approaches for detecting the various aspects of plagiarism. Despite the availability of these tools and techniques, their usage has mainly been employed in the detection of plagiarism and Copy-Paste situations. We propose the application of these tools and techniques in preventing the Copy-Paste Syndrome.

Table 1: Tools for Plagiarism Detection

Task	Tool
Manual Technique	Search Engines (Maurer et al., 2006)
Text-based Document Similarity Detection	Dedicated Software, Search and Web Databases (Maurer & Zaka, 2007)
Writing Style Detection	Stylometry software (Eissen & Stein, 2006)
Document Content Similarity	Semantic Analysis (Dreher & Williams, 2006; Ong & Kulathuramaiyer, 2006; Liu et al., 2006)
Denial of Plagiarism	Cloze Procedure (Standing & Gorassini, 1986)
Content Translation	Normalized Representation (Maurer & Zaka, 2006)
Multi-site Plagiarism	Distributed Plagiarism (Kulathuramaiyer & Maurer, 2007)

Comprehensively Addressing the Copy-Paste Syndrome
Rationale

In exploring a technological solution to comprehensively address the Copy-Paste Syndrome, the first question clearly is: Will it be ever be possible to comprehensively address the Copy-Paste Syndrome by software to check a paper submitted without any knowledge how the paper was compiled? Our answer is a clear "no." We have pointed out the existence of paper mills (Paper Mills, 2006) that even prepare papers to order (Kulathuramaiyer & Maurer, 2007). Documents also may contain large portions that are translations of some material in a not-so-common language making it nearly impossibly to find out if material is plagiarized. Furthermore, there are large collections of materials available in either closed databases or not in digitized form that are not available to any plagiarism checking software. As such a different approach is needed, we believe the key issue is to monitor the work of learners continuously.

We will discuss issues of an E-Learning ecosystem called ICARE[a]. We are trying out a number of components of a learning ecosystem at Graz University of Technology. We refer to the proposed suite of software and content as ICARE, aimed at controlling Copy-Paste situations.

The Main Concept of ICARE

ICARE stands for **I**dentify-**C**orrelate-**A**ssimilate-**R**ationalize-**E**xpress. ICARE denotes the five steps involved in the cultivation of academic reading and writing. These steps can be elaborated as:

- Identification: Identify key points (relevant) while reading a text document
- Correlate: Associate reading with concepts in the mind map of a learner
- Assimilate: Associate concepts learned with prior knowledge of learner
- Rationalize: Formulate ideas based on concepts arising from student learning
- Express: Express idea in learners own words

---

[a] To be read as 'I Care'

As opposed to the inadvertent (improper) Copy-Paste, ICARE enforces care on the part of the students' understanding of concepts, enabling them to apply learned concepts in the appropriate manner. The proposed approach to Copy-Paste will thus be seen as focusing on deeper appreciation and understanding ("care-why learning") as opposed to a less-diligent focusing on facts ("know-what learning"). Figure 1 contrasts these two forms of learning. Learning should not be based on a mere a collection of facts; it should rather be viewed as a connection to a learner's thoughts (Sathya Sai Baba, 2001). Support mechanisms are required to allow students to connect readings to the construction of knowledge. We believe that E-Learning systems should focus more on personal knowledge management activities and in fostering a deeper understanding. This practice will then effectively reduce the occurrence of improper Copy-Paste.
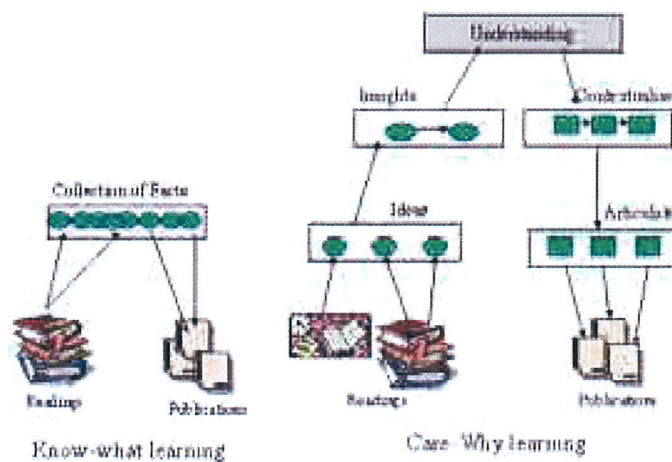

Figure 1: Types of Learning Modes

Practicing a constructive form of Copy-Paste supports a learner's ability of absorbing concepts, and consolidating and assimilating them before expressing ideas with a deeper understanding. The proposed ecosystem guides and allows students to become aware of the correct approach of reading, digesting and applying knowledge. At the same time, the platform fosters creativity in their associational and expressive ability. The proposed ecosystem allows an instructor to view and monitor the learning process of students, in observing and monitoring the rightful practice of "Copy-Paste skills." At the same time, creative expressions of students can be pinpointed, highlighted and recorded.

Toward a Holistic Learning Ecosystem

Although a variety of forms of E-Learning have been explored, the predominant form of E-Learning employs an E-Book paradigm. For the proposed ecosystem, however, multiple learning paradigms need to be incorporated. It will also need to enable pedagogical aspects of learning via technology-enhanced knowledge transfer (Helic, 2007).

Current E-Learning systems tend to employ a blended learning environment that involves the combination of instructional modalities or the instructional methods via the combination of online or face-to-face instruction (Graham, 2004). Options currently available in such learning systems (Graham, 2004) include self-study options such as Web-based courseware, simulations, systems and books together with live teaching options such as Web-casting, live video, conference calls, and instructor-led training. Each of these are often treated as standalone training objects delivered either via face-to-face (F2F) or computer mediated learning (CML) instruction (Valiathan, 2002). In this case, each training object represents a particular modality of learning where CML training objects are seen as alternative learning modes to classroom-based F2F approaches. The main weakness of this approach is that it does not allow composing training objects that contain both aspects of F2F and CML.

The realization of ICARE requires an E-Learning ecosystem that mixes F2F and CML within the context of a learning scenario that also minimizes the unwanted use of copy-and-paste by guiding the learner through the process. ICARE enables the complementary use of technology to harness the systematic development of both personal learning and collective intelligence. Table 2 describes the differences between blended learning in traditional E-Learning and the proposed learning ecosystem.

Table 2: Comparing the Proposed Learning System Functionalities Against a Typical E-Learning system

Teaching-Learning Activity	Typical E-Learning Environment	Proposed E-Learning Ecosystem
Announcements (Communicating timely messages to students)	Learning Management System or Email	Dynamically Activated from an Event database, RSS feeds
Overview session	Email, E-Books	Reading Scenario (E-Room)

Self-paced learning	Web-based tutorial. E-books simulations	Learning Specifications, Project-Rooms, E-Books
Student Question Answering	Email , Frequently Asked Questions	Active documents, schedule E-mentoring sessions
Assessment	Simulations, Online test, Submission system	Knowledge maps, Testing scenarios (can be personalized, collaborative, or peer-reviewed); Student Activity Logs and Reports
Collaborative Sessions	Discussion groups, Bulletin Boards, Chat	Brainstorming Scenario, Peer ranking
Feedback	Email	Examination Rooms
Continuous Assessment	Student Records	Student Portfolio, Learning Plans, Performance Monitoring Tool

Realization of the ICARE Ecosystem

Overall Design

ICARE will incorporate many of the experimental features in WBT-Master (WBT, 2006) coupled with the knowledge management capabilities found in Hyperwave Information Server (Mödritscher et al., 2005). It also will be augmented with a Copy-Paste detection and administration suite together with specifically prepared E-Learning modules to address both the issues mentioned earlier and anticipate future learning needs.

We propose additional functionalities to the E-Learning platform, (currently not available in any E-Learning system we know) for providing personalized guidance for administering academic reading and writing. Our previous works in the development of tools for fighting plagiarism and IPR violation has provided insights on the requirements of the proposed ecosystem (Kulathuramaiyer & Maurer, 2007).

Table 3 summarizes the functions to be incorporated describing the technological requirements for the ICARE ecosystem. The ecosystem employs an effective administration of E-Learning together with powerful tools for guiding and managing student learning and interaction.

The various learning functions together with well-designed assessments are crucial. In the next section components of an experimental system will address these issues. Tracking and analysis will be required to keep track of a variety of student works such as term papers, projects, examinations, etc. Tracking of activities also will be important in providing insights on learning and knowledge creation activities. The Copy-Paste handling suite of tools and techniques are required to assist and support the learner in the mastery of the rightful Copy-Paste skills. Specifically developed E-Learning modules enable the learners to master the fundamentals of academic reading and writing and promote an understanding of academic publishing culture.

Table 3: ICARE Ecosystem: Needs vs. Required Functionality

Supportive Environment Needed:	Functional Requirements of Ecosystem
Effective Administration of Learning	Ability to incorporate pedagogy in a learning environment combined with an ability to structure assessment; this includes the ability to discover and visualize student learning (knowledge maps) and integrate this with assessment. The management of capability-driven student learning, ability to manage and guide collaborative group-centered (project) work and flexible design of assessment tasks to manage learning as a series of steps
Guided Learning Process	Controlled environment for keeping track of learner activities, and workflow management and compliance checking
Tracking Learners' Copy-Paste Activity	Integrated Copy-Paste handling capability enabled by a suite of similarity checking software
Appreciation and Mastery of ICARE Principles and Process	To Incorporate E-Learning Modules on: • Western Scholarship • Academic Reading and Writing

	• Valuing Intellectual Property and Ethics

Key Components of ICARE Ecosystem

The following features from our past experimental developments in projects such as WBT-Master and Hyperwave will facilitate the realization of the learning ecosystem:

- Ability to define training scenarios as training objects or study rooms: A controlled environment can be established to track both the explorative and collaborative activities of students (Helic et al., 2004a).
- Pedagogy driven learning: A teaching scenario or environment can be built where a tutor works with a group of learners in both synchronous and asynchronous mode, leading them to achieve a particular learning goal.
- Project-oriented learning: A controlled learning environment can be built to allow a group of learners working together on a project, e.g., a software engineering project (Helic et al., 2003).
- Adaptive discovery of personalized background knowledge: A reading room paradigm can be created for enabling learners to chart their knowledge discovery process. This can be supported by the automated linking to related contents or background knowledge (Mödritscher et al., 2005).
- Annotations: Annotations allow the attachment of text segments, system or media objects or an URL to a learning object or material (Korica et al., 2005). It is possible to annotate any kind of material such as papers, parts of a digital library, other user contributions, etc.
- Active Documents: The idea of active documents presents an efficient way of students learning in a collaborative question-answering environment. Active documents present an innovative mean to demonstrate student learning and at the same time, an effective way for an instructor to direct knowledge discovery (Heinrich & Maurer, 2000).
- Visualisation as knowledge maps: The cluster of a document with documents containing similar concepts or ideas can be visualized via a knowledge map typical of knowledge management systems. A knowledge map with similar articles can be created and visualized (Helic et al., 2004b). "Knowledge cards" are used to describe a particular concept (i.e. *semantic entity*). Knowledge cards may be combined into a *semantic network*. For example,

the knowledge card "Student's Discovered concept" may be related as "is a part of" to the knowledge card "Course Domain Ontology."

- Workflow management and compliance checking capabilities: Learning can be visualized as a process flow of learning tasks. Non-compliance can then be automatically flagged by the system.

Controlled Environment for Pedagogy-Driven E-Learning

ICARE provides a controlled environment in which the instructor is able to track the usage of reference materials by students. Such a controlled environment makes it much easier to curtail unethical practices and also promotes constructivist learning among students. Furthermore, user tracking and user activity-logging facilities also can be used to enforce learners to read certain parts of a document before being allowed to annotate an article or ask questions about some part of it (Helic et al., 2004a).

An environment that closely monitors students' knowledge construction and collaborative activities can help the instructor to assess and guide students' ability to publish effectively. Process level support can be achieved via the workflow management and compliance checking capabilities of systems. The system can be trained to recognize non-conforming patterns to be able to flag instructors. Discovered patterns regarding a student's learning can then be captured and stored within a learner's profile. Knowledge profiling is supported in the acquisition, structuring, and reuse of extracted expert knowledge. By maintaining individual learner profiles, personalized learning can be supported. Personalized learning units then can be designed for each student as shown below:

A Read this unit (from 16.06 2000 till 31.12 2099)
Introduction to Databases
Resources:
☐ Learning Unit (allcoursescontent/bank/bank02.cif)
A Read also this document (from 16.06 2000 till 31.12 2099)
A Fill out questionnaire (from 16.06 2000 till 31.12 2099)
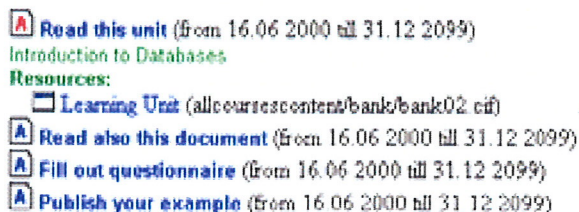A Publish your example (from 16 06 2000 till 31 12 2099)

Figure 2: Personalized Learning Units

Interactive collaborative scenarios (Helic, 2007) are employed to administer and respond directly to individual student learning activities. For example, active documents can then be employed to keep track of learner interactions and learning support within the context where learning occurs.

An explicit and implicit profiling of students has to be applied to keep track of the learning process of students. E-Portfolios (Alexander, 2006) enable the recording of student participation and contribution to support the profiling of students. E-Portfolios are important in allowing students to start valuing their own contributions and also other student contributions. An example of a student portfolio structure is shown in Figure 3. As shown here, the ecosystem provides a workspace for students to continuously expand their knowledge base while taking responsibility for their own learning. Records of student achievement then immediately become available to mentors and instructors for personalized evaluation and guidance.
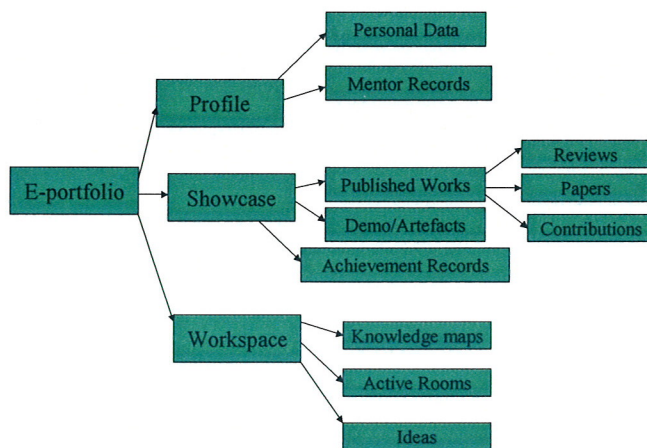

Figure 3: Student E-Portfolio

Incentive schemes can be tied to E-portfolios in order to acknowledge and highlight student achievement. Recommendation systems proposed to make explicit the valuations associated with each student's contribution. Recommendation systems play an important role in the development of rational impartial judgment among students. A combination of human and automated ranking of important topics, ideas, suggestions and contributions can further be applied to personalized interaction among students with a similar background and interests.

A number of tools are available for creating an environment for students to collaborate among themselves and with their instructors and mentors. These include peer-evaluation support, collaborative concept mapping, brainstorming and discussion forums. Brainstorming also incorporates mechanisms for specifying ranks and incorporating personal evaluation (see Figure 4). Annotations are again a key feature to represent and organize collective student learning. Annotations also have been proposed to represent links to Knowledge Cards to reflect the knowledge construction process of students.
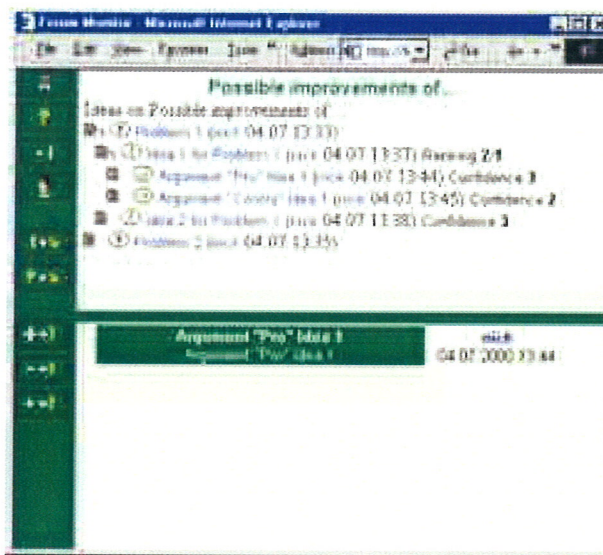

Figure 4: Brainstorming Support (extracted from WBT-Master Manual)

Integrated visual tools will be applied in the management and display of information in illustrating student learning and mastery of concepts. The tools allow the instructor to impart particular skills, to refine processes used for a specific task, or to organize information into a structured form. They also can be used by students to express their understanding of concepts. Knowledge visualization tools also will be applied as a form of assessment of students' incremental knowledge gain over a period of time. Learners also need to be supported by means of personalized knowledge retrieval facilities. Such a tool will be effective in identifying potential infringements by students and can be used to aid students in the mastery of useful skills. The visualization capability for concept maps further allows the incremental visualization of concepts formulated by students.

Knowledge Cards (K-Cards) enable the specification of concepts of is-a and instance-of links for ICARE knowledge maps The semantic relationships built upon K-Cards essentially define a semantic graph. The Knowledge Card mechanism is also used to automatically link to peer-learners and resource persons in collaborative mode. Two types of K-Card usage have been defined: personal Knowledge card attached to each learner, and context-based Knowledge Cards attached to assignments or scenarios. The use of K-Cards supports the creation of points of interests by students. A knowledge card also can be linked to other related readings that students may associate (if required). These K-Cards will then allow students to link to a concept map, which will demonstrate the students' understanding process.

Incorporating the Ability to Handle Copy-Paste into ICARE

ICARE benefits from the administration of academic reading procedures that can be integrated directly into the ICARE ecosystem. By enabling a business process model view of E-Learning (Helic et al., 2003), the learning process can be supported at each step.

E-Learning modules on effective Copy-Paste would then be embedded to educate students on the rightful procedure of academic publishing (reading and writing). Apart from employing a plagiarism or Copy-Paste detection suite for summative assessment of a breach of conduct, we propose the formative application of such tools for self-plagiarism checking and in cultivating constructive "Copy-Paste skills." For example, existing document similarity detection (as used in plagiarism detection tools) can be applied in conjunction with a learning scenario paradigm for facilitating students to master academic publishing. By consolidating the results from similarity search engines on local databases as well as the Internet, a plagiarism detection tool can be applied to assist students to teach them how and when to cite another publication.

Copy Paste Handling Software Suite

The Copy-Paste Handling Software Suite incorporates self-plagiarism checking tools and techniques to help students in mastering Copy-Paste. Both simple and advanced forms of Copy-Paste checking are supported. We propose the use of the plagiarism detection tools and techniques to achieve this task (see Table 1). This suite will be applied in two modes: closed world and open world modes. These modes will

allow the operation of the Copy-Paste handling in both a supervised mode (assisted by an instructor) and an unsupervised mode (self learning).

In the closed world mode, a student uses the Copy-Paste wizard as guide for the academic reading and writing process. This wizard is described in the next section. Here the text that students select for Copy-Paste will be used as a fingerprint and applied as query string to search the whole published text of the student for a weak or blatant Copy-Paste case. The similarity checking engine identifies the degree of similarity in determining the extent of paraphrasing (or the lack of it). The system also is able to check for compliance or negligence citation. A string similarity checking mechanism is applied for this purpose. In the case of identifying an improper Copy-Paste, the system presents its findings as an advice to students. The changes made by students are noted by the system and can be used in a mentoring session.

In the open world mode, students are not guided or restricted in terms of usage of specified references. Similarity detection is then applied to a larger collection of documents where it checks the Web for all possible improper Copy-Paste actions performed by the students. Student's past years papers also are checked for similar text strings to determine improper Copy-Paste and lack of citation. The system produces statistical information for the instructor to assess the mastery level of students.

A number of learning scenarios can be built by a selective application of one or more Copy-Paste handling tools. As described here, these scenarios could either be applied in a supervised manner assisted by an instructor or a mentor or the unsupervised manner with system inputs.

During the mentoring process, a manual selection approach for plagiarism detection may be employed checking with one or more search engines. This process can provide the system a set of constrained documents to be used for similarity checking. Specific tools to approve or disprove suspected plagiarism such as Cloze may also be applied when a dispute arises. A Cloze procedure (Maurer & Zaka, 2006) has been used to judge the originality of authorship of published works. As part of the Copy-Paste detection, alternative techniques such as stylometry can be applied to discover similar (or dramatically changing) stylistic patterns such as syntactic forms

usage, text structure of published works and the usage of key terms to indicate that some copying may have taken place.

Copy-Paste (Academic Reading and Writing) Wizard

This wizard has been proposed to enable learners to acquire the skills of academic reading and writing in a controlled environment. The wizard can be used by learners to perform the following:

- Highlight key points and annotate selected phrases, using the annotation feature of WBT-Master. A highlighting mechanism is supported to allow learners to highlight key points.
- Create a Knowledge Card for the point discovered, label it and link it to known concepts (or form a new concept).
- Review the internal concept map and assimilate new ideas found in reading. This may range from concept links to concept map restructuring. This stage involves substantiating the body of Knowledge Cards with links and metadata.
- Formulate an idea and add information to Knowledge Cards.
- Express an idea and present it as a descriptive text.

Annotations will be employed in linking original document to relevant information sources to perform the above steps. This enables the tracing of students' activities to check on process-flow of academic writing. Separate readings can be assigned to each student to track individual student activities and also to avoid plagiarism. At the same time, a single document may also be used for an entire class or a smaller group of students. In this way a comparative analysis of students' learning can be visualized and studied. These documents can then be constructed as active documents that allow collaborative learning to take place, built upon students' comprehension and ability. As with our previous experiments on active documents, we know that when 500-1,000 users have viewed a particular document, all possible questions that need experts become answered (Dreher & Maurer, 2000).

The technological support to prevent blatant copying by students is realized by imposing the use of annotations (through specially designed interface templates), which overcomes the need to duplicate content. Figure 5 illustrates the interface that allows students to express ideas, opinions, contribution to collaborative sessions, ask questions, etc. Additionally, the Copy-Paste interface further displays the highlighted text, representing key points with a ranking of

importance, paraphrased text, comments, etc. Students' published works will then be stored as annotations to the original text and visualized separately by the instructor for evaluation.

The use of annotations can be explored as a means of training students' use of the correct form of citations and referencing. By using annotations, a much simpler similarity checking system would suffice to overcome plagiarism to a large extent in ICARE. Annotations and its sophisticated communicational and collaborative features play an important role in the realization of a culture of Web-based reading and writing.
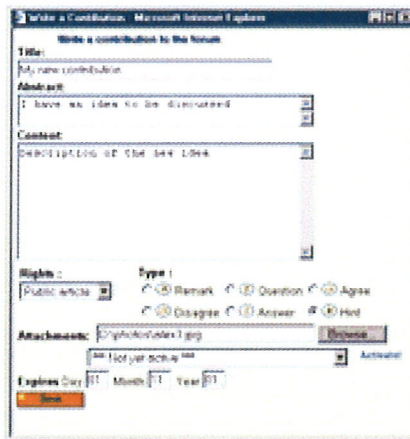

Figure 5: Interface for Learners to Annotate Documents

Design of Assessment

ICARE also includes mechanisms for the careful design and execution of assessments. The pedagogy driven learning together with the ability to define learning scenarios and rooms allow for highly personalized assessment design and curriculum development.

Beyond the features of the ICARE system as described, the ability to operate in the following modes is instrumental:

- Guided Mode: Interactive session (system auto-suggestions) with closed systems monitoring.
- Self-Learning Mode: Minimal non-interactive feedback, closed world systems monitoring but with feedback provided only on student request.

- Diagnostic Mode (formative): Closed world systems monitoring but with no feedback, results are archived for self-driven assessment.
- Evaluative Mode (summative): Open world mode, with text analysis performed (Copy-Paste analysis) and used as support for self-paced assessment.
- Mentor-Assisted Mode: Similar to diagnostic mode but with feedback sent to a mentor, who responds to students.
- Peer-Learning Mode: Open world learning mode, with the system tracking learner participation and contributions.

These modes of operation can be realized as scenarios (training objects) in WBT-Master. This system also allows assessments to be broken up into smaller parts as a means of supporting continuous assessment, and in the monitoring of student learning process.

As an example of the application of ICARE in a classroom, we propose the following illustration:

1. Students in a class are first asked to collaboratively construct a collective concept map for a domain of study.
2. Individual students are then required to construct a personalized concept map representing their personal learning space.
3. Subsequently, students are assigned selected reading material. An online copy of the reading material is placed in the reading room of each student (or a group of students).
4. Students are then required to identify key points by using the wizard in closed monitoring mode with all activities tracked by system. The highlighted text segments by students can be used to reflect their understanding. Both individual student learning and group learning can be highlighted.
5. The highlighted texts are then visualized for the instructor as annotations attached to the selected document. Statistical information is used to demonstrate student learning, e.g. common mistakes made by student, misunderstanding of text, etc.
6. Instructors' comments can either be placed in personal spaces of students or public spaces for the whole class.
7. Students are then requested to paraphrase the texts selected in guided mode.
8. A visualization of all student inputs is then made available for the instructor. Additional statistical information is presented to support student evaluation. Non-compliance in student learning workflows is visualized.

9. The next step involves a peer-learning mode, where student are requested to discuss the points selected by their peers in the brainstorming room. All points being discussed are referenced and the system links them together for visualization. The instructor or facilitator then provides interactive feedback in the brainstorming room.
10. Students are then required to update their personal concept maps, with the knowledge gained in step 9.
11. Statistics of popular concepts in knowledge-map, popularly selected key points, list of questions posed during brainstorming or during any other phase in the exercise are all presented to the classroom.
12. As the final task, students are asked to collaboratively construct a single concept map while continuing with discussions in the brainstorming rooms. All concepts in the knowledge map are uniquely identifiable as they are implemented using Knowledge Cards. Thus, students are able to discuss the addition of particular concepts or places for links and types of links as well.

The above hypothetical assessment has been defined to illustrate the various functions for the explorative employment in a classroom. A typical classroom usage may only require a subset of the tasks listed. This clearly highlights the power and potential of the proposed ecosystem, to serve as basis for the design of future E-Learning systems.

Conclusion

We have adopted the stand that Copy-Paste need not be entirely considered a wrong-doing. Students would then need to be educated and guided on the constructive use of Copy-Paste skills as a learning mechanism. We have presented an academic ecosystem with technological support to comprehensively address the Copy-Paste Syndrome.

We proposed the use of an advanced E-Learning system, together with carefully planned student assessments and the close monitoring of student learning to address the problem. Plagiarism and Copy-Paste Syndrome avoidance mechanisms and procedures are integrated into the ecosystem and applied throughout the program of study. E-Learning modules together with a suite of Copy-Paste handling tools enable the formative development of "effective Copy-Paste skills." A

complete suite of Copy-Paste detection and avoidance tools will need to be established in all educational institutions.

By effectively addressing the Copy-Paste Syndrome many of the social problems that we are likely to face (arising from the degradation of scientific quality and even possibly leading to quality of life) in future can be averted. Without the full institutional backing and commitment of academics, however, a culture that withstands and compensates the prevalent Copy-Paste culture cannot be achieved.

References

Alexander, B. (2006). Web 2.0: A new wave of innovation for teaching and learning? *EDUCAUSE Review, 41*(2), 32–44.

Bersin, J. (2004). *The Blended Learning Handbook: Best Practices, Proven Methodologies, and Lessons Learned (excerpt)*. San Francisco: Pfeiffer Wiley. **Retrieved DATE, from** http://media.wiley.com/product_data/excerpt/67/07879729/078 7972967.pdf

Downes, S. (2006). E-learning 2.0. *ACM eLearn Magazine*. http://www.elearnmag.org/subpage.cfm?section=articles&article =29-1

Dreher, H. V., Dreher, L. H. & McKaw, K. (1994). The active writing project: Small movements in the real world. *Proceedings of Asia Pacific Information Technology in Training and Education*, Brisbane, Australia.

Dreher, H. & Maurer, H. (2000). Active documents: Concept, implementation and applications. *Journal of Universal Computer Science, 6*(12), 1197-1202.

Dreher, H., Krottmaier, H. & Maurer, H. (2004). What we expect from digital libraries JUCS. *Journal of Universal Computer Science, 10(*9), 1110-1122.

Dreher, H. & Williams, R. (2006). Assisted query formulation using normalized word vector and dynamic ontological filtering. *Proceedings of 7th International Conference of Flexible Query Answering Systems*, 282–294, Milan, Italy.

Duff, A. H., Rogers, D. P. & Harris, M. B. (2006). International engineering students: avoiding plagiarism through understanding the western academic context of scholarship. *European Journal of Engineering Education, 31(6)*, 673.

Eissen, S. & Stein, B. (2006). Intrinsic plagiarism detection. Proceedings of the 28th European Conference on Information Retrieval. *Lecture Notes in Computer Science, 39(36)*, 565-569. Springer Publishing Company.

Graham, C. R. (2004). Blended learning systems: Definition, current trends, and future directions, in Bonk, C. J., Graham, C. R. (Eds.) *Handbook of Blended Learning: Global Perspectives, Local Designs.* San Francisco: Pfeiffer Wiley.
http://www.uab.edu/it/instructional/technology/docs/blended_learning systems.pdf

Harris, R. (2004, Nov. 17). Anti-plagiarism strategies for research papers. *Virtual Salt.* http://www.virtualsalt.com/antiplag.htm

Helic, D. (2007). Formal representation of e-learning scenarios: A methodology to reconfigure e-learning systems. *Journal of Universal Computer Science, 13(4)*, 504-530.

Helic, D., Krottmaier, H., Maurer, H. & Scerbakov, N. (2003). Implementing project-based learning in WBT systems. *Proceedings of E-Learn 2003*, 2189-2196, AACE, Charlottesville, USA.

Helic, D., Maurer, H. & Scerbakov, N. (2004a). Discussion forums as learning resources in Web-based education. *Advanced Technology for Learning, 1(1)*, 8-15.

Helic, D., Maurer, H. & Scerbakov, N. (2004b). Knowledge transfer processes in a modern WBT system. *Journal of Network and Computer Applications, 27(3)*, 163-190.

Kennedy, I. (2004). An assessment strategy to help forestall plagiarism problems. *Studies in Learning, Evaluation, Innovation and Development, 1(2)*, 1-8.
http://sleid.cqu.edu.au/viewissue.php?id=5#Refereed_Articles

Korica, P., Maurer, H. & Scerbakov, N. (2005). Extending annotations to make them truly valuable. *Proceedings of E-Learn 2005*, pp. 2149-2154, AACE, Vancouver, Canada.

Krottmaier, H. & Helic, D. (2002). More than passive reading: Interactive features in digital libraries. *Proceedings of E-Learn 2002*, 1734-1737, AACE, Charlottesville, USA.

Kulathuramaiyer, N. & Maurer, H. (2007). Why is fighting plagiarism and IPR violation suddenly of paramount importance? *Proceedings of International Conference on Knowledge Management*, in Stary, C., Baranchini, F., Hawamdeh, S. (Eds.) *Knowledge Management: Innovation, Technology and Cultures*, 363-372, World Scientific.

Liu, C., Chen, C., Han, J. & Yu, P. S. (2006). GPLAG: Detection of software plagiarism by program dependence graph analysis. *Proceedings of 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 872-881, Philadelphia, USA. http://www.ews.uiuc.edu/~chaoliu/papers/kdd06liu.pdf

Maurer, H., Kappe, F. & Zaka, B. (2006). Plagiarism: A survey. *Journal of Universal Computer Science*, 12(8), 1050-1084.

Maurer, H. & Zaka, B. (2007). Plagiarism: A problem and how to fight it. *Proceedings of Ed-Media 2007,* 4451-4458. AACE, Vancouver, Canada. http://www.iicm.tugraz.at/iicm_papers/plagiarism_ED-MEDIA.doc

Midolo, J. & Scott, S. (14 Aug. 2003). Teach them to copy and paste: Approaching plagiarism in the digital age. *Resourcing the Curriculum*.
http://www.det.wa.edu.au/education/cmis/eval/curriculum/copy right/islandjourneys/documents/paper.pdf

Mödritscher, F., García-Barrios, V. M. & Maurer, H. (2005). The use of a dynamic background library within the scope of adaptive e-learning. *Proceedings of eLearn 2005*, AACE, Vancouver, Canada.

Ong, S. C., Kulathuramaiyer, N. & Yeo, A. W. (2006). Automatic discovery of concepts from text. *Proceedings of the IEEE/ACM/WIC Conference on Web Intelligence,* 1046-1049.

Pannepacker, S. (2007). Is plagiarism on the rise? Why? http://muWeb.millersville.edu/~jccomp/acadintegrity/plagnotfin alnewsp.html

Paper Mills. (2006). Cheating 101: Internet paper mills. *Kimbel Library: Presentations*, Coastal Carolina University. http://www.coastal.edu/library/presentations/mills2.html

Plagiarism Statistics. (2007). http://www.plagiarism.org/plagiarism_stats.html

Sathya Sai Baba. (2001). *Proceedings of First Conference of Sri Sathya Sai Schools*, Prasanthi Nilayam, India. http://www.srisathyasai.org.za/programs/educare.asp

Standing, L. & Gorassini, D. (1986). *An Evaluation of the Cloze Procedure as a Test for Plagiarism Teaching of Psychology, 13(3)*, 130-132.

Valiathan, P. (2002). Blended Learning Models. http://www.learningcircuits.org/2002/aug2002/valiathan.html

WBT Master White Paper. (2006). http://www.coronet.iicm.edu

Weber, S. (2006). Das Google-copy-paste-syndrom. *Wie Netzplagiate Ausbildung und Wissen gefährden*, Heise, Hannover.

About the Authors

Narayanan Kulathuramaiyer
PhD
Associate Professor
Faculty of Computer Science and Information Technology
University of Malaysia
Sarawak City, Malaysia
nara@fit.unimas.my
Major research interests: Web intelligence, knowledge management, E-Learning

Hermann Maurer
PhD
Professor
Institute for Information Systems and Computer Media
Graz University of Technology
Graz, Austria
hmaurer@iicm.edu
Major research interests: Web intelligence, knowledge management, E-Learning