Carsten Ullrich, Martin Wessner (Eds.): Proceedings of DeLFI and GMW Workshops 2017
Chemnitz, Germany, September 5, 2017

Using Pseudonymous Identities to Support Learning Analytics in Open Learning Environments

René Röpke, Svenja Neitzel, Christoph Rensing

Abstract: Today, learners no longer use just one learning application but a variety of applications with very different functions. The use of different learning applications has a strong influence on Learning Analytics. It is necessary to collect data from different applications in order to obtain a holistic profile of a learner. This need forms a main motive of the xAPI specification and the establishment of centralized databases for learning data. Major challenges in this context are identity management and the preservation of learner's privacy. The creation of a collective learning data set requires an assignment of the identities of different applications. This contribution describes the concept of a pseudonymity provider as a new infrastructure component. An implementation of the concept and the use of pseudonyms in the learning management system Moodle and in a mobile learning application are presented. The strengths of the concept and its influence on Learning Analytics are demonstrated.

Carsten Ullrich, Martin Wessner (Eds.): Proceedings of DeLFI and GMW Workshops 2017
Chemnitz, Germany, September 5, 2017

Verwendung pseudonymer Identitäten zur Unterstützung von Learning Analytics in offenen Lernumgebungen

René Röpke[1], Svenja Neitzel[2], Christoph Rensing[3]

Abstract: In der heutigen Zeit verwenden Lernende nicht mehr nur eine Lernanwendung sondern eine Vielzahl von Anwendungen mit ganz unterschiedlichen Funktionen. Der Einsatz verschiedener Lernanwendungen hat einen starken Einfluss auf Learning Analytics. Es besteht die Notwendigkeit, Daten aus unterschiedlichen Anwendungen zu sammeln, um ein ganzheitliches Profil eines Lernenden zu erhalten. Diese Notwendigkeit bildet ein Hauptmotiv der xAPI-Spezifikation und die Einrichtung von zentralisierten Datenbanken für Lerndaten. Große Herausforderungen in diesem Zusammenhang sind das Identitätsmanagement und die Bewahrung der Privatheit der Lernenden. Die Erstellung eines gemeinsamen Lerndatensatzes erfordert eine Zuordnung der Identitäten von verschiedenen Anwendungen. In diesem Beitrag wird das Konzept eines Pseudonymity Providers als neue Infrastrukturkomponente beschrieben. Eine Implementierung des Konzepts und die Verwendung von Pseudonymen in dem Lernmanagementsystem Moodle und in einer mobilen Lernanwendung werden vorgestellt. Es werden die Stärken des Konzepts und der Einfluss auf Learning Analytics aufgezeigt

Keywords: Pseudonymität; Anonymität; xAPI; Learning Record Store; Learning Analytics

1 Einleitung

Das Potential von Learning Analytics für Lernende und Lehrende ist heutzutage weitgehend anerkannt. Es gibt jedoch viele offene Fragen und Forschungsinteressen auf diesem Gebiet. Eine Frage ergibt sich aus der Situation, dass die heutigen Lernenden während ihres Lernprozesses unterschiedliche Werkzeuge und Lernanwendungen verwenden. Neben einem Lernmanagementsystem (LMS) nutzen Lernende verschiedene Anwendungen auf dem PC und über mobile Endgeräte und greifen auf verschiedene Lernressourcen zu.
Im folgenden wird die individuelle Sammlung von Lernanwendungen, die von einem Lernenden verwendet wird als offene Lernumgebung (in Englisch: Open Learning Environment, kurz OLE) bezeichnet. Um Learning Analytics (wie definiert in [SP13]) in OLEs anzuwenden, müssen Daten aus den verschiedenen Anwendungen gesammelt werden. Nur Daten aus verschiedenen Anwendungen der OLE können ein umfassendes Profil des

[1] RWTH Aachen University, Informatik 9 (Learning Technologies), Ahornstraße 55, 52074 Aachen, Deutschland
roepke@informatik.rwth-aachen.de
[2] TU Darmstadt, Multimedia Communications Lab (KOM), Rundeturmstraße 10, 64283 Darmstadt, Deutschland
svenja.neitzel@kom.tu-darmstadt.de
[3] TU Darmstadt, Multimedia Communications Lab (KOM), Rundeturmstraße 10, 64283 Darmstadt, Deutschland
christoph.rensing@kom.tu-darmstadt.de

Lernenden bieten. Die xAPI-Spezifikation und Learning Record Stores (LRSs) sind beliebte Lösungsansätze zur zentralisierten Lerndatenspeicherung. Beides gewinnt mehr und mehr an Bedeutung in der Forschung und Praxis [KR16]. Mit xAPI werden Aktivitätsdaten durch drei Informationen erfasst: *Actor* beschreibt das Subjekt, welches die Aktivität durchführt (hier der Lernende). Das *Verb* beschreibt die Aktivität selbst und *Object* beschreibt das Objekt auf welches die Aktivität ausgeübt wird. Nach der Datensammlung und Speicherung folgt im nächsten Schritt die Datenanalyse und Datenvisualisierung. Dazu müssen die Daten wiederum anderen Anwendungen zur Verfügung gestellt werden.
Gegenüber dem Konzept der zentralisierten Lerndatenspeicherung gibt es einige Vorbehalte. Lernende sind besorgt wegen der Wahrung ihrer Privatsphäre [DG12]. Richtlinien für die Speicherung von Lerndaten stellen eine große Herausforderung dar [KR16]. Anonymisierung ist eine bekannte Methode um den Vorbehalten zu begegnen, doch erreicht diese ihre Grenzen, wenn der Lernende auf alle Daten zugreifen möchte, die über ihn gesammelt wurden oder wenn eine Anwendung die personalisierten Informationen benötigt, um entsprechend personalisierte, adaptive Dienste bereitzustellen.
Eine weitere Herausforderung besteht in der Notwendigkeit eindeutiger Nutzeridentitäten in der OLE. In der Regel verwenden heutige Anwendungen Nutzerkonten, die über E-Mail-Adressen oder andere Merkmale unterschieden werden. Die jeweiligen Identifikatoren werden als Actor-Elemente für xAPI-Anweisungen verwendet. Um einen Lernenden mit Learning Analytics auf Basis der Daten in der OLE zu unterstützen, wird eine Zuordnungen seiner Nutzerkonten benötigt. Ohne technische Unterstützung ist man auf eine manuelle Zuordnung durch die Lernenden angewiesen.
Der Beitrag ist wie folgt aufgebaut: Verwandte Arbeiten werden im Abschnitt 2 dargestellt. In Abschnitt 3 wird das vorgeschlagene Konzept mit einer detaillierten Beschreibung des Pseudonymity Providers und der Verwendung von Pseudonymen in Lernanwendungen erläutert. Details der Implementierung werden in Abschnitt 4 dargestellt. Abschnitt 5 schließt den Beitrag ab und gibt einen Ausblick.

2 Verwandte Arbeiten

Um den genannten Herausforderungen zu begegnen und um die Privatsphäre der Lernenden zu wahren, sind verschiedene Ansätze interessant. Hierzu werden Lösungen im Zusammenhang mit der Datenschutzfrage und des Identitätsmanagement vorgestellt.
Privatsphäre ist ein bedeutsames Thema, wenn es um Datensammlung und -speicherung geht [SP13]. Hierbei wird Privatsphäre kontextabhängig definiert [Gu16]. Im Kontext der offenen Lernumgebungen steht die datenschutzrechtliche Verwendung von Lerndaten im Fokus. Das Teilen von Lernerdaten in einem LRS und den Anwendungen, die daran angebunden sind, geschieht ohne Verletzung der Privatsphäre individueller Lerner.
Eine bekannte Maßnahme hierfür ist Anonymisierung. Im Anwendungsfall von Learning Analytics stellt Anonymisierung einen Kompromiss zwischen den Nutzungsmöglichkeiten der Daten und der Privatsphäre der Lernenden dar [Gu16]. Verschiedene Autoren

erwähnen die Notwendigkeit von Datenschutzerklärungsmechanismen in Learning Analytics [DG16, PS14]. [Gu16] bietet einen konkreten Ansatz für die Anonymisierung von Lernerdaten. Oftmals werden Transparenz und Zustimmung als Lösungskonzepte für die Herausforderung der Privatsphäre vorgeschlagen [SP13].
Eine vollständige Anonymisierung der Daten eines LRS macht es jedoch unmöglich, die Daten in personalisierten Funktionen und Anwendungen zu verwenden [Hi17]. Daher ist ein anderer Ansatz notwendig. Im folgenden wird das Konzept eines Pseudonymitätsanbieters als Alternative vorgestellt. Ein Pseudonym ist ein Identifikator eines Subjekts (hier eines Lerners), welcher abstrakt ist und keinen Rückschluss auf den Namen oder die wahre Identität ermöglicht.
In Lernanwendungen werden oftmals E-Mail-Adressen oder alternativ Nutzernamen verwendet um Nutzerkonten zu identifizieren. Beides repräsentiert dabei die Identität des Nutzers. Oftmals verwenden Nutzer unterschiedliche Identifikatoren in verschiedenen Anwendungen. Um die Daten im LRS an die Identität des Nutzers zu knüpfen wird eine Zuordnung von verschiedenen Anmeldeinformationen zu einer eindeutiger Identität des Nutzers benötigt. Auch aus der Sicht des Nutzers bedeutet die Verwendung von verschiedenen Anmeldeinformationen einen Mehraufwand, da er oder sie sich an alle verschiedenen Anmeldeinformationen erinnern muss. Das Konzept eines Single Sign-On Systems löst dieses Problem [Gr06]. Ein sehr vielversprechender Mechanismus zur Zuordnung der verschiedener Anmeldeinformationen eines Nutzers zu einer eindeutigen Identität ist OpenID. OpenID ist ein Open-Source-Authentifizierungsmechanismus, der eine dezentrale benutzerzentrische Identitätsverwaltung [RR06] bietet. Die Verwendung von OpenID ist auch eine mögliche Lösung für das Identitätsmanagement in offenen Lernumgebungen, welche voraussetzt, dass alle Anwendungen OpenID als Authentifizierungsmechanismus implementieren [Al11].

3 Konzept

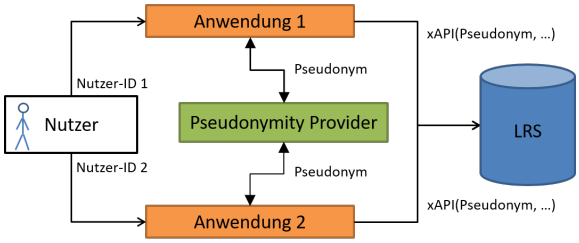

Abb. 1: Grundkonzept

Das Ziel des im folgenden vorgestellten Konzepts ist es, die Verwendung eines eindeutigen Identifikators zu ermöglichen, um den *Actor* in xAPI-Anweisungen zu kennzeichnen und die Privatsphäre der Lernenden zu wahren. So können alle Daten über einen Lernenden mit dem gleichen Bezeichner identifiziert werden und die kohärente Datensammlung über mehrere Anwendungen wird ermöglicht. Der LRS und die daran angebundenen Anwendungen

können nicht auf die echte Nutzeridentität zugreifen, aber dennoch personalisierte und adaptive Dienste unter Verwendung der Daten anbieten.
Die Grundidee besteht darin, Pseudonyme als eindeutige Identifikatoren für Lernende in den xAPI-Anweisungen zu verwenden. Diese Pseudonyme werden den unterschiedlichen Anwendungen unter Verwendung einer neuen Infrastrukturkomponente, dem Pseudonymity Provider, zur Verfügung gestellt. Der Lernende muss dabei aktiv der Bereitstellung des Pseudonyms für eine Lernanmeldung zustimmen und behält so Kontrolle darüber, welche Anwendung pseudonymisiert Daten sammeln kann. Abb. 1 veranschaulicht das Konzept.

3.1 Pseudonymity Provider

Der Pseudonymity Provider ist ein Service Provider, der als Identitätsanbieter fungiert. Er bietet registrierten Anwendungen den Zugriff auf pseudonyme Nutzeridentitäten. Im ersten Schritt muss jede Anwendung, die Pseudonyme beziehen möchte, beim Pseudonymity Provider registriert werden. Dadurch wird eine gültige und sichere Kommunikation zwischen den Anwendungen und dem Pseudonymity Provider sichergestellt. Der Pseudonymity Provider agiert als (dritte) vertrauenswürdige Partei neben den Anwendungen und dem Nutzer.
Authentifizierung ist einer der Schlüsselmechanismen, um die Identität des Nutzer zu schützen. Nur authentifizierte Nutzer haben Zugriff auf ihr Pseudonym. Um die Möglichkeiten der Authentifizierung zu erleichtern, können zusätzlich Authentifizierungsmechanismen externer Dienste eingebunden werden (z. B. Google SignIn, Twitter oder Facebook). Nutzer brauchen dann ein Konto bei dem entsprechenden Dienst, um sich zu authentifizieren.
Anwendungen, die bei einem Pseudonymity Provider registriert sind, können mittels Weiterleitung oder REST Schnittstelle Pseudonyme für Nutzer anfordern. Dabei wird vorausgesetzt, dass der Nutzer authentifiziert ist bzw. sich im Laufe der Weiterleitung anmeldet. Erfolgreiche Anfragen werden mit dem anfragten Pseudonym beantwortet. Fehlercodes werden zusätzlich genutzt, um die Antwort verarbeiten zu können.

3.2 Pseudonymitätsintegration in offenen Lernumgebungen

Mit der Verfügbarkeit des Pseudonymity Providers und der Pseudonyme besteht der nächste Schritt darin, die Pseudonyme in Lernanwendungen zu integrieren: Um ein Pseudonym in einer Lernanwendung zu registrieren, müssen Pseudonymitätsmanagement-Funktionen implementiert werden. Bei der Registrierung eines Pseudonyms sollte ein Nutzer an den Pseudonymity Provider weitergeleitet werden. Nur eine erfolgreiche Antwort auf die Anfrage schließt die Pseudonymregistrierung ab. Bei Time-outs, ungültigen Anfragen oder fehlerhaften Antworten wird der Registrierungsprozess abgebrochen. Kein Nutzer sollte in der Lage sein, ein Pseudonym für einen anderen Benutzer zu registrieren oder zu löschen. Pseudonyme können auch nicht zwischen Benutzern geteilt werden. Eine Lernanwendung sollte es einem Nutzer erlauben, jederzeit ein Pseudonym zu registrieren oder zu entfernen.

Sobald das Pseudonym der Lernanwendung zur Verfügung steht, sollte es als *Actor*-Information in den xAPI-Anweisungen verwendet werden. Welche Daten gesammelt werden und einem zentralen LRS gespeichert werden, ist abhängig von der Zustimmung des Nutzers. Die Frage nach Zustimmung ermöglicht es Nutzern, ein Pseudonym zu registrieren, aber dessen Verwendung innerhalb der Lernanwendung zu beschränken. Neben der Pseudonymität-Managementfunktionen und der Zustimmungsbehandlung sollte eine Konfigurationsschnittstelle zur Einrichtung aller erforderlichen Sicherheitsparameter und Eigenschaften für Administratoren bereitgestellt werden.

4 Implementierung

4.1 Implementierung des Pseudonymity Providers

Der Pseudonymity Provider, *PseudoProvider*, wurde als klassische Webanwendung implementiert[4]. Er bietet eine Nutzerauthentifizierung und die Konfiguration von registrierten Anwendungen. Für die Registrierung von Anwendungen spezifiziert der Nutzer den Namen der Anwendung, die URL für Pseudonymanfragen sowie entsprechende Sicherheitsparameter, sodass eine gesicherte Verbindung zwischen der registrierten Anwendung und dem Pseudonymity Provider aufgebaut werden kann. Nach erstmaliger Anmeldung des Nutzers beim Pseudonymity Provider wird ein Pseudonym für den Nutzer generiert. Dieses Pseudonym ist einzigartig in der Domäne des Pseudonymity Providers und ist damit eindeutig in allen Anwendungen, die Pseudonyme von dem selben Pseudonymity Provider beziehen.
Abb. 2 zeigt eine registrierte Moodle-Instanz. Die Details zur Installation werden nur dem Administrator der Lernanwendung angezeigt.

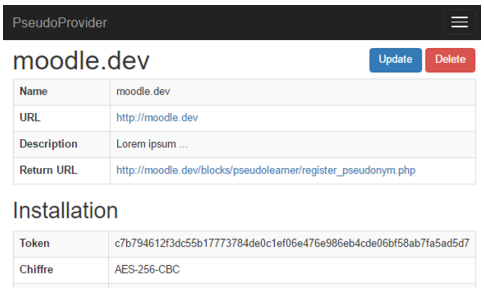

Abb. 2: Details zu einer registrierten Anwendung.

[4] https://github.com/rroepke/pseudoprovider

4.2 Verwendung des Pseudonyms in Moodle

Um Pseudonymität und die Zustimmung zur zentralisierten Datenspeicherung in Moodle zu ermöglichen, wurde das *PseudoLearner* Plugin für das LMS implementiert[5]. Nach Installation des Plugins muss es entsprechend mit den Installationsdetails (siehe Abb. 2) konfiguriert werden. Auf diese Weise können Pseudonymanfragen sicher an den Pseudonymity Provider übertragen werden. Das Plugin implementiert ein Blockelement, das Lehrenden zu einem Kurs hinzugefügt werden kann. Alle in den Kurs eingeschriebenen Nutzer können auf die Plugin-Funktionalität zugreifen (d.h. ein Pseudonym registrieren und pseudonymisierter Datenspeicherung zustimmen oder widersprechen).
Um ein Pseudonym vom Pseudonymity Provider zu erhalten, muss der Nutzer sich gegenüber dem Pseudonymity Provider authentifizieren. Nach erfolgreicher Authentifizierung kann der Nutzer der Anfrage zustimmen. Das Pseudonym wird dann an die Anwendung gesendet, von der die Anfrage gesendet wurde. Durch die gesichterte Verbindung ist garantiert, dass das Pseudonym nur für die Anwendung lesbar ist. Die sichere Verwahrung des Pseudonyms in der Lernanwendung wird vorausgesetzt. In Abb. 3 ist der Plugin-Inhalt nach erfolgreicher Registrierung des Pseudonyms abgebildet. Nutzer können als nächstes der pseudonymisierten Datenspeicherung zustimmen.

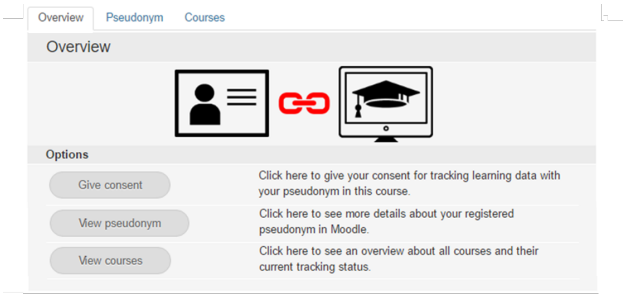

Abb. 3: Plugin-Inhalt nach Registrierung eines Pseudonyms.

4.3 Verwendung des Pseudonyms in mobilen Lernanwendungen

Smartphones sind heute ein ständiger Begleiter ihrer Nutzer und werden auch in Lernsituationen verwendet. Mobile Lernanwendungen können somit Informationen über Lernaktivitäten an einen LRS senden. Sie können bspw. auch Kontextinformationen (z.B. das Betreten der Bibliothek) erfassen. Durch die Entwicklung eines Moduls zur Anbindung von Pseudonymity Provider in Lernanwendungen für das mobile Betriebssystem Android, können auch in diesen Fällen Pseudonyme zur Beschreibung des *Actor*-Elements genutzt werden.
Der Entwickler registriert zunächst die mobile Anwendung beim Pseudonymity Provider

[5] https://github.com/rroepke/moodle-blocks-pseudolearner

und hinterlegt dessen URL und die ausgehandelten Sicherheitsparameter in der Anwendung. Will der Nutzer nun sein Pseudonym in der Anwendung registrieren, wird durch das implementierte Modul ein Login-Dialog für die Eingabe der Zugangsdaten zum Pseudonymity Provider zur Verfügung gestellt. Die Anwendung selbst erhält dabei keinen Zugriff auf die eingegebenen Daten. Nach erfolgreicher Authentifizierung gegenüber dem Pseudonymity Provider wird das Pseudonym der Anwendung zur Verfügung gestellt.
Mobile Anwendungen eignen sich aufgrund ihrer Allgegenwärtigkeit auch gut zur Schaffung von Awareness und für Interventionen, die das Lernverhalten positiv beeinflussen können. Eine an der TU Darmstadt entwickelte Lerntagebuch-App (siehe Abb. 4) visualisiert Lernaktivitäten seiner Nutzer [NR17]. Die Lernaktivitäten können im Lerntagebuch manuell eingetragen werden und werden entsprechend in einem LRS gespeichert. Gleichzeitig können Aktivitäten aus anderen Lernanwendungen (z. B. aus Moodle) visualisiert werden. In dieser Anwendung werden die pseudonymisierten Daten aus anderen System verwendet, aber auch neue Daten im LRS pseudonymisiert gespeichert.

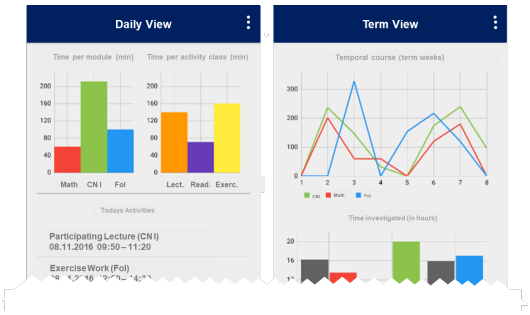

Abb. 4: Ausschnitt der Lerntagebuch-Anwendung mit Visualisierungen

5 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Konzept für Identitätsmanagement und Datenschutz präsentiert, um Learning Analytics in offenen Lernumgebungen zu verbessern. Ein Pseudonymity Provider ermöglicht es, Daten über Lernende in verschiedenen Anwendungen zu sammeln und ihre Nutzeridentitäten zu vereinen. Darüber hinaus können mit Verwendung des Pseudonyms personalisierte Anwendungen und Funktionen auf Basis der Daten implementiert und so Lernende gezielter unterstützt werden. Durch Pseudonymisierung wird nicht nur die Identität des Lernenden geschützt, Daten verschiedener Anwendungen können außerdem einer pseudonymen Identität zugeordnet werden.
Die Implementierungsdetails und vorgestellten Anwendungsfälle verdeutlichen die Stärken des Konzepts. Es wurde auch ein erster Ansatz zur Zustimmungsbehandlung in Lernanwendungen vorgestellt. Im Weiteren ist anvisiert ein ausgereifteres Zustimmungskonzept zu implementieren, um den Lernenden stärker zu informieren und einzubinden, z. B. über eine

manuelle Auswahl welche Daten gesammelt werden dürfen. Zusätzliche Möglichkeiten unter Verwendung des Konzepts liegen in elektronischen Befragungstools. Pseudonyme würden dort die bekannten Teilnehmercodes (Kombination aus Initialen und Geburtsdaten) ersetzen und eine eindeutige Zuordnung bei Wahrung der Privatsphäre der Lernenden ermöglichen. Es ist zudem interessant die Skalierbarkeit und Effektivität pseudonymisierter Identitäten in offenen Lernumgebungen zu untersuchen. Der Einfluss auf personalisiertes Feedback ist vielversprechend, da ein besseres Profil der Lernenden erfasst werden kann.

Literaturverzeichnis

[Al11] Alecu, Felician; Pocatilu, Paul; Stoica, George; Ciurea, Cristian; Capisizu, Sergiu: OpenID, a Single Sign-On Solution for E-learning Applications. Journal of Mobile, Embedded and Distributed Systems, 3(3):136–141, 2011.

[DG12] Drachsler, Hendrik; Greller, Wolfgang: The pulse of learning analytics understandings and expectations from the stakeholders. In: Proceedings of the 2nd international conference on learning analytics and knowledge. ACM, S. 120–129, 2012.

[DG16] Drachsler, Hendrik; Greller, Wolfgang: Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. ACM, S. 89–98, 2016.

[Gr06] Grandcolas, Michael L; Law, France; Doshi, Ashwin; Williams, Michael; Jang, Yeona; Merschen, Toni; Pan, Jack: , Method and system for single sign-on user access to multiple web servers, November 14 2006. US Patent 7,137,006.

[Gu16] Gursoy, Mehmet; Inan, Ali; Nergiz, Mehmet Ercan; Saygin, Yucel: Privacy-Preserving Learning Analytics: Challenges and Techniques. IEEE Transactions on Learning Technologies, 2016.

[Hi17] Hildebrandt, Mireille: Learning as a Machine. Crossovers Between Humans and Machines. Journal of Learning Analytics, 4(1):6–23, 2017.

[KR16] Kevan, Jonathan M; Ryan, Paul R: Experience API: Flexible, decentralized and activity-centric data collection. Technology, Knowledge and Learning, 21(1):143–149, 2016.

[NR17] Neitzel, Svenja; Rensing, Christoph: Automatische Sammlung von Aktivitäten Lernender in offenen Lernumgebungen und deren Nutzung in einer Lerntagebuchanwendung. In: Proceedings der 16. e-Learning Fachtagung Informatik (DeLFI 2017, accepted for publication). Springer, 2017.

[PS14] Pardo, Abelardo; Siemens, George: Ethical and privacy principles for learning analytics. British Journal of Educational Technology, 45(3):438–450, 2014.

[RR06] Recordon, David; Reed, Drummond: OpenID 2.0: a platform for user-centric identity management. In: Proceedings of the second ACM workshop on Digital identity management. ACM, S. 11–16, 2006.

[SP13] Slade, Sharon; Prinsloo, Paul: Learning analytics ethical issues and dilemmas. American Behavioral Scientist, 57(10):1510–1529, 2013.