

Mohamed Amine Chatti, Anna Lea Dyckhoff, Ulrik Schroeder, Hendrik Thüs

Forschungsfeld Learning Analytics

Learning Analytics Research Challenges

Keywords Learning Analytics_Academic Analytics_Educational Data Mining

Zusammenfassung. Das Thema Learning Analytics hat in letzter Zeit große Aufmerksamkeit erlangt. Es verspricht, vor allem selbstgesteuerte Lernprozesse zu unterstützen, die typisch für das E-Learning in Verbindung mit Web 2.0 und sozialen Medien sind. Dabei wird versucht, durch Aufzeichnung, Analyse und geeignete Präsentation von Lernaktivitäten und deren Kontext, die Reflexion und Optimierung eines Lehrangebots oder des Lernverhaltens zu fördern. In Johnson et al. (2011) wurde Learning Analytics als eine der Schlüsseltechnologien für zukünftige Lehr- und Lernansätze identifiziert. Einige der damit verbundenen Forschungsfragen werden in diesem Artikel vorgestellt.

Summary. Learning analytics has attracted a great deal of attention in technology enhanced learning (TEL) in recent years as educational institutions and researchers are increasingly seeing the potential that learning analytics has to support the learning process. Learning analytics has been identified as a possible key future trend in learning and teaching (Johnson et al., 2011). Analytics can be a powerful tool to support learning. There are, however, a number of issues that need to be addressed before starting analytics projects. In this paper, we identify various challenges and research opportunities in the emerging area of learning analytics.

1. Definition

In der Literatur sind verschiedene Definitionen des Forschungsgebiets Learning Analytics zu finden. Auf der ersten internationalen Konferenz¹ zu diesem Forschungsfeld (LAK11) wird es beschrieben als „measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs“. Elias (2011) hebt zusätzlich die Wichtigkeit von Analysetools hervor, während für Siemens (2010) die Nutzung von intelligenten und durch Nutzer produzierten Daten zur Informationsgewinnung und zum Auffinden von sozialen Verbindungen zur Unterstützung des Lernprozesses im Mittelpunkt steht. Die Definition der EDUCAUSE's Next Generation Learning Initiative sieht ebenfalls die Vorhersage des studentischen Lernfortschritts als zentralen Punkt (zitiert in Siemens, 2010). Obwohl es Unterschiede in einigen Details gibt, beziehen sich alle Definitionen auf die Sammlung und Auswertung von

Daten aus Lernprozessen, um diese in lernunterstützende Aktionen zu übertragen. Außerdem ist es Wert festzustellen, dass sich die Definitionen nicht auf die (voll-)automatische Analyse und Auswertung der Daten beschränken.

Wir verstehen in diesem Papier unter Learning Analytics dementsprechend ein aufstrebendes Forschungsfeld des E-Learning, das die Entwicklung von Methoden und Werkzeugen für die Analyse und Mustererkennung in Daten aus Lehr-/Lernkontexten erforscht, um Bildungsinstitutionen, Lehrenden und Lernenden zu ermöglichen, Lernprozesse zu reflektieren und damit die Optimierung von Lehrangeboten einerseits und Metalernen andererseits zu fördern.

Konkret bedeutet dies auf Seite von Lehrenden, dass sie Informationen über die Nutzung und Effekte bereitgestellter Lehrmaterialien und Betreuungsangebote nutzen können statt diese nur mittelbar aus Ergebnissen von Lernstandsüberprüfungen abzuleiten. Damit können sie nicht nur inhaltlich lehren, sondern auch als Coach und Mentor für Lernprozesse agieren und z. B. auch zu Lern- und Lösungsstrategien beraten. Insbesondere

für selbst-gesteuert Lernende bietet Learning Analytics die Chance, ihre Lernaktivitäten auch im Vergleich zu anderen zu bewerten und z. B. erfolgreiche Lernstrategien, wie das regelmäßige Bearbeiten von Übungsaufgaben oder das Arbeiten in Teams, von anderen zu übernehmen.

2. Forschungs Herausforderungen

Es gibt einige kritische Herausforderungen, die bei der Erforschung und Entwicklung von Learning-Analytics-Lösungen berücksichtigt werden müssen. Diese umfassen die Sammlung, Analyse und Verwaltung geeigneter Daten, die Unterstützung der Entwicklung aussagekräftiger Indikatoren und Metriken sowie das Profiling von Lernern. Ferner sind flexible Werkzeuge für die Integration in diverse Lernumgebungsarchitekturen zu entwickeln und deren Nutzbarkeit für eine Lernreflexion zu untersuchen. Da unterschiedliche Nutzergruppen von Resultaten einer Lernprozessanalyse profitieren können, sind die Bedarfe und Befürchtungen aller Beteiligten zu berücksichtigen und Da-

DOI 10.1524/icom.2012.0007

¹ <https://tekri.athabasca.ca/analytics/>

tenschutzfragen bei der Profilerfassung von Lernenden zu untersuchen.

Datenmanagement: Lernwerkzeuge und Ressourcen bewegen sich zunehmend in Richtung dezentral organisierter Dienste (Cloud), indem sie soziale Netzwerke integrieren. Daher wird es schwieriger, Rohdaten in diversen Formaten aus verschiedenen, heterogenen Quellen zu aggregieren und zu sinnvollen Datensätzen zu verknüpfen, die die verteilten Lernaktivitäten korrekt und aussagekräftig beschreiben und die Basis für solide und präzise Analysen liefern.

Darüber hinaus müssen riesige Massen an Daten verarbeitet werden. Dies ist einerseits eine technische Herausforderung, da geeignete Data Mining Methoden entwickelt und Werkzeuge implementiert werden müssen, die aussagekräftige Resultate in akzeptabler Zeit liefern, damit die an den Resultaten interessierten die Möglichkeit haben, auf die Erkenntnisse schnell zu reagieren. Auf der anderen Seite bedeutet die Menge an Daten nicht immer eine Verbesserung der Aussagekraft. Kop et al. (2011) weisen darauf hin, dass es für Lernanalyse vorteilhaft sein kann, geeignet eingeschränkte Datenquellen zu identifizieren und Sampling-Strategien anzuwenden, um die Reliabilität der Analyseergebnisse zu erhöhen.

Da sich die Datenbasis im Laufe von (bereits analysierten) Lernprozessen dynamisch stark verändern kann, z.B. durch den Beitritt eines Lerners in eine neue Lerngruppe und die Nutzung der dort verfügbaren, vernetzten Lernressourcen, benötigt man geeignete Wege einer Skalierung und inkrementellen Erweiterung der Analysemethoden und -werkzeuge und ihrer Funktionalität.

Derzeit sind viele Systeme reich an Daten aber arm an Information. Forscher müssen die Qualität der in der Praxis eingesetzten Analyseergebnisse evaluieren, um pädagogisch nützliche Indikatoren, Vorschläge und Empfehlungen zu entwickeln und zu optimieren. Im Vergleich zu anderen Bereichen von Empfehlungssystemen gibt es für Learning Analytics bislang keine Benchmarks, mit der sich die Güte von Algorithmen validieren ließe. Es stellt sich die Frage, wie eine empirisch fundierte Basis entstehen kann, die vertrauensvolle, validierte Entscheidungen ermöglicht.

Indikatoren und Metriken: Die Ziele einer Lernprozessanalyse sind manchmal schwierig zu bestimmen und müssen auf passende Leistungsindikatoren und -metriken abgebildet werden. Dabei müssen über Bepunktung und Benotung von Übungsaufgaben und Testergebnissen hinausgehende neue Leistungsindikatoren gefunden werden, um innovative selbstgesteuerte und vernetzte Lernformen und informelles, lebensbegleitendes, in Arbeitsprozesse und in die Freizeit eingebettetes Lernen adäquat zu berücksichtigen (vgl. Forschungsherausforderungen für Assessment und Feedback). Die Forschungsherausforderung besteht darin, die häufig iterativ erfolgende Definition geeigneter Ziel-Indikator-Metrik-Tripel zu unterstützen, die Lernprozessanalysen begleitet.

Profiling: Eine Kernaufgabe der Lernprozessanalyse ist das Erstellen geeigneter auswertender Lernerprofile (profiling). Die offene Frage ist, welche Informationen erhoben und/oder ermittelt und auf welche Weise korreliert werden, um eine effektive Intervention, eine Anpassung der Angebote oder eine Empfehlung für Lernwege, Lernmaterialien, Lernaktionen Lernwerkzeuge oder deren Parameter zu geben. Dies ist eine hohe Herausforderung, da sich Lernaktivitäten über längere Zeiträume und verschiedene Systeme verteilen können. Dabei entwickeln sich neue Fragen an Lehr- und Lernprozesse mit den Resultaten vorhergehender Analysen. Um ein detailliertes und profundes Bild von Lernereigenschaften für eine Analyse und Empfehlungen zu berücksichtigen, wäre es daher oft hilfreich, Daten, deren Nutzen nicht a-priori feststeht, aus einem weiten Feld von Lernaktivitäten hinzuzuziehen. Eine derartige Datenaggregation ist aber sowohl rechtlich und ethisch als auch technisch problematisch.

Lernumgebungen: Aktuelle Analysewerkzeuge sind oft eingebettet in zentrale, web-basierte Lernsysteme, wie adaptive, intelligente tutorielle Systeme (ITS) oder Lernmanagementsysteme (LMS). Wir erwarten, dass sich Lernsysteme entsprechend innovativer Lerntheorien wie Connectivism (Siemens, 2005) und Learning as a Network (Chatti, 2010) in Richtung offener, vernetzter, das lebens-

lange Lernen unterstützender Systeme entwickeln. Eine große Herausforderung ist es, die komplexen und sich schnell weiterentwickelnden personalisierten Lernumgebungen (PLE) geeignet an individuelles Lernverhalten und Lernbedarfe anzupassen. Hierzu bedarf es neuer Analysemodelle, die v. a. das selbstgesteuerte Lernen geeignet unterstützen.

Interaktivität: Eine weitere zukünftige Herausforderung besteht in der Entwicklung flexibler und effizienter Lernanalysewerkzeuge, die sich einerseits als Komponenten in Lernsysteme und -umgebungen einbinden lassen und andererseits die Zeit der Datenintegration und -analyse derart minimieren, dass sie unmittelbare Reaktionen auf die Erkenntnisse der Analyseergebnisse ermöglichen. Idealerweise sollten nur Sekunden zwischen Anfragen und Analyseergebnissen liegen, damit Nutzer quasi in Realzeit explorieren und Visualisierungen entsprechend der aktuellen Interessen anpassen können und interaktiv neue Fragestellungen generieren können.

Usability: Je nach aktuellen Analysezielen können unterschiedliche Visualisierungs- und Interaktionstechniken zum Einsatz kommen. Die Aufgabe besteht darin, einfach zu verwendende und nützliche statistische Visualisierungs-, Filter- und Mining-Werkzeuge zu entwickeln, die Lernende, Lehrende und Institutionen darin unterstützen, ihre Forschungsziele zu erreichen, ohne extensive fachliche Kenntnisse der zugrundeliegenden Technologien vorauszusetzen. Lernanalysewerkzeuge müssen derart entwickelt werden, dass sie einerseits in die Standard-Lehr- und Lernumgebungen integriert und durch Nicht-Experten genutzt werden können. Andererseits sollten sie dabei genauso personalisiert werden können wie personalisierte Lernumgebungen und damit zielorientierte Rückmeldungen für unterschiedliche Beteiligte geben, um eine Reflexion des Lernprozesses und Entscheidungen zu unterstützen und eine Selbsteinschätzung zu fördern. Um die Entwicklung nützlicher und nutzbarer Lernprozessanalysewerkzeuge zu unterstützen, wäre es hilfreich, Leitfäden und Design Patterns zu entwickeln. Dies setzt allerdings derzeit weitgehend noch fehlende Einsatzerfahrungen voraus. Es wird

entscheidend für eine Akzeptanz derartiger Werkzeuge sein, Erfahrungen Lehrender und Lernender auszuwerten und in die Entwicklung weitergehender Methoden und Werkzeuge einfließen zu lassen. Geeignete Visualisierungen könnten dabei eine zentrale Rolle spielen, die großen Mengen an Daten aus Lernprozessen besser zu verstehen und mehr Einsicht in Prozesse und Beziehungen von Lehre und Lernen zu gewinnen. Dies bildet eine Voraussetzung für das übergreifende Ziel von Learning Analytics, diese Prozesse schrittweise zu verbessern.

Forschungsmethoden: Eine weitere Herausforderung besteht in der Entwicklung gemischt quantitativer und qualitativer Methoden. Ein Lernanalyseprozess startet grundsätzlich mit einer Forschungsfrage und der Auswahl einer dafür geeigneten Untersuchungsmethode. Viele Forscher wählen dabei eine Kombination aus quantitativen und qualitativen Methoden. Dabei werden häufig sowohl verschiedene quantitative Methoden kombiniert, z. B. um die Robustheit der Forschungsergebnisse zu erhöhen, als auch ergänzend qualitative Methoden wie Interviewtechniken oder Fokusgruppen eingesetzt, z. B. um der Frage nach dem Warum nachzugehen, wenn ein Phänomen durch Analyse erkannt wurde (Dyckhoff, 2010). Während quantitative Methoden und Data Mining Techniken Trends, Korrelationen, Verbindungen, Cluster oder Strukturen aufzeigen können, benötigt man häufig qualitative Methoden um zusätzliche Informationen über mögliche Gründe hinzuzuziehen. Für eine fundierte und gültige Interpretation von Analysedaten ist es daher häufig sinnvoll, verschiedene Methoden zu kombinieren, insbesondere um komplexe Fragestellungen innerhalb offener und vernetzter Lernumgebungen zu beantworten. Eine Forschungsanforderung besteht demnach auch darin, verschiedene Untersuchungsmethoden bereitzustellen und deren sinnvolle Kombination (insbesondere für Nicht-Experten) geeignet zu unterstützen.

Balancieren der Anforderungen aller Beteiligten: Die unterschiedlichen Positionen und Interessen der Beteiligten in Lernprozessen bergen Konfliktpotenziale. Für die Ableitung guter Vorhersagen und

Empfehlungen ist es einerseits sinnvoll, möglichst viele Daten und Nutzungsgewohnheiten zu erheben und auszuwerten. Auf der anderen Seite können diese Daten dann aber auch zweckentfremdet werden, wie z. B. für die individuelle Beurteilung von Leistungen. Es ist notwendig, die unterschiedlichen Interessen der Beteiligten adäquat zu berücksichtigen. Dies ist eine komplexe Aufgabe, da sich viele Bedarfe und Befürchtungen in mehreren Dimensionen widersprechen. Zum Beispiel kann eine Lehrinstitution Resultate von Lernprozessanalysen heranziehen, um Best Practice Beispiele technologieunterstützter Lehre als Referenzen für Lehrende und den Nachweis guter Lehrqualität zu identifizieren. Auf der anderen Seite könnten sich dadurch Lehrende kontrolliert und überwacht fühlen. Das gleiche gilt auch auf Seite der Lernenden, die befürchten können, dass Analyseresultate nicht nur formativ zur Verbesserung der Lernprozesse genutzt werden, sondern auch in die Bewertung der Lernleistung einfließen. Dies kann dazu zu führen, dass sowohl Lehrende als auch Lernende nicht motiviert sind, Analysetechnologien zu nutzen. Es ist ein offenes Problem, Methoden und Werkzeuge sowohl technologisch und didaktisch als auch organisatorisch eingebettet so zu entwickeln, dass ein Vertrauen entsteht, das die Daten und Erkenntnisse daraus nur zum intendierten Zweck der Förderung des Lernprozesses ausgewertet werden.

Datenschutz und verantwortungsvoller Umgang: Eng mit dem vorherigen Bereich verbunden ist auch die Frage, wer Zugriff auf Lerner- und Lernaktivitätsdaten hat. Lassen sich diese auf individuelle Personen zurückführen oder können die Analysensysteme jederzeit garantieren, dass Daten und Visualisierungen jeweils nur die eigenen Daten in Relation zu einer anonymen Menge von Lernercharakteristika betrachtet werden können? Datenmissbrauch muss vermieden und falls nicht komplett verhindert dann zumindest entdeckt werden können. Es müssen klare Grenzen für Learning Analytics definiert werden, deren Einhaltung soweit möglich technisch unterstützt wird. Vertrauliche Lernerdaten und deren Identität müssen jederzeit geschützt werden. Z. B. sollte bei interaktiver Explorati-

on von Daten und Einstellung von Filtern (z. B. alle Teilnehmerinnen einer Lehrveranstaltung in der Informatik, ggf. noch gekoppelt mit weiteren Eigenschaften wie Migrationshintergrund etc.) niemals eine Mindestzahl möglicher Repräsentanten der spezifizierten genannten Teilmengen unterschritten werden, so dass eine eindeutige Zuordnung zu Personen nicht unmittelbar erfolgen kann. Da Datensätze potenziell aus einer Menge unterschiedlicher Systeme zusammengetragen werden, ist zu klären, ob und wie solche Daten persistiert, geschützt und für eine weitere Verwendung bereitgestellt werden können (Campbell et al. 2007, Campbell & Oblinger 2007).

3. Zusammenfassung

Learning Analytics ist ein junges Forschungsfeld, das Forscher verschiedener Disziplinen auf technischer, pädagogischer, organisatorischer und ethischer Ebene vor umfangreiche Herausforderungen stellt. Diese umfassen den Umgang mit wachsenden Datenvolumina, deren Heterogenität und Fragmentierung, die Interoperabilität oder Integration der am Prozess beteiligten Systeme und Komponenten, Performanz, Skalierung, Erweiterbarkeit, Echtzeitanforderungen, Verlässlichkeit, Nutzbarkeit und Akzeptanz, das Finden geeigneter Indikatoren und Metriken, geeignete Visualisierungstechniken, die Unterstützung von gemischt quantitativen und qualitativen Forschungsmethoden, Datenschutz und verantwortungsvoller Umgang mit Daten und die Integration von Learning Analytics in Alltagslernsituationen.

Literatur

- Campbell, J.P., DeBlois, P.B. & Oblinger, D.G. (2007). Academic Analytics: A New Tool for a New Area. *EDUCAUSE Review*, July/August 2007, 41–57.
- Campbell, J.P. & Oblinger, D.G. (2007). *Academic Analytics*. EDUCAUSE White Paper. Zugriff am 15. Juni 2011 unter <http://www.educause.edu/ir/library/pdf/PUB6101.pdf>
- Chatti, M.A. (2010). *Personalization in Technology Enhanced Learning: A Social Software Perspective*. Shaker Verlag, November 2010. Dissertation, RWTH Aachen University.
- Dyckhoff, A.L. (2010). *Towards Tools for Educational Action Research: A Meta-Analysis on*

Applied Research Questions. In: Proc. of the IADIS Int. Conf. E-Learning 2010 (pp. 235–242). Freiburg, Germany.

Elias, T. (2011). Learning Analytics: Definitions, Processes and Potential. Zugriff am 29. Juli 2011 unter <http://learninganalytics.net/Learning-AnalyticsDefinitionsProcessesPotential.pdf>

Johnson, L., Smith, R., Willis, H., Levine, A. & Haywood, K. (2011). The 2011 Horizon Report.

Austin, Texas: The New Media Consortium.

Kop, R., Fournier, H. & Sitlia, H. (2011). The Value of Learning Analytics to Networked Learning on a Personal Learning Environment. In: Proc. 1st Int. Conf. on Learning Analytics and Knowledge 2011. Banff, Alberta, Canada.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance

Learning, vol. 2, no. 1, 2005. Zugriff am 29. Juli 2011, unter http://www.itdl.org/Journal/Jan_05/article01.htm

Siemens, G. (2010). What are Learning Analytics? Zugriff am 29. Juli 2011, unter <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>



1

Dr. M.A. Chatti und Prof. Dr. U. Schroeder: siehe vorheriger Beitrag.

1 Anna Lea Dyckhoff ist wissenschaftliche Mitarbeiterin am Lehr- und Forschungsgebiet Informatik 9 (Learning Technologies) der RWTH Aachen und promoviert im Bereich „Learning Analytics“.

dyckhoff@cil.rwth-aachen.de



2

2 Hendrik Thüs promoviert im Bereich „Mobile and Social Learning“ im gleichen Lehr- und Forschungsgebiet.

thues@cs.rwth-aachen.de