Distributed Transcripts for Online Learning: Design Issues

Steven L. Tanimoto

Abstract:

A set of guiding principles is stated for the development of standards for representing student educational assessment information. These principles support the learner, rather than the academic institution, as the focus of an information system architecture. Unlike traditional academic transcripts, the items in portfolios, subjective written evaluations, self-assessments, and computer-based records of activity are complex representations of student achievement, involvement, or inclinations. Even more so than traditional grades, they depend upon a great deal of contextual information in order to be intepreted in useful and reliable ways. This paper also identifies the essential informational components of alternative assessment records and suggests a standard form for their representation. The consideration of evidence, judgment, context, and justification, as described in this paper, is relevant to the improvement of conventional (e.g., multiple-choice test) assessment methodologies, as well. Considerations of systemic educational reform issues such as equity, group learning, lifelong learning, locus of responsibility for learning, and privacy are briefly described insofar as they impact the design of assessment systems.

Commentaries:

All JIME articles are published with links to a commentaries area, which includes part of the article's original review debate. Readers are invited to make use of this resource, and to add their own commentaries. The authors, reviewers, and anyone else who has 'subscribed' to this article via the website will receive e-mail copies of your postings.

Steven L. Tanimoto, *Box 352350, Dept. of Computer Science and Engineering University of Washington, Seattle, WA 98195 USA tanimoto@cs.washington.edu*

1.   Introduction

How the records of learning are stored in archives, electronic learning environments, and academic information systems will have a profound effect on not only the technical aspects of educational systems, but on their social and political aspects as well. It is very important that emerging standards support the best interests of learners and not only those of academic institutions and commercial service and content providers.

1.1   Background

The efforts of several groups to enable interoperability among electronic learning environments include those of the IEEE Learning Technology Standards Committee (see for example, (Ritter and Suthers, 1997) and (LTSC, 2000)) and Educom's Instructional Management System (EDUCOM, 1998). These efforts are motivated in part by the opportunity for online learning presented by the Internet and the World Wide Web, and by various shortcomings of existing software. Indeed, one of the attractions of web-based learning is the possibility of synthesizing coherent educational experiences using distributed and heterogeneous materials (Murray, 1998). With the advent of electronic learning environments, including intelligent tutoring via computers and computer based construction environments, part of a student's getting started with a new piece of software involves the program learning about the student's current state of knowledge. This information allows the program to present material or suggest activities that are appropriate for the student. The time spent acquiring this information could be minimized if it were all available in one or more data files constructed during the student's earlier interactions with other educational programs. In addition, the existence of a thorough, machine-readable record of the student's education creates the possibility of having many separate (and possibly diverse) software programs that contribute to the student's education in a coordinated way.

The different programs may handle different aspects of the educational process. Some programs may be oriented towards the presentation of new material. Others may focus on assessment of the student's comprehension of certain material, while yet other programs may perform customized curriculum planning for the student. The complete record of a student's learning (let's call it the student's dossier) could be readily processed to produce resumes, career planning analyses, or progress reports for specific subjects or time periods. The CAI programs that utilize and add to the student dossier may utilize not only present and future computer technology, including future multimedia and virtual reality technologies, but also future pedagogical methodologies.

Not only does the student dossier represent the educational experiences of the student obtained

during sessions with CAI programs, but it may also contain assessments of non-computer-based learning experiences, obtained via experts (e.g., teachers) and interview programs. (See (Tanimoto, 1992) for additional justification for such dossiers.)

Current work on learner models captures many of the key ideas regarding the contents and purposes for dossiers (e.g., see Murphy and McTear, 1997)). However, little has yet been said about how non-test-based assessment items should be handled. (For a justification of non-test-based assessment see (Hoffman, 1962), and for a description of conventional and alternative assessment methods see (Linn, 1989) and (Broadfoot, 1986), respectively, and for a program-assessment level rationale for a broader panoply of assessment methods see (Haertel and Means, 2000).) The purposes of this paper are (1) to argue in favor of the learner as the focus of an information architecture for educational assessment, (b) to argue that any such architecture embrace assessment in its broadest form, and (3) to give the broad outline of a technical approach for handling of alternative assessment information within student learning databases.

Part of the philosophy behind this paper is that an evaluation of student learning should always be qualified in terms of the particular evidence for that evaluation, whether the evaluation is done by a human teacher, a computer testing program, or through a self- or peer-evaluation process. This approach to representing the student's education allows important decisions to be made where any possibilities for error can be appropriately taken into account, and additional evidence gathered when appropriate.

The primary purpose for the assessment information of concern in this paper is to improve the efficiency of the student learning process. This includes helping solve the "transfer problem" which occurs whenever a student moves from one educational environment to another; most of the information built up about the student in the first environment is either lost or simply not transferred to the second environment, so that time and energy must be wasted in rediscovering the precise needs of the student.

2   Scenarios

This subsection provides four fictional scenarios, in order to illustrate the manner in which assessment data for the dossier may be collected and used. The scenarios all follow a similar pattern: assessment information is collected in a learning/assessment activity and recorded in the student's dossier. Later, the student transitions to another learning environment or activity, and the information collected in the dossier is recalled in order to make the transfer as efficient as possible.

The scenarios vary in several respects. The first two involve traditional, offline learning or

assessment activities. The first requires an explicit capture phase in which documents are scanned and analyzed. The second involves video recording in a laboratory course. The third and fourth scenarios deal with online learning activities, one within a highly structured learning environment involving simulation software, and the other within a textual online discussion forum. In each case, once again, the pattern is a cycle that begins with capture and then follows that with recall to facilitate transfer of the student to another learning environment.

As they are intended to be futuristic, these scenarios suggest some questions, such as whether and how videotapes might be automatically analyzed, that are beyond the scope of this paper. The focus of the paper is on designing representational infrastructure and not on algorithms or specific assessment methods.

2.1   Scenario A. Traditional assessment and college transfer credit evaluation.

Summary: A college student enrolled in Combinations and Probability takes an offline test. Scores are entered by scanning the test and test answer sheet and using OCR to determine the ASCII text of the test and the answer choices selected. The correct answers are also scanned and recognized or they are entered manually. The context under which the test was taken is characterized by filling out an online form. The context includes information about time, place, testing environment, rules of conduct (were bathroom breaks permitted and if so, did students have access to reading materials of any kind during such breaks?). Two representations of the test information are added to the dossier: the scanned image of the test documents and the machine-recognized (or hand-entered) textual representation of the documents (e.g., the ASCII file of the test questions, suggested answers, and answers chosen by the student). This information would typically require further interpretation either by a human or by specialized software, before it could be used to update the mastery records for the student in the subject area of the test.

Later, the student transfers to another university and asks an advisor to approve transfer credit for a course named Probability and Statistics that is somewhat similar to the one taken earlier. The advisor, even with access to the student's dossier materials, cannot make a determination but forwards to a mathematics instructor an email message with a reference and permission code for the dossier. The instructor examines the record, including the scanned test and answers, and makes a placement recommendation, filling out a web form indicating a possible weakness on the part of the student in the area of statistical experiment design.

2.2   Scenario B. Live chemistry laboratory session, with video recording.

Summary: The student participates, with a partner in a chemistry laboratory activity such as titration of an acid. The student and partner are one pair of 12 in the laboratory at the same time. All participants have their own bench area, glassware, and chemicals. No electronic information technology is apparent in the classroom, except that there is a digital video camera high up in one corner of the room capturing most of the gross action in the room. The video is being compressed and stored on a hard drive. After the laboratory activity is done and the students have gone away, the instructor runs a computer program that scans the video and outputs a "gross activity report" for each student. Each of these reports shows a timeline with marks on it where possibly significant events were recognized - for example, the student moving away from the lab bench, bending over the pH meter, writing in a lab notebook, lifting a medicine dropper, etc. The instructor has collected the reports submitted by the students as they left the laboratory, and now correlates the students' reports with their gross activity reports. Then the instructor fills out an online form for each student, rating his or her performance in the activity on several scales: disengaged to fully engaged, unsuccessful to successful, etc. The online software being used by the instructor then posts updates to each student's dossier, including a copy of the gross activity report, an instructor evaluation, and a link to the compressed video file. The actual video file is not posted to the dossiers, by virtue of a policy decision by the school to keep those private to the instructor.

A year later, an advisor decides that the student has promise as a future physician and that the student should be groomed or mentored to get into medical school. However, consulting the dossier, the advisor finds that the student tended to be the follower, rather than the leader, in chemistry lab. The student is advised to take organic chemistry, and this course's instructor is sent a request by the advisor to assign a lab partner that will permit the student to easily take a leadership role in labs.

2.3   Scenario C. Individual online learning with a physics simulation applet.

Summary: The student constructs a configuration of springs and masses in order to study compound motion. The student succeeds in achieving the suggested goal of specifying the conditions under which nearly circular motion is achieved. The applet has gathered user interaction data throughout the session and uploaded it to the hosting server where it is processed by a program that makes inferences about the student's fluency with various features of the applet as well as the student's rate of progress and ultimate levels of achievement in the various components of the activity. This information is posted to (a) the database belonging to

the institution running the web site, and (b) the student's personal dossier which resides on a server operated by a third party.

Later, the student is studying differential equations in a mathematics course. A web site devoted to the generation of custom problems in mathematics based on a student's interests and prior knowledge scans the student's dossier, finds the records related to the physics activity, and produces a special problem in differential equations based upon compound motion of the sort constructed in the physics applet.

2.4   Scenario D. Online learning in a small-group discussion forum.

Summary: Four students enrolled in European History at a Distance debate the ethics and motivations of the French revolutionaries. Three of the students end up agreeing that the use of the guillotine was warranted under the circumstances, while the fourth doesn't fully agree. In the course of the debate, two of the students do most of the substantive posting, while the other two occasionally express agreement or disagreement with ideas written by the others. When they enrolled in the course, these students "signed" an agreement that permitted the institution running the course to collect and analyze all the postings of the students in this forum and to make copies of this data available to each of the participants for use in their own dossiers with the provision that the names of the other participants would be "laundered" and replaced by pseudonyms such as Joe Student, Mary Freshman, and Jean Historymajor. At the conclusion of the course, each student is given the advice that he or she have an analysis done by one of two or three companies that specialize in educational advising on the basis of data such as that gathered in this setting. These companies analyze both the group dynamics, the writing styles, and the content, and they make recommendations for ways to improve or to reach particular goals. One of the students follows this suggestion and then receives the recommendation to take a free web-based tutorial on effective writing in online forums.

Later the student enrolls in another online course using small-group discussions. The instructor of this course makes use of a group formation tool that scans the dossiers of the entering students and assigns them to groups in such as way as to maximize the expected learning efficiency for the class as a whole.

3   Discussion

The scenarios above illustrate major elements of how dossier information could be collected and used. They are not only fictional, but they are simplifications, because there are a number of issues that must be dealt with to make these transactions possible that were not mentioned in

the scenarios. To include them explicitly in the scenarios would lengthen them too much, and would tend to inject too many details into the presentation, many of which might seem contrived and artificial. However, it's worth mentioning some of these issues here in order to broaden the perspective from which these scenarios are considered.

First, each scenario assumes that those accessing the dossier have permission to do so, and that the costs in terms of student or institutional overhead to request and grant permissions are not overbearing. In a real system, it will be important to design agreement mechanisms and standards that make it easy for institutions to request appropriate permissions and for the student to grant such permissions without being forced to ask for or grant too much or too little.

We've also ignored the question of validation in these scenarios. When the college advisor forwards the Combinations and Probability test data to the mathematics instructor, the data is presumed to be valid. If the student provided this data, it might need to be validated by a server at the first school. The process and costs involved in this should be made simple and minimal. The information and permissions provided by the student can serve as an unofficial preview of assessment data plus the key to access validated data.

None of these scenarios involve either portfolios, or student self evaluations. One could easily come up with fictional scenarios to illustrate how portfolios could be created and presented. Students could identify certain parts of their dossiers as portfolio components, annotate them, and make them ready for presentation to various other parties. Self evaluations could be developed using special software tools that guide the student through a process of inspecting earlier work and evaluations and coming to new conclusions or observations. Self evaluations could sometimes be included in portfolios when they demonstrate the kinds of reflective insights that help others appreciate the record.

We have not discussed any details of ontologies that would be needed to support automatic translation of assessment data from one system to another. As one designs an ontology, a key issue is granularity. How large or small should an atomic assessment item be? How small a concept within a subject area is to be referred to by one assessment data item? If student misconceptions or partial conceptions are to be represented, roughly how many will be permitted or expected per topical concept? Various researchers have addressed such issues (e.g., (Minstrell, 1992)), and they continue to be a focus of research.

When students own their dossiers, there's a possibility that they could alter them unethically, not only in presenting an overly biased view of events in their educational histories but even completely falsifying events. There are some possible ways to limit such unethical activity. One would be to encourage the designers of the software tools for editing dossiers to build in

mechanisms to limit tampering. Another approach would be to make validation a regular component of the maintenance of dossiers, much as automobiles must pass annual inspections in many states. Queries on the dossier might be answered with a combination of assessment information and proof of recent validation.

Finally, there is a broad potential benefit of using comprehensive student-owned dossiers of the sort described in this paper. Distributed (de-centralized) matchmaking software, perhaps similar in some respects to the well-known music-sharing Napster software, could permit students to share learning resources in an opportunistic, low-cost, just-in-time manner that permits very specialized needs to be quickly satisfied by drawing upon a broad array of resources of a vast community of learners. In some cases, the students themselves would be the resources and the matchmaking software would help build peer learning relationships. In other cases, the software would match a student up with a document, a web site, an applet, or an online course. The important aspects of such a community of learners are the degree of specificity at which needs and resources could be expressed, the timeliness of the interactions thus facilitated, and the low per-match cost at which such interactions could be established.

The scenarios given above, along with the concerns just mentioned, give rise to some desirable features in systems to support learning in the coming decades. They are offered as principles, partly because they seem to form a foundation, and also because each one entails consequences that pose interesting design challenges.

4   Principles

The following list of six principles is offered to guide the development of standards for electronic records of assessment-related information in student records.

1. Multiple forms of assessment (e.g., multiple-choice tests, evaluations of students' prose, log-file analyses) are essential in developing an accurate model of a students' knowledge, skills, learning styles, experience, achievement, motivation, and goals.

2. Any single assessment item (e.g., test result, project evaluation, etc.) has a degree of unreliability and a degree of incompleteness that bears on its usefulness in analysis.

3. The effects of uncertainty and incompleteness in an analysis of the student can be reduced by recording and considering information about the context in which the assessment was performed.

4. When possible, assessment information should be evaluated in a manner similar to that used when evaluating evidence, using methodologies appropriate to handling of evidence, including using means to determine the reliability of evidence, and using means to make logically or mathematically valid inferences from the evidence.

5. Assessment information, like many other kinds of information, may be owned by people or institutions. However, the parties to an assessment event have various rights to the information about the event. The learner tends to be at the center of this activity and must retain essential rights of privacy. A system for representing assessment information must contain provisions both for the protection of ownership and for the exercise of these rights.

6. The cost of storage media, per unit of information, continues to fall. To the extent practicable, summaries or evaluations of student performance should be accompanied by detailed digital records of the student's actual activity. This may permit explanation, verification, and/or reworking of the assessments, if the objectives or methodologies for the assessments should change in the future.

5   Taxonomy of Assessment Items

The information pertaining to alternative assessment can be classified into the following three general categories: (1) educational history items, (2) portfolio items, and (3) evaluation items. The history items represent or reflect some actual activity that the student was involved with. The portfolio items are representations of constructions and accomplishments by the student that can be presented to other people as samples of the student's work or as parts of an explanation of the student's educational background. Although intended for viewing by people, portfolio items also contain machine-interpretable descriptions. Evaluation items include the results of inferences and judgments about the student's understanding, motivation, skill levels, etc. These items generally identify evidence, method of inference, and conclusions, including indications of doubt and degree of reliability.

5.1   Educational History Items

Actual student activity can be recorded in many ways. The following is a list of some different sorts of educational history items.

1. Registration record (statement that student registered for a particular activity at a particular time and date)

2. Activity Log (sequence of events)

3. Activity description (the learning activity and software)

4. Event description (meanings of events)

5. Communication log (email, newsgroup postings, audio calls, etc.)

6. Test, quiz, or exam and student answers

7. Video recording of a session, class, laboratory, interview, etc.

These records of learning activity can usually be distinguished from items in the next category, portfolio items, which are typically the intentional resultant products of student constructive activity, rather than by-products or traces of the means to achieve them.

5.2   Portfolio Items

The items in this category are explicit end-products of student expression, rather than representations of learning or creative processes.

1. Copy of project report

2. Pointer, URL or reference to actual project

3. Journal or notebook or pointer to such

These subcategories of portfolio items are intended to be general, and should admit textual documents (student writing), audio recordings of student explanations or of student musical compositions or performances, graphical material, video recordings (for example of student dramatic productions), and pieces of software or electronic constructions, intentionally created, of any sort.

5.3   Evaluation Items

In contrast to items of the first two types, the following are in a sense meta-expressions. They are expressions about student expressions. In any of the subcategories, there may be evaluations by instructors, by electronic tutors, by peers, self-evaluations by the student, or evaluations by any other agency such as a team or mixed human-and-computer group.

1.   Evaluation of project or activity

2.   Evaluation of understanding or skill

3.   Evaluation of motivation or learning styles

4.   Analysis of student performance on a test

5.   Evaluation of educational progress

6.   Result of an advising session

This taxonomy is organized according to degrees of summarization in the following sense. The details of participation in an activity are in a sense condensed and summarized by the final student product of the activity. For example, a student's participation in a chemistry experiment in an online simulated laboratory is represented by a log of the student's actions in each phase of the experiment. However, this activity is summarized in the student's lab report, which is a document that might form part of the portfolio. Finally, this document, after analysis by a teacher or agent, is represented in an even more condensed form in an evaluation. In this case, and ideally in all cases, for each activity in the student's educational history, there is one or more corresponding portfolio item, and in turn, corresponding to that or those, evaluation items. Thus there are relationships that connect assessment records across the top level classes of the taxonomy.

Another distinction that one can see in this taxonomy is that the educational history items and portfolio items constitute "raw material" for the evaluation items. The portfolio items can be thought of as student products, while the educational history items are typically by-products of the student activity. The evaluations will commonly be products of teachers and programs, but in the case of self-evaluation, one would have an evaluation item that also happens to be a product of the student.

Of course, various categories of activity could be identified and the taxonomy thus refined.

However, the intention of this paper is to provide only the general outline of this structure.

6   Components of Assessment Records and Files

Now that a general taxonomy has been established, let us proceed with some more concrete forms of representation.

A representation method is described having two levels: the file and the record. An assessment file is intended to represent the results of any specific assessment process such as computer analysis of an essay by the student or one or more computer-assisted-instruction sessions with the student. An assessment file contains a context header, a collection of assessment records, and a list of "post requests" to a summary file. The context header provides information that applies or that may apply to all of the records in the file. Each record describes either a history item, a portfolio item, or an evaluation item. The assessment file may contain references to other files or information resources. These may be local (within the same educational profile) or they may be external (for example in the database of an academic institution at a remote web site).

The list of post requests is a representation of summary data that derives from the other information in this assessment file and which should be used to update an overall summary file for the student. While it may seem redundant to include this information here rather than to simply transmit it to the summary record, there are two advantages to storing it here: (a) the influence of the activities represented in this file upon the overall summary record is clear and can be interrogated, and (b) it is possible to conveniently create this file while disconnected from the central summary file and yet have the influence communicated at such time as the connection may be made.

6.1   Context Header

An assessment file begins with a context header. This contains information that applies or potentially applies to all of the records in the file.

1. Date and time created

2. Id of root file (reference to core file of student's dossier that contains student id., info access policies, etc.)

3. Name and version of application program (e.g., intelligent tutor program) writing this record

4.   Name and version code of assessment standard to which this file conforms.

5.   Name(s) and version(s) of domain model(s) referenced, if any.

6.   Date and time of last update to this file.

7.   Number of assessment items in this file.

8.   Special conditions known to be in force during creation/updating of this file. E.g., student not working alone, student using various tools, noisy or nonideal environment, etc. Information here applies to all records in file unless explicitly overridden in individual records.

6.2   Assessment Records

Within an assessment file are zero or more assessment records. Each record describes a history item (participation by the student in an activity), a portfolio item (a product of a learning activity), or an evaluation (set of judgments about the student's learning).

The number of these records is given in the context header above.

Each assessment record contains some or all of the following components:

1.   type of record (e.g., activity log, activity description, portfolio entry, etc.)

2.   reference to file, URL, or external database entry that contains the actual log file, portfolio document, etc.

3.   a domain-based concept reference. e.g., domain 1 (mathematics). algebra. linearfn. intercept. computing (applicable to evaluations of understanding and skills).

4.   a learning category (exposed to concept, has basic understanding of concept, is able to use in problem solving, is skilled in use, etc.).

5.   possible misconceptions and attitude problems (e.g., has misconception: intercept = slope, has aversion or phobia).

6.   category of evidence (answered multiple-choice question correctly, incorrectly; gave clear oral explanation, etc.)

7.   detail of evidence (e.g., including time spent, testimonial, etc., the text of the multiple choice question, reference to recorded speech of oral answer, etc.)

8.   inferred skill level, level of understanding, facet of concept, etc.

9.   assessor's degree of confidence that inference is valid.

10.   any special assumptions required for this inference.

11.   representation of the reasoning behind this judgment, if available.

12.   information about circumstances of this judgment (e.g., limitations on the assessor's time or knowledge, references consulted during the assessment, etc.).

13.   assessor's identity.

14.   date and time of creation of this record.

With the above elements, an assessment record can be seen to respect the view that information about student learning be regarded as evidence. Evidence must carry with it enough description of its context that its over or under-interpretation does not unnecessarily lead to inappropriate judgments. (An example of a computer-mediated assessment system that respects assessment as evidence is described in (Tanimoto et al, 2000).)

6.3   List of Post Requests

When assessment is performed, the results should be recorded in two ways. First, either an assessment record is created and added to an existing assessment file, or a new assessment file is produced and one or more assessment records inserted. Then, summary information from the new or updated file should be carried over to a summary file for the student. The summary file may or may not be directly accessible to the system handling this assessment. For example, if the summary file is on the student's home computer but this assessment is being performed on a university computer, the updates may have to be downloaded at a later time. The post-requests part of the assessment file contains the updates that represent the influence of this assessment file upon the student's summary file.

Here are a couple of examples of possible post requests.

```
Post 3.2 additional hours spent on math.algebra
```

```
Post  to  math.summary.algebra.mastery  at  5.7  with  confidence  3.2
by  NetBasedAlgebra3.2,  based  on  3.2  hours  ending  14  April  1995.
```

If this file is created in a session that is disconnected from the summary file, a means is needed to make the update to the summary file as soon as a connection becomes established (which may be any time after the session). The actual mechanism for this is beyond the scope of this paper. However, we note that if this assessment file is changed after an update to the summary file has been performed, another update will be required to correct the original (and any other previous updates).

7   Design Issues

Aside from many questions about the details of implementing distributed transcripts, two fundamental and potentially problematical issues must be confronted. One has to do with ownership and rights to transcript information in contexts where multiple parties are participants in educational activities. The other has to do with systematic approaches to achieving agreement in meaning for assessment representations, and sometimes goes by the name "ontologies."

The ownership and privacy issue grows directly out of the notion of distributed transcripts, where the information pertinent to one student's learning exists partly in the student's dossier core and partly in databases owned by educational service providers.

The ontology issue is in part a debate between the position of permitting multiple ontologies to be accommodated ("ontological diversity") versus insisting on one ontology in each subject area (representing "ontological convergence").

7.1   Ownership and Privacy

Distributed transcripts are composed of multiple parts. The core of the dossier is kept by the student on a computer under his or her control, or it is on a server that permits each student to exercise ownership and maximum control of the dossier. Other components include temporary ones residing on other systems as well as permanent ones. When a student participates in a tutorial with a service provider, that provider keeps its own record of the activity, and it may choose to retain ownership of that record. However, that record forms a part of the student's dossier nonetheless, as long as it can be consulted by, or with the permission of the student. In the core dossier would be found a record identifying that the tutorial session took place, a summary of what took place or the results, and a pointer to the provider's record, possibly in the form of a URL. The provider would be capable and willing to validate certain particulars about the session, either as part of a previous agreement with the student, or for a new fee to be

paid either by the student or a new party, such as a prospective employer.

A standard agreement between a student and a tutorial provider then might involve the following elements: a single fee paid by the student to support the relationships, a package of learning services to the student including a number of lessons or an amount of time or number of allowed interactions or credits to be spent, and a number of validation transactions allowed without additional charge, when authorized by the student. It is assumed that the provider would maintain the records of learning involving the student for a minimum period such as, say, 10 years, in order to be ready to perform related validations. However, the provider would agree to let the student have a copy of the record, so that after expiration, if any, of the record, the information would not actually be lost, but simply would not be officially checkable.

7.2   Handling Assessments of Group Projects

Another situation in which the distribution of transcript information is particularly problematical and interesting is where records of group work by students is concerned. When students work together on a project, there may be evaluations of each student's contribution, and there may be evaluations of the progress or achievements of the group as a whole. In either case, an understanding of the significance of an individual's role in the project is dependent upon knowing the context for the role, and, to some extent, knowing what each of the other students in the group contributed. A perfect understanding of one student's performance in a group would seem to require a perfect understanding of each other group member's performance, too.

In order to make transcripts of group learning as valuable as possible yet protect the privacy of the group members, it is necessary to plan for appropriate compromises.

The very fact that group work is assessed and that transcripts are supposed to be comprehensive together create a requirement that groups agree to be assessed/monitored and that they agree within certain limits that assessment information about the group can be rightfully used by the members of the group. The potential complexity of this agreement is likely to prompt the development of standard agreements for students engaging in group projects, discussions, etc. (Another point where the legal profession may have something to tell educational technologists!)

The key elements of a possible agreement between a student and a service provider for a group activity might be the following: description of the activity, time period, and expected student role; fee or consideration to the provider; student's permission to be mentioned, using a pseudonym, in the released transcript records of the activity in response to requests for the descriptions of other students' participation; agreement by the provider to deliver descriptions and/or assessments of the activity on behalf of this student and to provide validation service on

these for a given period of time, such as for at least 10 years after the beginning of the activity.

These agreements might well have many details and variations from one provider to another. When a student enters a long-term relationship with a provider (for example, when matriculating in a four-year undergraduate program), a highly particular agreement could be justified, since the investment of attention to understand it might be repaid over time. However, an agreement with a provider for a short-term activity, such as a two-day online group discussion about algebra problems, should follow a standard template so that the investment required to simply understand the agreement does not get in the way of benefitting from the agreement.

7.3   Ontological Diversity and Convergence

The value of a transcript in facilitating transfer of a student from one learning environment to another depends upon (a) the detail and accuracy of the relevant information in the transcript, and (b) the extent to which this information can be correctly understood by the agencies within the learning environments. The latter capability is dependent upon the extent to which the agencies share compatible systems of meaning or "ontologies."

This paper assumes that that ontologies will be developed, since they are necessary for interpretation of transcripts, whether they are implicit or explicit. On the other hand, it is neither necessary nor strictly desirable that we assume that all service providers share a common ontology. We may hope that providers move in the direction of adopting a common ontology -- something called "ontological convergence" -- but that convergence may be some time off. Let's consider this issue further.

While the acceptance of standard ontologies would be desirable in order to achieve wide interoperability of educational systems, detailed educational records could still be valuable without that acceptance, provided there exist appropriate translation mechanisms. While the efforts to obtain standard ontologies should be applauded, to build a system that depends completely upon them is perhaps to build a system that is overly committed to a particular view of a field.

Different schools of pedagogy may ascribe to different ontologies for the same subject fields, and, while we might want to encourage these parties to join forces and reduce their differences, our representation system should not be completely dependent upon their doing so. Thus, an assessment record that happens to refer to the entities in some ontology must identify that ontology and not assume that there exists only one. The identification need not be burdensome; the naming of the ontology may be done in the context header of the assessment file. That naming should include, either explicitly or implicitly, the author(s) and version of the ontology.

7.4   Accommodating Ontological Diversity

It is not the case that having one accepted ontology for a subject area is a fundamental assumption for distributed transcripts, even though that agreement might indeed be nice to have. Given this, it becomes highly desirable to anticipate and to encourage the prospects for automatic translation across ontologies. Let's consider those prospects.

Differing ontologies certainly may complicate interpretations of raw data. Creating translators is likely to be problematic, in general, since concepts allowed in one ontology might have no corresponding representation in another. However, humans cope with such difficulties when they translate texts from one foreign language to another. Although we don't want to design a system that goes out of its way to encourage a new Tower of Babel, we also do not want the use of distributed transcripts to be held hostage until, say, all academic communities agree to view the world in the same way.

Let us consider an ontology to be a system of meaning, in which various component objects or concepts have been identified and related to one another. Some of those relations may be hierarchical, and others may have an arbitrary structure. Typically, natural language text may be used to explain what objects, concepts, and relations represent. An early ontology project was Cyc (Lenat et al, 1990), which consisted of a collection of micro-theories encoded in a computer language. A more recent ontology scheme has had the goal of supporting semantically accurate information retrieval (DCMI, 2001). Whereas most effort at ontologies in education have been directed at a system of metadata for cataloging learning resources, in the assessment of learning, ontologies are needed for both the detailed subject matter and for the forms and concepts of evidence of learning.

Two ontologies may differ in several respects. They may cover different subjects. They may cover similar subject matter at different levels of granularity. They may cover similar subject matter from differing conceptual bases. Or, they may describe similar notions with different language.

The easiest case for translation is when two ontologies are alike except that each uses different language to describe the same notions. Translation here requires little more than an enumeration of the correspondences between the pairs and a means for applying the implied mapping. It might be simply a matter of performing textual substitutions on XML tags to perform translations in such a case.

When two ontologies cover the same subject matter, and they divide up the matter in the same manner but to varying levels of granularity, translation methods can work systematically, taking advantage of the fact that the two ontologies are like two different refinements of a common

tree structure. Whereas in one ontology, a notion *A* may have child notions *B, C* and *D*, while in the other ontology, a notion *A'* corresponding to *A* may have no child notions, a translation of *C* may be performed in either of two ways. One is simply to map it to *A'*. For example, if *A* represents Newton's laws of motion as a group, and *C* represents the second law, then a provider's report of successful student activity involving *C* could be simply translated into a report of a successful activity involving *A'*. A second approach is to give in the translation an effective construction of a version of *C*, assuming that the second ontology permits it. Then the translation of the reference to *C* would be of the form "Some child of *A'*." The translation might report that the student "successfully solved a problem dealing with an aspect of Newton's kinematic laws." This is a more precise statement than the alternative translation that the student "had some success in problem solving with Newton's kinematic laws." But with either scheme, it should be clear that some reasonable translation is possible.

A translation in the other direction, say from *A'* to *A*, is less problematical, provided *A* is an explicit notion in its ontology; if not, then it may be necessary to translate the reference to *A'* either as a joint reference to *B, C* or *D*, or as a reference to a constructed parent "the parent of *C*." Clearly, it will be easier to translate to an ontology that does provide mechanisms for identifying groups of notions and constructing new children and parents of existing notions.

If two ontologies are fundamentally different, and the axes of conceptualization in one are incompatible with those in the other, or the overlap of their covered subjects is too small to be of use in a given situation, then little translation can be done. One would hope that some primitive kinds of translation would still be possible; for example, that a student spent so much time in a particular activity and that the activity falls under the general subject of physics.

If a student has invested a substantial effort or amount of time in an activity whose records later prove difficult to translate into a form suitable for supporting some desired transfer, then one solution is to request a manual or computer-assisted interpretation of the earlier records. For example, a student who participated in a summer-camp construction of a treehouse might have only a cursory report in her transcript on the educational significance of the project with entries that might read as "experienced the laws of physics when hoisting boards up to the big maple branch." In order to determine how this experience should affect a plan for the student's formal physics laboratory schedule, a questionnaire could be generated for the student that helps the translation and fills in missing entries dealing with specific laws of motion.

While much of the difficulty of translation could be avoided with uniform agreement on a standard ontology for each discipline, there may still be a need for translation because of (a) overlapping disciplines, and (b) evolution of ontologies over time, as with other software and standards. So ontologies should be designed with translation in mind, for example, providing

embedded facilities for expressing notions beyond those in their core domains. Facilities for referring to unspecified, new children of existing nodes is one example. The use of categories such as "other" or "set of notions" is another.

8   Relation to Systemic Educational Reform

The fundamental design assumptions we've made here are that (1) the overall transcript database for a student is owned by that student, although particular objects referenced in it may be owned by institutions or by other individuals, (2) the transcript is comprehensive and accumulates much as is economically practical of the potentially relevant data pertaining to the learning experiences of this student, (3) inferences and summative evaluations are evidence based, and the reasoning processes employed in the assessments are represented explicitly and honestly. (4) All judgments are signed by the responsible parties, and (5) the student takes responsibility for the care and protection of the transcript (with the help of properly designed systems).

These assumptions are somewhat at odds with the old-fashioned model of education as the responsibility of institutions. The notion that the student owns the primary transcript will require that institutions acknowledge that they do not have absolute control over the educational records of their students, even though they may have complete control of records of their own relationships to the student. The completeness of the record requires that all parties to an educational experience strive to assist in the accumulation of records of the experience and turn them in to whatever agent is assisting the student with the transcript. The requirement of justification and citing of evidence for summative evaluations will require a more sophisticated notion of assessment than the one currently taken by most schools, teachers, and educational testing services; assessors will be held accountable for their judgments.

Adoption of a framework for educational assessment based upon these assumptions will require radical changes in teacher training. However, it should be possible to make the transition gradually, with transcripts collecting more and more kinds of evidence as time goes on. Students will have to be trained in the proper care of their transcripts, and they must be held accountable by parents, institutions, and online learning systems for the manner in which they carry out this responsibility. As with driving a car, they may need to take a course, get a temporary novice's permit, and then get a license. They could hurt themselves and cause trouble for others if they are negligent with their transcripts.

One of the possible consequences of a greater reliance on electronic learning systems for education is a widening of the so-called digital divide. Students from affluent families with excellent access to high-technology facilities may become smarter, and perhaps richer, while

children without much access to technology get left even further behind than poor children were before the age of electronic learning environments. The assessment paradigm given here may help to counter the digital divide, provided all students can be empowered with the means to take control of their own education. Mismatches between resources and individual needs can be better avoided with the kinds of information in the full transcript. Access to appropriate materials could become as taken for granted as access to the global telephone network.

When transcripts become rich enough in valuable information, it will become easier to facilitate the formation and maintenance of online communities that can communicate tightly and meaningfully with each other, even when they are new.

Privacy of educational data must be respected. All parties to the information must honor the right of privacy. When the student takes ownership of the principal records, a big step towards privacy is taken. Then he or she controls the access to that information, with the assistance of trusted software agents.

9   Concluding Remarks

The Internet is changing many aspects of formal and informal education from the ways college professors distribute their handouts to the ways that students research a topic. Many aspects of learning could work more efficiently if the learners themselves were empowered by the technology to set their own goals and find their own means of achieving them. Current standards efforts in electronic learning can be expected to have a positive influence on the efficiency of the markets for educational materials and activities. However, these standards groups have tended to be industry and institution oriented, rather than learner oriented. The greatest efficiencies connecting learners with the resources they need will occur when the learners themselves are fully empowered and engaged, with primary responsibility for setting and achieving their educational goals and primary ownership of their own comprehensive educational histories. An assessment information architecture that facilitates the capture, control and exchange of detailed educational histories and evaluations, fully in service of the learners, is a key step in maximizing the efficiency of learning in the 21st century and beyond.

10   Acknowledgements

11   References

Broadfoot, Patricia (ed). 1986. Profiles and Records of Achievement. London: Holt, Rinehart and Winston.

DCMI (2001). Dublin Core Metadata Initiative. [dublincore.org]

EDUCOM (1998). Instructional Management System. [www.imsproject.org]

Haertel, G., and Means, B. (2000). Stronger Designs for Research on Educational Uses of Technology: Conclusion and Implications. Center for Innovative Learning Technologies, SRI International [www.sri.com/policy/designkt/synthe1b.pdf]

Hoffman, B. (1962). The Tyranny of Testing. NY: Crowell-Collier.

Lenat, D. B., Guha, R. V., Pittman, K., Pratt, D., and Shepherd, M. (1990). Cyc: Toward Pograms with Common Sense. Communications of the ACM, Vol. 33, No. 8 (August), pp. 30 - 49.

Linn, R. L. 1989, (Ed.). Educational Measurement, Third Edition. London: Collier Macmillan Publishers.

LTSC (2000). IEEE Learning Technology Standards Committee website. [ltsc.ieee.org]

Minstrell, J. (1992). Facets of Students' Knowledge and Relevant Instruction. In: Duit, R., Goldberg, F., and Niedderer, H. (Eds.), Research in Physics Learning: Theoretical Issues and Empirical Studies. Kiel, Germany: Kiel University, Institute for Science Education.

Murphy, M., and McTear, M. (1997). Learner Modelling for Intelligent CALL. In: Jameson, A; Paris, C.; and Tasso, C. (Eds.) User Modeling: Proceedings of the Sixth International Conference, UM97. Vienna, New York: Springer Wien New York. pp. 301-312.

Murray, T. (1998). A Model for Distributed Curriculum on the World Wide Web. Journal of Interactive Media in Education, 98 (5) [www-jime.open.ac.uk/98/5]

Ritter, S., and Suthers, D. (1997). Technical Standards for Education. Working Paper, Educational Object Economy web site. [www.eoe.org]

Tanimoto, S. L. (1992). Beyond the Naivety of Grades: Educational Record Keeping for the Twenty-First Century. Technical Report 92-07-09, Dept. of Computer Science and Engineering, Univ. of Washington, Seattle WA, USA, July.

Tanimoto, S. L., Carlson, A., Hunt, E., Madigan, D., and Minstrell, J. (2000). Computer Support for Unobtrusive Assessment of Conceptual Knowledge as Evidenced in Newsgroup Postings. Proceedings of ED-MEDIA 2000, Montreal, Canada.