Chapter 12
Users' Data

Trails Analysis

Christophe Choquet, Sébastien Iksal, Mark Levene and Judith Schoonenboom

Abstract With the development of Web-based distance learning environments, acquiring and analysing trails has become a very important issue for the technology-enhanced learning (TEL) community. We consider a trail (or a track or a trace) as the digital or non-digital record that learners – or more generally, the different actors within a learning session in a TEL system – leave behind. This chapter addresses the life cycle of such trails from a computer science point of view. In particular, we elaborate on the engineering and usage of the different kinds of trails by highlighting the main scientific issues raised by the trails analysis process and by presenting research findings from the Kaleidoscope Network of Excellence in this field.

Keywords Curricular activities · Design pattern · Indicator · Instructional design · Trails analysis · Trails classification

12.1 Using Trails for Supporting Curricular Activities

12.1.1 The Relevance of Trails Analysis

It is the very nature of distance learning and teaching applications to provide a multitude of user data which can be used for perceiving and understanding the users' activity. Analysing this data will enable learners to reflect on their activity for the purpose of self-assessing their progress and measuring the suitability of their curriculum with respect to their learning objectives. Analysing this data will also enable tutors and teachers to regulate the learning session and/or evaluate the learners' activities. On the other hand, designers need session feedback for evaluation purposes as well as for improving the quality of their learning environments.

With the development of Web-based distance learning environments, acquiring and analysing trails has become a very important issue for the technology-enhanced

C. Choquet (✉)
LIUM, University of Maine, Le Mans, France
e-mail: christophe.choquet@univ-lemans.fr

N. Balacheff et al. (eds.), *Technology-Enhanced Learning*,
DOI 10.1007/978-1-4020-9827-7_12,

learning (TEL) community. We consider a trail (or a track or a trace) as the digital or non-digital record that learners, or more generally, the different actors within a learning session in a TEL system, leave behind. There are different kinds of trails depending on (1) the nature of the pedagogical situation, (2) the possibilities for interaction and the learning context and (3) the purpose for which the system is used (i.e. knowledge acquisition, assessment, session regulation, reflection, system reengineering).

This chapter considers the life cycle of trails from a computer science perspective. In particular, the chapter will elaborate on the engineering and usage of different kinds of trails by highlighting the main scientific issues raised by the trails analysis process as well as presenting research findings of the Kaleidoscope Network of Excellence in this field.

12.1.2 Trails Analysis in Tracking Problems: An Example

The trails analysis process and its relevance will be demonstrated in two tracking problems (i.e. how trails can be used to support those involved in education in solving educational design problems) that could occur when designing an educational system where learners must perform learning activities by navigating through digital learning materials (called here “learning objects” or “Los”).

Both tracking problems involve a system that supports students in studying English grammar, as described by Turcsányi-Szabó, Kaszás, and Pluhár (2004). The system supports the students in selecting Web resources (i.e. LOs) that match their topic of interest and proficiency level by advising them about the trail to follow. The system was created on the basis of materials in five free, good quality, intermediate level English language grammar teaching portals. In the system, the materials are organised into *topical units* that contain explanations, exercises and self-assessments for each exercise, and a topical unit test for the entire unit. The topical units are classified into three difficulty levels (beginner/re-starter, pre-intermediate, intermediate). Students’ results are stored on several occasions to enable personalisation. At the very beginning, students’ knowledge level is determined using a general grammar test with more than 50 items. After taking this test the student is presented with a *knowledge map* (see Fig. 12.1).

The knowledge map shows a set of trails containing the topics that have already been mastered by the student (i.e. mastered nodes, where all test items corresponding to these sub-topics have been successfully completed) and do not need to be revisited. Mastered nodes appear in grey, indicating an existing route, but not a recommended route. Those topics where the student’s knowledge is unsatisfactory are distinguished using a different colour (black), indicating an unmastered node. This colouring indicates to the student those topics in the map where deficiencies were recorded (as well as providing the outline of an advised route for visiting the topics of the material). When visiting topical units, students can also take a test on that specific unit, from which the system receives feedback. After the student completes any topical test successfully, the colouring of that particular topical unit node in the map turns white to indicate that it is now *semi-mastered*.

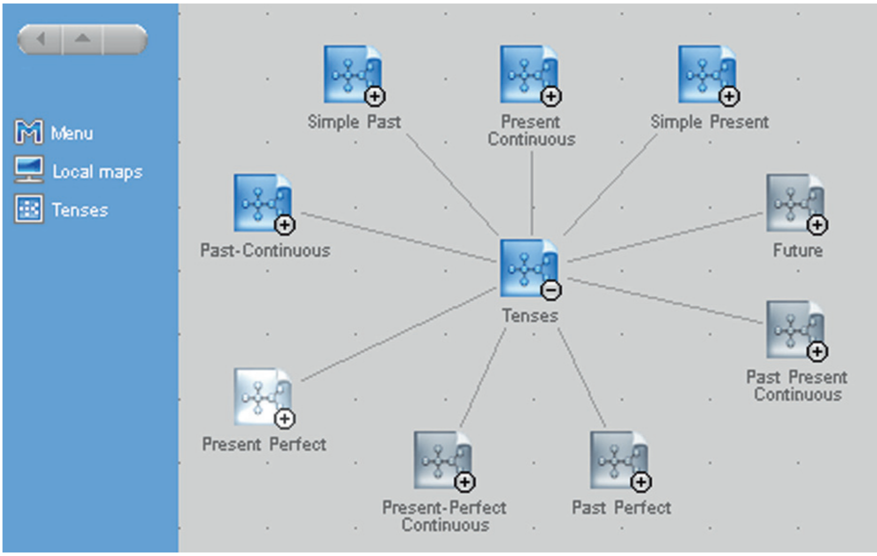

Fig. 12.1 Knowledge map of a semi-mastered node (adapted from Turcsányi-Szabó et al., 2004, p. 33, by permission of the authors)

Both the TRAILS (personalised and collaborative TRAILS of digital and non-digital learning objects) project and the DPULS (Design Patterns for recording and analysing Usage of Learning Systems) project – actions of the Kaleidoscope Network of Excellence to address tracking problems in TEL – have provided means to better understand and support the use of trails in solving tracking problems. In the above example, the main tracking problem can be formulated as follows: "how can students be presented with those explanations and exercises on English grammar that best match their interests and current proficiency level at each point in time during the learning process". But in such an educational system, several tracking problems usually occur simultaneously. An example of such an additional problem would be as follows: "how can the system detect when the learner is just playing around with LOs instead of visiting them conscientiously, in order to alert the learner and/or the tutor, or to adapt the interaction with the system".

12.1.3 Trails Analysis as a Process of Deriving Indicators

Analysis of trails for solving such multi-faceted tracking problems consists of the deduction of meaningful indicators, based on the existence of trails data acquired during a learning session, which will assist the actor (human or not) in his/her task using the outcome of the analysis. An indicator highlights a relation between a trail and a significant envisaged event, which could be interpreted as characterising the activity of the actors within a learning session. With the help of Fig. 11.1, an indicator can be defined from two different points of view:

1. An indicator “stems from what is important” from a pedagogical and/or psychological point of view.
2. An indicator is a numeric or symbolic representation of what is important, a significant (structured or not) datum which supports the analysis of the learning activity.

This chapter focuses on the second, technical and computational definition of an indicator. The process of trails analysis includes first a modelling phase, where the main design questions are as follows:

- What is required for acquiring and understanding a trail?
- What are the indicators, and how can they be deduced from trails?
- What are the technical requirements for acquiring the data needed to shape a trail and to define or evaluate indicators?

The acquisition of data constitutes the second phase, where the main design questions are the following:

- What kind of acquisition techniques can coexist within a TEL system?
- Is the acquisition obtrusive (e.g., tests and questionnaires) or not?
- When does the acquisition need to occur (before the session for profiling a learner, during the session, or after the session, as for instance a debriefing)?

The analysis itself is the third phase. Based on raw data (e.g. directly collected from the learning environment; see Fig. 11.1 and the narrative for details), how can one – human or machine – extract and construct indicators and thus characterise the activity of the actors in a learning session?

Finally, delivering the results of the analysis to the end user is the last phase:

- Who is the end user (the learner, the tutor or the designer)?
- What tasks are supported by the analysis fed back to the user?
- What kinds of representations of trails are well suited for the analysis of trail results?

12.1.4 The Issues of Trails Analysis in this Chapter

Figure 12.2 illustrates how the analysis of trails can support curricular activities. Through their learning activities, learners create trails, and feeding these trails back to the learner in a suitable format can help learners, as well as tutors, teachers or designers, in reflecting on their activities.

Figure 12.2 also provides the basis for the questions that this chapter will address. These are the following:

For what types of learning and at what stages in the learning process can the use of trails provide support? The TRAILS project developed the trails cycle of learning, including the stages of planning, navigation, learning activities and analysis/reflection, which will be presented in Section 12.2.

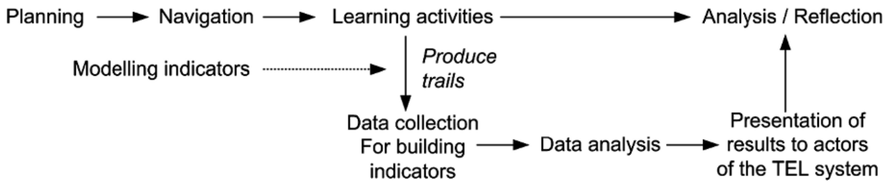

Fig. 12.2 Supporting curricular activities through the analysis of trails

The second question is related to the more low-level questions of whom we are supporting using trails, which actors performing which activities at which level of education. Actors, activities, levels and other relevant factors can vary widely from context to context and the TRAILS project produced a classification describing different contexts (see Section 12.3).

From Section 12.4 onwards, the focus is on the analysis of trails. Section 12.4 presents a typology of the kind of data that can be collected for trails analysis. This typology helps the designers of a TEL system model how an indicator could be derived from other data.

With the help of the example of the second tracking problem presented above, Section 12.5 presents a Design Pattern approach for supporting the designer who is facing concrete tracking problems for which trails analysis can be beneficial.

The last section discusses future directions we believe the community should follow that will shape the Kaleidoscope vision on trails.

The trails cycle of learning, the trails classification, the typology of trails analysis data and the design patterns are the results of two Kaleidoscope projects, DPULS and TRAILS. The TRAILS project formalised the concept of a trail and proposed a trails analysis process and a trails taxonomy based on a trail's use and content. The DPULS project focused on the issue of how to support the designer of a TEL system during the modelling phase. The project has proposed an open set of design patterns which provide instructional designers, teachers and tutors with improved and possibly reusable solutions to support them in solving recurrent problems when tracking students' activity. Together these projects have initiated a comprehensive approach for developing, capitalising, sharing and using trails analysis techniques which could be very valuable for the TEL community.

12.2 Learning Types and Stages Supported by Trails Analysis

The analysis of trails can provide support for various types of learning, which were identified and described in the TRAILS project and the subsequent edited book documenting the main results from the project (Schoonenboom, Levene, Heller, Keenoy, & Turcsányi-Szabó, 2007). The first type of learning supported is *navigational learning* (Peterson & Levene, 2003). With the advent of digital learning materials in general and the mass usage and growth of the Internet in particular, the volume of learning materials available to the learner has multiplied. As a consequence both learners and teachers must navigate through learning materials. The second type of learning supported is *personalised learning*. Because of the increasing amount of learning materials, both teachers and learners need to create their own trails for navigating through an overload of materials. Learners need to be able to follow the trails that best match their needs and capacities. (Note that these two types of learning are not mutually exclusive. Personalised learning often, but not always, involves navigation.)

From a trails perspective, navigation occupies a very prominent position in learning. Trails are created when learners navigate through learning materials. For this reason, members of the TRAILS project concluded that traditional views of the learning curriculum were no longer sufficient, as they did not do justice to this prominent position of navigation. From a traditional curriculum perspective, there are three curricula: the intended curriculum envisioned in curriculum documents, the curriculum-in-action, interpreted by its users and consisting of the actual process of teaching and learning; and the attained curriculum consisting of the learning experiences as perceived by the learners and the resulting learning outcomes (van den Akker, 2003). From these curricula, three major phases in the learning process can be deduced: learning starts with *planning* the intended curriculum. In the

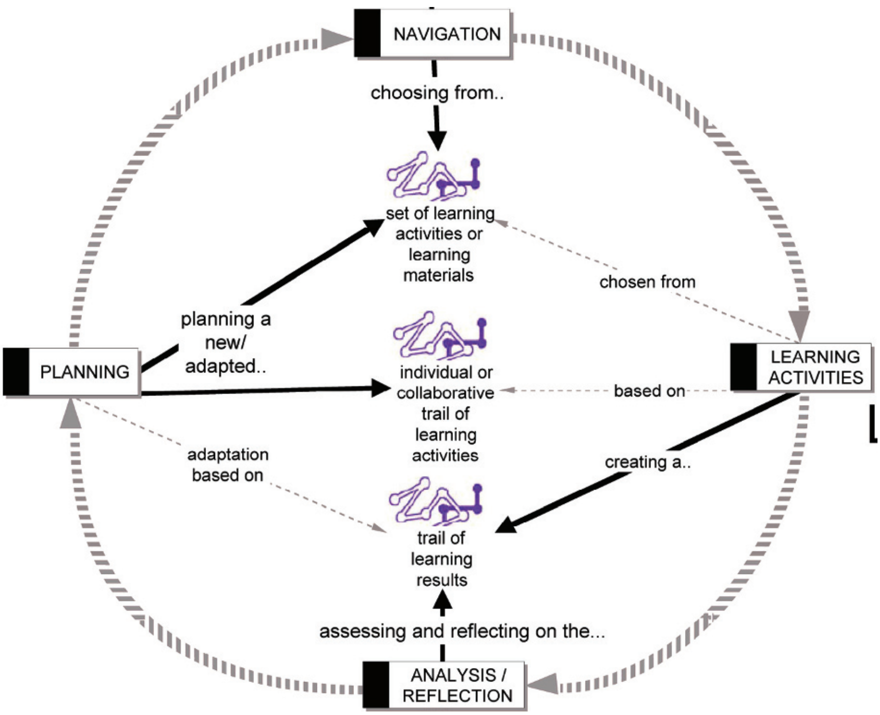

Fig. 12.3 The trails cycle of learning showing the creation and use of trails in each of the four stages of learning

curriculum-in-action, *learning activities* are performed. The learner interacts with the learning materials producing a trail of learning results. The learning results of the attained curriculum can be *assessed* or consulted for *reflection*.

The problem with this tripartite division of the learning process is that navigation does not have a place. Navigation in the sense of finding one's own path through learning materials is not a part of the planning phase, as it is not planned. But it occurs before performing learning activities, as it is the search for learning activities that fit one's needs. In fact, *navigation* constitutes a phase of its own, which is located in between planning and performing learning activities. Thus, the TRAILS project developed a four-part division of the learning process, which they called the "TRAILS cycle of learning" (see the introductory chapter of Schoonenboom et al., 2007).

In Fig. 12.3, learning starts with *planning*, which consists of either planning a fixed trail of learning activities or selecting a set of learning activities the learner can choose from. In the latter case, planning is followed by a *navigation* phase in which the learner chooses from the set of learning activities. Next, *learning activities* are performed in which the learner interacts with the learning materials and thereby produces a trail of learning results. These learning results can be *assessed* or consulted for *reflection*. After that, a new cycle can be started, which may be wholly or partly based on the trail of learning results.

12.3 Actors and Activities Supported by Trails Analysis: A Classification of Trails

As stated above, the actors and the activities to be supported in solving tracking problems can be very diverse. The TRAILS project developed a classification of trails for determining whom and what exactly to support. This section describes this classification in brief; a more detailed description is reported in Schoonenboom et al. (2007).

To start with a concrete example, Table 12.1 shows how the trails classification can be applied to the English grammar teaching case described in Section 12.1 of this chapter. This classification of trails is based on the curriculum classification of van den Akker (2003) and the preliminary taxonomy of trails of Keenoy and Levene (2004). Six elements can be used in classifying trails, as shown in Table 12.1.

The *stage in the trails cycle of learning* refers to the four stages of planning, navigation, learning activities and reflection. The *level of the trail* indicates the part of curriculum that the trail as a whole covers. Not surprisingly, the level can be very diverse, covering the whole range from a small part of a lesson, a lesson, a task, a module, a course, to an entire degree or school curriculum. Various *actors* can be involved in the learning process, all of whom might need support; actors include learners, teachers, researchers, managers and designers. *Activities supported* include, but are not limited to, such diverse activities as goal setting, timing of activities, locating activities, choosing from relevant learning activities, choosing learning materials and resources, assigning activities to specific learner roles, analysis and

Table 12.1 Classification of a trail used for the adaptive navigation support of pupils exploring materials on English grammar tracking problem

Classifier	Value
Stage in the trails cycle of learning	Navigation
Level of the trail	Course
Actors involved	Learner
Activities supported	Choosing from relevant learning activities
Units of the trail	Learning activities, materials and resources
Rationale for ordering the elements of the trail	The trail is a selection of elements that belong to a specific topic and level of proficiency

reflection. The *units that make up the trail* category refer to the type of curriculum elements that are connected within the trail. Four types were identified within the TRAILS project: aims and objectives, learning activities, materials and resources and learning outcomes. Finally, there is always a *rationale for ordering the elements of the trail*, a reason why the elements of the trail are put together in the way that they are. With respect to the rationale, Keenoy and Levene (2004) make a top-level distinction between temporal links and conceptual links. Temporal links allow LOs within a sequence to interact. Temporal links could be as follows:

- Hyperlinks between Web pages or pages in any hypertextual learning environment.
- Physical adjacency, such as exhibits in a museum being next to one another, or one chapter of a book following another.

In extending this definition beyond the scope of learning objects, a temporal link can also be formed, as in the case of learning objectives that must be mastered in a certain sequence. One frequently occurring rationale for the ordering of the elements of a temporal trail is that the ordering is the path that has been followed or is to be followed by a learner or by a group of learners (e.g. learners with specific roles).

Conceptual links, according to Keenoy and Levene (2004), reflect connections between LOs based on their content. Conceptual links could indicate the following:

- When LO_a covers prerequisite knowledge for being able to interact with LO_b.
- When one LO contains similar content to another LO, such as
  - LOs on the same topic.
  - LOs teaching the same competency.
  - LOs with the same learning objective.
  - LOs containing examples demonstrating the same principle.

12.4 Trails Analysis: From Data to Indicators

When the intended use of a trail is the analysis of (or the reflection on) the user's activity (the learner, the tutor, etc.) in a TEL system, the designer should a priori model the trail and its components. Most of the existing systems build a trail by first manually or automatically collecting heterogeneous raw data (Champin, Prié, & Mille, 2003; Jermann, Soller, & Mühlenbrock, 2001), and then structuring them to establish "learning indicators" that are meaningful for a specific analysis purpose. The methods used for establishing these learning indicators are multiple but should be explicitly modelled, especially when the indicator is automatically inferred or calculated (see, for instance, Laflaquière, Settouti, Prié, & Mille (2006) and Mostow (2004) for details on data transformation for a tracking purpose). Explicit modelling is also needed when data are collected from heterogeneous sources such as manual, audio or digital records that need to be combined (see Marty, Héraud, France, & Carron (2007) for examples).

12.4.1 The Users' Data Typology

Based on the existing literature and on Kaleidoscope project results, essentially established by the ICALTS (Interaction & Collaboration Analysis' supporting Teachers & Students' self-regulation) and TRAILS projects, the DPULS project proposed a user's data typology, where types are defined in accordance with the intended use of the data and their provenance (see Fig. 12.4).

The *primary data* are not calculated or elaborated with the help of other data or knowledge. They could be raw data, additional data or subjective data.

Raw data are recorded before, during or after the learning session by the learning environment, for instance in a log file recorded by the system, a videotape of the learner recorded during the session, a questionnaire acquired before or after the session, or the sets of posts in a forum.

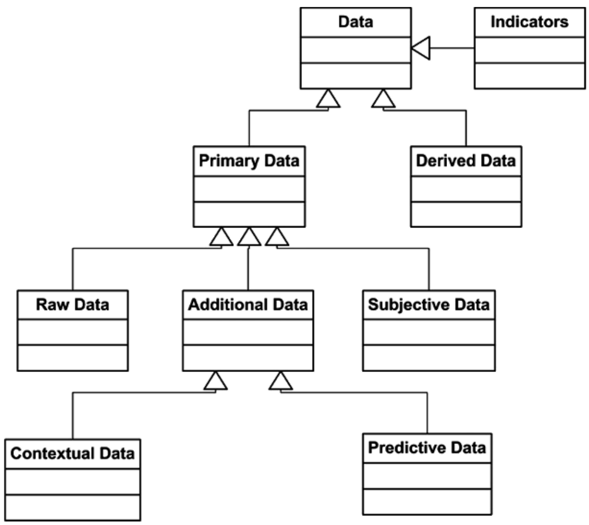

Fig. 12.4 The user's data typology of DPULS

The *additional data* type describes data that are linked to the learning situation and could be involved in the usage analysis. Additional data can be further classified as contextual or predictive. *Contextual data* can be picked from the learning materials, such as the metadata of a learning object, the formally planned scenario for the pedagogical situation or any information which is directly accessible. *Predictive data* on the other hand refer to the outcomes provided by the learning session actors (learners, tutors, teachers). This kind of data is mainly produced by the learners and is intended to be assessed, but could also come in the form of a tutor's report on the activity of a learner or the use of learning material.

The *subjective data* type refers to primary data which are a priori defined by an actor in the learning situation (a learner, a teacher, a tutor) or part of the analysis (output by an analyst, a designer, any learning staff member who is involved or concerned in the analysis).

The *derived data* are calculated or inferred from *primary data* or other *derived data*.

The *indicators* type refers to *derived* or *primary data* having pedagogical significance. Thus, an indicator is always relevant to a pedagogical context, and it is always defined for at least one useful purpose (e.g. validating the learning materials, assessing, reflecting, regulating). Based on the computationally oriented definition of an interaction analysis indicator (see Chapter 11), the DPULS project adopted the following definition: an *indicator* is a variable, calculated or inferred with the help of collected users' data, that describes something related to the quality of the interaction, the activity and/or the learning process of actors acting in the frame of a social context formed via the technology-based learning environment. The next section provides a short example in which this typology is used for modelling two indicators.

12.4.2 Playing Around with Learning Objects Example

In the context of individual learning, learners often play around with the system – especially at the beginning of a learning activity – by rapidly browsing the learning objects. It could be pertinent to detect this behaviour for regulation or a learner's purposes of reflection. The DPULS project proposed a generic solution to this tracking problem that is based on two indicators. The dependencies between data involved in their calculation are shown in Fig. 12.5.

It is assumed that every LO is described by LOM (Learning Object Metadata; LOM, 2007) or, at least, the typical learning time for each LO (the time needed for learners to correctly use the LO; see LOM specification for a formal definition of the "EducationalTypicalLearningTime" resource descriptor) can be estimated. It is also assumed that the sequence and time spent by a learner in consulting a LO can be recorded, for instance in log files.

The solution consists in recording for each LO the login time (date of connection in a log file) and the logout time (date of logout in a log file), and calculating the effective duration of use (the difference between logout and login times). The

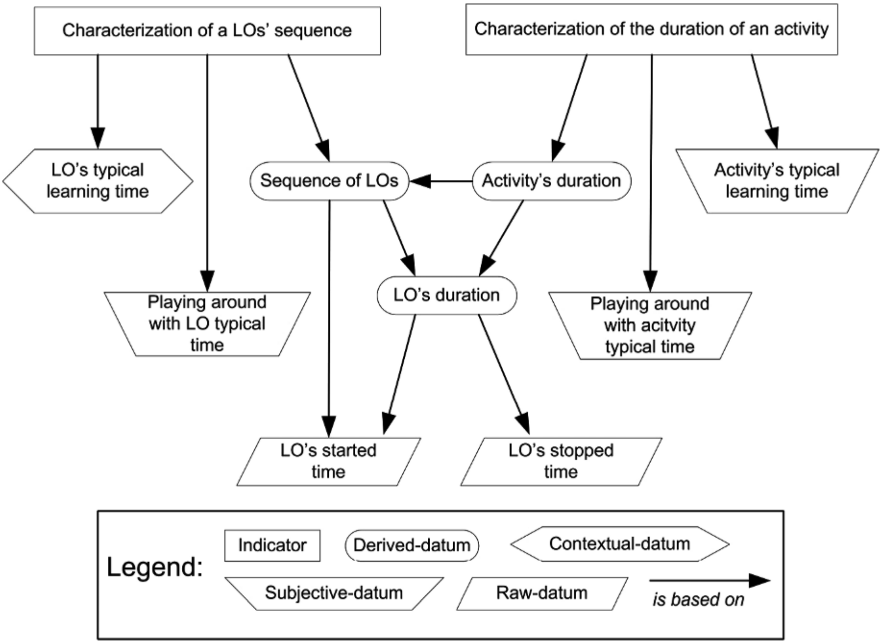

Fig. 12.5 Dependencies between data involved in "playing around" detection

sequence of LOs attempted by a learner could be labelled as "non significant" if the time spent on each LO in the sequence is less than the relevant Playing Around Typical Learning Time (a fraction of the typical learning time of a LO, typically 10%, which defines the duration under which a LO cannot really be consulted, but only browsed).

If such a sequence is detected at the beginning of a learning activity, one could presume that the learner is playing around with the system. If its effectiveness is proven by experimentation, this kind of solution for a tracking problem could be abstracted and capitalised and shared to support educational systems designers. The DPULS project chose a Design Pattern approach for doing this.

12.5 The Designer's Support for Modelling the Use of a Trail: A Design Pattern Approach

The DPULS project focused on the know-how required for acquiring, modelling and analysing trails. The main aim of this project was to address the following question:

> Considering usage analysis with a specific aim (e.g., a learning context, a pedagogical analysis purpose, or a considered trail's end user – the learner, the tutor, the designer, etc.), what are the indicators one needs to collect, how could one analyse the usage, and what existing techniques or tools are well suited for this usage analysis purpose?

The DPULS project aimed at capitalising on the know-how of trails analysis by providing the TEL community with a structured set of Design Patterns that allow for sharing the users' data acquisition and analysis expertise. Each pattern addresses an acquisition or analysis problem in an identified context and proposes a concrete solution for it.

Each pattern is formatted within a template, and the Design Pattern language constituted by the entire set is accessible through a Web browser referenced by the TeLearn Open Archive (http://telearn.noe-kaleidoscope.org/). This set is open source and should be considered as a bootstrap for creating a wider set of patterns, fed and used by the TEL community.

The Design Patterns have a common framework for their representation. This framework is composed of a pattern template (displayed in Fig. 12.6), vocabularies for possible values of its fields and types of links which could be drawn between patterns.

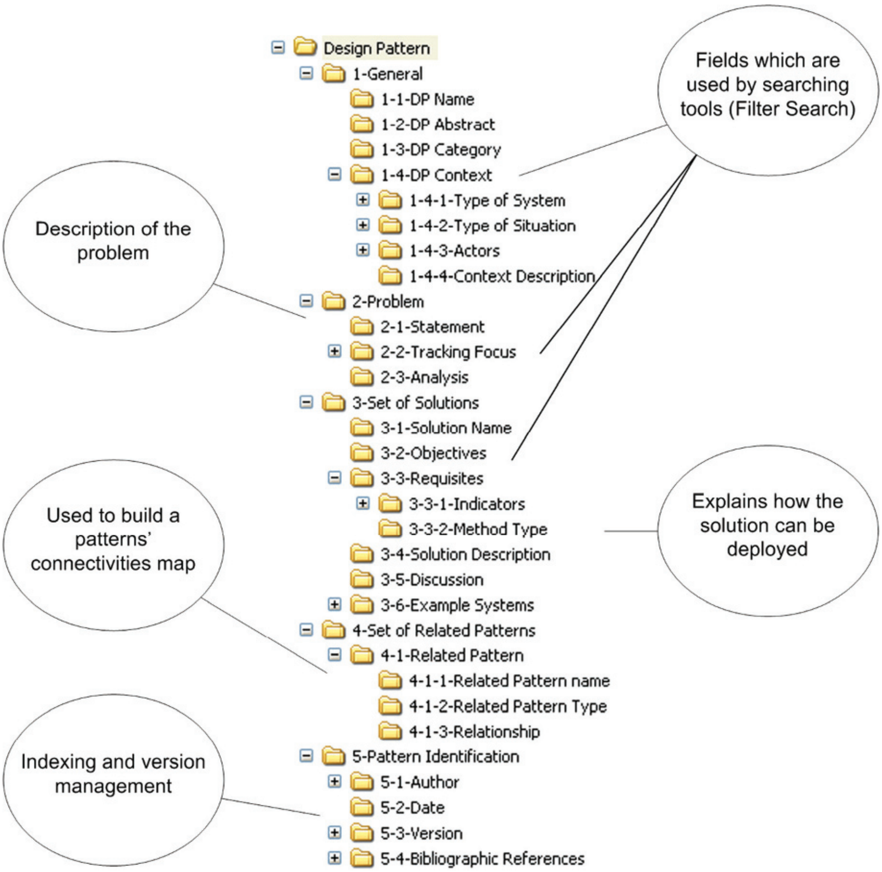

Fig. 12.6 The DPULS design pattern template

Table 12.2 General section of the design pattern "Detecting Playing Around with the System"

General			
	Name	Detecting Playing Around with the System	
	Abstract	This pattern provides an approach to detect playing around with the system by a student, at the beginning of an activity.	
	Category	Course usage	
	Context	Type of system	LMS (Learning Management System)
		Type of situation	Each student is alone in front of the machine. The teacher intervenes only when he is solicited.
		Actors	– Instructional designer – Tutor
		General Description	You can record and analyse tracks of resources used. It can be very valuable for describing the resources with metadata.

The Design Pattern template is structured in five sections. The first one, *General*, is composed of fields that concern the *Identification* of the pattern. Each of the 40 Design Patterns defined by the project is indexed with a *Category* describing the learning *Context* where the pattern is relevant. The four categories tackled by the project are "Collaboration", "Learner's Assessment", "Material Validation" and "Tutoring/Regulation of Learning". The *General* section of the Design Pattern "Detecting Playing Around with the System" is presented in Table 12.2.

The second section deals with the *Description of the Usage Analysis Problem* that is addressed by the pattern. This section indexes the pattern by means of the *Tracking Focus* which helps to determine the kind of tracking addressed by the pattern: the Actors' Behaviour, the Actor's Performance, the System, the Contents, the Resources or the Tasks. The *Problem* section of the Design Pattern "Detecting Playing Around with the System" is presented in Table 12.3.

The third section details the *Solution* proposed in the pattern for tackling the problem. A synthesised version of the *Solution* section of the Design Pattern "Detecting Playing Around with the System" is presented in Table 12.4. The fourth section references the links drawn to other DPULS or external existing patterns. These links are relationships between patterns such as "More General",

Table 12.3 Problem section of the design pattern "Detecting Playing Around with the System"

Problem		
	Statement	You want to know if the student plays around with the system.
	Tracking Focus	Actor's behaviour/performance
	Analysis	At the beginning of an activity, when the learner discovers the learning environment, he could play around with it, starting the LOs without really engaging in the activity. It could be problematic if it is not detected: the activity is not really engaged in and tutor's and system's reactions could need to take this into account, especially for the user's profile evolution and for assessment.

Table 12.4 Solution section of the design pattern "Detecting Playing Around with the System"

Solution			
	Solution name	LOs' sequence characterisation	
	Requisites	Indicators	Characterisation of a LOs sequence Characterisation of the duration of an activity
		Methods	Analysing the sequence of LOs consulted by the learner (semi-automatic/automatic)
	Description	*[See Section 12.4 of this chapter]*	
	Discussion	This solution is facilitated when the system records log files and when designers have described each LO with metadata.	
	Examples		

"More Specific", "Part Of", "Has As A Part", "Can Use", "Can Be Used", "Similar", "Incompatible", "Temporal Successor" and "Temporal Predecessor". And lastly, the fifth section contains documentation about the pattern, such as its authors, the date of its creation and its version.

A complex design problem may require a large number of inter-linked patterns to solve it. Individual patterns do not stand alone, and the connectivity between patterns plays an important role in achieving a system design that meets the design goals and objectives for a complex problem (Deng, Kemp, & Todd, 2005).

DPULS design patterns follow the recommendation of Meszaros and Doble (1997) and are named by a "Noun Phrase Name" referring to the result implied by the name of the patterns. The problem and solution summary are in the abstract field to help the reader find the right solution.

The DPULS Design Patterns Browser is used for navigating inside the set of patterns, and, in fact, for sharing them. The browser contains all functionalities needed to manage, publish and share design patterns.

12.6 Discussion and Scientific Issues

The DPULS and TRAILS projects within Kaleidoscope have provided several results regarding the formulation of what trails are and how they can be useful in an educational setting. The trails structure and the analysis process have both been shaped in these two projects that we have discussed. We have also defined common vocabularies for naming the different types of trails, as well as all the data with which they are constructed.

The community now needs to embrace automated support for the trails analysis process. In order to realise this, two important additional issues need to be considered: (1) standardisation, which will allow us to capitalise, share and reuse existing and well-known techniques and (2) development of support tools for all phases of the trails analysis process.

Defining standards is crucial if a scientific community wants to spread its results and to foster wider research and experimentation in its field of study. If we take the example of learning design with the recent specifications proposals, such as IMS Learning Design (IMS-LD, 2007), SCORM (SCORM, 2007) or LOM (LOM, 2007), we notice that, even if – or perhaps because – these proposals are not perfect, they have caused the research community to enter into a debate arguing for or against them, thus catalysing the research effort.

Standards aim to enable data sharing and interoperability of tools. In the field of user's data analysis, the following issues should be addressed in working towards standardisation:

1. *Enabling data sharing*. Each trail collected and each indicator constructed should be expressed so that it could be shared. The DPULS and TRAILS projects proposed a classification of trails, a user's data typology and some vocabularies that could form the basis of common formats. Some specifications have already been proposed, such as the common format proposed in Chapter 11 of this book for representing data and allowing their analysis by a variety of analysis tools. The Usage Tracking Language (Choquet & Iksal, 2006, 2007) is also an example of such a specification. This language is proposed for modelling user's data collected by different TEL environments and indicators constructed by different analysis tools in a unified format. Of course, these research outcomes need to be tested on a wide scale and improved through experience, but we think they constitute a fruitful approach for further work to define a standard for enabling data sharing.
2. *Allowing interoperability for analysis tools*. Addressing this issue would allow the community to define a common repository of analysis tools that could be used in different learning systems. Moreover, with interoperable tools one could combine these tools in order to define a new one. Here again, the interaction analysis projects described in Chapter 11 have proposed some solutions. We should also mention the "Track-Based System" approach (e.g. Laflaquière et al., 2006) that proposes a framework architecture in which collection, transformation, visualisation and query systems could be combined.

Since the beginning of research on technology-enhanced learning, a vast number of techniques and tools have been proposed for modelling, collecting, analysing and visualising users' data. Most of these were defined for a specific purpose, in a specific context, and only a few have been studied from an engineering point of view. When it matures, each scientific discipline must consider the possibilities for how to develop engineering methods and processes for spreading its results and engaging the research and practice community in rational and concerted growth. Some research teams are now engaged in this approach as, for instance, the LISTEN project (Mostow & Beck, 2006), and we think that further thought needs to be given to trails analysis techniques in order to better support the whole process of users' data analysis. When a tool, a technique, a model or a language is proposed for supporting an activity in this process, it should be studied from an engineering point of view. For facilitating reuse, these proposals should be characterised by answering the following questions:

- What is its general purpose? Is it a support for modelling, collecting, analysing or visualising trails? Is it done for a reengineering purpose, for assessing a learner's knowledge, for reflecting on the user's activity, for helping to regulate the activity?
- What is its application field and its expected results? We must develop a method for evaluating the quality of our proposals: to systematically test them in different contexts, to study their limits and their potentials closely. For instance, Beck (2007) has chosen this approach for analysing the knowledge-tracing model. Even if he takes his distances with this model, he points out the expected results when this model is used, while depicting its limitations, as the possibility of local and multiple global maxima – see Corbett and Anderson (1995) to learn more about this model.
- From a technological point of view, what are its reuse possibilities? Does it require a specific technology or a specific data format?
- From an educational point of view, for which learning framework is it well suited?

Addressing all of these issues will stimulate cooperation and collaboration within the trails analysis research community, as well as sharing of its results with the communities that concentrate on more practical issues.

In conclusion, the research community for users' data analysis should engage itself in a process where each effort is analysed from an engineering point of view, in order to bring the theory into practice. We think that consideration of the question of engineering of trails analysis is useful as such, but also and mainly because it would enhance the research in this field: working on users' data engineering will require us to define the proper place and the roles of theoretic proposals in the user's data analysis process, and we believe that it could enhance their quality.

Acknowledgments We would like to thank all of the collaborators in the DPULS and TRAILS projects for their contribution to the scientific results presented here.

References

van den Akker, J., (2003). Curriculum perspectives: An introduction. In J. van den Akker, W. Kuiper & U. Hameyer, (Eds.), *Curriculum landscapes and trends* (pp. 1–10). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Beck, J. E. (2007). Difficulties in inferring student knowledge from observations (and why you should care). In: C. Heiner, N. Heffernan & T. Barnes (Eds.), *Educational Data Mining, Supplementary Proceedings AIED'07* (pp. 21–30). Amsterdam: IOS Press.

Champin, P. -A., Prié, Y., & Mille, A. (2003). MUSETTE: Modeling uses and tasks for tracing experience. In K. Ashley & D. G. Bridge (Eds.), *Case-based Reasoning Research and Development: Proceedings of the 5th International Conference on Case-Based Reasoning, ICCBR 2003* (pp. 279–286). Berlin, Germany: Springer.

Choquet, C., & Iksal, S. (2006). Usage tracking language: A meta-language for modeling tracks in TEL systems. In: J. Filipe, M. Helfert & B. Shishkov (Eds.), *Proceedings of the First International Conference on Software and Data Technologies, ICSOFT'06* (pp. 133–138). Berlin, Germany: Springer.

Choquet, C., & Iksal, S. (2007). Modeling tracks for the model driven reengineering of a TEL system. *Journal of Interactive Learning Research, 18*, 161–184.

Corbett, A., & Anderson, J. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253–278.

Deng, J., Kemp, E., & Todd, E. G. (2005). Managing UI pattern collections. In B. Plimmer & R. Amor (Eds.), *Proceedings of the 6th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Making CHI Natural* (pp. 31–38). New York: ACM press.

IMS-LD (2007). *Learning design specification*. Retrieved December 20, 2007, from http://www.imsglobal.org/learningdesign/index.html.

Jermann, P., Soller, A., & Mühlenbrock, M. (2001). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. In: P. Dillenbourg, A. Eurelings & K. Hakkarainen (Eds.), European perspectives on computer-supported collaborative learning. *Proceedings of the first European Conference on Computer-Supported Collaborative Learning* (pp. 324–331). Maastricht, The Netherlands: University of Maastricht.

Keenoy, K., & Levene, M. (2004). A taxonomy of trails of learning objects. In J. Schoonenboom, M. Turcsányi-Szabó, P. Kaszás, Z. Pluhár, A. Jones, C. Blake et al. (Eds.), *Trails of digital and non-digital LOs*. (pp. 97–106) (Tech. Rep. No. D-21-01-01). [n.p.]: Kaleidoscope Network of Excellence. Retrieved December 20, 2007, from http://www.dcs.bbk.ac.uk/trails/docs/D22-02-01-F.pdf

Laflaquière, J., Settouti, L. S., Prié, Y., & Mille, A. (2006). Trace-based system framework for experience management and engineering. In G. Bogdan, J. H. Robert, & C. J. Lakhmi (Eds.), *Knowledge-based Intelligent Information and Engineering Systems: Proceeding of the 10th international conference, KES 2006* (pp. 1171–1178). New York: Springer.

LOM (2007). *IEEE standards association*. Retrieved December 20, 2007, from http://www.ieee.org/web/standards/home/index.html.

Marty, J. C., Héraud, J. -M., France, L., & Carron, T. (2007). Matching the performed activity on an educational platform with a recommended pedagogical scenario: A multi source approach. *Journal of Interactive Learning Research, 18*, 267–283.

Meszaros, G., & Doble, J. (1997). A pattern language for pattern writing. In R. C. Martin, D. Riehle, F. Buschmann (Eds.), *Pattern languages of program design 3* (pp. 529–574). Reading, MA: Addison Wesley Longman.

Mostow, J. (2004). Some useful design tactics for mining ITS data. In: J. C. Lester, R. M. Vicari & F. Paraguaçu (Eds.), *Proceedings of the Analysing Student-Tutor Interaction Logs to Improve Educational Outcomes Workshop, ITS' 04* (pp. 20–28). Berlin, Germany: Springer.

Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering, 12*, 195–208.

Peterson, D., & Levene, M. (2003). Trail records and navigational learning. *London Review of Education, 1*, 207–216.

Schoonenboom, J., Levene, M., Heller, J., Keenoy, K., & Turcsányi-Szabó, M. (2007). *Trails in education: Technologies that support navigational learning*. Rotterdam, The Netherlands: Sense publishers.

SCORM (2007). *SCORM 2004 (3rd edn.)*. Retrieved December 20, 2007, from http://www.adlnet.gov/scorm/index.cfm.

Turcsányi-Szabó, M., Kaszás, P., & Pluhár, Z. (2004). Adaptive navigation support. In J. Schoonenboom, M. Turcsányi-Szabó, P. Kaszás, Z. Pluhár, A. Jones, C. Blake et al. (Eds.), *Trails of digital and non-digital LOs*. (pp. 25–55) (Tech. Rep. No. D-21-01-01). [n.p.]: Kaleidoscope Network of Excellence. Retrieved December 20, 2007, from http:// www.dcs.bbk.ac.uk/trails/docs/D22-02-01-F.pdf.