Chapter 9

USERS' DATA: COLLABORATIVE AND SOCIAL ANALYSIS

Andreas Harrer, Alejandra Martínez-Monés, and Angelique Dimitracopoulou
University of Duisburg-Essen, University of Valladolid, University of the Aegean

Abstract: Interaction analysis has been used in computer-mediated settings for approximately two decades. Its main purpose has been to understand and identify the traits of electronic communication, collaboration, and coordination. In recent years, however, its scope has expanded to include the support of students and teachers during online learning activities. This chapter documents the findings from three European projects that focussed on this novel, supportive role of interaction analysis. Following a definition of interaction analysis indicators and their computation, the use of unified data formats and interfaces is considered as means to utilise tools and data beyond their original scope or setting. Finally, the issue of visualisation of analysis results is discussed.

Key words: Interaction Analysis, tools, data formats, network visualisations

1. INTRODUCTION

Interaction Analysis has been used in computer-mediated settings for approximately two decades (Harasim, 1990). Its main usage was by researchers to understand and identify the properties of electronic communication, collaboration, and coordination and the conditions of success and failure of its usage. Some early examples of Interaction Analysis tackled patterns of email exchange in scientific discussion and resulted in diagrams showing the discussion threads evolving between the discussants, a

representation that helps to follow the flow of communication in the discussion. Recent remarkable use cases of interaction analysis can be found in the observation of social web phenomena, such as blogospheres, wikipedia etc.

Besides this analytical use in recent years the direct support of the immediate participants of learning activities, i.e. the students and the teachers, by means of automated computer based interaction analysis has been introduced. One example of a teacher tool that helps in analysing student's actions is the *Exxelant* tool (Granger *et al.* 2007 & Chapter 6 of this book), that supports the teacher in finding errors in the students' writing skills.

Computer-based Interaction Analysis is an emerging field aiming at supporting directly or indirectly the participants in every kind of technology-mediated activities. The automated analysis of participants' interactions is taken into account in two main directions that are distinguished but also complementary ones. A first direction is that the system, based on the output of interactions analysis, takes into account the profiles and the cognitive processes of individuals or collaborating groups, in order to adapt the learning environments to their own needs and preferences, or even to provide appropriate messages guiding them. A second direction is to provide information directly to the participants based on the automated interaction analysis, so as to self-regulate their decisions, actions and behaviors, supporting them in a level of awareness and metacognition (Dimitracopoulou & Bruillard 2006). In the first case, it is the system that makes the decisions (leading to the so-called Intelligent Tutoring Systems), while in the second one the locus of control is on the human actors' side. Concerning the participants in the technology-mediated activities, they can be typically the teachers, moderators and students (in a wide variety of roles in the range of learning situations), or also the members of a scientific network, such as Kaleidoscope.

The topic of computer based Interaction Analysis, intended to support participants in highly interactive learning environments, was a particularly strong research field within the Kaleidoscope Network of Excellence. Activities elaborating on this issue existed throughout the lifetime of Kaleidoscope: Starting with the projects ICALTS in 2004 (Interaction & Collaboration Analysis' supporting Teachers & Students' Self-regulation) and IA in 2005 (Interaction Analysis' supporting Teachers & Students' Self-regulation), and following with the European Research Team CAViCoLA (Computer-based Analysis and Visualisation of Collaborative Learning Activities) from the third year on, a continuous line of research activities has been present in the Network. Related to these activities, but more oriented towards an analysis of general tracks of users in learning applications (i.e.

also navigating through hypermedia learning courses), is the approach described in Chapter 10. Some of our findings and approaches have been transferred also into a community service for the Kaleidoscope network to make members of the network aware about interactions happening in the KAL Communication Infrastructure.

2. MOTIVATION

The development of appropriate interaction analysis methods is a major research topic in many areas, among these in Computer Supported Collaborative Learning. In the first place these analysis methods support the understanding of the collaborative learning processes. Yet, this does not have to be the end of the line, but it is additionally the basis for different functionalities that might be offered by an enhanced CSCL environment: students' and teachers' self-regulation can be supported by giving feedback about the current state of the learning situation (e.g. if the students have a balanced level of contributions between each others), as well as the assessment of the learning experiences (Soller, Martínez, Jermann, Mühlenbrock, 2005; Dimitracopoulou et al., 2004) by providing some metrics to the teacher, such as average level of activity, continuity etc. Both functionalities are important for the design of enhanced learning environments that go beyond the mere communication and information sharing tools that are normally provided in CSCL.

Analysis methods can be totally or partially supported by computer-based tools that provide for better and more efficient procedures. The development and practical usage of these computer-supported analysis tools is increasingly attracting researchers, but current research in the field is mainly based on unstable prototypes applied to isolated experiences. As they are not thought for general use, their usability is normally limited.

Thus we developed a joint conceptualisation of interaction analysis indicators and the process how to compute indicators based on the learning data. This initial conceptual integration laid the foundation for terminology, methodology and comparability of interaction analyses. To allow also the usage of interaction analysis in throughout evaluation processes of TEL environments, we elaborated on aspects of interoperability between different interaction analysis methods and associated tools by means of a proposed common framework (the CAViCoLA analysis process Model) for researchers supported by unified data formats and interfaces; this enables the cross-usage of analysis tools and research data beyond its initial limited scope.

Besides the computational challenges, i.e. how to formally define characteristics of interaction and collaboration (called an indicator) so that a machine can detect these automatically or interactively with researcher's input, a major aspect is also how to enable the researcher to interpret the analysis results: One possibility is to use research results from the field of information visualisation (Card et al., 1999) and provide the results to the user in diagrammatic and graphical representations that support a first-glimpse comprehension.

Among the diagrammatic representations that have been used in the area of Technology-enhanced Learning for collaborative scenarios are sociograms (Moreno, 1951), that represent persons and their relations as a graph, statistical representations (e.g. bar charts for absolute number of contributions, pie charts for relative shares of participation), and metaphorical representations, such as botanical representations (e.g. the "Wattle Tree", Kay et el., 2006), that show activity by flourishing or withering of plants.

3. TOWARDS A CONCEPTUALISATION OF INTERACTION ANALYSIS INDICATORS AND METHODS

The analysis of student interactions is usually driven by some sort of hypothesis which shall be proven or rejected by certain observation. So the first question is "What are the important questions to ask?", "What do I want to analyse?". The intention for these questions will be determined by the role or the perspective of the questioner. The answer to this question influences (indicated by the large arrow in Figure 1) the choice of indicators used to conduct further analysis. The questioner will choose indicators able to express the concept to be analysed, i.e. the choice of an indicator is influenced by the target group and will vary with the interest of the questioner.

In the area of CSCL the potential of awareness through analysis, representation, and especially visualisation of collaborative and social traits is discussed for several target groups:

- Researchers
- Teachers
- Students

Researchers can use the results and visualisations as means to support other methods of analysis (Martinéz-Mones et al. 2006; Harrer et al. 2006a), such as statistical and qualitative analysis, since triangulation research designs are common within the interdisciplinary CSCL field.

Teachers can be supported to understand the group structures in their computer supported classes and courses (e. g. for school classes using discussion forums or blogs and university courses with blended scenarios) and potentially can use this information to guide and advice the students, e. g. when participation of specific students is extraordinarily high or low.

Finally, students can use the visual feedback for self-reflection and self-regulation in reaction to the information and the style in which it is represented. Studies (cp. Sun &Vassileva 2006; Bratitsis & Dimitracopoulou 2008) have shown that the presentation of analysis results can have an substantial impact on participation and behaviour.

The choice of an indicator determines certain constraints a learning environment has to fulfill (indicated by the dashed arrow in Figure 1). So every indicator determines "what should be available to compute the indicator's values?": e.g. to measure social structures or patterns of interactions it is necessary to capture the information, who is sending messages to whom, an information that is not readily available when coordination is managed via a phone line (in contrast to a chat system). Since the learning environment usually exists beforehand and is not specifically implemented for one experiment, either it can be adapted to meet the constraints given by the indicator or it is not a candidate to be used for the analysis.

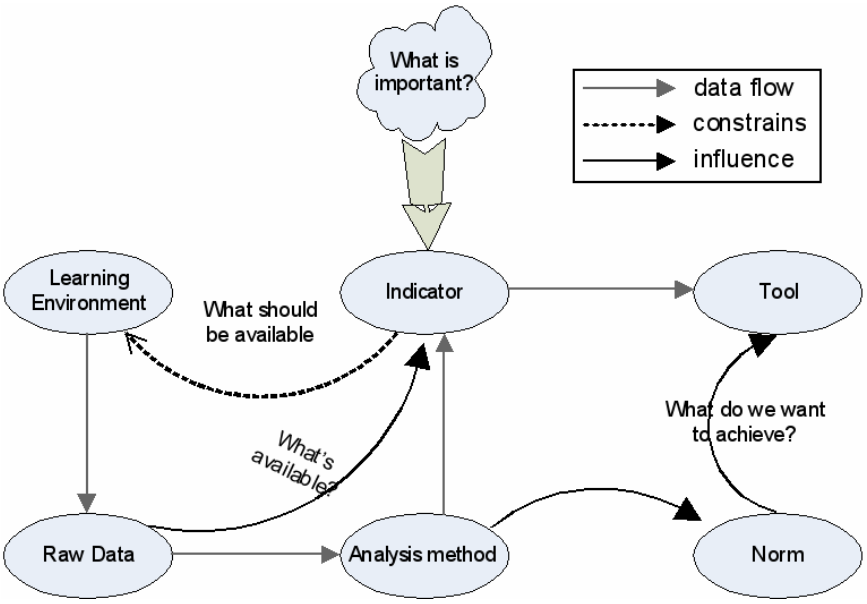


Figure #9-1. Schema of Interaction Analysis

One remark on the term indicator: There are two meanings of indicator in the diagram: First, the indicator that stems from "what is important". So for instance, "following a well-established argumentative schema" is a pedagogically and psychologically oriented indicator for learning. The analysis (e.g. if an argument is met by a counterargument) leads to a numeric or symbolic representation of these criteria which we also call indicator, but this time with a technical-computational meaning. The large arrow from “what's important” to indicator represents the pedagogical-psychological aspect of the indicator.

The learning environment is responsible for generating the raw data (e.g. log files) used in further analysis steps to compute the chosen indicators. As stated before, sometimes the data required by an indicator cannot be supplied by the used learning system causing the set of indicators to be narrowed or re-assembled. So the available raw data influences the choice of indicators as well (shown by the curved arrow).

The raw data is put into the analysis method and is processed there. The analysis produces the indicators [for a list cp. (Dimitracopoulou et al., 2004)]. In the process of analysis the raw data is usually filtered because it has to be purified and some of the analysis methods need special input formats. The method can be simple (e.g. applying filters only) or complex (e.g. plan recognition, pattern recognition etc.)

In the end the analysis method relays the indicators to a certain tool that uses these indicators and eventually presents them to the intended target users. These tools can be the learning environment itself or a dedicated tool for this purpose (e.g. visualising communication networks, participation diagrams, etc.).

For the concrete utilisation of the indicator a norm can be applied. This norm is used for meta-cognition or guiding (Jermann, 2004). That means, the norm defines desired values and behaviour, such as “less than 10% participation of one student is too low for good collaboration”.

As the norm has to be expressed in terms of the analysis output format the method influences the types of norms. In the same way the choice of norms influences (shown by the curved arrow) the choice of tools or the kind of representation in the tool. Usually the representation of analysis results against norms is done via metaphors (such as green for desired behavior - red for undesired).

Looking at the Schema of Interaction Analysis (figure 1) developed during the ICALTS project, there are many decisions to be made in the basic process of interaction analysis. Our aim here is to enable users of this schema to apply the proposed analysis methods more easily or recognise new opportunities of analysis. With the given schema it was possible to

provide an overview of analysis methods, indicators and norms and their correspondent pre-requisites, possible outputs and interpretations respectively. While the selection of one specific indicator from the inventory of available indicators (Dimitracopoulou et al., 2005) is supported with the Schema of Interaction Analysis (Harrer et al. 2004), this is usually constrained to one analytical activity. In the CAViCoLA project we extended the scope by defining a general model for the whole process of computer-supported collaboration analysis and connected this with technical tools/formats to make a seamless analysis feasible.

4. THE CAVICOLA PROCESS MODEL

For the purpose of providing a common framework for researchers to model the analysis process in CSCL activities, the European Research Team "Computer-based Analysis and Visualisation of Collaborative Learning Activities" (CAViCoLA) within the Kaleidoscope research network devised a common analysis process model. It has been derived from four empirical research designs which have been conducted by its four research groups in Germany, Greece and Spain (Martínez et al., 2006; Meier et al., 2007; Harrer et al., 2006a), one of them conjointly between two research teams (Harrer et al., 2006b). These teams collaborated on the conceptual and technical integration of their research approaches. A graphical overview of the combination of different analysis methods and their facilitation by a unified data format (CAViCoLA Common Format) can be seen in Figure 2.

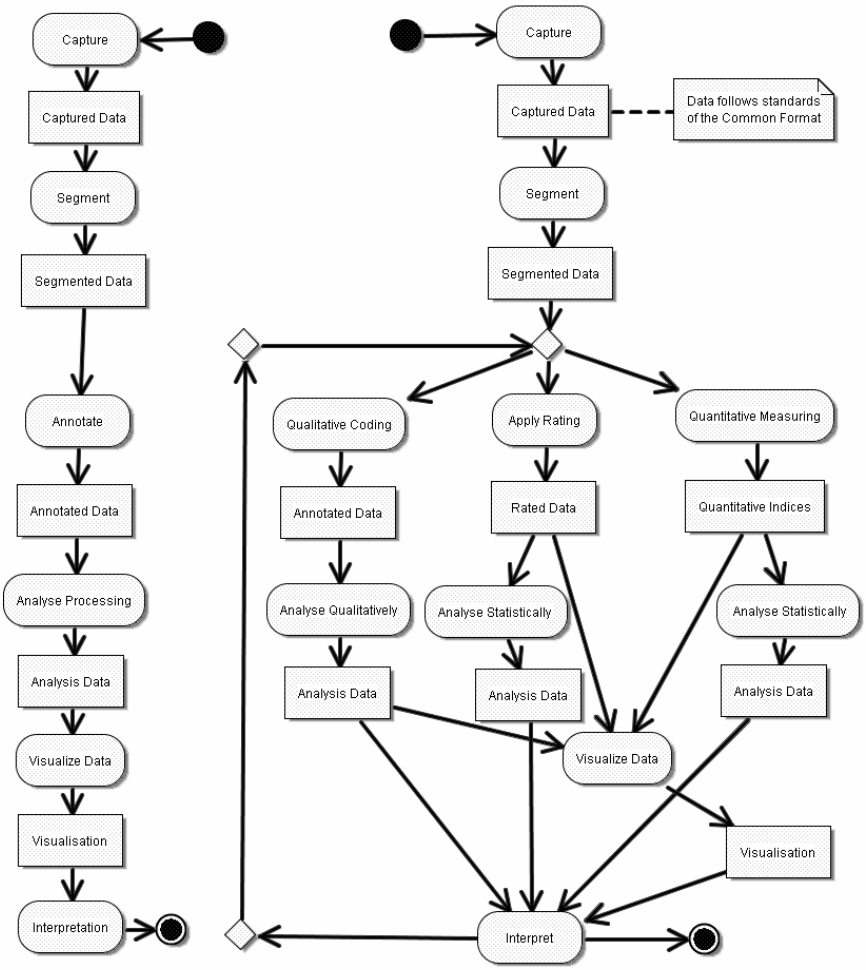


Figure #9-2. Graphical Representation of the CAViCoLA Process Model

The left side of Figure 2 shows the generic process sequence which has been used within the European Research Team. On the right side the CAViCoLA process model combines several quantitative approaches, such as interaction analysis of the participants' actions in time, analysis of group structures in learning communities (Social Network Analysis and statistics; Harrer et al., 2006a; Martínez et al., 2006), and a rating scheme for assessing the quality of the collaboration process (Meier et al., 2007). This is complemented by qualitative methods, such as content analysis,

observations, questionnaires, focus groups and category building (Harrer et al., 2006b; Martínez et al., 2006).

All these analysis methods follow the classical procedure of

- *data capturing*
- *data segmentation*
- *preprocessing* (e.g. annotation and measuring)
- qualitative, statistical and social network *analysis*
- and potentially *visualisation* to support
- *interpretation*.

The overall approach follows the idea of the triangulation of results (Denzin, 1989) in a multi-method approach which is represented in the figure as confluence of several branches of analysis into a joint interpretation.

To facilitate the flexible combination of different analysis tools during the process, we defined a standardised data format (cf. next section) that captures the relevant information of collaborative learning activities. This allows the analysis of several types of captured data, such as the different learning environments used by the partners, e.g. Synergo (http://www.synergo.gr), FreeStyler (http://www.collide.info), Discussion Forums (e.g. phpBB), BSCL-Synergeia shared workspaces (http://bscl.fit.fraunhofer.de/), with the same interoperable set of analysis tools. Among these analysis tools are applications for the qualitative coding of observation data captured by video, for the generation of logfiles capturing user interactions in CSCL systems, and for gathering sociometric data.

5. ON INTEROPERABILITY BETWEEN MULTIPLE LEARNING ENVIRONMENTS AND ANALYSIS METHODS

As a step to complement the conceptual integration - manifested in the analysis indicators and the process model - with a technical integration and interoperability between learning environments and tools, we worked in the Kaleidoscope projects towards a joint format usable by all partners' available analysis tools and other collaborators in the international research community. This common format facilitates the idea of defining a library of tools that can be used for a variety of different purposes of analysis. Essentially this idea targets at having one format, that is produced by all the learning support environments of the project and that is consumed by all the

analysis tools. This mediating format enables on the one hand the analysis of the same data set with multiple analysis approaches in a triangulation approach and on the other hand the use of an analysis tool with several data sets coming from different learning environments, which can facilitate cross-discipline and cross-cultural studies (that otherwise would have to be performed by exclusively manually comparison; with the joint format the same measurements and processing can be done using the same analysis tool). Similar thoughts for interoperability between different learning tools are also relevant for the area of inquiry learning, where usually for different learning phases various tools are suitable and developed specifically; Chapter 2 presents some insights in the methods and solutions chosen for that area of TEL.

Based on pre-existing work on explicit representation of user actions in XML-representations in the project consortium (Gaßner, Jansen, Harrer, Herrmann, Hoppe, 2003) (Martínez et al., 2003) (Harrer et al., 2005) a structured format using XML was chosen as the basis for the common format. This unified data format was designed and refined throughout the Kaleidoscope activities and has been tested for replaying, interpreting, and annotating collaborative workspace activities in the projects; besides the Kaleidoscope-internal use the work had also an impact on another European research project ARGUNAUT (IST 027728) that uses the standardised data format to support the moderator of electronic discussions in analysis and annotation.

5.1 Description of the Format

Defining a common format is influenced by two competing desiderata: on the one hand it should be well-defined and structured, on the other hand it has to be flexible enough to allow additional information created by specific learning environments, which might be useful for some analysis techniques. Thus it was our challenge to define a reliable common denominator for the minimally required information, but also to have a mechanism to mix in important features flexibly. This is reflected in the required and optional elements of our Document Type Description (DTD) for the Common Format.

The required elements are based on our previous approaches and experiences and also other approaches described in the literature (Avouris, Dimitracopoulou, Komis, 2003). As most important information captured for analytic purposes in potentially collaborative scenarios we see:

- *Action type*: a description of the precise type of action according to the typology used within the learning environment. For more general

categories an attribute “classification” is available with standard categories, such as read/write/modify etc.

- *User*: for collaboration analysis we need information about the actors involved in the action and the roles assumed by them with respect to the action, such as originator and receiver of the action
- *Time*: a timestamp that allows ordering and measuring intervals of the captured information traces.

Since most user actions, especially in environments with Direct Manipulation interfaces, target object(s), an optional *object* can be specified. The object description can be enriched as needed by *properties*, such as the object’s attributes and associated values. This provides a very flexible mechanism, but on the backside it lacks the strictness of definition that would be needed to check if all needed elements for some specific analysis are available.

Usually a description of the general setup of a learning situation, such as participants, formal roles of the participants, subgroups, external resources available and so on, help to understand the scenario and the analysis processes suited for it. Consequently we added a section called “PREAMBLE” that provides the information to define the learning scenario.

A graphical overview of the elements and attributes in the common format is represented in Figure 3. The technical specification of the format is available through the Kaleidoscope reports (Martínez, Harrer, Barros, Vélez, 2005) and is currently being published as a format proposal in the international TEL community.

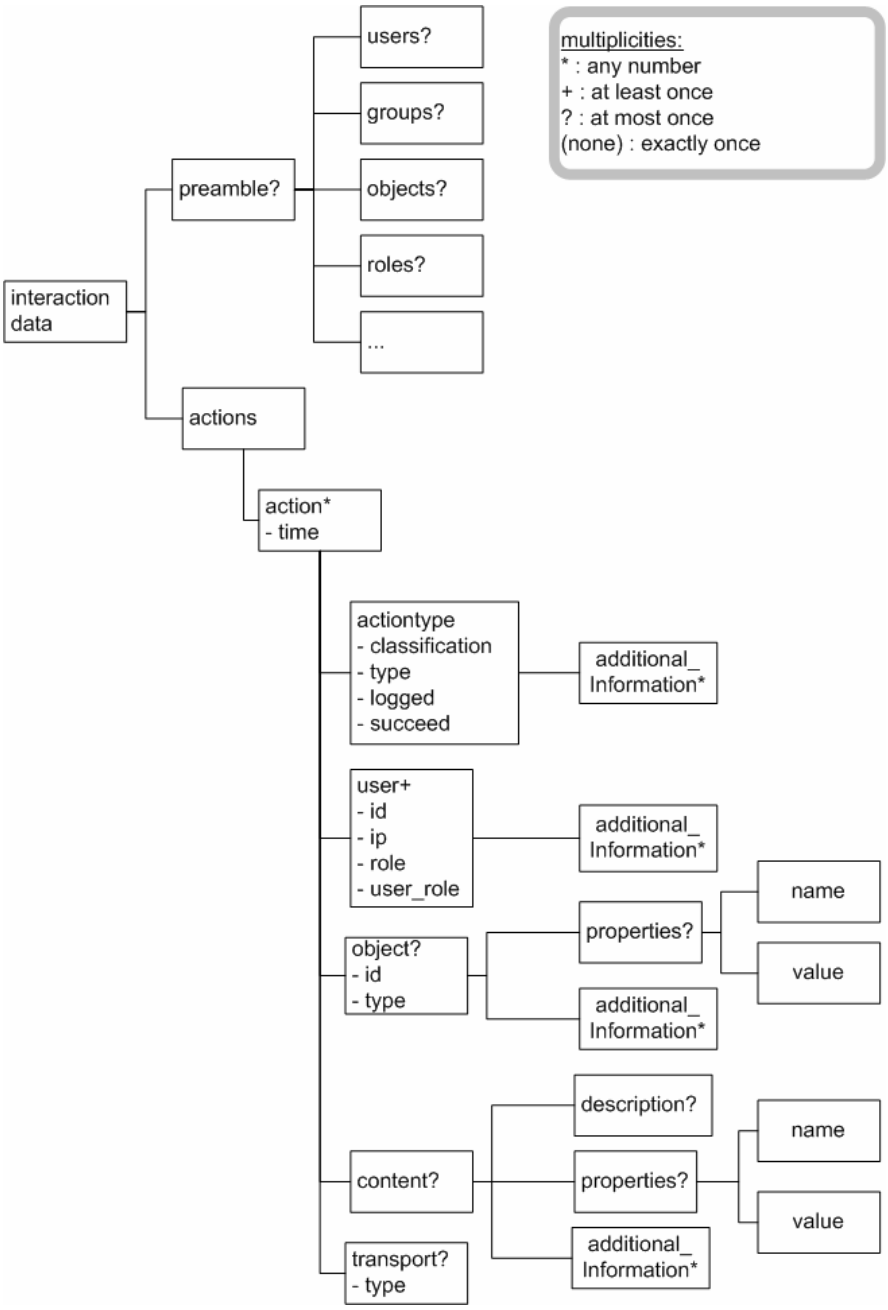


Figure #9-3. Diagram of the basic structure of the Common Format

As a general comment the common format is meant to be a flexible representation preserving the data of all the potential input format. To achieve this we had to declare a lot of elements as optional. This makes the checking of integrity of documents more complicated than with a fixed structure of elements, but this cannot be helped because of the potential difference in information richness of the original input formats. On the other hand not all information needed by some analysis tools can be enforced by the common format (e.g. links between postings are necessary for SNA and posting pattern analysis). As a means to cope with this in some cases the information needed by an analysis tool could be derived by additional scripts/preprocessors in intermediate steps; all the intermediate products should use the common format to enable analysis chains and multiple analyses for each step.

5.2 Practical Use

To put our common format approach to practice with existing learning environments and tools there are different options for realisation with respect to touching the existing code. Ideally each of our learning environments should produce the common format while each of the analysis tools accepts this common format as the input for analysis.

As an alternative for existing learning environments it is possible to use the environment as it is, with the output it produces originally, and to define a mapping from this proprietary output format to the common format. With an exact conceptual mapping the definition of an appropriate XSL transformation or other means of transformation (such as an intermediary object structure written out as common format XML) can be implemented. Similar arguments can be raised for the side of the analysis tools whether accepting input in the common format or in its original one.

In the practical work conducted in our KAL activities we collected experience with all the different variants, to give some examples:

- The export from a discussion forum (that holds the postings in a database) was realised directly to the common format to be used by arbitrary analysis tools.
- The DIAS (Discussion Interaction Analysis System, Bratitsis & Dimitracopoulou (2006)) and the PatternDiscovery tool (Harrer et al., 2005) were modified to accept the common format as input.
- For the data that is used by the SAMSA tool (Martínez et al., 2003) with an XML-based format, both the transformation to the common format to allow other analyses and the transformation from the common format to

the SAMSA format was realised, to make SAMSA usable with input from different learning environments. The integration of the project is well reflected in the fact, that the transformation was defined in cooperation between two partners.

The interoperability achieved by means of the common format was tested practically by analysing the discussion forum logs of the DIAS tool (provided in the common format) with the PatternDiscovery analysis tool (accepting the common format), that can detect patterns a researcher specified beforehand or that produces frequently occurring patterns along a mining approach. Mappings of the common format to related specifications, such as the more general data mining standard PMML (Predictive Model Markup Language - http://www.dmg.org/products.html), are feasible and will be explored in our future work.

6. VISUALISATION OF ANALYSIS RESULTS

The problem of making use of analysis results and facilitate the meaning-making on the user's side has been another aspect of the interaction analysis projects. It led us to the question how to support users to interpret results from logfile analysis and how to utilise social network analysis techniques.

For the work conducted in the KAL activities we took up ideas from the field of information visualisation (Card et al., 1999) and provide the results to the user in diagrammatic and graphical representations that support a first-glimpse comprehension.

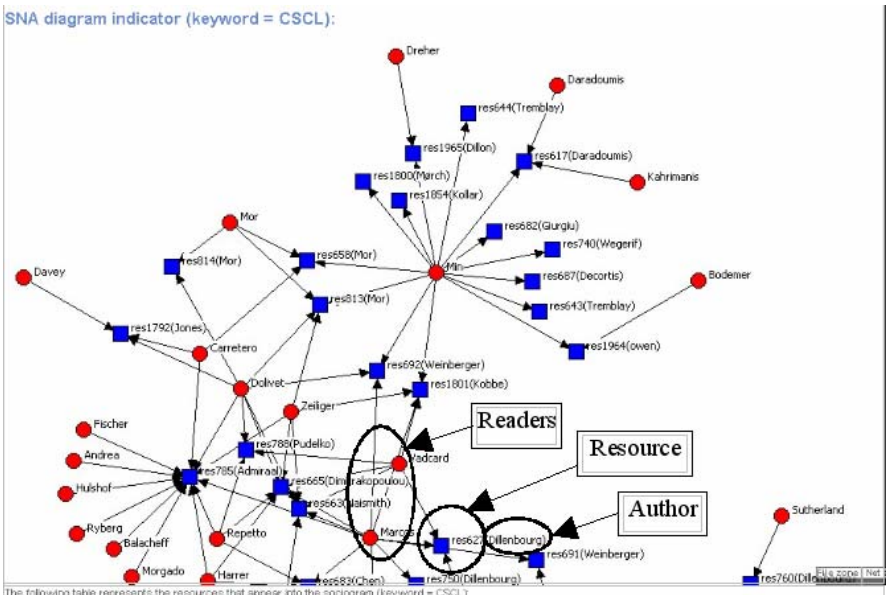
With respect to visualisation of logfile analysis several techniques have been developed, compared, and discussed. One of the challenges here is the integration of temporal information with actor-related information and the representations in the learning environment. Synergo, a collaborative shared workspace tool (Avouris et al., 2004) uses an augmented replay functionality of the workspace content with superimposed information about the workspace objects. The Pattern Discovery Tool (Harrer et al., 2007a) can be used with all common format logfiles to search for user-specified interaction patterns or for supervised automatic mining of characteristic interaction patterns and display the results in a two-dimensional timeline representation considering the initiators of actions.

With the widespread proliferation of web communities and "social software" visualisation techniques for online collaborations have been a

popular means to make interactions explicit. These techniques are based on work in the field of Social Network Analysis (Wassermann&Faust, 1994), that have been used in TEL scenarios for several years now (e.g. Palonen&Hakkarainen, 2000; Reffay&Chanier, 2003). Two approaches addressing the operationalisation of social network analysis for TEL users have been used in our projects: The first within the Interaction Analysis activity in the Kaleidoscope Communication and Collaboration Infrastructure (CCI-IA) is to provide interactivity, by allowing the users to inspect the data about the nodes displayed in the networks. The other approach enriches social networks with additional actor properties and measurements to integrate multiple dimensions of information into a network. These two approaches are presented in the following, as two elaborated extensions of Social Network Analysis.

6.1 Augmentation of Social Networks through navigability

The first approach to enhance the visualisation of social networks was offered as one of the IA indicators offered to the Kal community through the mentioned CCI-IA service. Figure 4 shows the output presented to network leaders when inspecting the social and topic networks that are created by user interactions in the KAL TeLearn Archive. This social network indicator is useful to find resources and people related to a set of research interests. Besides the official structure of a network, a user can be interested in the emerging communities based on the shared use of resources. Social networks built on the use of these shared resources can point out to these communities and thus, help to identify and join people with common research interests.



The following table represents the resources that appear into the sociogram (keyword = CSCL):

Resource	Authors	Title
Res617	Daradoumis, Xhafa, Marqu?s Puig.	A Methodological Framework for Project-based Collaborative Learning in a Networked Environment
Res627	Dillenbourg	Framework for integrated learning
Res638	Pinkwart, Harrer, Lohmann, Vetter.	Integrating Portal Based Support Tools to Foster Learning Communities in University Courses
Res643	Tremblay	Industrial Districts and Innovation Business Networks: New Theories and Concepts for Entrepreneurship Development
Res644	Tremblay	New Types of Careers in the Knowledge Economy ? Networks and boundaryless jobs as a career strategy in the ICT and multimedia sector
Res658	Mor, Hoyles, Kahn, Noss, Simpson.	Constructing structure in number sequences
Res663	Naismith, Longdale, Vavoula, Sharples.	Literature Review in Mobile Technologies and Learning
Res665	Dimitrakopoulou	State of the art on Interaction and Collaboration Analysis (Telearn archive)
Res682	Glurglu, Treichel.	SMECTRA - An European network for online counselling on the application of new media in enterprise training
Res683	Chen	Supporting Teachers Intervention in Collaborative Knowledge Building
Res687	Decortis, Rizzo, Saudelli.	Mediating effects of active and distributed instruments on narrative activities
Res691	Weinberger	Scripts for Computer-Supported Collaborative Learning. Effects of social and epistemic cooperation scripts on collaborative knowledge construction

Figure #9-4. Social Networks in Kaleidoscope around the keyword "CSCL"

This indicator shows the *resources* associated to a concrete keyword and the *visitors* that have accessed these resources. The user can observe the information shown in the sociogram and he/she can visit the resources or people that he/she considers most interesting (e.g.: the resources visited by some relevant member, the most popular resources with this keyword, the profiles of those researchers connected to a particular resource, etc.) using the links provided by the indicator.

For example, Figure 4 represents the social network of researchers and resources for the keyword "CSCL". The squares represent the publications

related to this keyword and the circles represent the users that have visited those publications. The table matches the short name used in the sociogram for the resource (e.g., res627) with its complete name (e.g., "Framework for integrated learning") and includes links to the resources.

The service also permits to identify the resources visited by relevant researchers in the network and access to these resources (e.g.: In Figure 4, if "Marcos" is a relevant researcher for the user, he/she can access the resources visited by him). Also if the user knows a specific document that appears in the sociogram and it is relevant for him/her, he /she can access the resources visited by members that also visited this relevant resource. With this information it is possible to identify people with common research interests and choose the resources visited by them. This suggests the idea of using this social network tool as a kind of recommendation system.

The CCI-IA service was evaluated by Kaleidoscope members by means of a questionnaire placed at the CCI-IA web site. The social network indicator was rated high by the users, who also provided suggestions for its improvement. For example, the features discussed in this section were included in the service after the feedback received from the users of the service in this questionnaire.

6.2 Augmentation of Social Networks with measurements and properties

The other approach tested is to embed advanced measurements into the visualisation. According to (Krempel, 2005) techniques exist to integrate structural properties of networks in the display. For this purpose we created the Weaver application, a 3D visualiser for social networks, which arranges and renders the nodes according to properties such as degree, centrality, or other externally defined properties (e.g. formal roles in a learning scenario) into specific interaction spaces. From the user view this means, that he or she is able to perceive the properties of a node immediately (e.g. what is the most central topic in the network).

One important means to understand and support interaction in computer-based scenarios is the identification of relations between the different actors. Because of their high availability web-based communities tend to grow to large network structures, which might inhibit both researchers and the community members themselves to observe the aspects relevant for them. The identification of substructures and decomposition of large networks into smaller constituents is one of the methods to provide overviews and selectively focus on specific aspects (de Nooy et al., 2005).

This is especially important for visualisation of networks in diagrams and requires information about common attributes to define subgroups in a network (Krempel, 2005).

In the literature of SNA several proposals have been made to define subgroups and substructures of sociograms (Wassermann&Faust, 1994). The coarse differentiation of these proposals can be made by the criterion if the grouping provides a disjoint split into subsets, i.e. a partition, or if the smaller sets may overlap. Depending on the purpose of the analysis, either of the two perspectives might be the preferable one: the partition is desirable, when separate parts of the network should be analysed in isolation; overlapping structures can show the relations between two sets of actors, such as overlap, embedment into larger subgroups etc. Thus in the Weaver application (Harrer et al., 2007b) we implemented both grouping concepts (disjoint or inclusive sets) to be used as needed by analysts.

The identification of subgroups within communities and of core actors is especially important for coordinators and reviewers of the community. It might also be used as a feedback mechanism to let the community members reflect about their own roles and positions in such a network.

To illustrate and test these ideas we conducted an analysis of the Special Interest Group Artificial Intelligence in Education (AIED) of the Kaleidoscope network. As natural data we used the participation in project planning activities facilitated by the SIG (see Figure 5). Squares represent the eight activities initiated by SIG members in the first two years of Kaleidoscope. Teams are represented by triangles and links show the participation of the teams resp. members of the teams in the planned activities.

The resulting network based on this data is visualised without specific network analytical aspects in Figure 5. The arrangement that separates the different node types into circles (projects in the inner circle, teams in the outer) is used to improve readability.

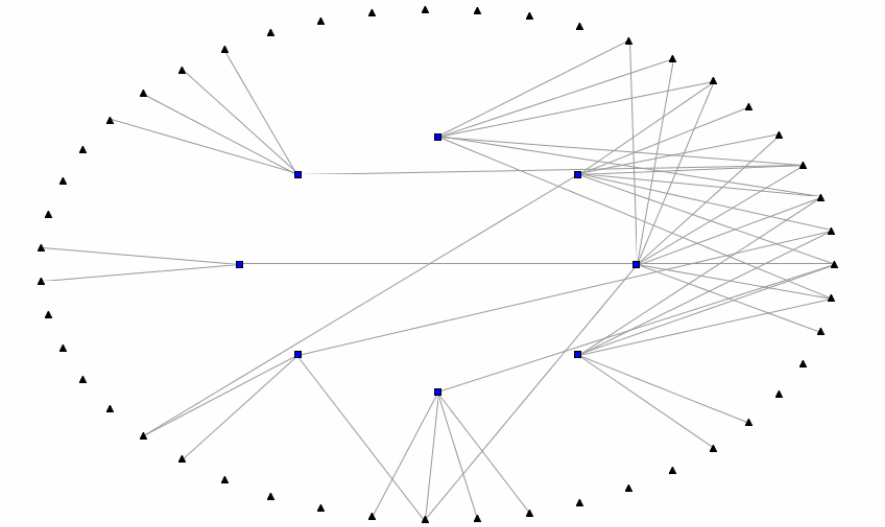


Figure #9-5. Network diagram showing projects (squares) prepared by SIG AIED teams (triangles)

The use of network analysis and visualisation of the analytical results helps the researcher in the meaning making of the network structures. The detection of cycles in the network allows the identification of collaborations between teams.

By using our grouping algorithm we get as the result Figure 6. The teams visualised in the circle at the bottom form a group, which can be interpreted as that each pair of teams belonging to it have either been involved in 2 joint projects directly or they have been collaborating with a mediating team that also collaborated with the other team of the pair.

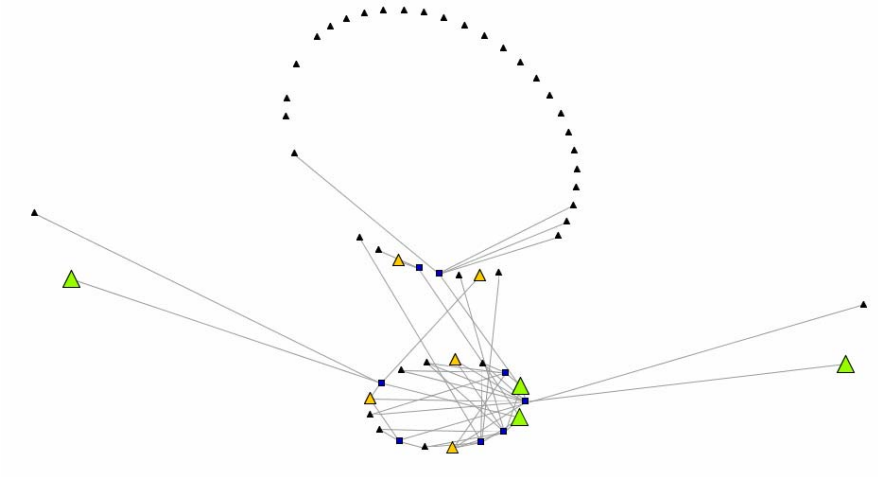


Figure #9-6. Diagram augmented with information about publication activities

As Lothar Krempel points out (Krempel, 2005) the augmentation of networks with additional information gives an improved visual perception of the networks for the user, both at expert and non-expert level. Very often actors of a network have relevant attributes that can be independent of SNA properties. The visualisation of these attributes can be an important feature to enhance the interpretation of these networks.

To identify the main actors in an international research network the impact on the scientific community by publishing at relevant conferences is one of the crucial aspects. Thus we reviewed all papers and posters published at the - in our opinion - most relevant conferences for this field, namely AIED 2005 and ITS 2006, and counted the publications of the members of SIG AIED for each team. Thus, in Figure 6 the grouping is combined with visually marking teams participating at both conferences (large green triangles), either one (medium sized triangles), or neither of both (small black triangles). The interpretation of the diagram indicates that all of the teams publishing at the main conferences are either part of the group or at least collaborated with parts of the group.

By means of inspecting and interpreting such diagrams a deeper insight in participation patterns and activity within a community can be gained, which is a valuable instrument for coordinators/moderators of the network, reviewers of the activity, and – when fed back to the participants themselves – also to inform the participants, potentially influencing their behaviour in a self-regulative manner.

7. CONCLUSION

This chapter presented work conducted within the Kaleidoscope network on the topic of interaction and collaboration analysis. Several facets of interaction analysis have been discussed here, starting with a conceptualisation of interaction analysis indicators and the process how to compute indicators based on the learning data. This conceptual integration built the foundation for a uniform terminology, methodology and comparability of interaction analyses between the different research teams involved.

To complement the conceptual integration with a technical equivalent we elaborated on aspects of interoperability between different interaction analysis methods and tools by means of unified data formats and interfaces; this enables the cross-usage of analysis tools and research data beyond its initial scope, which is usually tied to one specific learning environment and one analysis tool.

To make the results of interaction analysis methods understandable to both experts and laypeople, the representation of results is a decisive factor. Based on principles of information visualisation, we presented two characteristic examples, where we used and designed visualisations for social networks that carry the most important properties of the collaboration, thus enabling the analysts to interpret the nature and intensiveness of collaboration more easily than by using plain numbers, measurements, and data tables. It is to be noted that the different information needs of students, teachers, and researchers, and their capabilities of understanding these representations have to be considered to choose the appropriate level of detail and data richness.

ACKNOWLEDGEMENTS

We would like to thank all the collaborators in the ICALTS, IA, and CAViCoLA projects, the co-authors of our joint publications, as well as the student workers and student volunteers involved in the activities described.

8. REFERENCES

Avouris, N., Dimitracopoulou, A., Komis, V. (2003) On analysis of collaborative problem solving: An object-oriented approach. Computers in Human Behavior, 19 (2), 147-167

Avouris N., Margaritis M., and Komis V. (2004). Modelling interaction during small-group synchronous problem-solving activities: The Synergo approach, 2nd International Workshop on Designing Computational Models of Collaborative Learning Interaction, ITS2004, 7th Conference on Intelligent Tutoring Systems, Maceio, Brasil.

Bratitsis T. & Dimitracopoulou A. (2008), Monitoring and Analysing Group Interactions in Asynchronous Discussions with DIAS system, Special Issue, Interactive and Interdisciplinary Perspectives on Groupware Research, International Journal of e-Collaboration (IJeC), 4(1).

Bratitsis, T., Dimitracopoulou, A. (2006). Monitoring and Analyzing Group Interactions in Asynchronous Disucussions with the DIAS system, Groupware: design, Implementation, and Use; Proc. of CRIWG 2006, LNCS 4154, Y. Dimitriadis, I. Zigurs, E. Gomez-Sanchez (eds.). 54-61, Springer, Berlin.

Card K. S., Mackinlay J. D., Shneiderman B. (1999), Readings in Information Visualization, using Vision to think, Morgan Kaufmann, San Francisco, USA, 1999.

Denzin, N. (1989): The Research Act. A Theoretical Introduction to Sociological Methods, Prentice Hall, Englewood Cliffs, NJ.

Dimitracopoulou *A.* & Bruillard E. (2006). Interfaces de Forum enrichies par la visualization d'analyses automatiques des interactions et du contenu. (Guest Editors) E.Bruillard & M. Sidir,. Special Issue on 'Forum de Discussion en Education', *Sciences et Technologies de l'Information et la Communication pour l'Education et la Formation, STICEF*, 13.

Dimitracopoulou, A et al. (2005): State of the art of interaction analysis for Metacognitive Support & Diagnosis. Deliverable D.31.1.1. Kaleidoscope NoE,. Available online at: www.noe-kaleidoscope.org

Dimitracopoulou, A. et al. (2004). State of the Art on Interaction Analysis: "Interaction analysis indicators". Deliverable D.26.1,1 Kaleidoscope NoE. Available online at: www.noe-kaleidoscope.org

Gaßner, K. (2003). Diskussionen als Szenario zur Ko-Konstruktion von Wissen. Doctoral Dissertation. Faculty of Engineering Sciences, University Duisburg-Essen.

Gaßner K., Jansen M., Harrer A, Herrmann K & Hoppe H.U. (2003), Analysis methods for collaborative models and Activities, In Designing for Change in Networked Learning Environments, Proc. CSCL 2003, B. Wasson, S. Ludvigsen, U. Hoppe (Eds.), 411-420, Kluwer Academic Publ., Dordrecht.

Harasim, L.M. (1990), Online Education - Perspectives on a new environment, Praeger Publishers. New York.

Harrer, A. et al. (2004). Unified Framework on Interaction Analysis. Deliverable D.26.2,1. Kaleidoscope NoE. Available online at: www.noe-kaleidoscope.org

Harrer, A., Hever, R., Ziebarth, S. (2007a). Empowering researchers to detect interaction patterns in e-collaboration, Artificial Intelligence in Education - Building Technology Rich Learning Contexts that Work, R. Luckin, K.R. Koedinger J. Greer (eds.), IOS Press, Amsterdam, 503-510.

Harrer, A., Kahrimanis, G., Zeini, S., Bollen, L., Avouris, N. (2006b) Is there a way to e-Bologna? Cross-National Collaborative Activities in University Courses, Proceedings 1st European Conference on Technology Enhanced Learning EC-TEL, LNCS 4227, 140-154, Springer Berlin

Harrer, A., Vetter, M., Thür, S., Brauckmann, J. (2005). Discovery of patterns in learner actions. In: Chee-Kit Looi, Gord McCalla, Bert Bredeweg, & Joost Breuker (Eds.),

Artificial Intelligence in Education - Supporting Learning through Intelligent and Socially Informed Technology, 816-818, Amsterdam: IOS Press.

Harrer, A., Zeini, S., Pinkwart, N. (2006a). Evaluation of communication in web-supported learning communities – an analysis with triangulation research design. In: International Journal of Web Based Communities 2006, 2(4), 428 - 446, Inderscience.

Harrer, A., Zeini, S., Ziebarth, S., Münter, D. (2007b): Visualisation of the Dynamics of Computer-mediated Community Networks, Paper presented at International Sunbelt Social Network Conference 2007.

Jermann, P. (2004). Computer Support for Interaction Regulation in Collaborative Problem-Solving, Doctoral Dissertation. Universite de Geneve.

Kay, J., Maisonneuve, N., Yacef, K., Reimann, P. (2006) The Big Five and Visualisations of Team Work Activity. In: Intelligent Tutoring Systems: Proc. ITS 2006, Ikeda, M, K D Ashley and T Chan (eds), 197 – 206, Springer.

Krempel, L (2005). Visualisierung komplexer Strukturen. Grundlagen der Darstellung mehrdimensionaler Netzwerke. Frankfurt a. M.: Campus.

Martínez, A., Dimitriadis, Y., Gómez-Sánchez, E., Rubia-Avi, B., Jorrín-Abellán, I., Marcos, J. A. (2006). Studying participation networks in collaboration using mixed methods in three case studies. In: International Journal of Computer-Supported Collaborative Learning, 1(3), march 2006. Springer.

Martínez, A., de la Fuente, P., & Dimitriadis, Y. (2003). Towards an XML-Based Representation of Collaborative Interaction. In B. Wasson, S. Ludvigsen, U. Hoppe (Eds.), Designing for Change in Networked Learning Environments, Proc. CSCL 2003, 379-388, Kluwer Academic Publ., Dordrecht.

Martínez, A., Harrer, A., Barros, B., & Vélez, J. (2005). Library of Interaction Analysis Methods. Deliverable D.31.2,1 , Kaleidoscope NoE, Available online at: www.noe-kaleidoscope.org

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. The International Journal of Computer-Supported Collaborative Learning, 2(1), 63-86.

Moreno J.L. (1951). Sociometry, Experimental Method and The Science of Society, Beacon House, Inc., New York.

de Nooy, W., Mrvar, A. & Batagelj, V. (2005). Exploratory Social Network Analysis with Pajek, Cambridge University Press.

Palonen, T. & Hakkarainen, K. (2000). Patterns of interaction in computer-supported learning: A social network analysis. In: Fishman, B. & O'Conner-Divelbiss, S. (Eds.), Proc. of the Fourth International Conference on the Learning Sciences. Mahwah, NJ. Lawrence Erlbaum.

Reffay, C., & Chanier, T. (2003). How Social Network Analysis can hep to measure cohesion in Collaborative Distance Learning. In: Proceedings of the International Conference on Computer Support for Collaborative Learning. Wasson, B. Ludvigsen, S & Hoppe, H. U. (Eds.) Dordrecht: Kluwer Academic Publishers, 343-352.

Soller, A., Martínez Monés, A., Jermann, P., Muehlenbrock, M. (2005) From mirroring to guiding: a review on the state of the art technology for supporting collaborative learning, International Journal of Artificial Intelligence in Education (ijAIED). 15:261-290.

Sun, L & Vassileva, J. (2006). Social Visualization Encouraging Participation in Online Communities. CRIWG 2006: 349-363

Wassermann, S., & Faust, K. (1994). Social Network Analysis: Methods and Application, Cambridge: University Press.

View publication stats