Chapter 6 The big problems – What does the research say about how Artificial Intelligence and Big Data can close the achievement gap?

Benedict du Boulay[1], Alexandra Poulovassilis[2], Wayne Holmes[3], Manolis Mavrikis[4]
1. Department of Informatics, University of Sussex, Brighton
2. Birkbeck Knowledge Lab, Birkbeck, University of London
3. Institute of Educational Technology, The Open University
4. UCL Institute of Education, University College London

Introduction

> Currently, we are failing to meet the needs of all learners. The gap between those who achieve the most and those who achieve the least is a challenge that teachers, school leaders, administrators, and government officials face every day, in every country. Globally, students from poorer backgrounds perform worse than students from richer backgrounds (Conroy and Rothstein, 2013). The results of this achievement gap impacts upon a country's economy as well as the social well-being of their population (Hanushek and Woessmann, 2010). The reasons behind the achievement gaps in different countries vary, but the fact remains that not all learners are achieving their potential at school.
>
> (Luckin, Holmes, Griffiths, and Forcier, 2016, p. 42)

We observe achievement gaps even in rich western countries, such as the UK, which in principle have the resources as well as the social and technical infrastructure to provide a better deal for all learners. The reasons for such gaps are complex and include the social and material poverty of some learners with their resulting other deficits, as well as failure by government to allocate sufficient resources to remedy the situation. On the supply side of the equation, a single teacher or university lecturer, even helped by a classroom assistant or tutorial assistant, cannot give each learner the kind of one-to-one attention that would really help to boost both their motivation and their attainment in ways that might mitigate the achievement gap.

This chapter argues that we now have the technologies to assist both educators and learners, most commonly in science, technology, engineering and mathematics subjects (STEM), at least some of the time. We present case studies from the fields of Artificial Intelligence in Education (AIED) and Big Data. We look at how they can be used to provide personalised support for students and demonstrate that they are not designed to replace the teacher. In addition, we also describe tools for teachers to increase their awareness and, ultimately, free up time for them to provide nuanced, individualised support even in large cohorts.

Artificial Intelligence and Big Data in Education

The name "Artificial Intelligence" (AI) can be a little scary, especially at present where the notion of (an) AI taking over the world to the detriment of society is a popular contemporary nightmare.

Deleted: is

Artificial intelligence in education is not about educational robots taking away jobs from teachers and brainwashing children. It is much more prosaic and consists of programs running on tablets and laptops that help teach learners on a one-to-one basis in a way that adapts the tasks, assistance and the feedback to the capabilities and progress of the individual learner.

Artificial Intelligence in Education is a computer-based technology that tries to provide insightful, adaptive and personalised teaching, at the level of competence of an expert human tutor, for individuals and groups. In particular, such computer-based systems attempt to choose appropriate tasks for the learner to work on and then react dynamically to how they go about dealing with these tasks. These reactions can take the form of specific hints on individual steps taken and on requests for help, as well as providing general assistance (or “scaffolding”). Note that the reactions of the system are not only provided once the learner has submitted an answer but can also be provided in response to individual steps towards that answer. Such systems are also known as “intelligent tutoring systems” (ITS). Other systems are more open-ended and sometimes less individually adaptive but provide opportunities for a learner to explore a domain; they are referred to as Exploratory or Open-Ended Learning Environments (ELE). We will refer to all of these here as “AIED systems”.

The term “Big Data” is also a little scary, especially where corporations and governments hoover up huge amounts of personal data, where there seems to be endless breeches of privacy and data hacking, and the boundary between secure and insecure data is very porous. In an educational context, Big Data can be beneficial in that it can collect information about how a cohort of learners are interacting with a learning environment and making progress with their learning. This information can be used by teachers and instructional designers to improve the environment and the support it offers to students. So Big Data enables learning

Deleted: seem

Deleted: is

environments to be adapted by showing where they work well and where they do not.

This chapter is organised as follows. In the next section we describe in more detail what an AIED system is and provide some examples. We then describe some exploratory learning environments. We then move to Big Data, both to describe it and give examples. In the final part of the chapter we examine the evidence for the educational value of AIED systems and Big Data.

The key parts of AIED systems

The capability to individualise its teaching and assist even with partial answers depends on an AIED system having the following four components:

1. The **Domain Knowledge Model** is the component that provides the capability of the system to complete the tasks that it sets the students and to judge which steps contribute towards a solution, or which parts of an answer are correct. In other words the system needs to understand the material that it is teaching, unlike a book or a website that can merely present that material. Because STEM subjects lend themselves much more readily to having their domains represented in ways that can be automatically reasoned about, most AIED systems have been built to teach these areas.

Formatted: Font:Not Bold

2. The **Student Model** is the component that provides a representation of the learner in terms of their developing knowledge and skills. This is needed so that tasks of an appropriate complexity and difficulty can be set. As the learner works through various tasks, the system builds up a "student model" of what the learner can reliably get right, what they seem to partially understand, and what they seem to be as yet very poor at. This model can never be exact, but is a best guess and can be used, for example, to select the next task for the learner or to give a little bit of challenge in areas not yet mastered and also practice in areas that seem well understood.

Formatted: Font:Not Bold

3. The **Model of Pedagogy** is the component that represents the teaching capability of the system. This is used to make decisions about how best to present new material, how best to deal with requests for help, how best to deal with incorrect steps and answers and so on. This might also include an understanding of how to motivate the learners if they become demotivated and tactics to deal with students who try to "game the system" (Baker et al., 2008) by demanding so much help that the system might otherwise give them all the answers.

Formatted: Font:Not Bold

4. The **Interface** is the component that provides the channel through which the learner and the system communicate. This channel might be through spoken dialogue, or text and diagrams provided either by the learner or the system (see Figure 2). Such an interface may also include an animated pedagogical agent taking the role of a tutor or of a fellow student.

Formatted: Font:Not Bold

We present below some examples of AIED systems illustrating their Domain Model, Student Model, Pedagogical Model and Interface. We chose a variety of systems and their potential to support students in different contexts.

Examples of AIED systems

Procedural skills – the Cognitive Algebra Tutor

Our first example is an older and ‘traditional’ intelligent tutoring system. It teaches algebraic skills such as equation solving. In the Pittsburgh Algebra Tutor, and other similar systems derived from it, the overall form of interaction is that the system chooses an individualised sequence of algebraic problems for the learner to solve and then monitors each step that the learner takes in solving each problem. The system has gone through several iterations. The interface shown in Figure 1, taken from a much-cited early paper (Koedinger, Anderson, Hadley, and Mark, 1997), offers a problem specific worksheet for the learner to fill out their partial answers to each sub-step of the problem. Later versions, such as the one used in a large evaluation described later, provide a more modern look and feel with access to a number of other tools to help the learner and find information (Koedinger and Aleven, 2016).

Deleted: The
Deleted: first, and perhaps
Deleted: more
Deleted: ,
Deleted: changing

<<FIGURE 1 about here>>

Figure 1. Interface for the Pittsburgh Algebra Tutor, taken from Koedinger et al. (1997)

The system uses its domain model and student model to sequence the problems for each learner on an individual basis, depending on their rate of progress in mastering the various algebraic subskills needed for each problem (see bottom right of the interface). They are also used to reason about partial answers in various representations, such as a graph, spreadsheet and equation solver, to decide when a partial answer is a step in the right direction to solving the overall problem and when it is not. The pedagogical model makes the system react quickly to any mistake made by the learner so as to reduce the chance that they stray too far from the solution and get muddled. It also dynamically assesses what is the best next problem for the learner to work on so as to ensure that new skills are encountered and old skills practiced.

Exploratory Learning Environments for conceptual understanding - BETTY's Brain

Our second example is Betty's Brain, a system designed to teach scientific *conceptual* understanding of river ecosystems (Leelawong and Biswas, 2008). In particular, it aims to help learners appreciate the complexity of the causal and other relationships between different processes occurring in such ecosystems; for example, that fish produce waste and that this waste is food for bacteria (see Figure 2). There is also skill building in the learner's interactions with Betty's Brain, such as following causal chains of reasoning and developing generic study skills, but the main focus is still on conceptual understanding.

Deleted: is

<<FIGURE 2 about here >>

Figure 2. Betty's Brain interface, taken from Leelawong and Biswas (2008)

The heart of the system is the Concept Map Editor pane in the top right of the interface. This is where the learner builds up a conceptual map of the river ecosystem, using nodes and links via the Editor on the top left of the screen. The conceptual map can also be understood by the system and this enables it to answer questions based on it.

The narrative behind the interaction is that the human learner is attempting to teach fellow student, Betty, seen on the bottom left of the screen in Figure 2. The conceptual map is a record of what the learner has so far taught Betty – hence "Betty's Brain". The learner can test the adequacy of what Betty has learnt by asking her to take a quiz administered by Mr Davis, the teacher. Mr Davis assesses Betty's answers to the quiz questions and provides feedback to the learner, who then has the chance to edit the conceptual map in an attempt to help Betty get a better quiz score. Mr Davis assesses Betty's answers to the quiz by reasoning from the conceptual map created by the human learner. This slightly indirect way of learning has a particular advantage for the human learner in terms of somewhat forestalling any negative reactions from the learner to mistakes in the quiz, as they are Betty's mistakes.

The learner can test the adequacy of the conceptual map directly by asking Betty such questions as "If macroinvertebrates increase what happens to bacteria?" Betty can answer such a question and explain that answer by following the causal reasoning indicated in the conceptual map using qualitative reasoning techniques.

The system also provides learning materials that the learner is encouraged to use, see the lower part of the screen in Figure 2. In addition to feedback about the domain of river ecosystems, Mr Davis also makes suggestions at the meta-cognitive level, for example about making better use of the reading materials, in an effort to help the learner develop good study skills.

In terms of the four components mentioned in the previous section we note that domain knowledge of the system is its ability to reason using the conceptual maps produced by the learner. Its student model is made up of a record of the various actions taken by the learner and the partial but growing understanding of the domain as exemplified in the conceptual map. In pedagogical terms the system is driven by the actions of the learner, although the overall educational goal of having Betty pass all the quizzes is clearly provided by the system. The system does have a model of pedagogy that drives how and when it makes comments at the metacognitive level, for example when Betty reacts to being asked to take a second quiz even though there has been no change to the conceptual map. Finally, the interface is key to the interaction as the conceptual map built by the learner is both an expression of their evolving understanding and can be reasoned about by the system (even if the map is wrong or partial).

eXpresser and the MiGen system

Another example of an exploratory environment is a mathematical microworld called eXpresser that aims to support 11-14 year olds' learning of algebraic generalization, as part of a system called MiGen (Noss et al.

2012)[1]. Using eXpresser, students are asked to construct two-dimensional tiled models and associated algebraic rules. The algebraic rules relate to the number of tiles of each colour required to paint each pattern and their model overall (see Figure 3).

<<FIGURE 3 here >>

Figure 3. The eXpresser microworld. Letters highlight the main features: (A) An 'unlocked' number (i.e. variable) is given the name 'reds' and signifies the number of red (dark grey) tiles in the pattern. (B) Building block to be repeated to make a pattern. (C) Number of repetitions (in this case, the value of the variable 'reds'). (D,E) Number of grid squares to translate B to the right and down after each repetition. (F) Units of colour required to paint the pattern. (G) General expression that gives the total number of units of colour required to paint the whole pattern.

Figure 4 illustrates a feedback message given by the eXpresser to a student who has constructed a correct pattern and a correct colouring rule for it, nudging the student towards "unlocking" a number (i.e. turning it into a variable) so as to now generalise their pattern and rule. Figure 5 shows a message of encouragement but also a stronger prompt to guide the student towards generalising their construction.

<<FIGURE 4 here >>

Figure 4. A 'nudge' from the eXpresser

[1] eXpresser is one of a set of tools making up the MiGen system, which was developed through funding from the ESRC/EPSRC Technology Enhanced Learning programme, award no. RES-139-25-0381

<<FIGURE 4 here >>

Figure 5. A message of encouragement and a stronger 'prompt' from the eXpresser

In terms of the four components of AIED systems mentioned earlier, the domain knowledge of the system is its internal model of mathematical concepts relating to algebraic generalisation. The student model records the learner's gradual mastery of these concepts as the learner works through successively harder tasks, as well as a history of the learner's constructions and interactions with the system. In pedagogical terms, each task comprises a set of learning goals that the learner needs to achieve as they work on the task using eXpresser. The system provides adaptive support based on how the student is approaching the task and how they are interacting with the system. Again, the eXpresser interface is key to the student-system interactions and the student's growing conceptual knowledge, as their construction of models and rules can be reasoned about by the system in order to provide appropriate support and also demonstrates the student's evolving understanding of the domain.

Combining ITS and ELE – the case of iTalk2Learn

Intelligent Tutoring Systems like Cognitive Algebra Tutor and Exploratory Environments like eXpresser do not have to exist in isolation. The iTalk2learn project developed an adaptive digital learning platform for primary school mathematics that allows interaction via direct manipulation and speech to provide intelligent interventions and individualized task sequences.[2] Importantly, for the discussion here, iTalk2Learn combines structured and exploratory activities to improve learners' procedural as well as conceptual knowledge (Rummel et al. 2016). It does so by offering activities from a commercial intelligent tutoring system (Math-Whizz, http://www.whizz.com/) to support procedural knowledge, and from a microworld called Fractions Lab to improve students' conceptual knowledge of fractions (Hansen et al. 2016). In Fractions Lab, students are asked to construct one or more fractions and, using the affordances of the system, to compare, add or subtract fractions. In Figure 6, for example, the student has been asked to create a fraction, and then to create four equivalent fractions with increasingly larger denominators. So far the student has created their first fraction, but has not yet created any equivalent ones. The glowing lightbulb at the top of the screen indicates that there is help currently available from the system (Grawemeyer et al. 2015). Clicking on the lightbulb results in the feedback message shown in Figure 7, which is aiming to nudge the student towards the next step. Figure 8 shows that the student has indeed made their first equivalent function. After a period of inactivity, Figure 9 shows a message of encouragement and also an unsolicited prompt (Grawemeyer et al. 2015) to guide the student towards the next step.

As students are undertaking tasks, they are encouraged to talk aloud. A speech recognition system extracts keywords which are combined with prosodic features also extracted from the speech signal and used as input to methods for the classification of students' sentiment and cognitive load. The outcomes of this emotion and affect recognition serve as input for providing

[2] The iTalk2Learn system was developed through co-funding from the EU FP7 programme, ref. no. ICT-318051.

intelligent support to the student and automatic selection of interventions. This relies on large amounts of data and student modelling, as described in the next section on Big Data.

<<FIGURE 6 here >>

Figure 6. FractionsLab microworld, showing the availability of low-interruption feedback.

<<FIGURE 7 here >>

Figure 7. FractionsLab microworld, showing the elective display of low-interruption feedback.

<<FIGURE 8 here >>

Figure 8. FractionsLab microworld, showing that the student has progressed to the next step.

<<FIGURE 9 here >>

Figure 9. FractionsLab microworld, showing a message of encouragement and also a stronger prompt to guide the student towards accomplishing the subsequent step.

In terms of the four components of AIED systems, these are similar in functionality to those of eXpresser, except that in Fractions Lab the student model also includes information about the student's evolving affective state as the student works on a task using the microworld.

Two ways in which AIED systems might be used

There are two main ways that AIED systems can be used effectively in schools. First, such systems can be deployed as classroom assistants in the following sense. While whole group teaching or small group teaching by a human teacher continues to be the norm, it is commonplace for an individual or a small group to be handed over to a human classroom assistant. This might be to provide individual help for pupils who are not doing so well, or it might be to assist pupils who have already mastered the material ahead of the rest of the class and who need a bit more of a challenge. The idea here is that in addition to the human classroom assistant, an AIED system could be used by an individual pupil or a group of pupils who need extra practice or who need exposure to more challenging material. The ability of such systems to monitor the individual problem-solving steps of the pupil and to provide help, hints and scaffolding specifically appropriate to the individual could be a valuable extra tool in the classroom.

There are also potential benefits for a group of pupils working with such a system to discuss and argue about different possible answers to problem-solving steps, as well as the meaning and intent of feedback received from the system on their errors. For the more able pupils running ahead of the rest of the class, such systems can provide more challenging problems, possibly with less help and scaffolding, thus maintaining their motivation.

The second way that AIED systems can be used is as assistants in after-school classes, revision classes or for homework. In these situations the classroom teacher is typically less available, but the pupil will still need the kind of detailed assistance that such systems are able to provide. Just as we have mentioned that groups of pupils can discuss an ongoing interaction with an AIED system to help create better understanding, so a child and a parent at home can have a similarly fruitful discussion together in the context of using an AIED system.

Note that our use cases have the AIED system working in tandem with the classroom teacher and not as a replacement. Those visions of future education involving simply computer-based instruction without the social and pedagogic support of human teachers are barren indeed. It is instructive to note the high drop-out rates when college-level courses are delivered solely via Massive Open Online Courses (MOOCs) direct to the individual learner, with little in the way of face-to-face interaction with either the teacher or with fellow students (Liyanagunawardena, Adams, and Williams, 2013).

Big Data in Education

Emergent web, mobile, and pervasive digital technologies are generating data at unprecedented scales and speeds in virtually all areas of human activity. Across industry, commerce and the public sector this Big Data is being digitally collected and computationally analysed in order to gain better understanding of providers' services and products, consumers' needs and preferences, and, more fundamentally, to expand human knowledge across the sciences, social sciences and humanities.

Originally, Big Data was taken to mean data sets that are beyond the management and analysis capabilities of traditional software tools. The generation of such data sets led to the development of new data storage and data processing paradigms, such as NoSQL data stores (Cattell 2011), massively data-parallel distributed processing frameworks (Dean & Ghemawat 2008, EMC 2015) and cloud computing platforms (Armbrust et al 2010).

Big data is distinguished from other data by exhibiting the so-called 'V' attributes. These include:

- *volume* – the size of the datasets;
- *velocity* – the rapid rate at which the data may generated;
- *variety* – different types of data being generated from multiple sources, needing to be cross-referenced and combined in order to be fully exploited;
- *veracity* – the incompleteness of the data being collected, and the imprecision of inferences being made from it; and
- *volatility* – data being collected or inferred may become less relevant over time.

More recently, there is a recognition that these 'V' attributes are not the whole story and that what is most important is the ability to extract *value* from such data while also complying with given time, human and technical resource constraints.

Learning Analytics and Educational Data Mining

Big data in the education sector is the focus of two complementary academic fields: **Learning Analytics** and **Educational Data Mining**.

The field of Learning Analytics (LA) is concerned with *gathering, analysing and visualising* data about learners and learning processes, so as to increase stakeholders' understanding of these and hence *to improve learning and the environments in which it occurs* (Siemens 2012, Drachsler and Greller 2012, Ferguson 2012). This data may be collected from many different sources:

- virtual learning environments (VLEs) that track and support students' activities, interactions, reflections and progress through learning tasks;
- students' assessment activities – both formative and summative;
- students' personal records and records of prior achievement;
- learner profiling and learner modelling software;
- software supporting social networking, peer support, and collaboration;
- audio and video recordings;
- gesture and physiological sensor recordings (e.g. heart rate, galvanic skin response, blood pressure, EEG readings); and
- mobile learning apps, gathering large-scale user-centred and context-aware data.

This exceptionally broad range of data sources is allowing increasingly *individualised*, *detailed* and *longitudinal* data to be collected and analysed, bringing with it the potential to derive new insights and to provide more effective support to learners and tutors.

The field of **Educational Data Mining** (EDM) was established a few years earlier than the LA field and it, too, is concerned with gathering and analysing data so as to understand, support and improve students' learning.

However, the LA and EDM fields have somewhat different emphases (Siemens and Baker 2012):

- LA focuses on tools to aid users in their roles, whereas EDM focuses on tools for automated knowledge discovery.
- LA focuses on understanding learning processes as a whole, whereas EDM focuses on understanding specific aspects of learning and the relationships between them.
- LA focuses on tools that empower students, learners, teachers and other stakeholders to make decisions, whereas EDM focuses on automated personalisation and adaptation of learning environments.

Nonetheless, there is also much commonality between LA and EDM and they can indeed be regarded as parts of a larger interdisciplinary continuum of research and practice involving disciplines such as computer science, education and psychology, as well as teachers, learners, learning designers, policy makers and other stakeholders in learning processes from across the public and private sectors.

There is also commonality in the computing techniques developed and applied in the LA and EDM fields, which include: data modelling; data cleansing, transformation and integration; knowledge representation and reasoning; data mining, analytics and visualisation; learner modelling; recommender systems; predictive modelling; social network analysis; and discourse analysis. We refer readers to (Poulovassilis 2016) for a more detailed discussion of these different techniques, their applications, and references to the relevant technical literature.

The Sources and Design Process of Big Data in Education

Collection and analysis of learning-related data has been used in Technology Enhanced Learning research and practice for many years. Big data, however, start playing a particular role when considering data from systems such as the AIED ones presented in the previous section. We can see from the description of these systems that the data they generate include:

- Event-based data: log data of students' interactions with the system; students' responses, ranging from simple answers to a question to more complex reflections, e.g. through text (in MiGen) or speech (in iTalk2Learn); occurrences of key indicators as students interact with the system; generation and provision of feedback by the system.
- Students' constructions: the diagrams in Betty's brain, or the models and mathematical expressions being constructed by students in eXpresser, including a full history of how each was constructed.
- Task information: task descriptions, task learning goals, common solution approaches to each task.
- Learner models: information about students' level of attainment of concepts and skills, recent history of interactions with the system, progress with tasks set, achievement of learning goals, affective states.

We can see that this data exhibits all of the 'V' attributes that we discussed earlier. As well as its evident volume and velocity, under the 'variety' attribute we have unstructured data (e.g. the students' reflections), semi-structured data (e.g. the log data, task information, and students' constructions) and structured data (e.g. the learner models and indicator data). Under 'veracity' there is the inherent imprecision of the inferences being made by the system's intelligent components, e.g. in the detection of task-dependent indicators (Gutierrez-Santos et al 2012) or students' affective states (Grawemeyer et al 2015). Under 'volatility', a student's history of interactions, inferred indicators and affective states may become less relevant with time.

The rich range of data that can be collected by an AIED system provides not only the possibility to generate personalised feedback for the learner, but also the opportunity to design visualisation and notification tools for the teacher. The provision of such tools can help the teacher to formulate her own interventions to support both individual students and the class as a whole.

To be fully effective in the classroom, such tools need to be designed by multi-disciplinary teams involving teachers, pedagogical experts and computer scientists. In our own work in this area, we have used an iterative participatory methodology, comprising successive phases of prototyping, requirements elicitation, incremental development and evaluation (Gutierrez-Santos et al. 2012, Mavrikis et al. 2016, Gutierrez-Santos et al. in press). The next section illustrates this through examples.

Examples of applications of Big Data in Education

Student modelling from big data – the case of affective learning

Perhaps the most common use of data from digital learning environments is to inform the system's internal conception of the learner and its learner modelling, as mentioned already in the previous section. One of the most innovative applications of such data is for the detection of a student's *affective state*. Such information can be used to enhance learning by means of nudges that move students out of negative states such as boredom or frustration that inhibit learning into positive states such as engagement or enjoyment. Affective states can be detected through computational analysis of data extracted from speech, facial expressions, eye tracking, body language, physiological signals, or combinations of these (D'Mello and Kory 2015). In the iTalk2Learn system, for example, a student's affective state is determined through detection of keywords and prosodic features in their speech as they talk aloud when interacting with the system (Grawemeyer et al 2015b). Such detailed student modelling can enable affect-aware support for the student, which has been shown to contribute to reducing boredom and off-task behavior, with promising effects on learning (Grawemeyer et al, in press).

Deleted: students'

In addition, rich data from such systems can be used by designers and researchers to investigate the system's performance and efficacy and to identify areas requiring further development. For example, the system-student interaction data arising from iTalk2Learn have been recently remodelled using graph-based methods so as to more easily investigate the effectiveness of the intelligent support being provided by the system. Figure 10 illustrates one possible visualisation of how a student's affective state changes during a learning task.

<<FIGURE 10 here >>

Figure 10. Graph-based modelling and visualisation of students' interactions; the figure illustrates how one student's affective state changes between states of Engagement (green), Frustration (amber) and Confusion (red). Successive events are shown in blue and are connected by red edges.

Teacher tools for Exploratory Learning Environments

We described earlier the eXpresser mathematical microworld, which is one of the tools making up the MiGen system. Figures 11 and 12 illustrate two of that system's Teacher Assistance tools, each of which draws on the data generated by students' use of eXpresser: the Classroom Dynamics (CD) tool and the Goal Achievements (GA) tool. In the CD tool, each student present in the classroom is represented by a circle containing their initials. At the outset of the lesson, the teacher can drag-and-drop these circles so that their positions on the screen reflect the students' spatial positioning in the classroom. The colour of a student's circle reflects the student's current activity status, as inferred by the system. Green indicates a student working productively on the task set. Amber indicates a student who has not interacted with eXpresser for some time (by default, five minutes). Red indicates a student who has requested help from the system in a situation where the intelligent support cannot help any further: in such cases, the eXpresser displays the message "The teacher will come to help you now" to the student, and the student's circle becomes coloured red to attract the attention of the teacher.

Deleted: s

Most of the time, the teacher will have the CD tool selected for display on her handheld computer. When students show as amber, she can approach them and encourage them to resume working on the task set. If students who are not showing as red call out for help she can encourage them to first seek help from the system, knowing that if the intelligent support cannot help the student's circle will automatically appear as red in the CD tool. If a student does appear as red, the teacher can click on the student's circle on her way over to the student so as to see their current model and rule, which helps her to prepare her feedback for the student.

From time to time, the teacher will also consult the GA tool, which again visualises part of the big data being generated by the system (in this case, indicators inferring the current status of the student's achievement of the expected learning goals of the task). The GA tool presents a tabular display of students and task goals. Each row of the table shows the progress of one student (identified by their initials) in completing the task goals. A white cell indicates a goal that has not yet been achieved by the student. A green cell indicates that the goal is currently being achieved by the student's construction. An amber cell indicates that the goal was achieved at some point, but is not currently being achieved by the student's construction. Knowing which students have accomplished all the task goals allows the teacher to set them additional activities, for example comparing their construction approach with that of a peer (see below). If the GA tool shows that many students are not achieving a particular task goal, the teacher can interrupt the lesson to help all the students at the same time.

Deleted: also

<<FIGURE 11 here >>

Figure 11. MiGen's Classroom Dynamics tool. On the left, a classroom with the students sitting at benches in rows. On the right, the teacher has clicked on the 'red' student to see their construction and rule on the way over to help them.

<<FIGURE 12 here >>

Figure 12. MiGen's Goal Achievements tool. We see that some students have achieved all or most task goals, some students have not made any progress yet, and some students are moving back and forth.

Another of MiGen's teacher tools – the Grouping Tool (GT) (Gutierrez-Santos et al, in press) – supports the teacher in managing group discussion activities after students have finished their individual construction activities, by automating the pairing of students based on their constructions. Identifying appropriate pairs would be very time-consuming for the teacher to do manually during a lesson: it would require the teacher to investigate every student's construction, identify pairs of constructions that are sufficiently dissimilar to lead to fruitful student discussions and reflections, and then put the students into pairs, taking also into account interpersonal factors. The GT generates an initial set of pairings, aiming to minimise the overall similarity across all pairings. The proposed pairings are presented visually to the teacher, who can then confirm or change each pairing – see Figure 13 (we note that in the case of an odd number of students, one of the 'pairings' generated will be a triplet!). In the GT, students are represented by their initials within a circle. The degree of similarity between pairs of constructions is represented by a small green rectangle for low similarity; medium-sized yellow rectangle for moderate similarity; or large red rectangle for high similarity. The teacher can select students' circles and drag them into different groups in order to change the pairings suggested by the system so as to take into account factors that are beyond the system's knowledge, such as students' interpersonal relationships.

<<FIGURE 13 here >>

Figure 13. MiGen's Grouping Tool.

The immediacy of the big data presented through MiGen's teacher tools can help teachers formulate their interventions during the current lesson, set additional homework, plan the next lesson, as well as adjust the design of future tasks to be set for a given class of students. The availability of such tools allows teachers to use ELEs in the classroom in new ways because they provide a greater sense of awareness than is possible with general-purpose student monitoring tools. Moreover, such tools can support teachers in providing evidence of students' learning, even in a context that is less subject to formal assessment, and to engage in their own enquiry into more conceptual student learning.

Tools for planning and reflecting on learning

So far, we have seen examples of educational software in which data volume and velocity arise from the fact that the majority of the data are being generated by the system as users interact with it. There are other categories of system (most notably, social networking and collaboration software) in which high data volume and velocity arise from the numbers of users and where the majority of the data are user-generated. Research in the L4All and

MyPlan projects[3] provides an example of this latter category of system. The prototype L4All system developed by these projects aimed to support adult learners in exploring learning opportunities and in planning and reflecting on their learning. The system allows users to create and maintain a chronological record of their learning, work and personal episodes—their timelines. Users' timelines are encoded as RDF triples, compliant with an RDFS ontology[4]. There are some 20 types of episode, each belonging to one of four categories: Educational, Occupational, Personal, and Other. Figure 14 illustrates a fragment of the overall L4All ontology.

<<FIGURE 14 here >>

Figure 14. Fragment of the L4All ontology. Each instance of the Episode class is: linked to other episode instances by edges labelled `next' or `prereq' (indicating whether the earlier episode simply preceded, or was necessary in order to be able to proceed to, the later episode; linked either to an Occupation or to an educational qualification (Subject) by means of an edge labelled 'job' or 'qualif'. Each occupation is linked to an instance of the Industry Activity Sector class by an edge labelled 'sector'. Each qualification is linked to an instance of the National Qualification Framework (NQF) class by an edge labelled 'level'. The Occupation, Subject, Industry Activity Sector and NQF hierarchies are drawn from standard United Kingdom occupational and educational taxonomies (see Labour Force Survey User Guide, Vol 5, http://www.ons.gov.uk/ons/guide-method/method-quality/speci_c/labour-market/labour-market-statistics/index.html).

[3] L4All – Lifelong Learning in London for All; MyPlan – Personal Planning for Learning throughout life. Funded by JISC Distributed e-learning Pilot Call, 2005 – 2008.

[4] See https://www.w3.org/standards/semanticweb/ for information about RDF and RDFS.

Users can choose to make their timelines ‘public’ and thus accessible by other users. This sharing of timelines exposes future learning and work possibilities that may otherwise not have been considered, positioning successful learners as role models to inspire confidence and a sense of opportunity. The system’s interface provides screens for the user to enter their personal details, to create and maintain their timeline (see Figure 15), and to search over the timelines of other users based on a variety of search criteria.

<<FIGURE 15 here >>

Figure 15. The main screen of the L4All system. At its centre is a visual representation of the user's timeline, and the system functionalities are organised around this. Each episode of learning or work is displayed in chronological order, depicted by an icon specific to its type and a horizontal block representing its duration. Details of an episode can be viewed by clicking on the block representing it, which pops-up more detailed information about the episode (dates, description), as well as access to edit and deletion functions.

Van Labeke et al. (2009, 2011) describe two of the search facilities provided by the system, one to search for “people like me” and another to find recommendations of “what to do next”. The latter is illustrated in Figure 16 where we see one of the recommended timelines being displayed beneath the user’s own, for easy visual comparison.

<<FIGURE 16 here >>

Figure 16. The “What Next” user interface. Episodes in the recommended (lower) timeline that match episodes in the user’s own (upper) timeline are shown in blue; episodes that start

after all blue episodes are shown in orange – these are deemed by the system to be relevant for this user as they occur after the matching episodes, and thus represent possible choices that the user may be inspired to explore further for their future learning and career development; episodes that occur earlier than all blue episodes or have no matches within the user's own timeline, are shown in grey.

The technical basis for both the "people like me" and the "what to do next" facilities is the users' annotation of their episodes with concepts drawn from the L4All ontology. The availability of this metadata allows similarity algorithms to be used to compare the user's own timeline with all other timelines (see Van Labeke et al. 2009, 2011, Poulovassilis et al. 2012).

In terms of the four components of AIED systems, the domain knowledge of the system is represented in the L4All ontology. Its 'student model' is the timeline that is created and annotated by the user. The pedagogical model is encapsulated in the "people like me" and "what to do next" functionalities offered to users, to help them explore possible future learning and career options and to plan and reflect on their lifelong learning. Again, the system's interface is key to the user's growing knowledge and confidence as they interact with their peers' timelines.

Evidence of Effectiveness

Over the last 35 years or so a great variety of AIED systems have been developed and evaluated in the laboratory and in schools, colleges and universities. Such evaluations have compared AIED systems against more traditional teaching methods, such as whole class teaching by an individual human teacher, one-to-one tutoring by a human teacher, or the use of a text-book on its own, or some blend of these and other teaching methods. The evaluations have usually looked at either comparative learning gains or the study time needed to reach some mastery criterion. To date, there have been few comparative evaluations of big data-enabled interventions (although see Ferguson et al. 2016 for a recent review of the use of Learning Analytics in education), so our scope in this chapter is AIED systems in general.

There has now been a sufficient body of work published to allow a number of meta-reviews to be created. These are reviews that look at a large number of individual evaluations and try to draw general conclusions, typically by computing an average of the comparative learning gains. This chapter focuses on the meta-review evaluations of AIED systems, comparing them either against one-to-one human tutoring or against whole class teaching by a single instructor. These include using an AIED system blended into whole class teaching as compared to simply whole class teaching by an individual teacher.

Table 1 shows the results from six meta-reviews as well as a large study that evaluated a single AIED system, the Cognitive Algebra Tutor described earlier, in a large number of schools in the USA. Some meta-reviews involved more than one kind of comparison. In the table positive effect sizes and percentile rank changes indicate that the AIED system produced better learning outcomes than the human method it was compared with. Negative effect sizes and percentile rank changes indicate the opposite.

Table 1. Six meta-reviews and a large scale study adapted from du Boulay (2016)

<< TABLE 1 about here >>

Column 2 in the table shows the kind of comparison being made and column 3 shows the number of such comparisons collected in that meta-review. Column 4 shows the mean effect size across the comparisons (bigger indicates a larger effect). Column 5 shows the standard error of the mean effect size (smaller indicates reduced disparity between the individual studies examined). The effect size measures how far the mean of the experimental group is from the mean of the control group measured in terms of the standard deviation of the control group scores, with effects above 0.4 "worth having" (Hattie, 2008). Note that although most of the effect sizes in Table 1 are positive, some are negative. A negative effect size indicates that the AIED systems produced worse learning outcomes than human tutoring (see rows 1, 2, 6 and 9).

Column 6 shows the equivalent increase/decrease in percentile rank as a result of using the AIED system in the comparison. For example, a change in percentile rank of 10 would mean that on average students using the AIED system would have increased their ranking by 10 percentage points compared to the control group.

The final study, Row 11 in Table 1, was different from the others. This was an evaluation of a single system, The Cognitive Tutor for Algebra (this is a successor to the Pittsburgh Algebra Tutor, see earlier) across a large number of matched pairs of schools in the USA (Pane et al., 2014). The comparisons were between schools that included the AIED system "blended" into their algebra teaching *versus* schools that carried on teaching in a traditional manner. There were four comparisons, see Row 9 of Table 1. The study was conducted over two years in both middle schools and high schools.

The most positive result (an effect size of 0.21) was in the second year of the study in the high schools. The other results were more mixed, but still broadly positive with respect to the utility of AIED system used in a blended fashion.

The overall picture from the meta-reviews is positive with respect to the use of AIED systems compared to whole class teaching. The weighted mean (weighted by number of comparisons) of the effect size from the meta-reviews is 0.47 (see row 10 of Table 1). When AIED systems have been compared to one-to-one human teaching they do not do so well, with a weighted mean of -0.19 (see row 9 of Table 1). This is hardly surprising at this stage of the development of such systems.

The authors of these meta-reviews made the following comments. For example, VanLehn found that AIED systems were, within the limitations of his review, 'just as effective as adult, one-to-one tutoring for increasing learning gains in STEM topics' (VanLehn, 2011, p. 214). While Nesbit et al. found, 'a significant advantage of ITS over teacher-led classroom instruction and non-ITS computer-based instruction' (Nesbit et al., 2014, p. 99). Likewise, Kulik and Fletcher concluded that,

> This meta-analysis shows that ITSs can be very effective instructional tools . . . Developers of ITSs long ago set out to improve on the success of CAI tutoring and to match the success of human tutoring. Our results suggest that ITS developers have already met both of these goals.
>
> (Kulik and Fletcher, 2016, p. 67).

Steenbergen-Hu and Cooper found that,

Deleted: 9

Deleted: 10

Deleted: argued that

Deleted: '

Deleted: '

Formatted: Indent: First line: 0", Line spacing: 1.5 lines

> ITS have demonstrated their ability to outperform many [human led] instructional methods or learning activities in facilitating college level students' learning of a wide range of subjects, although they are not as effective as human tutors. ITS appear to have a more pronounced effect on college-level learners than on K-12 students.
>
> (Steenbergen-Hu and Cooper, 2014, p. 344).

Two points are of special note. First, there is some double counting in that there is some overlap in the papers that the meta-reviews examined. Second, most of the comparisons concerned STEM subjects, as it is these kinds of domain that are best suited to the development of AIED systems (see the earlier section on What is Artificial Intelligence in Education).

Conclusions

This chapter has described, on the one hand, the nature of AIED systems in terms of their four major components and provided examples of such systems and, on the other hand, examples of some of the opportunities that Big Data brings to children's and adults' learning.

Deleted: one side

Deleted: side

We have argued that AIED systems have been sufficiently evaluated through a number of meta-reviews to demonstrate their effectiveness as part of blended learning in STEM subjects. These meta-reviews have shown that AIED systems do rather better than conventional classroom teaching, though a bit worse than one-to-one human tutoring. We have also made the case that the provision of personalised and adaptive feedback to students can enhance students’ engagement, motivation and self-confidence, leading to improved learning outcomes.

No argument in favour of replacing teachers by AIED systems has been offered or is implied by these results. Human teachers are still the essential factor in any classroom to take control of the overall learning trajectory of the students, to motivate the unmotivated and the demotivated and to answer queries from students, particularly those who do not exactly know what it is that they do not understand. Indeed it is acknowledged that some students may not have the study skills and reasoning powers to take advantage of such systems (Biswas, Segedy, & Bunchongchit, 2016) and so need support beyond what the system itself can provide.

However, we do argue that provision of individual automated feedback to students for common occurrences can free up time for the teacher to formulate more complex or nuanced support for students, particularly in larger classes. In addition, the rich data generated by such systems are being used to design visualisation and notification tools for the teacher. Such tools can increase the teacher's awareness of the classroom state and of individual students' progress on the task set, and hence help the teacher in supporting both individual students and the class as a whole.

Despite these opportunities, there are still many challenges to fully exploiting the potential of AIED and big data in education. For example, this chapter has not addressed the issue of the cost-effectiveness of such systems. They are time-consuming to create and, for them to be effective, multi-disciplinary teams of pedagogical experts, learning designers and computer scientists must work together to understand what information is useful to whom and in what learning contexts, and to design computational techniques for detecting or inferring such information and generating appropriate feedback for users. Also, many AIED systems cover only a small part of the curriculum. However, both these factors are changing for the better, as authoring tools emerge that allow more cost-effective design of intelligent systems without the need for specialist computing expertise. Moreover, as AIED systems are increasingly used, the data they collect can be analysed so as to design improvements to them.

There are also wider socio-technical challenges. As we have already argued, the design of AIED systems and of methods for collecting, managing, integrating, analysing and visualising their big data needs to be both practically feasible and pedagogically meaningful. Moreover, it requires teachers, learners and other stakeholders to be sufficiently empowered, involved, and trained to make effective use of these systems and the information that can be obtained from them. Lastly, agreements need to be framed between different educational stakeholders so as to allow sharing of learning-related data for the benefit of learners. This exposes numerous ethical questions, such as: What data about an individual should require their explicit consent in order to be collected, combined, used and shared? Likewise, what knowledge should be allowed to be inferred from the data, and what uses of such knowledge should be permitted? What levels of information and explanation are needed so that individuals can make fully informed decisions? What are appropriate anonymization, privacy, authorisation and preservation policies for both data and inferred knowledge in different contexts of usage? From the opposite perspective, what inequalities may be faced by students (for example from less advantaged backgrounds) whose learning-related data is *not* being collected and used to offer them enhanced educational opportunities? Some of these ethical issues are explored by Manca et al. (2016)

focussing specifically on the information being gathered by large-scale web-based learning platforms and social media applications.

In our own research projects, we aim to address these challenges through close collaboration between researchers, developers, students, teachers, and other stakeholders. We draw on multi-disciplinary expertise from across computer science, the learning sciences and education. In the absence as yet of sufficiently broad and robust ethical frameworks, we address ethical challenges on a project-by-project basis, fully engaging with our institutions' processes for ethical review of research, and also aiming to inform and shape these anticipating an era where Artificial Intelligence and Big Data are pervasive.

Deleted: going forwards into

What the research says

- Globally, students from poorer backgrounds perform worse than students from richer backgrounds. Artificial Intelligence in Education (AIED) and Big Data in Education are technologies that can help with this problem.
- AIED might be used both in classrooms (to support teachers much as human classroom assistants support teachers) and at home (enabling students to build on what they have learned in the classroom while being given personalised support).
- The AIED known as Intelligent Tutoring Systems (ITS) have been shown in many classroom studies to be more effective than group tuition but not (yet) quite as effective as individual tuition.
- AIED is most effective when it is working in tandem with the classroom teacher.

- Digital educational systems, such as VLEs, ITS and ELEs, are generating data at unprecedented scale and speed. Computational techniques can extract value from such data.
- The two complementary fields of Learning Analytics and Educational Data Mining are devising multiple computational techniques to gather, analyse and visualise data about learners and processes of learning.
- Visualisation and other tools can help teachers integrate AIED systems in the classroom, increase their awareness and, ultimately, free up time to provide nuanced support to students, beyond what is possible through the system.
- There remains a range of challenges – pedagogical, technical, socio-technical and ethical – that need to be addressed by multidisciplinary teams.

Acknowledgements: This chapter is based on earlier papers by du Boulay (2016) and Poulovassilis (2016). We would like to thank Margarita Steinberg for her comments on parts of this chapter. We also thank the MiGen, LIBE, iTalk2learn and GRASIDELE team members for our very fruitful collaborations, invaluable help and insights.

Table 1

	1. Meta-review	**2. Comparison**	**3. No. of Compar-isons**	**4. Mean Effect Size**	**5. Stand-ard Error**	**6. Approx. percentile rank change**
1	VanLehn (2011)	AIED system vs one-to-one human tutoring	10	-0.21	0.19	-8
2	Ma, Adesope, Nesbit, and Liu (2014)	AIED system *vs* one-to-one human tutoring	5	-0.11	0.10	-4
3		AIED system *vs* "large group human instruction"	66	0.44	0.05	17
4	Nesbit, Adesope, Liu, and Ma (2014) *examining systems to teach computer science*	AIED system *vs* "teacher led group instruction"	11	0.67	0.09	25
5	Kulik and Fletcher (2016)	AIED system *vs* "conventional classes"	63	0.65	0.07	24
6	Steenbergen-Hu and Cooper (2014)	AIED system vs one-to-one human tutoring	3	-0.25	0.24	-10
7		AIED system vs "traditional	16	0.37	0.07	15

	examining college level use	classroom instruction”				
8	Steenbergen-Hu and Cooper (2013) *examining school level use*	AIED or CAI system vs “traditional classroom instruction”	26	0.09	0.01	3
9	**Overall weighted mean**	**AIED system vs one-to-one human tutoring**	**18**	**-0.19**		**18**
10	**Overall weighted mean**	**AIED system vs conventional classes**	**182**	**0.47**		**-7**

11	Pane, Griffin, McCaffrey, and Karam (2014) *Examining the Algebra Tutor*	Blended learning including a AIED system *vs* traditional classroom instruction	147 schools	-0.1	0.10	-4
				0.21	0.10	8
				0.01	0.11	0
				0.19	0.14	7

References

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A. et al. (2010). A view of cloud computing. *Communications of the ACM*, *53*(4), 50-58.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., and Koedinger, K. (2008). 'Why students engage in "gaming the system" behaviours in interactive learning environments'. *Journal of Interactive Learning Research*, 19(2), 185-224.

Biswas, G., Segedy, J. R., and Bunchongchit, K. (2016). 'From design to implementation to practice in a learning by teaching system: Betty's brain'. *International Journal of Artificial Intelligence in Education*, 26(1), 350-364.

Cattell, R. (2011). Scalable SQL and NoSQL data stores. *ACM SIGMOD Record*, *39*(4), 12-27.

Conroy, M., and Rothstein, R. (2013). What do international tests really show about U.S. student performance? Economic Policy Institute.

Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, *51*(1), 107-113.

D'Mello, S. K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys (CSUR)*, *47*(3), 43.

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 120-129). ACM.

du Boulay, B. (2016). 'Artificial intelligence as an effective classroom assistant'. *IEEE Intelligent Systems*. 31(6) 76-81.

EMC Education Services. (2015). Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. John Wiley & Sons.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5-6), 304-317.

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., & Vuorikari, R. 2016. Research Evidence on the Use of Learning Analytics.

Grawemeyer, B., Holmes, W., Gutiérrez-Santos, S., Hansen, A., Loibl, K., & Mavrikis, M. (2015). Light-bulb moment?: towards adaptive presentation of feedback based on students' affective state. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 400-404). ACM.

Grawemeyer, B., Gutierrez-Santos, S., Holmes, W., Mavrikis, M., Rummel, N., Mazziotti, C. J., & Janning, R. (2015 b). Talk, tutor, explore, learn: intelligent tutoring and exploration for robust learning. In *Proceedings of the 17th International Conference on Artificial Intelligence in Education (AIED)*.

Grawemeyer, B., Mavrikis, M., Holmes, W., Gutierrez-Santos, S., Wiedmann, M., Rummel, N., (In Press). Improving engagement and enhancing learning with affect-aware feedback.

Gutierrez-Santos, S., Geraniou, E., Pearce-Lazard, D., & Poulovassilis, A. (2012). Design of teacher assistance tools in an exploratory learning environment for algebraic generalization. *IEEE Transactions on Learning Technologies*, *5*(4), 366-376.

Gutierrez-Santos, S., Mavrikis, M., Geraniou, E., & Poulovassilis, A. (In Press). Similarity-based Grouping to Support Teachers on Collaborative Activities in Exploratory Learning Environments. *IEEE Transactions on Emerging Topics in Computing.*

Hansen, A., Mavrikis, M., & Geraniou, E. (2016). Supporting teachers' technological pedagogical content knowledge of fractions through co-designing a virtual manipulative. *Journal of Mathematics Teacher Education*, *19*(2), 205–226. https://doi.org/10.1007/s10857-016-9344-0

Hattie, J. (2008). Visible Learning. 1st Edition. Routledge.

Hanushek, E. A., and Woessmann, L. (2010). The high cost of low educational performance: The long-run economic impact of improving Pisa outcomes. Paris: Organisation For Economic Co-Operation And Development.

Koedinger, K. R., and Aleven, V. (2016). 'An interview reflection on "Intelligent tutoring goes to school in the big city"'. *International Journal of Artificial Intelligence in Education*, 16(1), 13-24.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., and Mark, M. A. (1997). 'Intelligent tutoring goes to school in the big city'. *International Journal of Artificial Intelligence in Education*, 8(1), 30-43.

Kulik, J. A., and Fletcher, J. D. (2016). 'Effectiveness of intelligent tutoring systems: A meta-analytic review'. *Review of Educational Research*, 86(1), 42-78.

Leelawong, K., and Biswas, G. (2008). 'Designing learning by teaching agents: The Betty's brain system'. *International Journal of Artificial Intelligence in Education*, 18(3), 181-208.

Liyanagunawardena, T. R., Adams, A. A., and Williams, S. A. (2013). 'MOOCs: A systematic study of the published literature 2008-2012'. *International Review of Research in Open and Distance Learning*, 14(3), 202-227.

Luckin, R., Holmes, W., Griffiths, M., and Forcier, L. B. (2016). *Intelligence unleashed. An argument for AI in education*. London: Pearson.

Ma, W., Adesope, O. O., Nesbit, J. C., and Liu, Q. (2014). 'Intelligent tutoring systems and learning outcomes: A meta-analysis'. *Journal of Educational Psychology*, 106(4), 901-918.

Manca, S., Caviglione, L. and Raffaghelli, J.E., 2016. Big data for social media learning analytics: potentials and challenges. *Journal of e-Learning and Knowledge Society*, *12*(2).

Mavrikis, M., Gutierrez-Santos, S., & Poulovassilis, A. (2016). Design and evaluation of teacher assistance tools for exploratory learning environments. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 168-172). ACM.

Nesbit, J. C., Adesope, O. O., Liu, Q., and Ma, W. (2014). How Effective are Intelligent Tutoring Systems in Computer Science Education? Paper presented at the IEEE 14th International Conference on Advanced Learning Technologies (ICALT), Athens, Greece.

Noss, R., Poulovassilis, A., Geraniou, E., Gutierrez-Santos, S., Hoyles, C., Kahn, K., Magoulas, G.D., & Mavrikis, M. (2012). The design of a system to support exploratory learning of algebraic generalisation. *Computers & Education*, *59*(1), 63-81.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., and Karam, R. (2014). 'Effectiveness of Cognitive Tutor Algebra I at scale'. *Educational Evaluation and Policy Analysis*, 36(2), 127-144.

Poulovassilis, A. (2016). Big Data and Education. Technical Report BBKCS-16-01, Birkbeck, University of London, December 2016. Available at http://www.dcs.bbk.ac.uk/research/techreps/2016/bbkcs-16-01.pdf

Poulovassilis, A., Selmer, P., & Wood, P. T. (2012). Flexible querying of lifelong learner metadata. *IEEE Transactions on Learning Technologies*, *5*(2), 117-129.

Rummel, N; Mavrikis, M; Wiedmann, M; Loibl, K; Mazziotti, C; Holmes, W; Hansen, A; (2016) Combining Exploratory Learning with Structured Practice to Foster Conceptual and Procedural Fractions Knowledge. In: Looi, CK and Polman, J and Cress, U and Reimann, P, (eds.) Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016, Volume 1. (pp. pp. 58-65). International Society of the Learning Sciences (ISLS): Singapore.

Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 4-8). ACM.

Siemens, G., & d Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the*

2nd international conference on learning analytics and knowledge (pp. 252-254). ACM.

Steenbergen-Hu, S., and Cooper, H. (2013). 'A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning'. *Journal of Educational Psychology*, 105(4), 970-987.

Steenbergen-Hu, S., and Cooper, H. (2014). 'A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning'. *Journal of Educational Psychology*, 106(2), 331-347.

Van Labeke, N., Magoulas, G. D., & Poulovassilis, A. (2009). Searching for "people like me" in a lifelong learning system. In *European Conference on Technology Enhanced Learning* (pp. 106-111). Springer Berlin Heidelberg.

Van Labeke, N., Magoulas, G. D., & Poulovassilis, A. (2011). Personalised search over ligelong learners' timelines using string similarity measures. Technical Report BBKCS-11-01, Birkbeck, University of London. Available at http://www.dcs.bbk.ac.uk/research/techreps/2011/bbkcs-11-01.pdf, 2011.

VanLehn, K. (2011). 'The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221.

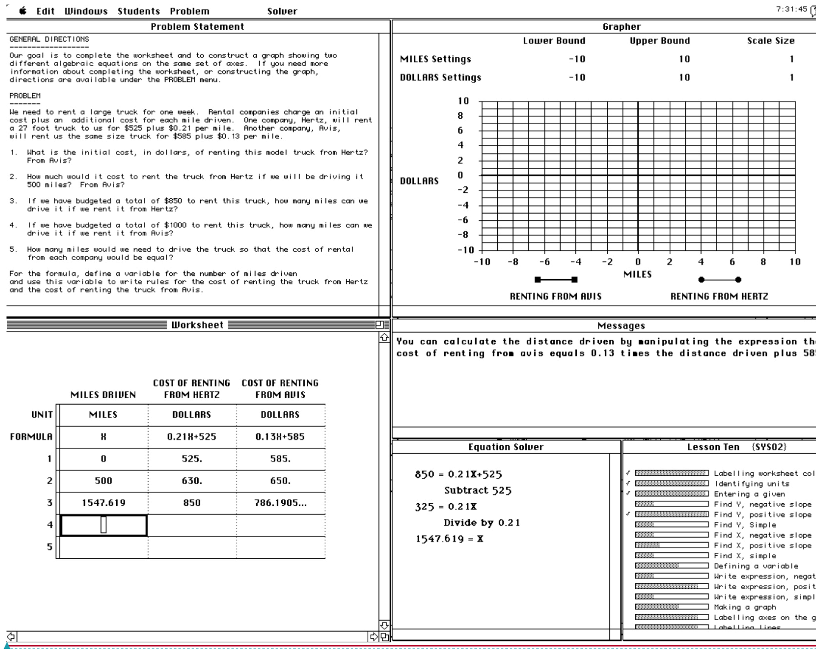

Figure 1. Interface for the Pittsburgh Algebra Tutor, taken from Koedinger et al. (1997)

Formatted: Font:+Theme Headings (Calibri), 12 pt, Not Bold

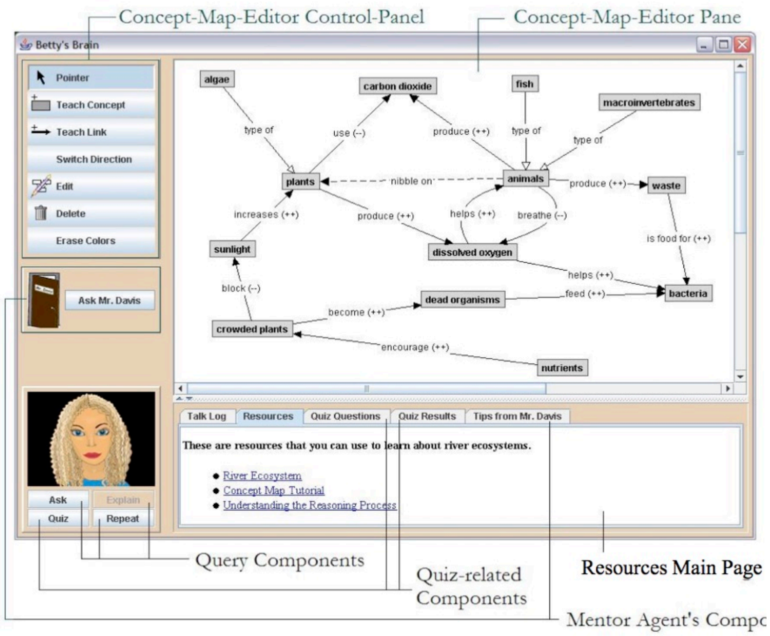

Figure 2. Betty's Brain interface, taken from Leelawong and Biswas (2008)

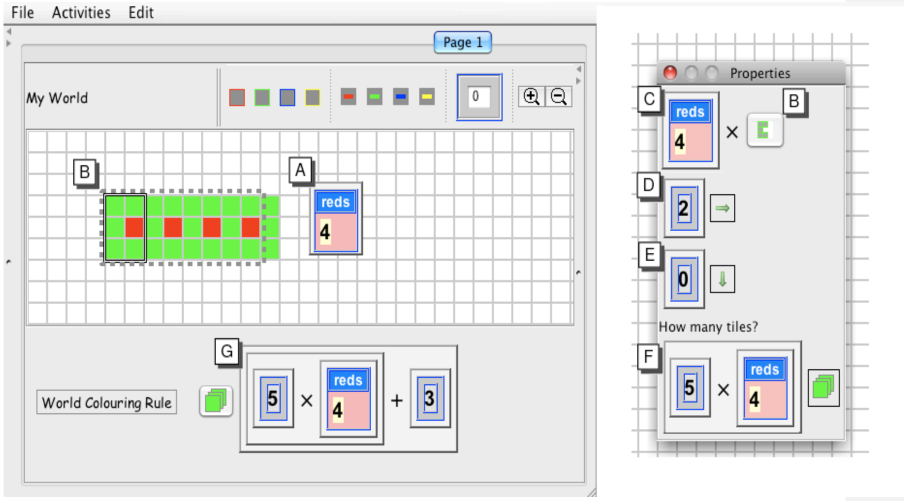

Figure 3. The eXpresser microworld. Letters highlight the main features: (A) An ‘unlocked’ number is given the name ‘reds’ and signifies the number of red (dark grey) tiles in the pattern. (B) Building block to be repeated to make a pattern. (C) Number of repetitions (in this case, the value of the variable ‘reds’). (D,E) Number of grid squares to translate B to the right and down after each repetition. (F) Units of colour required to paint the pattern. (G)

General expression that gives the total number of units of colour required to paint the whole pattern.

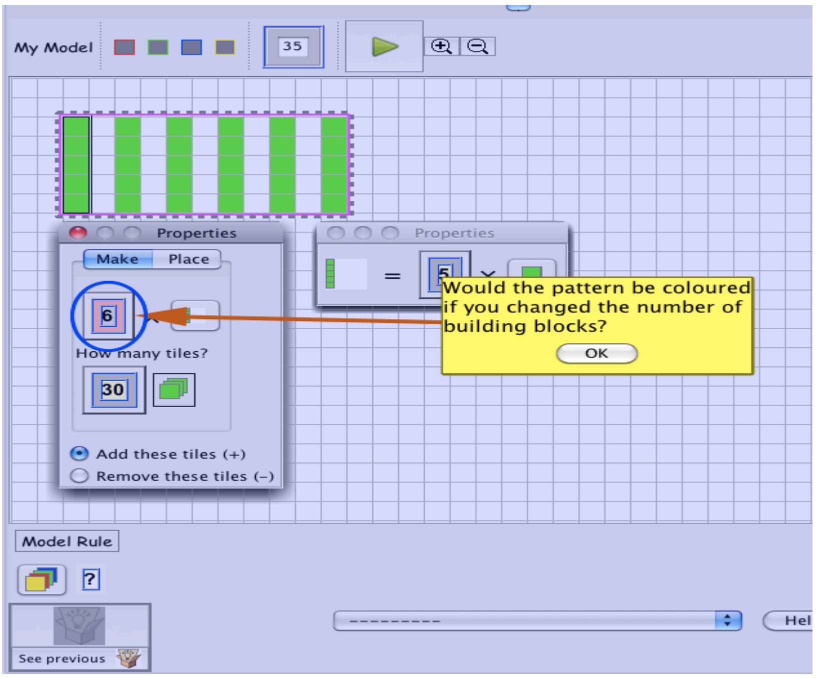

Figure 4. A ‘nudge’ from the eXpresser

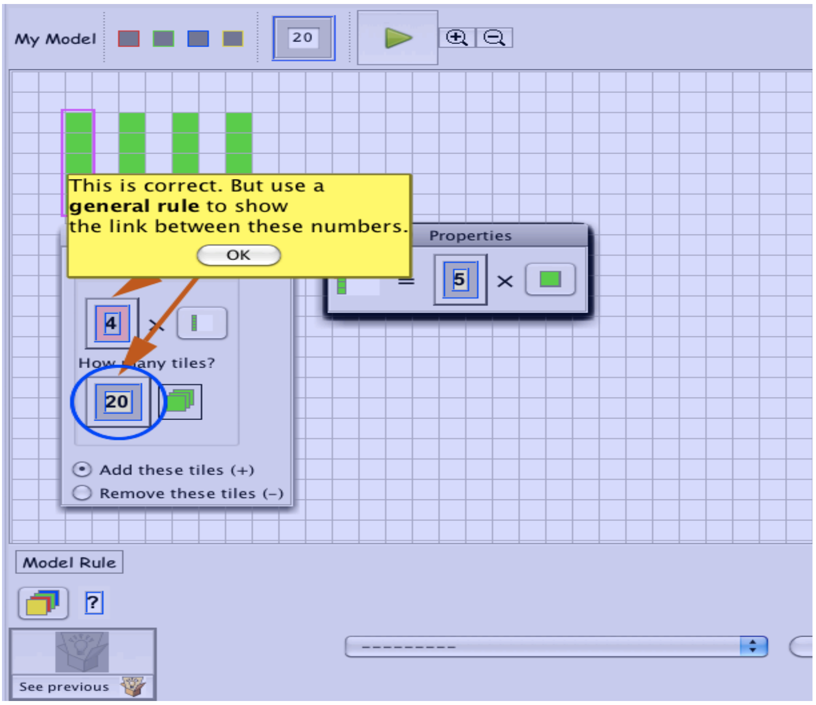

Figure 5. A message of encouragement and a stronger 'prompt' from the eXpresser

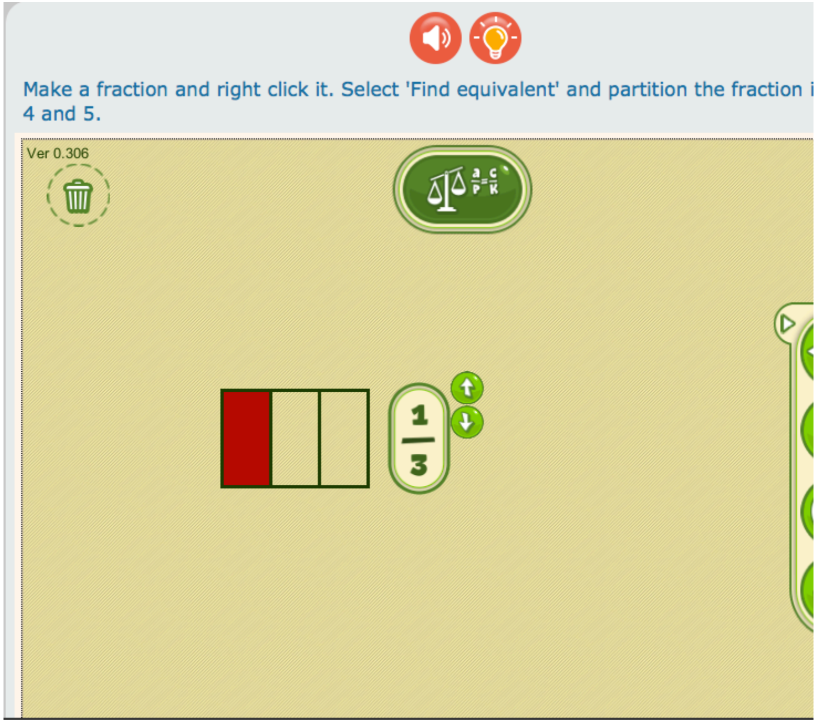

Figure 6. FractionsLab microworld, showing the availability of low-interruption feedback.

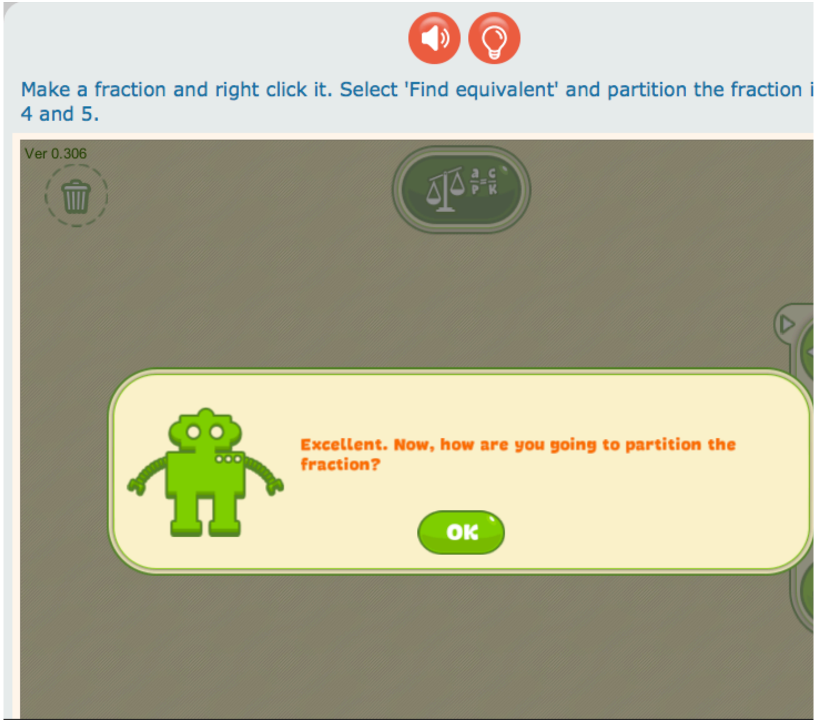

Figure 7. FractionsLab microworld, showing the elective display of low-interruption feedback.

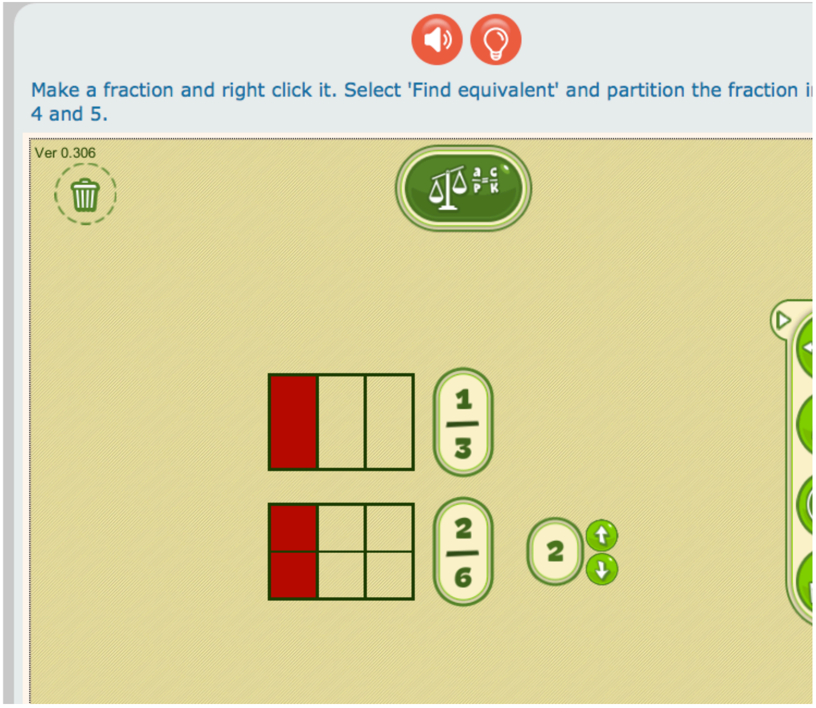

Figure 8. FractionsLab microworld, showing that the student has progressed to the next step.

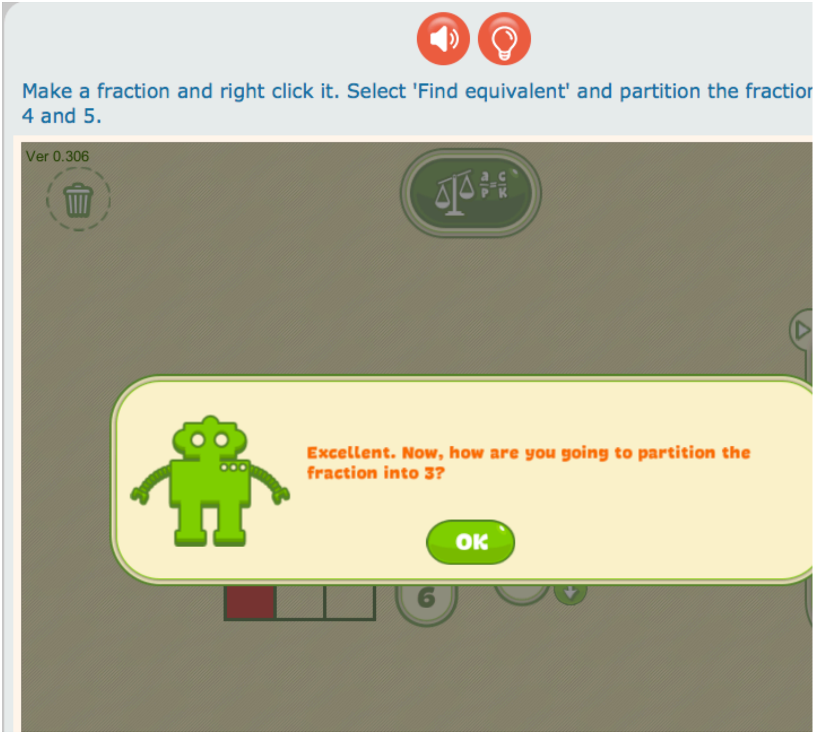

Figure 9. FractionsLab microworld, showing a message of encouragement and also a stronger prompt to guide the student towards accomplishing the subsequent step.

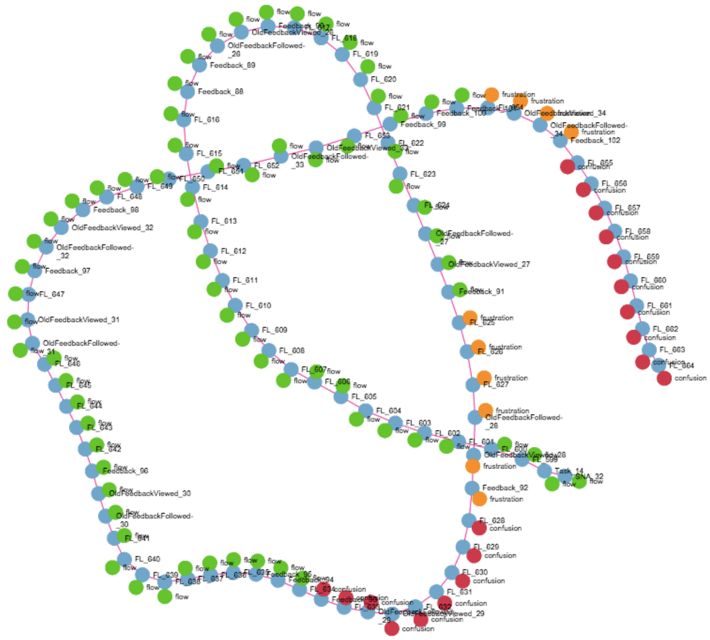

Figure 10. Graph-based modelling and visualisation of students' interactions, illustrating how a student's affective state changes between states of Engagement (green), Frustration (amber) and Confusion (red). Successive events are shown in blue and linked to each other by red edges.

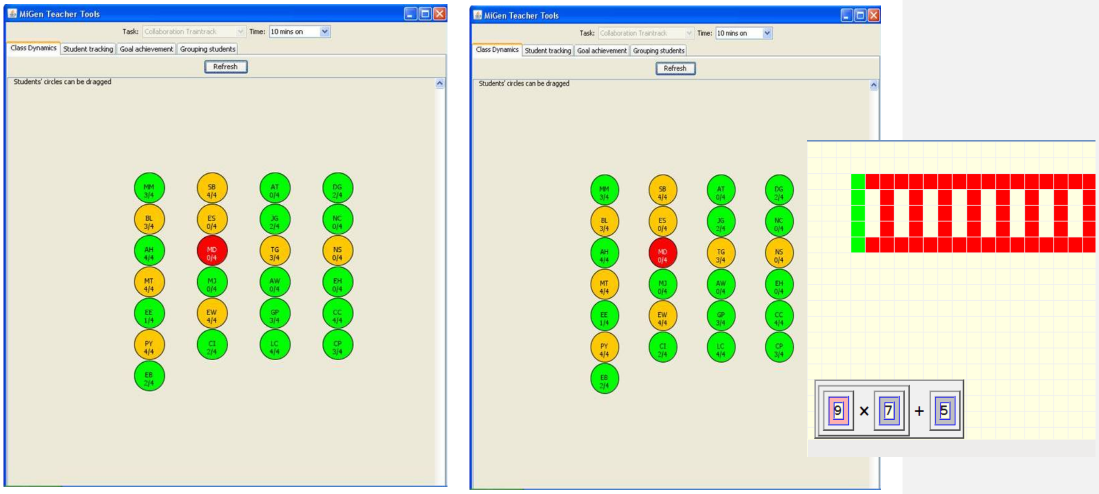

Figure 11. MiGen's Classroom Dynamics tool. On the left, a classroom with the students sitting at benches in rows. On the right, the teacher has clicked on the 'red' student to see their construction and rule on the way over to help them.

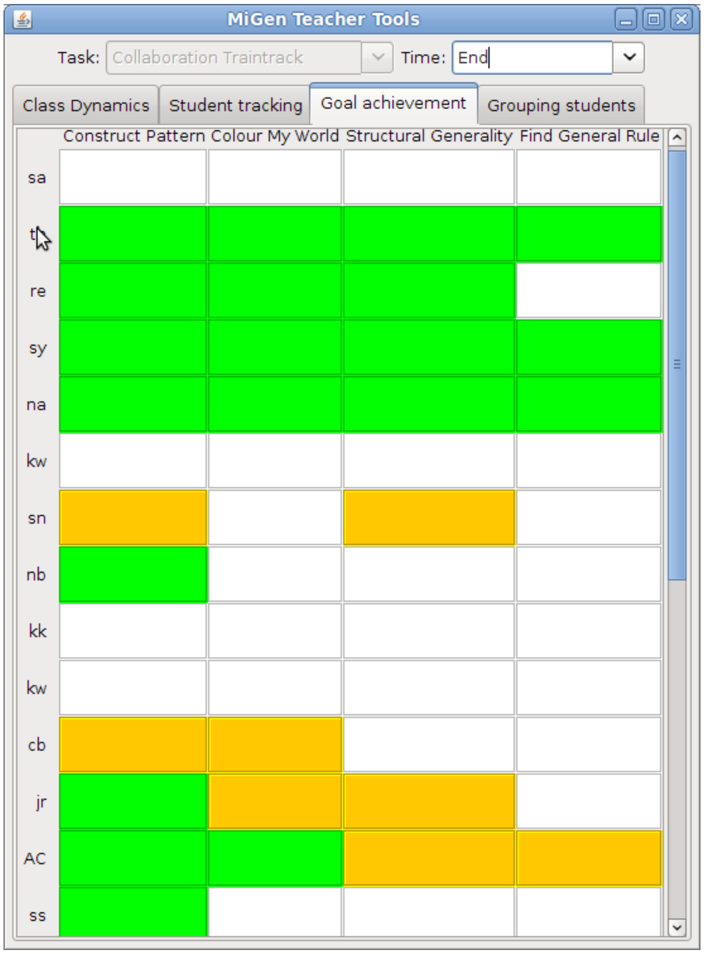

Figure 12. MiGen's Goal Achievements tool. We see that some students have achieved all or most task goals, some students have not made any progress yet, and some students are moving back and forth.

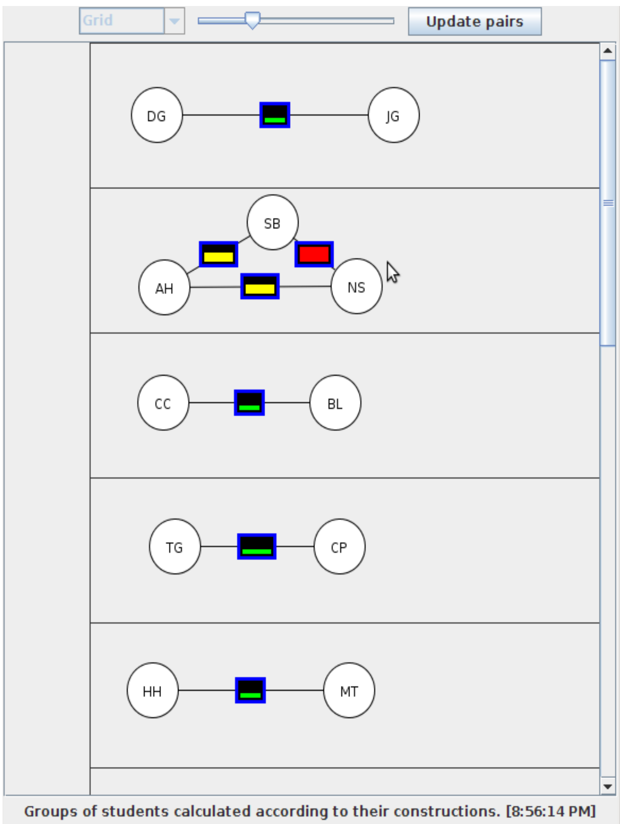
Grid
Update pairs
DG
JG
SB
AH
NS
CC
BL
TG
CP
HH
MT
Groups of students calculated according to their constructions. [8:56:14 PM]

Figure 13. MiGen’s Grouping Tool.

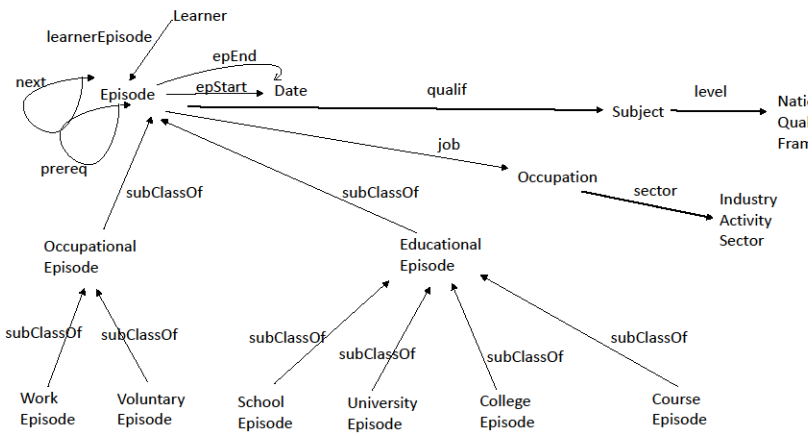

Figure 14. Fragment of the L4All ontology. Each instance of the Episode class is: linked to other episode instances by edges labelled `next' or `prereq' (indicating whether the earlier episode simply preceded, or was necessary in order to be able to proceed to, the later episode; linked either to an Occupation or to an educational qualification (Subject) by means of an edge labelled ‘job’ or ‘qualif’. Each occupation is linked to an instance of the Industry Activity Sector class by an edge labelled ‘sector’. Each qualification is linked to an instance of the National Qualification Framework (NQF) class by an edge labelled ‘level’. The Occupation, Subject, Industry Activity Sector and NQF hierarchies are drawn from standard United Kingdom occupational and educational taxonomies (see Labour Force Survey User Guide, Vol 5, http://www.ons.gov.uk/ons/guide-method/method-quality/speci_c/labour-market/labour-market-statistics/index.html).

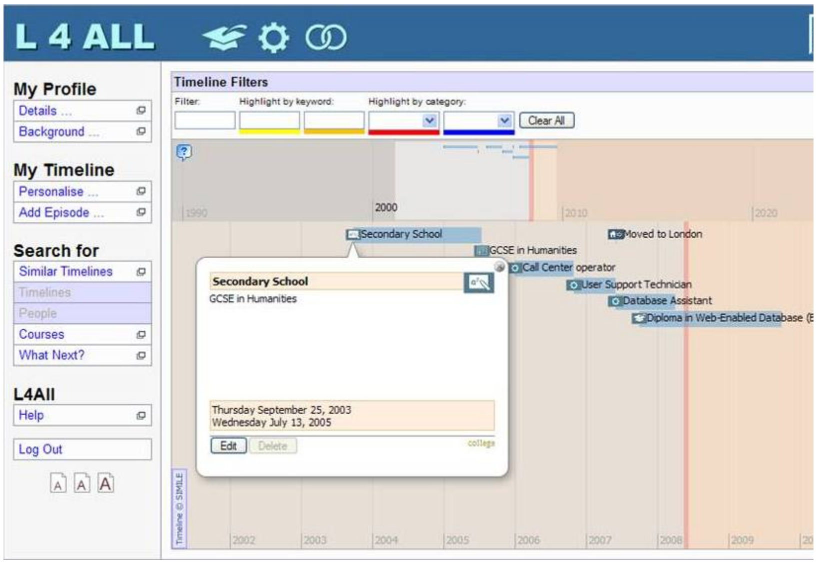

Figure 15. The main screen of the L4All system. At its centre is a visual representation of the user's timeline, and the system functionalities are organised around this. Each episode of learning or work is displayed in chronological order, depicted by an icon specific to its type and a horizontal block representing its duration. Details of an episode can be viewed by clicking on the block representing it, which pops-up more detailed information about the episode (dates, description), as well as access to edit and deletion functions.

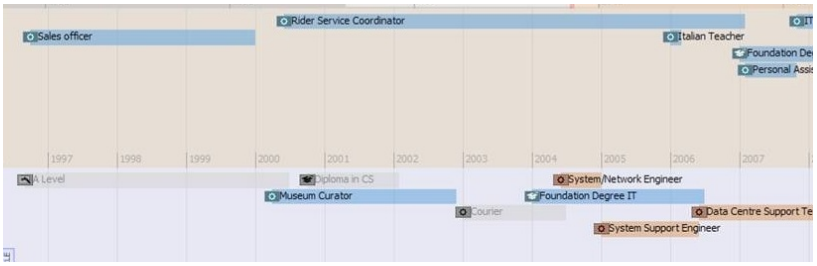

Figure 16. The “What Next” user interface. Episodes in the recommended (lower) timeline that match episodes in the user’s own (upper) timeline are shown in blue; episodes that start after all blue episodes are shown in orange – these are deemed by the system to be relevant for this user as they occur after the matching episodes, and thus represent possible choices that the user may be inspired to explore further for their future learning and career development; episodes that occur earlier than all blue episodes or have no matches within the user’s own timeline, are shown in grey.